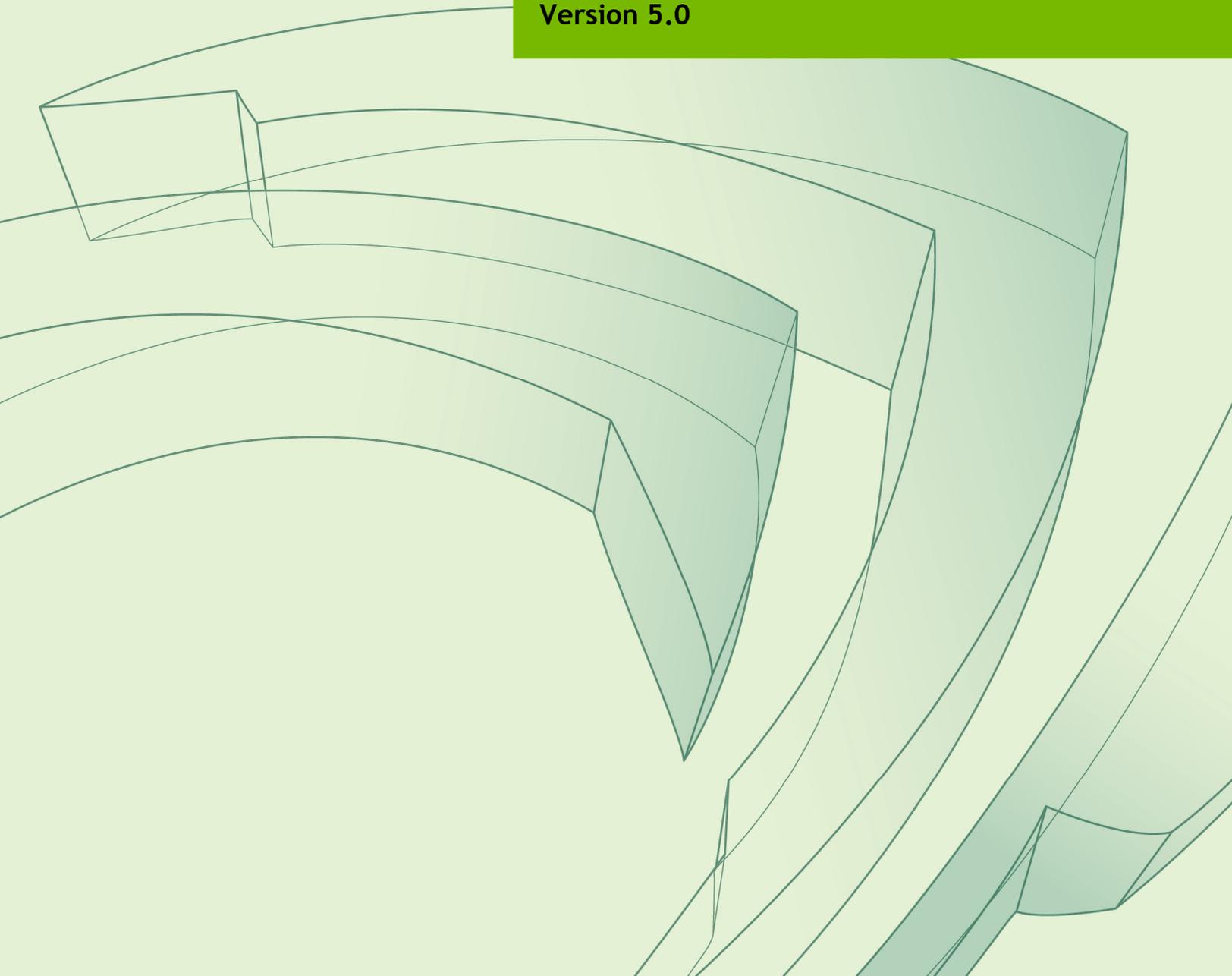




CUDA API REFERENCE MANUAL

October 2012

Version 5.0



Contents

1	API synchronization behavior	1
1.1	Memcpy	1
1.1.1	Synchronous	1
1.1.2	Asynchronous	1
1.2	Memset	2
1.3	Kernel Launches	2
2	Deprecated List	3
3	Module Index	11
3.1	Modules	11
4	Data Structure Index	13
4.1	Data Structures	13
5	Module Documentation	15
5.1	Device Management	15
5.1.1	Detailed Description	16
5.1.2	Function Documentation	16
5.1.2.1	cudaChooseDevice	16
5.1.2.2	cudaDeviceGetAttribute	17
5.1.2.3	cudaDeviceGetByPCIBusId	20
5.1.2.4	cudaDeviceGetCacheConfig	20
5.1.2.5	cudaDeviceGetLimit	21
5.1.2.6	cudaDeviceGetPCIBusId	21
5.1.2.7	cudaDeviceGetSharedMemConfig	22
5.1.2.8	cudaDeviceReset	22
5.1.2.9	cudaDeviceSetCacheConfig	23
5.1.2.10	cudaDeviceSetLimit	23
5.1.2.11	cudaDeviceSetSharedMemConfig	24

5.1.2.12	cudaDeviceSynchronize	25
5.1.2.13	cudaGetDevice	25
5.1.2.14	cudaGetDeviceCount	26
5.1.2.15	cudaGetDeviceProperties	26
5.1.2.16	cudaIpcCloseMemHandle	30
5.1.2.17	cudaIpcGetEventHandle	30
5.1.2.18	cudaIpcGetMemHandle	30
5.1.2.19	cudaIpcOpenEventHandle	31
5.1.2.20	cudaIpcOpenMemHandle	31
5.1.2.21	cudaSetDevice	32
5.1.2.22	cudaSetDeviceFlags	33
5.1.2.23	cudaSetValidDevices	34
5.2	Thread Management [DEPRECATED]	35
5.2.1	Detailed Description	35
5.2.2	Function Documentation	35
5.2.2.1	cudaThreadExit	35
5.2.2.2	cudaThreadGetCacheConfig	36
5.2.2.3	cudaThreadGetLimit	36
5.2.2.4	cudaThreadSetCacheConfig	37
5.2.2.5	cudaThreadSetLimit	38
5.2.2.6	cudaThreadSynchronize	39
5.3	Error Handling	40
5.3.1	Detailed Description	40
5.3.2	Function Documentation	40
5.3.2.1	cudaGetErrorString	40
5.3.2.2	cudaGetLastError	40
5.3.2.3	cudaPeekAtLastError	41
5.4	Stream Management	42
5.4.1	Detailed Description	42
5.4.2	Typedef Documentation	42
5.4.2.1	cudaStreamCallback_t	42
5.4.3	Function Documentation	43
5.4.3.1	cudaStreamAddCallback	43
5.4.3.2	cudaStreamCreate	43
5.4.3.3	cudaStreamCreateWithFlags	44
5.4.3.4	cudaStreamDestroy	44
5.4.3.5	cudaStreamQuery	45

5.4.3.6	cudaStreamSynchronize	45
5.4.3.7	cudaStreamWaitEvent	45
5.5	Event Management	47
5.5.1	Detailed Description	47
5.5.2	Function Documentation	47
5.5.2.1	cudaEventCreate	47
5.5.2.2	cudaEventCreateWithFlags	48
5.5.2.3	cudaEventDestroy	48
5.5.2.4	cudaEventElapsedTime	49
5.5.2.5	cudaEventQuery	49
5.5.2.6	cudaEventRecord	50
5.5.2.7	cudaEventSynchronize	50
5.6	Execution Control	51
5.6.1	Detailed Description	51
5.6.2	Function Documentation	51
5.6.2.1	cudaConfigureCall	51
5.6.2.2	cudaFuncGetAttributes	52
5.6.2.3	cudaFuncSetCacheConfig	52
5.6.2.4	cudaFuncSetSharedMemConfig	53
5.6.2.5	cudaLaunch	54
5.6.2.6	cudaSetDoubleForDevice	54
5.6.2.7	cudaSetDoubleForHost	55
5.6.2.8	cudaSetupArgument	55
5.7	Memory Management	56
5.7.1	Detailed Description	59
5.7.2	Function Documentation	60
5.7.2.1	cudaArrayGetInfo	60
5.7.2.2	cudaFree	60
5.7.2.3	cudaFreeArray	60
5.7.2.4	cudaFreeHost	61
5.7.2.5	cudaFreeMipmappedArray	61
5.7.2.6	cudaGetMipmappedArrayLevel	62
5.7.2.7	cudaGetSymbolAddress	62
5.7.2.8	cudaGetSymbolSize	63
5.7.2.9	cudaHostAlloc	63
5.7.2.10	cudaHostGetDevicePointer	64
5.7.2.11	cudaHostGetFlags	65

5.7.2.12	<code>cudaHostRegister</code>	65
5.7.2.13	<code>cudaHostUnregister</code>	66
5.7.2.14	<code>cudaMalloc</code>	66
5.7.2.15	<code>cudaMalloc3D</code>	67
5.7.2.16	<code>cudaMalloc3DArray</code>	67
5.7.2.17	<code>cudaMallocArray</code>	69
5.7.2.18	<code>cudaMallocHost</code>	70
5.7.2.19	<code>cudaMallocMipmappedArray</code>	70
5.7.2.20	<code>cudaMallocPitch</code>	72
5.7.2.21	<code>cudaMemcpy</code>	73
5.7.2.22	<code>cudaMemcpy2D</code>	73
5.7.2.23	<code>cudaMemcpy2DArrayToArray</code>	74
5.7.2.24	<code>cudaMemcpy2DAsync</code>	75
5.7.2.25	<code>cudaMemcpy2DFromArray</code>	75
5.7.2.26	<code>cudaMemcpy2DFromArrayAsync</code>	76
5.7.2.27	<code>cudaMemcpy2DToArray</code>	77
5.7.2.28	<code>cudaMemcpy2DToArrayAsync</code>	78
5.7.2.29	<code>cudaMemcpy3D</code>	79
5.7.2.30	<code>cudaMemcpy3DAsync</code>	80
5.7.2.31	<code>cudaMemcpy3DPeer</code>	81
5.7.2.32	<code>cudaMemcpy3DPeerAsync</code>	82
5.7.2.33	<code>cudaMemcpyArrayToArray</code>	82
5.7.2.34	<code>cudaMemcpyAsync</code>	83
5.7.2.35	<code>cudaMemcpyFromArray</code>	84
5.7.2.36	<code>cudaMemcpyFromArrayAsync</code>	84
5.7.2.37	<code>cudaMemcpyFromSymbol</code>	85
5.7.2.38	<code>cudaMemcpyFromSymbolAsync</code>	86
5.7.2.39	<code>cudaMemcpyPeer</code>	86
5.7.2.40	<code>cudaMemcpyPeerAsync</code>	87
5.7.2.41	<code>cudaMemcpyToArray</code>	88
5.7.2.42	<code>cudaMemcpyToArrayAsync</code>	88
5.7.2.43	<code>cudaMemcpyToSymbol</code>	89
5.7.2.44	<code>cudaMemcpyToSymbolAsync</code>	90
5.7.2.45	<code>cudaMemGetInfo</code>	90
5.7.2.46	<code>cudaMemset</code>	91
5.7.2.47	<code>cudaMemset2D</code>	91
5.7.2.48	<code>cudaMemset2DAsync</code>	92

5.7.2.49	cudaMemset3D	92
5.7.2.50	cudaMemset3DAsync	93
5.7.2.51	cudaMemsetAsync	94
5.7.2.52	make_cudaExtent	94
5.7.2.53	make_cudaPitchedPtr	94
5.7.2.54	make_cudaPos	95
5.8	Unified Addressing	96
5.8.1	Detailed Description	96
5.8.2	Overview	96
5.8.3	Supported Platforms	96
5.8.4	Looking Up Information from Pointer Values	96
5.8.5	Automatic Mapping of Host Allocated Host Memory	96
5.8.6	Direct Access of Peer Memory	97
5.8.7	Exceptions, Disjoint Addressing	97
5.8.8	Function Documentation	97
5.8.8.1	cudaPointerGetAttributes	97
5.9	Peer Device Memory Access	99
5.9.1	Detailed Description	99
5.9.2	Function Documentation	99
5.9.2.1	cudaDeviceCanAccessPeer	99
5.9.2.2	cudaDeviceDisablePeerAccess	99
5.9.2.3	cudaDeviceEnablePeerAccess	100
5.10	OpenGL Interoperability	101
5.10.1	Detailed Description	101
5.10.2	Enumeration Type Documentation	101
5.10.2.1	cudaGLDeviceList	101
5.10.3	Function Documentation	102
5.10.3.1	cudaGLGetDevices	102
5.10.3.2	cudaGraphicsGLRegisterBuffer	102
5.10.3.3	cudaGraphicsGLRegisterImage	103
5.10.3.4	cudaWGLGetDevice	104
5.11	OpenGL Interoperability [DEPRECATED]	105
5.11.1	Detailed Description	105
5.11.2	Enumeration Type Documentation	105
5.11.2.1	cudaGLMapFlags	105
5.11.3	Function Documentation	106
5.11.3.1	cudaGLMapBufferObject	106

5.11.3.2	<code>cudaGLMapBufferObjectAsync</code>	106
5.11.3.3	<code>cudaGLRegisterBufferObject</code>	107
5.11.3.4	<code>cudaGLSetBufferObjectMapFlags</code>	107
5.11.3.5	<code>cudaGLSetGLDevice</code>	108
5.11.3.6	<code>cudaGLUnmapBufferObject</code>	108
5.11.3.7	<code>cudaGLUnmapBufferObjectAsync</code>	109
5.11.3.8	<code>cudaGLUnregisterBufferObject</code>	110
5.12	Direct3D 9 Interoperability	111
5.12.1	Detailed Description	111
5.12.2	Enumeration Type Documentation	111
5.12.2.1	<code>cudaD3D9DeviceList</code>	111
5.12.3	Function Documentation	112
5.12.3.1	<code>cudaD3D9GetDevice</code>	112
5.12.3.2	<code>cudaD3D9GetDevices</code>	112
5.12.3.3	<code>cudaD3D9GetDirect3DDevice</code>	113
5.12.3.4	<code>cudaD3D9SetDirect3DDevice</code>	113
5.12.3.5	<code>cudaGraphicsD3D9RegisterResource</code>	114
5.13	Direct3D 9 Interoperability [DEPRECATED]	116
5.13.1	Detailed Description	117
5.13.2	Enumeration Type Documentation	117
5.13.2.1	<code>cudaD3D9MapFlags</code>	117
5.13.2.2	<code>cudaD3D9RegisterFlags</code>	117
5.13.3	Function Documentation	117
5.13.3.1	<code>cudaD3D9MapResources</code>	117
5.13.3.2	<code>cudaD3D9RegisterResource</code>	118
5.13.3.3	<code>cudaD3D9ResourceGetMappedArray</code>	119
5.13.3.4	<code>cudaD3D9ResourceGetMappedPitch</code>	120
5.13.3.5	<code>cudaD3D9ResourceGetMappedPointer</code>	120
5.13.3.6	<code>cudaD3D9ResourceGetMappedSize</code>	121
5.13.3.7	<code>cudaD3D9ResourceGetSurfaceDimensions</code>	122
5.13.3.8	<code>cudaD3D9ResourceSetMapFlags</code>	123
5.13.3.9	<code>cudaD3D9UnmapResources</code>	123
5.13.3.10	<code>cudaD3D9UnregisterResource</code>	124
5.14	Direct3D 10 Interoperability	125
5.14.1	Detailed Description	125
5.14.2	Enumeration Type Documentation	125
5.14.2.1	<code>cudaD3D10DeviceList</code>	125

5.14.3	Function Documentation	125
5.14.3.1	cudaD3D10GetDevice	125
5.14.3.2	cudaD3D10GetDevices	126
5.14.3.3	cudaGraphicsD3D10RegisterResource	127
5.15	Direct3D 10 Interoperability [DEPRECATED]	129
5.15.1	Detailed Description	130
5.15.2	Enumeration Type Documentation	130
5.15.2.1	cudaD3D10MapFlags	130
5.15.2.2	cudaD3D10RegisterFlags	130
5.15.3	Function Documentation	130
5.15.3.1	cudaD3D10GetDirect3DDevice	130
5.15.3.2	cudaD3D10MapResources	131
5.15.3.3	cudaD3D10RegisterResource	131
5.15.3.4	cudaD3D10ResourceGetMappedArray	133
5.15.3.5	cudaD3D10ResourceGetMappedPitch	133
5.15.3.6	cudaD3D10ResourceGetMappedPointer	134
5.15.3.7	cudaD3D10ResourceGetMappedSize	135
5.15.3.8	cudaD3D10ResourceGetSurfaceDimensions	135
5.15.3.9	cudaD3D10ResourceSetMapFlags	136
5.15.3.10	cudaD3D10SetDirect3DDevice	137
5.15.3.11	cudaD3D10UnmapResources	137
5.15.3.12	cudaD3D10UnregisterResource	138
5.16	Direct3D 11 Interoperability	139
5.16.1	Detailed Description	139
5.16.2	Enumeration Type Documentation	139
5.16.2.1	cudaD3D11DeviceList	139
5.16.3	Function Documentation	139
5.16.3.1	cudaD3D11GetDevice	139
5.16.3.2	cudaD3D11GetDevices	140
5.16.3.3	cudaGraphicsD3D11RegisterResource	141
5.17	Direct3D 11 Interoperability [DEPRECATED]	143
5.17.1	Detailed Description	143
5.17.2	Function Documentation	143
5.17.2.1	cudaD3D11GetDirect3DDevice	143
5.17.2.2	cudaD3D11SetDirect3DDevice	143
5.18	VDPAU Interoperability	145
5.18.1	Detailed Description	145

5.18.2	Function Documentation	145
5.18.2.1	cudaGraphicsVDPAURegisterOutputSurface	145
5.18.2.2	cudaGraphicsVDPAURegisterVideoSurface	146
5.18.2.3	cudaVDPAUGetDevice	146
5.18.2.4	cudaVDPAUSetVDPAUDevice	147
5.19	Graphics Interoperability	148
5.19.1	Detailed Description	148
5.19.2	Function Documentation	148
5.19.2.1	cudaGraphicsMapResources	148
5.19.2.2	cudaGraphicsResourceGetMappedMipmappedArray	149
5.19.2.3	cudaGraphicsResourceGetMappedPointer	149
5.19.2.4	cudaGraphicsResourceSetMapFlags	150
5.19.2.5	cudaGraphicsSubResourceGetMappedArray	151
5.19.2.6	cudaGraphicsUnmapResources	151
5.19.2.7	cudaGraphicsUnregisterResource	152
5.20	Texture Reference Management	153
5.20.1	Detailed Description	153
5.20.2	Function Documentation	153
5.20.2.1	cudaBindTexture	153
5.20.2.2	cudaBindTexture2D	154
5.20.2.3	cudaBindTextureToArray	155
5.20.2.4	cudaBindTextureToMipmappedArray	155
5.20.2.5	cudaCreateChannelDesc	156
5.20.2.6	cudaGetChannelDesc	157
5.20.2.7	cudaGetTextureAlignmentOffset	157
5.20.2.8	cudaGetTextureReference	157
5.20.2.9	cudaUnbindTexture	158
5.21	Surface Reference Management	159
5.21.1	Detailed Description	159
5.21.2	Function Documentation	159
5.21.2.1	cudaBindSurfaceToArray	159
5.21.2.2	cudaGetSurfaceReference	159
5.22	Texture Object Management	161
5.22.1	Detailed Description	161
5.22.2	Function Documentation	161
5.22.2.1	cudaCreateTextureObject	161
5.22.2.2	cudaDestroyTextureObject	165

5.22.2.3	cudaGetTextureObjectResourceDesc	165
5.22.2.4	cudaGetTextureObjectResourceViewDesc	166
5.22.2.5	cudaGetTextureObjectTextureDesc	166
5.23	Surface Object Management	167
5.23.1	Detailed Description	167
5.23.2	Function Documentation	167
5.23.2.1	cudaCreateSurfaceObject	167
5.23.2.2	cudaDestroySurfaceObject	167
5.23.2.3	cudaGetSurfaceObjectResourceDesc	168
5.24	Version Management	169
5.24.1	Function Documentation	169
5.24.1.1	cudaDriverGetVersion	169
5.24.1.2	cudaRuntimeGetVersion	169
5.25	C++ API Routines	170
5.25.1	Detailed Description	172
5.25.2	Function Documentation	172
5.25.2.1	cudaBindSurfaceToArray	172
5.25.2.2	cudaBindSurfaceToArray	172
5.25.2.3	cudaBindTexture	173
5.25.2.4	cudaBindTexture	173
5.25.2.5	cudaBindTexture2D	174
5.25.2.6	cudaBindTexture2D	175
5.25.2.7	cudaBindTextureToArray	175
5.25.2.8	cudaBindTextureToArray	176
5.25.2.9	cudaBindTextureToMipmappedArray	176
5.25.2.10	cudaBindTextureToMipmappedArray	177
5.25.2.11	cudaCreateChannelDesc	177
5.25.2.12	cudaEventCreate	178
5.25.2.13	cudaFuncGetAttributes	179
5.25.2.14	cudaFuncSetCacheConfig	179
5.25.2.15	cudaGetSymbolAddress	180
5.25.2.16	cudaGetSymbolSize	180
5.25.2.17	cudaGetTextureAlignmentOffset	181
5.25.2.18	cudaLaunch	181
5.25.2.19	cudaMallocHost	182
5.25.2.20	cudaMemcpyFromSymbol	183
5.25.2.21	cudaMemcpyFromSymbolAsync	183

5.25.2.22	cudaMemcpyToSymbol	184
5.25.2.23	cudaMemcpyToSymbolAsync	185
5.25.2.24	cudaSetupArgument	185
5.25.2.25	cudaUnbindTexture	186
5.26	Interactions with the CUDA Driver API	187
5.26.1	Primary Contexts	187
5.26.2	Initialization and Tear-Down	187
5.26.3	Context Interoperability	187
5.26.4	Interactions between CUstream and cudaStream_t	188
5.26.5	Interactions between CUevent and cudaEvent_t	188
5.26.6	Interactions between CUarray and cudaArray_t	188
5.26.7	Interactions between CUgraphicsResource and cudaGraphicsResource_t	188
5.27	Profiler Control	189
5.27.1	Detailed Description	189
5.27.2	Function Documentation	189
5.27.2.1	cudaProfilerInitialize	189
5.27.2.2	cudaProfilerStart	190
5.27.2.3	cudaProfilerStop	190
5.28	Data types used by CUDA Runtime	191
5.28.1	Define Documentation	198
5.28.1.1	CUDA_IPC_HANDLE_SIZE	198
5.28.1.2	cudaArrayCubemap	198
5.28.1.3	cudaArrayDefault	198
5.28.1.4	cudaArrayLayered	199
5.28.1.5	cudaArraySurfaceLoadStore	199
5.28.1.6	cudaArrayTextureGather	199
5.28.1.7	cudaDeviceBlockingSync	199
5.28.1.8	cudaDeviceLmemResizeToMax	199
5.28.1.9	cudaDeviceMapHost	199
5.28.1.10	cudaDeviceMask	199
5.28.1.11	cudaDevicePropDontCare	199
5.28.1.12	cudaDeviceScheduleAuto	199
5.28.1.13	cudaDeviceScheduleBlockingSync	199
5.28.1.14	cudaDeviceScheduleMask	200
5.28.1.15	cudaDeviceScheduleSpin	200
5.28.1.16	cudaDeviceScheduleYield	200
5.28.1.17	cudaEventBlockingSync	200

5.28.1.18	<code>cudaEventDefault</code>	200
5.28.1.19	<code>cudaEventDisableTiming</code>	200
5.28.1.20	<code>cudaEventInterprocess</code>	200
5.28.1.21	<code>cudaHostAllocDefault</code>	200
5.28.1.22	<code>cudaHostAllocMapped</code>	200
5.28.1.23	<code>cudaHostAllocPortable</code>	200
5.28.1.24	<code>cudaHostAllocWriteCombined</code>	200
5.28.1.25	<code>cudaHostRegisterDefault</code>	200
5.28.1.26	<code>cudaHostRegisterMapped</code>	201
5.28.1.27	<code>cudaHostRegisterPortable</code>	201
5.28.1.28	<code>cudaIpcMemLazyEnablePeerAccess</code>	201
5.28.1.29	<code>cudaPeerAccessDefault</code>	201
5.28.1.30	<code>cudaStreamDefault</code>	201
5.28.1.31	<code>cudaStreamNonBlocking</code>	201
5.28.2	Typedef Documentation	201
5.28.2.1	<code>cudaArray_const_t</code>	201
5.28.2.2	<code>cudaArray_t</code>	201
5.28.2.3	<code>cudaError_t</code>	201
5.28.2.4	<code>cudaEvent_t</code>	201
5.28.2.5	<code>cudaGraphicsResource_t</code>	201
5.28.2.6	<code>cudaIpcEventHandle_t</code>	202
5.28.2.7	<code>cudaIpcMemHandle_t</code>	202
5.28.2.8	<code>cudaMipmappedArray_const_t</code>	202
5.28.2.9	<code>cudaMipmappedArray_t</code>	202
5.28.2.10	<code>cudaOutputMode_t</code>	202
5.28.2.11	<code>cudaStream_t</code>	202
5.28.2.12	<code>cudaSurfaceObject_t</code>	202
5.28.2.13	<code>cudaTextureObject_t</code>	202
5.28.2.14	<code>cudaUUID_t</code>	202
5.28.3	Enumeration Type Documentation	202
5.28.3.1	<code>cudaChannelFormatKind</code>	202
5.28.3.2	<code>cudaComputeMode</code>	203
5.28.3.3	<code>cudaDeviceAttr</code>	203
5.28.3.4	<code>cudaError</code>	205
5.28.3.5	<code>cudaFuncCache</code>	209
5.28.3.6	<code>cudaGraphicsCubeFace</code>	209
5.28.3.7	<code>cudaGraphicsMapFlags</code>	210

5.28.3.8	<code>cudaGraphicsRegisterFlags</code>	210
5.28.3.9	<code>cudaLimit</code>	210
5.28.3.10	<code>cudaMemcpyKind</code>	210
5.28.3.11	<code>cudaMemoryType</code>	211
5.28.3.12	<code>cudaOutputMode</code>	211
5.28.3.13	<code>cudaResourceType</code>	211
5.28.3.14	<code>cudaResourceViewFormat</code>	211
5.28.3.15	<code>cudaSharedMemConfig</code>	212
5.28.3.16	<code>cudaSurfaceBoundaryMode</code>	212
5.28.3.17	<code>cudaSurfaceFormatMode</code>	212
5.28.3.18	<code>cudaTextureAddressMode</code>	213
5.28.3.19	<code>cudaTextureFilterMode</code>	213
5.28.3.20	<code>cudaTextureReadMode</code>	213
5.29	Data types used by CUDA driver	214
5.29.1	Define Documentation	222
5.29.1.1	<code>CU_IPC_HANDLE_SIZE</code>	222
5.29.1.2	<code>CU_LAUNCH_PARAM_BUFFER_POINTER</code>	222
5.29.1.3	<code>CU_LAUNCH_PARAM_BUFFER_SIZE</code>	223
5.29.1.4	<code>CU_LAUNCH_PARAM_END</code>	223
5.29.1.5	<code>CU_MEMHOSTALLOC_DEVICEMAP</code>	223
5.29.1.6	<code>CU_MEMHOSTALLOC_PORTABLE</code>	223
5.29.1.7	<code>CU_MEMHOSTALLOC_WRITECOMBINED</code>	223
5.29.1.8	<code>CU_MEMHOSTREGISTER_DEVICEMAP</code>	223
5.29.1.9	<code>CU_MEMHOSTREGISTER_PORTABLE</code>	223
5.29.1.10	<code>CU_PARAM_TR_DEFAULT</code>	223
5.29.1.11	<code>CU_TRSA_OVERRIDE_FORMAT</code>	223
5.29.1.12	<code>CU_TRSF_NORMALIZED_COORDINATES</code>	223
5.29.1.13	<code>CU_TRSF_READ_AS_INTEGER</code>	224
5.29.1.14	<code>CU_TRSF_SRGB</code>	224
5.29.1.15	<code>CUDA_ARRAY3D_2DARRAY</code>	224
5.29.1.16	<code>CUDA_ARRAY3D_CUBEMAP</code>	224
5.29.1.17	<code>CUDA_ARRAY3D_LAYERED</code>	224
5.29.1.18	<code>CUDA_ARRAY3D_SURFACE_LDST</code>	224
5.29.1.19	<code>CUDA_ARRAY3D_TEXTURE_GATHER</code>	224
5.29.1.20	<code>CUDA_VERSION</code>	224
5.29.2	Typedef Documentation	224
5.29.2.1	<code>CUaddress_mode</code>	224

5.29.2.2	CUarray	224
5.29.2.3	CUarray_cubemap_face	225
5.29.2.4	CUarray_format	225
5.29.2.5	CUcomputemode	225
5.29.2.6	CUcontext	225
5.29.2.7	CUctx_flags	225
5.29.2.8	CUDA_ARRAY3D_DESCRIPTOR	225
5.29.2.9	CUDA_ARRAY_DESCRIPTOR	225
5.29.2.10	CUDA_MEMCPY2D	225
5.29.2.11	CUDA_MEMCPY3D	225
5.29.2.12	CUDA_MEMCPY3D_PEER	225
5.29.2.13	CUDA_POINTER_ATTRIBUTE_P2P_TOKENS	225
5.29.2.14	CUDA_RESOURCE_DESC	226
5.29.2.15	CUDA_RESOURCE_VIEW_DESC	226
5.29.2.16	CUDA_TEXTURE_DESC	226
5.29.2.17	CUdevice	226
5.29.2.18	CUdevice_attribute	226
5.29.2.19	CUdeviceptr	226
5.29.2.20	CUdevprop	226
5.29.2.21	CUevent	226
5.29.2.22	CUevent_flags	226
5.29.2.23	CUfilter_mode	226
5.29.2.24	CUfunc_cache	226
5.29.2.25	CUfunction	226
5.29.2.26	CUfunction_attribute	227
5.29.2.27	CUgraphicsMapResourceFlags	227
5.29.2.28	CUgraphicsRegisterFlags	227
5.29.2.29	CUgraphicsResource	227
5.29.2.30	CUipcEventHandle	227
5.29.2.31	CUipcMem_flags	227
5.29.2.32	CUipcMemHandle	227
5.29.2.33	CUjit_fallback	227
5.29.2.34	CUjit_option	227
5.29.2.35	CUjit_target	227
5.29.2.36	CULimit	227
5.29.2.37	CUmemorytype	227
5.29.2.38	CUmipmappedArray	228

5.29.2.39	CUmodule	228
5.29.2.40	CUpointer_attribute	228
5.29.2.41	CUresourcetype	228
5.29.2.42	CUresourceViewFormat	228
5.29.2.43	CUresult	228
5.29.2.44	CUsharedconfig	228
5.29.2.45	CUstream	228
5.29.2.46	CUstream_flags	228
5.29.2.47	CUstreamCallback	228
5.29.2.48	CUsurfObject	229
5.29.2.49	CUsurfref	229
5.29.2.50	CUtexObject	229
5.29.2.51	CUtexref	229
5.29.3	Enumeration Type Documentation	229
5.29.3.1	CUaddress_mode_enum	229
5.29.3.2	CUarray_cubemap_face_enum	229
5.29.3.3	CUarray_format_enum	229
5.29.3.4	CUcomputemode_enum	230
5.29.3.5	CUctx_flags_enum	230
5.29.3.6	cudaError_enum	231
5.29.3.7	CUdevice_attribute_enum	233
5.29.3.8	CUevent_flags_enum	236
5.29.3.9	CUfilter_mode_enum	236
5.29.3.10	CUfunc_cache_enum	236
5.29.3.11	CUfunction_attribute_enum	237
5.29.3.12	CUgraphicsMapResourceFlags_enum	237
5.29.3.13	CUgraphicsRegisterFlags_enum	237
5.29.3.14	CUipcMem_flags_enum	237
5.29.3.15	CUjit_fallback_enum	238
5.29.3.16	CUjit_option_enum	238
5.29.3.17	CUjit_target_enum	239
5.29.3.18	CUlimit_enum	239
5.29.3.19	CUmemorytype_enum	239
5.29.3.20	CUpointer_attribute_enum	240
5.29.3.21	CUresourcetype_enum	240
5.29.3.22	CUresourceViewFormat_enum	240
5.29.3.23	CUsharedconfig_enum	241

5.29.3.24 CUstream_flags_enum	241
5.30 Initialization	242
5.30.1 Detailed Description	242
5.30.2 Function Documentation	242
5.30.2.1 cuInit	242
5.31 Version Management	243
5.31.1 Detailed Description	243
5.31.2 Function Documentation	243
5.31.2.1 cuDriverGetVersion	243
5.32 Device Management	244
5.32.1 Detailed Description	244
5.32.2 Function Documentation	244
5.32.2.1 cuDeviceGet	244
5.32.2.2 cuDeviceGetAttribute	245
5.32.2.3 cuDeviceGetCount	248
5.32.2.4 cuDeviceGetName	249
5.32.2.5 cuDeviceTotalMem	249
5.33 Device Management [DEPRECATED]	250
5.33.1 Detailed Description	250
5.33.2 Function Documentation	250
5.33.2.1 cuDeviceComputeCapability	250
5.33.2.2 cuDeviceGetProperties	250
5.34 Context Management	252
5.34.1 Detailed Description	253
5.34.2 Function Documentation	253
5.34.2.1 cuCtxCreate	253
5.34.2.2 cuCtxDestroy	254
5.34.2.3 cuCtxGetApiVersion	254
5.34.2.4 cuCtxGetCacheConfig	255
5.34.2.5 cuCtxGetCurrent	256
5.34.2.6 cuCtxGetDevice	256
5.34.2.7 cuCtxGetLimit	256
5.34.2.8 cuCtxGetSharedMemConfig	257
5.34.2.9 cuCtxPopCurrent	258
5.34.2.10 cuCtxPushCurrent	258
5.34.2.11 cuCtxSetCacheConfig	258
5.34.2.12 cuCtxSetCurrent	259

5.34.2.13	cuCtxSetLimit	260
5.34.2.14	cuCtxSetSharedMemConfig	261
5.34.2.15	cuCtxSynchronize	262
5.35	Context Management [DEPRECATED]	263
5.35.1	Detailed Description	263
5.35.2	Function Documentation	263
5.35.2.1	cuCtxAttach	263
5.35.2.2	cuCtxDetach	264
5.36	Module Management	265
5.36.1	Detailed Description	265
5.36.2	Function Documentation	265
5.36.2.1	cuModuleGetFunction	265
5.36.2.2	cuModuleGetGlobal	266
5.36.2.3	cuModuleGetSurfRef	266
5.36.2.4	cuModuleGetTexRef	267
5.36.2.5	cuModuleLoad	267
5.36.2.6	cuModuleLoadData	268
5.36.2.7	cuModuleLoadDataEx	268
5.36.2.8	cuModuleLoadFatBinary	270
5.36.2.9	cuModuleUnload	270
5.37	Memory Management	272
5.37.1	Detailed Description	276
5.37.2	Function Documentation	276
5.37.2.1	cuArray3DCreate	276
5.37.2.2	cuArray3DGetDescriptor	279
5.37.2.3	cuArrayCreate	280
5.37.2.4	cuArrayDestroy	281
5.37.2.5	cuArrayGetDescriptor	282
5.37.2.6	cuDeviceGetByPCIBusId	282
5.37.2.7	cuDeviceGetPCIBusId	283
5.37.2.8	cuIpcCloseMemHandle	283
5.37.2.9	cuIpcGetEventHandle	283
5.37.2.10	cuIpcGetMemHandle	284
5.37.2.11	cuIpcOpenEventHandle	284
5.37.2.12	cuIpcOpenMemHandle	285
5.37.2.13	cuMemAlloc	285
5.37.2.14	cuMemAllocHost	286

5.37.2.15 cuMemAllocPitch	287
5.37.2.16 cuMemcpy	288
5.37.2.17 cuMemcpy2D	288
5.37.2.18 cuMemcpy2DAsync	291
5.37.2.19 cuMemcpy2DUnaligned	293
5.37.2.20 cuMemcpy3D	295
5.37.2.21 cuMemcpy3DAsync	298
5.37.2.22 cuMemcpy3DPeer	301
5.37.2.23 cuMemcpy3DPeerAsync	301
5.37.2.24 cuMemcpyAsync	301
5.37.2.25 cuMemcpyAtoA	302
5.37.2.26 cuMemcpyAtoD	303
5.37.2.27 cuMemcpyAtoH	303
5.37.2.28 cuMemcpyAtoHAsync	304
5.37.2.29 cuMemcpyDtoA	305
5.37.2.30 cuMemcpyDtoD	305
5.37.2.31 cuMemcpyDtoDAsync	306
5.37.2.32 cuMemcpyDtoH	307
5.37.2.33 cuMemcpyDtoHAsync	307
5.37.2.34 cuMemcpyHtoA	308
5.37.2.35 cuMemcpyHtoAAsync	308
5.37.2.36 cuMemcpyHtoD	309
5.37.2.37 cuMemcpyHtoDAsync	310
5.37.2.38 cuMemcpyPeer	310
5.37.2.39 cuMemcpyPeerAsync	311
5.37.2.40 cuMemFree	311
5.37.2.41 cuMemFreeHost	312
5.37.2.42 cuMemGetAddressRange	312
5.37.2.43 cuMemGetInfo	313
5.37.2.44 cuMemHostAlloc	314
5.37.2.45 cuMemHostGetDevicePointer	315
5.37.2.46 cuMemHostGetFlags	315
5.37.2.47 cuMemHostRegister	316
5.37.2.48 cuMemHostUnregister	317
5.37.2.49 cuMemsetD16	317
5.37.2.50 cuMemsetD16Async	318
5.37.2.51 cuMemsetD2D16	319

5.37.2.52	cuMemsetD2D16Async	319
5.37.2.53	cuMemsetD2D32	320
5.37.2.54	cuMemsetD2D32Async	321
5.37.2.55	cuMemsetD2D8	321
5.37.2.56	cuMemsetD2D8Async	322
5.37.2.57	cuMemsetD32	323
5.37.2.58	cuMemsetD32Async	323
5.37.2.59	cuMemsetD8	324
5.37.2.60	cuMemsetD8Async	325
5.37.2.61	cuMipmappedArrayCreate	325
5.37.2.62	cuMipmappedArrayDestroy	327
5.37.2.63	cuMipmappedArrayGetLevel	328
5.38	Unified Addressing	329
5.38.1	Detailed Description	329
5.38.2	Overview	329
5.38.3	Supported Platforms	329
5.38.4	Looking Up Information from Pointer Values	329
5.38.5	Automatic Mapping of Host Allocated Host Memory	329
5.38.6	Automatic Registration of Peer Memory	330
5.38.7	Exceptions, Disjoint Addressing	330
5.38.8	Function Documentation	330
5.38.8.1	cuPointerGetAttribute	330
5.39	Stream Management	333
5.39.1	Detailed Description	333
5.39.2	Function Documentation	333
5.39.2.1	cuStreamAddCallback	333
5.39.2.2	cuStreamCreate	334
5.39.2.3	cuStreamDestroy	334
5.39.2.4	cuStreamQuery	335
5.39.2.5	cuStreamSynchronize	335
5.39.2.6	cuStreamWaitEvent	336
5.40	Event Management	337
5.40.1	Detailed Description	337
5.40.2	Function Documentation	337
5.40.2.1	cuEventCreate	337
5.40.2.2	cuEventDestroy	338
5.40.2.3	cuEventElapsedTime	338

5.40.2.4	cuEventQuery	339
5.40.2.5	cuEventRecord	339
5.40.2.6	cuEventSynchronize	340
5.41	Execution Control	341
5.41.1	Detailed Description	341
5.41.2	Function Documentation	341
5.41.2.1	cuFuncGetAttribute	341
5.41.2.2	cuFuncSetCacheConfig	342
5.41.2.3	cuFuncSetSharedMemConfig	343
5.41.2.4	cuLaunchKernel	344
5.42	Execution Control [DEPRECATED]	346
5.42.1	Detailed Description	346
5.42.2	Function Documentation	346
5.42.2.1	cuFuncSetBlockShape	346
5.42.2.2	cuFuncSetSharedSize	347
5.42.2.3	cuLaunch	347
5.42.2.4	cuLaunchGrid	348
5.42.2.5	cuLaunchGridAsync	349
5.42.2.6	cuParamSetf	349
5.42.2.7	cuParamSeti	350
5.42.2.8	cuParamSetSize	350
5.42.2.9	cuParamSetTexRef	351
5.42.2.10	cuParamSetv	351
5.43	Texture Reference Management	353
5.43.1	Detailed Description	354
5.43.2	Function Documentation	354
5.43.2.1	cuTexRefGetAddress	354
5.43.2.2	cuTexRefGetAddressMode	355
5.43.2.3	cuTexRefGetArray	355
5.43.2.4	cuTexRefGetFilterMode	356
5.43.2.5	cuTexRefGetFlags	356
5.43.2.6	cuTexRefGetFormat	356
5.43.2.7	cuTexRefGetMaxAnisotropy	357
5.43.2.8	cuTexRefGetMipmapFilterMode	357
5.43.2.9	cuTexRefGetMipmapLevelBias	358
5.43.2.10	cuTexRefGetMipmapLevelClamp	358
5.43.2.11	cuTexRefGetMipmappedArray	358

5.43.2.12	cuTexRefSetAddress	359
5.43.2.13	cuTexRefSetAddress2D	359
5.43.2.14	cuTexRefSetAddressMode	360
5.43.2.15	cuTexRefSetArray	361
5.43.2.16	cuTexRefSetFilterMode	361
5.43.2.17	cuTexRefSetFlags	362
5.43.2.18	cuTexRefSetFormat	362
5.43.2.19	cuTexRefSetMaxAnisotropy	363
5.43.2.20	cuTexRefSetMipmapFilterMode	363
5.43.2.21	cuTexRefSetMipmapLevelBias	364
5.43.2.22	cuTexRefSetMipmapLevelClamp	364
5.43.2.23	cuTexRefSetMipmappedArray	365
5.44	Texture Reference Management [DEPRECATED]	366
5.44.1	Detailed Description	366
5.44.2	Function Documentation	366
5.44.2.1	cuTexRefCreate	366
5.44.2.2	cuTexRefDestroy	366
5.45	Surface Reference Management	368
5.45.1	Detailed Description	368
5.45.2	Function Documentation	368
5.45.2.1	cuSurfRefGetArray	368
5.45.2.2	cuSurfRefSetArray	368
5.46	Texture Object Management	370
5.46.1	Detailed Description	370
5.46.2	Function Documentation	370
5.46.2.1	cuTexObjectCreate	370
5.46.2.2	cuTexObjectDestroy	374
5.46.2.3	cuTexObjectGetResourceDesc	374
5.46.2.4	cuTexObjectGetResourceViewDesc	375
5.46.2.5	cuTexObjectGetTextureDesc	375
5.47	Surface Object Management	376
5.47.1	Detailed Description	376
5.47.2	Function Documentation	376
5.47.2.1	cuSurfObjectCreate	376
5.47.2.2	cuSurfObjectDestroy	376
5.47.2.3	cuSurfObjectGetResourceDesc	377
5.48	Peer Context Memory Access	378

5.48.1	Detailed Description	378
5.48.2	Function Documentation	378
5.48.2.1	cuCtxDisablePeerAccess	378
5.48.2.2	cuCtxEnablePeerAccess	378
5.48.2.3	cuDeviceCanAccessPeer	379
5.49	Graphics Interoperability	380
5.49.1	Detailed Description	380
5.49.2	Function Documentation	380
5.49.2.1	cuGraphicsMapResources	380
5.49.2.2	cuGraphicsResourceGetMappedMipmappedArray	381
5.49.2.3	cuGraphicsResourceGetMappedPointer	381
5.49.2.4	cuGraphicsResourceSetMapFlags	382
5.49.2.5	cuGraphicsSubResourceGetMappedArray	383
5.49.2.6	cuGraphicsUnmapResources	383
5.49.2.7	cuGraphicsUnregisterResource	384
5.50	Profiler Control	385
5.50.1	Detailed Description	385
5.50.2	Function Documentation	385
5.50.2.1	cuProfilerInitialize	385
5.50.2.2	cuProfilerStart	386
5.50.2.3	cuProfilerStop	386
5.51	OpenGL Interoperability	387
5.51.1	Detailed Description	387
5.51.2	Typedef Documentation	387
5.51.2.1	CUGLDeviceList	387
5.51.3	Enumeration Type Documentation	388
5.51.3.1	CUGLDeviceList_enum	388
5.51.4	Function Documentation	388
5.51.4.1	cuGLGetDevices	388
5.51.4.2	cuGraphicsGLRegisterBuffer	389
5.51.4.3	cuGraphicsGLRegisterImage	389
5.51.4.4	cuWGLGetDevice	390
5.52	OpenGL Interoperability [DEPRECATED]	392
5.52.1	Detailed Description	392
5.52.2	Typedef Documentation	392
5.52.2.1	CUGLmap_flags	392
5.52.3	Enumeration Type Documentation	393

5.52.3.1	CUGLmap_flags_enum	393
5.52.4	Function Documentation	393
5.52.4.1	cuGLCtxCreate	393
5.52.4.2	cuGLInit	393
5.52.4.3	cuGLMapBufferObject	394
5.52.4.4	cuGLMapBufferObjectAsync	394
5.52.4.5	cuGLRegisterBufferObject	395
5.52.4.6	cuGLSetBufferObjectMapFlags	395
5.52.4.7	cuGLUnmapBufferObject	396
5.52.4.8	cuGLUnmapBufferObjectAsync	397
5.52.4.9	cuGLUnregisterBufferObject	397
5.53	Direct3D 9 Interoperability	399
5.53.1	Detailed Description	399
5.53.2	Typedef Documentation	400
5.53.2.1	CUD3d9DeviceList	400
5.53.3	Enumeration Type Documentation	400
5.53.3.1	CUD3d9DeviceList_enum	400
5.53.4	Function Documentation	400
5.53.4.1	cuD3D9CtxCreate	400
5.53.4.2	cuD3D9CtxCreateOnDevice	401
5.53.4.3	cuD3D9GetDevice	401
5.53.4.4	cuD3D9GetDevices	402
5.53.4.5	cuD3D9GetDirect3DDevice	402
5.53.4.6	cuGraphicsD3D9RegisterResource	403
5.54	Direct3D 9 Interoperability [DEPRECATED]	405
5.54.1	Detailed Description	406
5.54.2	Typedef Documentation	406
5.54.2.1	CUD3d9map_flags	406
5.54.2.2	CUD3d9register_flags	406
5.54.3	Enumeration Type Documentation	406
5.54.3.1	CUD3d9map_flags_enum	406
5.54.3.2	CUD3d9register_flags_enum	406
5.54.4	Function Documentation	406
5.54.4.1	cuD3D9MapResources	406
5.54.4.2	cuD3D9RegisterResource	407
5.54.4.3	cuD3D9ResourceGetMappedArray	408
5.54.4.4	cuD3D9ResourceGetMappedPitch	409

5.54.4.5	cuD3D9ResourceGetMappedPointer	410
5.54.4.6	cuD3D9ResourceGetMappedSize	410
5.54.4.7	cuD3D9ResourceGetSurfaceDimensions	411
5.54.4.8	cuD3D9ResourceSetMapFlags	412
5.54.4.9	cuD3D9UnmapResources	413
5.54.4.10	cuD3D9UnregisterResource	413
5.55	Direct3D 10 Interoperability	414
5.55.1	Detailed Description	414
5.55.2	Typedef Documentation	414
5.55.2.1	CUd3d10DeviceList	414
5.55.3	Enumeration Type Documentation	414
5.55.3.1	CUd3d10DeviceList_enum	414
5.55.4	Function Documentation	415
5.55.4.1	cuD3D10GetDevice	415
5.55.4.2	cuD3D10GetDevices	415
5.55.4.3	cuGraphicsD3D10RegisterResource	416
5.56	Direct3D 10 Interoperability [DEPRECATED]	418
5.56.1	Detailed Description	419
5.56.2	Typedef Documentation	419
5.56.2.1	CUD3D10map_flags	419
5.56.2.2	CUD3D10register_flags	419
5.56.3	Enumeration Type Documentation	419
5.56.3.1	CUD3D10map_flags_enum	419
5.56.3.2	CUD3D10register_flags_enum	419
5.56.4	Function Documentation	419
5.56.4.1	cuD3D10CtxCreate	419
5.56.4.2	cuD3D10CtxCreateOnDevice	420
5.56.4.3	cuD3D10GetDirect3DDevice	421
5.56.4.4	cuD3D10MapResources	421
5.56.4.5	cuD3D10RegisterResource	422
5.56.4.6	cuD3D10ResourceGetMappedArray	423
5.56.4.7	cuD3D10ResourceGetMappedPitch	424
5.56.4.8	cuD3D10ResourceGetMappedPointer	424
5.56.4.9	cuD3D10ResourceGetMappedSize	425
5.56.4.10	cuD3D10ResourceGetSurfaceDimensions	426
5.56.4.11	cuD3D10ResourceSetMapFlags	426
5.56.4.12	cuD3D10UnmapResources	427

5.56.4.13	cuD3D10UnregisterResource	428
5.57	Direct3D 11 Interoperability	429
5.57.1	Detailed Description	429
5.57.2	Typedef Documentation	429
5.57.2.1	CUD3d11DeviceList	429
5.57.3	Enumeration Type Documentation	429
5.57.3.1	CUD3d11DeviceList_enum	429
5.57.4	Function Documentation	430
5.57.4.1	cuD3D11GetDevice	430
5.57.4.2	cuD3D11GetDevices	430
5.57.4.3	cuGraphicsD3D11RegisterResource	431
5.58	Direct3D 11 Interoperability [DEPRECATED]	433
5.58.1	Detailed Description	433
5.58.2	Function Documentation	433
5.58.2.1	cuD3D11CtxCreate	433
5.58.2.2	cuD3D11CtxCreateOnDevice	434
5.58.2.3	cuD3D11GetDirect3DDevice	434
5.59	VDPAU Interoperability	435
5.59.1	Detailed Description	435
5.59.2	Function Documentation	435
5.59.2.1	cuGraphicsVDPAURegisterOutputSurface	435
5.59.2.2	cuGraphicsVDPAURegisterVideoSurface	436
5.59.2.3	cuVDPAUCtxCreate	437
5.59.2.4	cuVDPAUGetDevice	437
5.60	Mathematical Functions	439
5.61	Single Precision Mathematical Functions	440
5.61.1	Detailed Description	444
5.61.2	Function Documentation	445
5.61.2.1	acosf	445
5.61.2.2	acoshf	445
5.61.2.3	asinf	445
5.61.2.4	asinhf	445
5.61.2.5	atan2f	446
5.61.2.6	atanf	446
5.61.2.7	atanhf	446
5.61.2.8	cbrtf	446
5.61.2.9	ceilf	447

5.61.2.10 copysignf	447
5.61.2.11 cosf	447
5.61.2.12 coshf	447
5.61.2.13 cospif	448
5.61.2.14 erfcf	448
5.61.2.15 erfcinvf	448
5.61.2.16 erfcxf	448
5.61.2.17 erff	449
5.61.2.18 erfinvf	449
5.61.2.19 exp10f	449
5.61.2.20 exp2f	449
5.61.2.21 expf	450
5.61.2.22 expm1f	450
5.61.2.23 fabsf	450
5.61.2.24 fdimf	450
5.61.2.25 fdividef	451
5.61.2.26 floorf	451
5.61.2.27 fmaf	451
5.61.2.28 fmaxf	452
5.61.2.29 fminf	452
5.61.2.30 fmodf	452
5.61.2.31 frexpf	453
5.61.2.32 hypotf	453
5.61.2.33 ilogbf	453
5.61.2.34 isfinite	454
5.61.2.35 isinf	454
5.61.2.36 isnan	454
5.61.2.37 j0f	454
5.61.2.38 j1f	454
5.61.2.39 jnf	455
5.61.2.40 ldexpf	455
5.61.2.41 lgammaf	455
5.61.2.42 llrintf	455
5.61.2.43 llroundf	456
5.61.2.44 log10f	456
5.61.2.45 log1pf	456
5.61.2.46 log2f	456

5.61.2.47	logbf	457
5.61.2.48	logf	457
5.61.2.49	lrintf	457
5.61.2.50	lroundf	458
5.61.2.51	modff	458
5.61.2.52	nanf	458
5.61.2.53	nearbyintf	458
5.61.2.54	nextafterf	459
5.61.2.55	normcdf	459
5.61.2.56	normcdfinvf	459
5.61.2.57	powf	459
5.61.2.58	rcbrtf	460
5.61.2.59	remainderf	460
5.61.2.60	remquof	461
5.61.2.61	rintf	461
5.61.2.62	roundf	461
5.61.2.63	rsqrtf	461
5.61.2.64	scalblnf	462
5.61.2.65	scalbnf	462
5.61.2.66	signbit	462
5.61.2.67	sincosf	462
5.61.2.68	sincospif	463
5.61.2.69	sinf	463
5.61.2.70	sinhf	463
5.61.2.71	sinpif	463
5.61.2.72	sqrtf	464
5.61.2.73	tanf	464
5.61.2.74	tanhf	464
5.61.2.75	tgammaf	464
5.61.2.76	truncf	465
5.61.2.77	y0f	465
5.61.2.78	y1f	465
5.61.2.79	ynf	466
5.62	Double Precision Mathematical Functions	467
5.62.1	Detailed Description	471
5.62.2	Function Documentation	471
5.62.2.1	acos	471

5.62.2.2	acosh	472
5.62.2.3	asin	472
5.62.2.4	asinh	472
5.62.2.5	atan	473
5.62.2.6	atan2	473
5.62.2.7	atanh	473
5.62.2.8	cbrt	473
5.62.2.9	ceil	474
5.62.2.10	copysign	474
5.62.2.11	cos	474
5.62.2.12	cosh	474
5.62.2.13	cospi	474
5.62.2.14	erf	475
5.62.2.15	erfc	475
5.62.2.16	erfcinv	475
5.62.2.17	erfcx	476
5.62.2.18	erfinv	476
5.62.2.19	exp	476
5.62.2.20	exp10	476
5.62.2.21	exp2	477
5.62.2.22	expm1	477
5.62.2.23	fabs	477
5.62.2.24	fdim	477
5.62.2.25	floor	478
5.62.2.26	fma	478
5.62.2.27	fmax	478
5.62.2.28	fmin	479
5.62.2.29	fmod	479
5.62.2.30	frexp	479
5.62.2.31	hypot	480
5.62.2.32	ilogb	480
5.62.2.33	isfinite	480
5.62.2.34	isinf	480
5.62.2.35	isnan	480
5.62.2.36	j0	481
5.62.2.37	j1	481
5.62.2.38	jn	481

5.62.2.39 ldexp	481
5.62.2.40 lgamma	482
5.62.2.41 llrint	482
5.62.2.42 llround	482
5.62.2.43 log	482
5.62.2.44 log10	483
5.62.2.45 log1p	483
5.62.2.46 log2	483
5.62.2.47 logb	484
5.62.2.48 lrint	484
5.62.2.49 lround	484
5.62.2.50 modf	484
5.62.2.51 nan	485
5.62.2.52 nearbyint	485
5.62.2.53 nextafter	485
5.62.2.54 normcdf	485
5.62.2.55 normcdfinv	486
5.62.2.56 pow	486
5.62.2.57 rcbrt	487
5.62.2.58 remainder	487
5.62.2.59 remquo	487
5.62.2.60 rint	487
5.62.2.61 round	488
5.62.2.62 rsqrt	488
5.62.2.63 scalbln	488
5.62.2.64 scalbn	488
5.62.2.65 signbit	489
5.62.2.66 sin	489
5.62.2.67 sincos	489
5.62.2.68 sincospi	489
5.62.2.69 sinh	490
5.62.2.70 sinpi	490
5.62.2.71 sqrt	490
5.62.2.72 tan	490
5.62.2.73 tanh	491
5.62.2.74 tgamma	491
5.62.2.75 trunc	491

5.62.2.76	y0	491
5.62.2.77	y1	492
5.62.2.78	yn	492
5.63	Single Precision Intrinsic	493
5.63.1	Detailed Description	495
5.63.2	Function Documentation	495
5.63.2.1	__cosf	495
5.63.2.2	__exp10f	495
5.63.2.3	__expf	496
5.63.2.4	__fadd_rd	496
5.63.2.5	__fadd_rn	496
5.63.2.6	__fadd_ru	496
5.63.2.7	__fadd_rz	497
5.63.2.8	__fdiv_rd	497
5.63.2.9	__fdiv_rn	497
5.63.2.10	__fdiv_ru	497
5.63.2.11	__fdiv_rz	498
5.63.2.12	__fdivdef	498
5.63.2.13	__fmaf_rd	498
5.63.2.14	__fmaf_rn	498
5.63.2.15	__fmaf_ru	499
5.63.2.16	__fmaf_rz	499
5.63.2.17	__fmul_rd	499
5.63.2.18	__fmul_rn	500
5.63.2.19	__fmul_ru	500
5.63.2.20	__fmul_rz	500
5.63.2.21	__frcp_rd	500
5.63.2.22	__frcp_rn	501
5.63.2.23	__frcp_ru	501
5.63.2.24	__frcp_rz	501
5.63.2.25	__frsqrt_rn	501
5.63.2.26	__fsqrt_rd	502
5.63.2.27	__fsqrt_rn	502
5.63.2.28	__fsqrt_ru	502
5.63.2.29	__fsqrt_rz	502
5.63.2.30	__log10f	503
5.63.2.31	__log2f	503

5.63.2.32	__logf	503
5.63.2.33	__powf	503
5.63.2.34	__saturatef	504
5.63.2.35	__sincosf	504
5.63.2.36	__sinf	504
5.63.2.37	__tanf	504
5.64	Double Precision Intrinsics	505
5.64.1	Detailed Description	506
5.64.2	Function Documentation	506
5.64.2.1	__dadd_rd	506
5.64.2.2	__dadd_rn	507
5.64.2.3	__dadd_ru	507
5.64.2.4	__dadd_rz	507
5.64.2.5	__ddiv_rd	507
5.64.2.6	__ddiv_rn	508
5.64.2.7	__ddiv_ru	508
5.64.2.8	__ddiv_rz	508
5.64.2.9	__dmul_rd	508
5.64.2.10	__dmul_rn	509
5.64.2.11	__dmul_ru	509
5.64.2.12	__dmul_rz	509
5.64.2.13	__drep_rd	509
5.64.2.14	__drep_rn	510
5.64.2.15	__drep_ru	510
5.64.2.16	__drep_rz	510
5.64.2.17	__dsqrt_rd	510
5.64.2.18	__dsqrt_rn	511
5.64.2.19	__dsqrt_ru	511
5.64.2.20	__dsqrt_rz	511
5.64.2.21	__fma_rd	511
5.64.2.22	__fma_rn	512
5.64.2.23	__fma_ru	512
5.64.2.24	__fma_rz	512
5.65	Integer Intrinsics	513
5.65.1	Detailed Description	514
5.65.2	Function Documentation	514
5.65.2.1	__brev	514

5.65.2.2	<code>__brevll</code>	514
5.65.2.3	<code>__byte_perm</code>	514
5.65.2.4	<code>__clz</code>	515
5.65.2.5	<code>__clzll</code>	515
5.65.2.6	<code>__ffs</code>	515
5.65.2.7	<code>__ffsll</code>	515
5.65.2.8	<code>__mul24</code>	515
5.65.2.9	<code>__mul64hi</code>	515
5.65.2.10	<code>__mulhi</code>	516
5.65.2.11	<code>__popc</code>	516
5.65.2.12	<code>__popc11</code>	516
5.65.2.13	<code>__sad</code>	516
5.65.2.14	<code>__umul24</code>	516
5.65.2.15	<code>__umul64hi</code>	516
5.65.2.16	<code>__umulhi</code>	517
5.65.2.17	<code>__usad</code>	517
5.66	Type Casting Intrinsic	518
5.66.1	Detailed Description	522
5.66.2	Function Documentation	522
5.66.2.1	<code>__double2float_rd</code>	522
5.66.2.2	<code>__double2float_rn</code>	522
5.66.2.3	<code>__double2float_ru</code>	522
5.66.2.4	<code>__double2float_rz</code>	523
5.66.2.5	<code>__double2hiint</code>	523
5.66.2.6	<code>__double2int_rd</code>	523
5.66.2.7	<code>__double2int_rn</code>	523
5.66.2.8	<code>__double2int_ru</code>	523
5.66.2.9	<code>__double2int_rz</code>	523
5.66.2.10	<code>__double2ll_rd</code>	524
5.66.2.11	<code>__double2ll_rn</code>	524
5.66.2.12	<code>__double2ll_ru</code>	524
5.66.2.13	<code>__double2ll_rz</code>	524
5.66.2.14	<code>__double2loint</code>	524
5.66.2.15	<code>__double2uint_rd</code>	524
5.66.2.16	<code>__double2uint_rn</code>	525
5.66.2.17	<code>__double2uint_ru</code>	525
5.66.2.18	<code>__double2uint_rz</code>	525

5.66.2.19	__double2ull_rd	525
5.66.2.20	__double2ull_rn	525
5.66.2.21	__double2ull_ru	525
5.66.2.22	__double2ull_rz	526
5.66.2.23	__double_as_longlong	526
5.66.2.24	__float2half_rn	526
5.66.2.25	__float2int_rd	526
5.66.2.26	__float2int_rn	526
5.66.2.27	__float2int_ru	526
5.66.2.28	__float2int_rz	527
5.66.2.29	__float2ll_rd	527
5.66.2.30	__float2ll_rn	527
5.66.2.31	__float2ll_ru	527
5.66.2.32	__float2ll_rz	527
5.66.2.33	__float2uint_rd	527
5.66.2.34	__float2uint_rn	528
5.66.2.35	__float2uint_ru	528
5.66.2.36	__float2uint_rz	528
5.66.2.37	__float2ull_rd	528
5.66.2.38	__float2ull_rn	528
5.66.2.39	__float2ull_ru	528
5.66.2.40	__float2ull_rz	529
5.66.2.41	__float_as_int	529
5.66.2.42	__half2float	529
5.66.2.43	__hiloInt2double	529
5.66.2.44	__int2double_rn	529
5.66.2.45	__int2float_rd	529
5.66.2.46	__int2float_rn	530
5.66.2.47	__int2float_ru	530
5.66.2.48	__int2float_rz	530
5.66.2.49	__int_as_float	530
5.66.2.50	__ll2double_rd	530
5.66.2.51	__ll2double_rn	530
5.66.2.52	__ll2double_ru	531
5.66.2.53	__ll2double_rz	531
5.66.2.54	__ll2float_rd	531
5.66.2.55	__ll2float_rn	531

5.66.2.56	__ll2float_ru	531
5.66.2.57	__ll2float_rz	531
5.66.2.58	__longlong_as_double	532
5.66.2.59	__uint2double_rn	532
5.66.2.60	__uint2float_rd	532
5.66.2.61	__uint2float_rn	532
5.66.2.62	__uint2float_ru	532
5.66.2.63	__uint2float_rz	532
5.66.2.64	__ull2double_rd	533
5.66.2.65	__ull2double_rn	533
5.66.2.66	__ull2double_ru	533
5.66.2.67	__ull2double_rz	533
5.66.2.68	__ull2float_rd	533
5.66.2.69	__ull2float_rn	533
5.66.2.70	__ull2float_ru	534
5.66.2.71	__ull2float_rz	534
6	Data Structure Documentation	535
6.1	CUDA_ARRAY3D_DESCRIPTOR_st Struct Reference	535
6.1.1	Detailed Description	535
6.1.2	Field Documentation	535
6.1.2.1	Depth	535
6.1.2.2	Flags	535
6.1.2.3	Format	535
6.1.2.4	Height	535
6.1.2.5	NumChannels	536
6.1.2.6	Width	536
6.2	CUDA_ARRAY_DESCRIPTOR_st Struct Reference	537
6.2.1	Detailed Description	537
6.2.2	Field Documentation	537
6.2.2.1	Format	537
6.2.2.2	Height	537
6.2.2.3	NumChannels	537
6.2.2.4	Width	537
6.3	CUDA_MEMCPY2D_st Struct Reference	538
6.3.1	Detailed Description	538
6.3.2	Field Documentation	538

6.3.2.1	dstArray	538
6.3.2.2	dstDevice	538
6.3.2.3	dstHost	538
6.3.2.4	dstMemoryType	538
6.3.2.5	dstPitch	538
6.3.2.6	dstXInBytes	539
6.3.2.7	dstY	539
6.3.2.8	Height	539
6.3.2.9	srcArray	539
6.3.2.10	srcDevice	539
6.3.2.11	srcHost	539
6.3.2.12	srcMemoryType	539
6.3.2.13	srcPitch	539
6.3.2.14	srcXInBytes	539
6.3.2.15	srcY	539
6.3.2.16	WidthInBytes	539
6.4	CUDA_MEMCPY3D_PEER_st Struct Reference	540
6.4.1	Detailed Description	540
6.4.2	Field Documentation	540
6.4.2.1	Depth	540
6.4.2.2	dstArray	540
6.4.2.3	dstContext	540
6.4.2.4	dstDevice	541
6.4.2.5	dstHeight	541
6.4.2.6	dstHost	541
6.4.2.7	dstLOD	541
6.4.2.8	dstMemoryType	541
6.4.2.9	dstPitch	541
6.4.2.10	dstXInBytes	541
6.4.2.11	dstY	541
6.4.2.12	dstZ	541
6.4.2.13	Height	541
6.4.2.14	srcArray	541
6.4.2.15	srcContext	541
6.4.2.16	srcDevice	542
6.4.2.17	srcHeight	542
6.4.2.18	srcHost	542

6.4.2.19	srcLOD	542
6.4.2.20	srcMemoryType	542
6.4.2.21	srcPitch	542
6.4.2.22	srcXInBytes	542
6.4.2.23	srcY	542
6.4.2.24	srcZ	542
6.4.2.25	WidthInBytes	542
6.5	CUDA_MEMCPY3D_st Struct Reference	543
6.5.1	Detailed Description	543
6.5.2	Field Documentation	543
6.5.2.1	Depth	543
6.5.2.2	dstArray	543
6.5.2.3	dstDevice	543
6.5.2.4	dstHeight	544
6.5.2.5	dstHost	544
6.5.2.6	dstLOD	544
6.5.2.7	dstMemoryType	544
6.5.2.8	dstPitch	544
6.5.2.9	dstXInBytes	544
6.5.2.10	dstY	544
6.5.2.11	dstZ	544
6.5.2.12	Height	544
6.5.2.13	reserved0	544
6.5.2.14	reserved1	544
6.5.2.15	srcArray	544
6.5.2.16	srcDevice	545
6.5.2.17	srcHeight	545
6.5.2.18	srcHost	545
6.5.2.19	srcLOD	545
6.5.2.20	srcMemoryType	545
6.5.2.21	srcPitch	545
6.5.2.22	srcXInBytes	545
6.5.2.23	srcY	545
6.5.2.24	srcZ	545
6.5.2.25	WidthInBytes	545
6.6	CUDA_POINTER_ATTRIBUTE_P2P_TOKENS_st Struct Reference	546
6.6.1	Detailed Description	546

6.7	CUDA_RESOURCE_DESC_st Struct Reference	547
6.7.1	Detailed Description	547
6.7.2	Field Documentation	547
6.7.2.1	devPtr	547
6.7.2.2	flags	547
6.7.2.3	format	547
6.7.2.4	hArray	547
6.7.2.5	height	547
6.7.2.6	hMipmappedArray	547
6.7.2.7	numChannels	548
6.7.2.8	pitchInBytes	548
6.7.2.9	resType	548
6.7.2.10	sizeInBytes	548
6.7.2.11	width	548
6.8	CUDA_RESOURCE_VIEW_DESC_st Struct Reference	549
6.8.1	Detailed Description	549
6.8.2	Field Documentation	549
6.8.2.1	depth	549
6.8.2.2	firstLayer	549
6.8.2.3	firstMipmapLevel	549
6.8.2.4	format	549
6.8.2.5	height	549
6.8.2.6	lastLayer	549
6.8.2.7	lastMipmapLevel	549
6.8.2.8	width	550
6.9	CUDA_TEXTURE_DESC_st Struct Reference	551
6.9.1	Detailed Description	551
6.9.2	Field Documentation	551
6.9.2.1	addressMode	551
6.9.2.2	filterMode	551
6.9.2.3	flags	551
6.9.2.4	maxAnisotropy	551
6.9.2.5	maxMipmapLevelClamp	551
6.9.2.6	minMipmapLevelClamp	551
6.9.2.7	mipmapFilterMode	551
6.9.2.8	mipmapLevelBias	552
6.10	cudaChannelFormatDesc Struct Reference	553

6.10.1 Detailed Description	553
6.10.2 Field Documentation	553
6.10.2.1 f	553
6.10.2.2 w	553
6.10.2.3 x	553
6.10.2.4 y	553
6.10.2.5 z	553
6.11 cudaDeviceProp Struct Reference	554
6.11.1 Detailed Description	555
6.11.2 Field Documentation	555
6.11.2.1 asyncEngineCount	555
6.11.2.2 canMapHostMemory	555
6.11.2.3 clockRate	555
6.11.2.4 computeMode	555
6.11.2.5 concurrentKernels	555
6.11.2.6 deviceOverlap	555
6.11.2.7 ECCEnabled	555
6.11.2.8 integrated	555
6.11.2.9 kernelExecTimeoutEnabled	556
6.11.2.10 l2CacheSize	556
6.11.2.11 major	556
6.11.2.12 maxGridSize	556
6.11.2.13 maxSurface1D	556
6.11.2.14 maxSurface1DLayered	556
6.11.2.15 maxSurface2D	556
6.11.2.16 maxSurface2DLayered	556
6.11.2.17 maxSurface3D	556
6.11.2.18 maxSurfaceCubemap	556
6.11.2.19 maxSurfaceCubemapLayered	556
6.11.2.20 maxTexture1D	556
6.11.2.21 maxTexture1DLayered	557
6.11.2.22 maxTexture1DLinear	557
6.11.2.23 maxTexture1DMipmap	557
6.11.2.24 maxTexture2D	557
6.11.2.25 maxTexture2DGather	557
6.11.2.26 maxTexture2DLayered	557
6.11.2.27 maxTexture2DLinear	557

6.11.2.28	maxTexture2DMipmap	557
6.11.2.29	maxTexture3D	557
6.11.2.30	maxTextureCubemap	557
6.11.2.31	maxTextureCubemapLayered	557
6.11.2.32	maxThreadsDim	557
6.11.2.33	maxThreadsPerBlock	558
6.11.2.34	maxThreadsPerMultiProcessor	558
6.11.2.35	memoryBusWidth	558
6.11.2.36	memoryClockRate	558
6.11.2.37	memPitch	558
6.11.2.38	minor	558
6.11.2.39	multiProcessorCount	558
6.11.2.40	name	558
6.11.2.41	pciBusID	558
6.11.2.42	pciDeviceID	558
6.11.2.43	pciDomainID	558
6.11.2.44	regsPerBlock	558
6.11.2.45	sharedMemPerBlock	559
6.11.2.46	surfaceAlignment	559
6.11.2.47	tccDriver	559
6.11.2.48	textureAlignment	559
6.11.2.49	texturePitchAlignment	559
6.11.2.50	totalConstMem	559
6.11.2.51	totalGlobalMem	559
6.11.2.52	unifiedAddressing	559
6.11.2.53	warpSize	559
6.12	cudaExtent Struct Reference	560
6.12.1	Detailed Description	560
6.12.2	Field Documentation	560
6.12.2.1	depth	560
6.12.2.2	height	560
6.12.2.3	width	560
6.13	cudaFuncAttributes Struct Reference	561
6.13.1	Detailed Description	561
6.13.2	Field Documentation	561
6.13.2.1	binaryVersion	561
6.13.2.2	constSizeBytes	561

6.13.2.3	localSizeBytes	561
6.13.2.4	maxThreadsPerBlock	561
6.13.2.5	numRegs	561
6.13.2.6	ptxVersion	561
6.13.2.7	sharedSizeBytes	562
6.14	cudaIpcEventHandle_st Struct Reference	563
6.14.1	Detailed Description	563
6.15	cudaIpcMemHandle_st Struct Reference	564
6.15.1	Detailed Description	564
6.16	cudaMemcpy3DParms Struct Reference	565
6.16.1	Detailed Description	565
6.16.2	Field Documentation	565
6.16.2.1	dstArray	565
6.16.2.2	dstPos	565
6.16.2.3	dstPtr	565
6.16.2.4	extent	565
6.16.2.5	kind	565
6.16.2.6	srcArray	565
6.16.2.7	srcPos	565
6.16.2.8	srcPtr	566
6.17	cudaMemcpy3DPeerParms Struct Reference	567
6.17.1	Detailed Description	567
6.17.2	Field Documentation	567
6.17.2.1	dstArray	567
6.17.2.2	dstDevice	567
6.17.2.3	dstPos	567
6.17.2.4	dstPtr	567
6.17.2.5	extent	567
6.17.2.6	srcArray	567
6.17.2.7	srcDevice	567
6.17.2.8	srcPos	568
6.17.2.9	srcPtr	568
6.18	cudaPitchedPtr Struct Reference	569
6.18.1	Detailed Description	569
6.18.2	Field Documentation	569
6.18.2.1	pitch	569
6.18.2.2	ptr	569

6.18.2.3	xsize	569
6.18.2.4	ysize	569
6.19	cudaPointerAttributes Struct Reference	570
6.19.1	Detailed Description	570
6.19.2	Field Documentation	570
6.19.2.1	device	570
6.19.2.2	devicePointer	570
6.19.2.3	hostPointer	570
6.19.2.4	memoryType	570
6.20	cudaPos Struct Reference	571
6.20.1	Detailed Description	571
6.20.2	Field Documentation	571
6.20.2.1	x	571
6.20.2.2	y	571
6.20.2.3	z	571
6.21	cudaResourceDesc Struct Reference	572
6.21.1	Detailed Description	572
6.21.2	Field Documentation	572
6.21.2.1	array	572
6.21.2.2	desc	572
6.21.2.3	devPtr	572
6.21.2.4	height	572
6.21.2.5	mipmap	572
6.21.2.6	pitchInBytes	572
6.21.2.7	resType	572
6.21.2.8	sizeInBytes	573
6.21.2.9	width	573
6.22	cudaResourceViewDesc Struct Reference	574
6.22.1	Detailed Description	574
6.22.2	Field Documentation	574
6.22.2.1	depth	574
6.22.2.2	firstLayer	574
6.22.2.3	firstMipmapLevel	574
6.22.2.4	format	574
6.22.2.5	height	574
6.22.2.6	lastLayer	574
6.22.2.7	lastMipmapLevel	574

6.22.2.8	width	575
6.23	cudaTextureDesc Struct Reference	576
6.23.1	Detailed Description	576
6.23.2	Field Documentation	576
6.23.2.1	addressMode	576
6.23.2.2	filterMode	576
6.23.2.3	maxAnisotropy	576
6.23.2.4	maxMipmapLevelClamp	576
6.23.2.5	minMipmapLevelClamp	576
6.23.2.6	mipmapFilterMode	576
6.23.2.7	mipmapLevelBias	576
6.23.2.8	normalizedCoords	577
6.23.2.9	readMode	577
6.23.2.10	sRGB	577
6.24	CUdevprop_st Struct Reference	578
6.24.1	Detailed Description	578
6.24.2	Field Documentation	578
6.24.2.1	clockRate	578
6.24.2.2	maxGridSize	578
6.24.2.3	maxThreadsDim	578
6.24.2.4	maxThreadsPerBlock	578
6.24.2.5	memPitch	578
6.24.2.6	regsPerBlock	578
6.24.2.7	sharedMemPerBlock	578
6.24.2.8	SIMDWidth	579
6.24.2.9	textureAlign	579
6.24.2.10	totalConstantMemory	579
6.25	CUipcEventHandle_st Struct Reference	580
6.25.1	Detailed Description	580
6.26	CUipcMemHandle_st Struct Reference	581
6.26.1	Detailed Description	581
6.27	surfaceReference Struct Reference	582
6.27.1	Detailed Description	582
6.27.2	Field Documentation	582
6.27.2.1	channelDesc	582
6.28	textureReference Struct Reference	583
6.28.1	Detailed Description	583

6.28.2	Field Documentation	583
6.28.2.1	addressMode	583
6.28.2.2	channelDesc	583
6.28.2.3	filterMode	583
6.28.2.4	maxAnisotropy	583
6.28.2.5	maxMipmapLevelClamp	583
6.28.2.6	minMipmapLevelClamp	583
6.28.2.7	mipmapFilterMode	583
6.28.2.8	mipmapLevelBias	584
6.28.2.9	normalized	584
6.28.2.10	sRGB	584

Chapter 1

API synchronization behavior

1.1 Memcpy

The API provides memcpy/memset functions in both synchronous and asynchronous forms, the latter having an "Async" suffix. This is a misnomer as each function may exhibit synchronous or asynchronous behavior depending on the arguments passed to the function. In the reference documentation, each memcpy function is categorized as *synchronous* or *asynchronous*, corresponding to the definitions below.

1.1.1 Synchronous

1. For transfers from pageable host memory to device memory, a stream sync is performed before the copy is initiated. The function will return once the pageable buffer has been copied to the staging memory for DMA transfer to device memory, but the DMA to final destination may not have completed.
2. For transfers from pinned host memory to device memory, the function is synchronous with respect to the host.
3. For transfers from device to either pageable or pinned host memory, the function returns only once the copy has completed.
4. For transfers from device memory to device memory, no host-side synchronization is performed.
5. For transfers from any host memory to any host memory, the function is fully synchronous with respect to the host.

1.1.2 Asynchronous

1. For transfers from pageable host memory to device memory, host memory is copied to a staging buffer immediately (no device synchronization is performed). The function will return once the pageable buffer has been copied to the staging memory. The DMA transfer to final destination may not have completed.
2. For transfers between pinned host memory and device memory, the function is fully asynchronous.
3. For transfers from device memory to pageable host memory, the function will return only once the copy has completed.
4. For all other transfers, the function is fully asynchronous. If pageable memory must first be staged to pinned memory, this will be handled asynchronously with a worker thread.
5. For transfers from any host memory to any host memory, the function is fully synchronous with respect to the host.

1.2 Memset

The `cudaMemset` functions are asynchronous with respect to the host except when the target memory is pinned host memory. The *Async* versions are always asynchronous with respect to the host.

1.3 Kernel Launches

Kernel launches are asynchronous with respect to the host. Details of concurrent kernel execution and data transfers can be found in the CUDA Programmers Guide.

Chapter 2

Deprecated List

Global [cudaThreadExit](#)

Global [cudaThreadGetCacheConfig](#)

Global [cudaThreadGetLimit](#)

Global [cudaThreadSetCacheConfig](#)

Global [cudaThreadSetLimit](#)

Global [cudaThreadSynchronize](#)

Global [cudaGLMapBufferObject](#) This function is deprecated as of CUDA 3.0.

Global [cudaGLMapBufferObjectAsync](#) This function is deprecated as of CUDA 3.0.

Global [cudaGLRegisterBufferObject](#) This function is deprecated as of CUDA 3.0.

Global [cudaGLSetBufferObjectMapFlags](#) This function is deprecated as of CUDA 3.0.

Global [cudaGLSetGLDevice](#) This function is deprecated as of CUDA 5.0.

Global [cudaGLUnmapBufferObject](#) This function is deprecated as of CUDA 3.0.

Global [cudaGLUnmapBufferObjectAsync](#) This function is deprecated as of CUDA 3.0.

Global [cudaGLUnregisterBufferObject](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9MapResources](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9RegisterResource](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9ResourceGetMappedArray](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9ResourceGetMappedPitch](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9ResourceGetMappedPointer](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9ResourceGetMappedSize](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9ResourceGetSurfaceDimensions](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9ResourceSetMapFlags](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9UnmapResources](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D9UnregisterResource](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10GetDirect3DDevice](#) This function is deprecated as of CUDA 5.0.

Global [cudaD3D10MapResources](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10RegisterResource](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10ResourceGetMappedArray](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10ResourceGetMappedPitch](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10ResourceGetMappedPointer](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10ResourceGetMappedSize](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10ResourceGetSurfaceDimensions](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10ResourceSetMapFlags](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10SetDirect3DDevice](#) This function is deprecated as of CUDA 5.0.

Global [cudaD3D10UnmapResources](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D10UnregisterResource](#) This function is deprecated as of CUDA 3.0.

Global [cudaD3D11GetDirect3DDevice](#) This function is deprecated as of CUDA 5.0.

Global [cudaD3D11SetDirect3DDevice](#) This function is deprecated as of CUDA 5.0.

Global [cudaErrorPriorLaunchFailure](#) This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

Global [cudaErrorAddressOfConstant](#) This error return is deprecated as of CUDA 3.1. Variables in constant memory may now have their address taken by the runtime via [cudaGetSymbolAddress\(\)](#).

Global [cudaErrorTextureFetchFailed](#) This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

Global [cudaErrorTextureNotBound](#) This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

Global [cudaErrorSynchronizationError](#) This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

Global [cudaErrorMixedDeviceExecution](#) This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

Global [cudaErrorNotYetImplemented](#) This error return is deprecated as of CUDA 4.1.

Global [cudaErrorMemoryValueTooLarge](#) This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

Global [cudaErrorProfilerNotInitialized](#) This error return is deprecated as of CUDA 5.0. It is no longer an error to attempt to enable/disable the profiling via [cudaProfilerStart](#) or [cudaProfilerStop](#) without initialization.

Global [cudaErrorProfilerAlreadyStarted](#) This error return is deprecated as of CUDA 5.0. It is no longer an error to call [cudaProfilerStart\(\)](#) when profiling is already enabled.

Global [cudaErrorProfilerAlreadyStopped](#) This error return is deprecated as of CUDA 5.0. It is no longer an error to call [cudaProfilerStop\(\)](#) when profiling is already disabled.

Global [cudaErrorApiFailureBase](#) This error return is deprecated as of CUDA 4.1.

Global [cudaDeviceBlockingSync](#) This flag was deprecated as of CUDA 4.0 and replaced with [cudaDeviceScheduleBlockingSync](#).

Global `CU_CTX_BLOCKING_SYNC` This flag was deprecated as of CUDA 4.0 and was replaced with `CU_CTX_SCHED_BLOCKING_SYNC`.

Global `CUDA_ERROR_PROFILER_NOT_INITIALIZED` This error return is deprecated as of CUDA 5.0. It is no longer an error to attempt to enable/disable the profiling via `cuProfilerStart` or `cuProfilerStop` without initialization.

Global `CUDA_ERROR_PROFILER_ALREADY_STARTED` This error return is deprecated as of CUDA 5.0. It is no longer an error to call `cuProfilerStart()` when profiling is already enabled.

Global `CUDA_ERROR_PROFILER_ALREADY_STOPPED` This error return is deprecated as of CUDA 5.0. It is no longer an error to call `cuProfilerStop()` when profiling is already disabled.

Global `CUDA_ERROR_CONTEXT_ALREADY_CURRENT` This error return is deprecated as of CUDA 3.2. It is no longer an error to attempt to push the active context via `cuCtxPushCurrent()`.

Global `cuDeviceComputeCapability`

Global `cuDeviceGetProperties`

Global `cuCtxAttach`

Global `cuCtxDetach`

Global `cuFuncSetBlockShape`

Global `cuFuncSetSharedSize`

Global `cuLaunch`

Global `cuLaunchGrid`

Global `cuLaunchGridAsync`

Global `cuParamSetf`

Global `cuParamSeti`

Global [cuParamSetSize](#)

Global [cuParamSetTexRef](#)

Global [cuParamSetv](#)

Global [cuTexRefCreate](#)

Global [cuTexRefDestroy](#)

Global [cuGLCtxCreate](#) This function is deprecated as of Cuda 5.0.

Global [cuGLInit](#) This function is deprecated as of Cuda 3.0.

Global [cuGLMapBufferObject](#) This function is deprecated as of Cuda 3.0.

Global [cuGLMapBufferObjectAsync](#) This function is deprecated as of Cuda 3.0.

Global [cuGLRegisterBufferObject](#) This function is deprecated as of Cuda 3.0.

Global [cuGLSetBufferObjectMapFlags](#) This function is deprecated as of Cuda 3.0.

Global [cuGLUnmapBufferObject](#) This function is deprecated as of Cuda 3.0.

Global [cuGLUnmapBufferObjectAsync](#) This function is deprecated as of Cuda 3.0.

Global [cuGLUnregisterBufferObject](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9MapResources](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9RegisterResource](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9ResourceGetMappedArray](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9ResourceGetMappedPitch](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9ResourceGetMappedPointer](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9ResourceGetMappedSize](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9ResourceGetSurfaceDimensions](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9ResourceSetMapFlags](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9UnmapResources](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D9UnregisterResource](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10CtxCreate](#) This function is deprecated as of Cuda 5.0.

Global [cuD3D10CtxCreateOnDevice](#) This function is deprecated as of Cuda 5.0.

Global [cuD3D10GetDirect3DDevice](#) This function is deprecated as of Cuda 5.0.

Global [cuD3D10MapResources](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10RegisterResource](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10ResourceGetMappedArray](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10ResourceGetMappedPitch](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10ResourceGetMappedPointer](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10ResourceGetMappedSize](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10ResourceGetSurfaceDimensions](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10ResourceSetMapFlags](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10UnmapResources](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D10UnregisterResource](#) This function is deprecated as of Cuda 3.0.

Global [cuD3D11CtxCreate](#) This function is deprecated as of Cuda 5.0.

Global [cuD3D11CtxCreateOnDevice](#) This function is deprecated as of Cuda 5.0.

Global [cuD3D11GetDirect3DDevice](#) This function is deprecated as of Cuda 5.0.

Chapter 3

Module Index

3.1 Modules

Here is a list of all modules:

Device Management	15
Thread Management [DEPRECATED]	35
Error Handling	40
Stream Management	42
Event Management	47
Execution Control	51
Memory Management	56
Unified Addressing	96
Peer Device Memory Access	99
OpenGL Interoperability	101
OpenGL Interoperability [DEPRECATED]	105
Direct3D 9 Interoperability	111
Direct3D 9 Interoperability [DEPRECATED]	116
Direct3D 10 Interoperability	125
Direct3D 10 Interoperability [DEPRECATED]	129
Direct3D 11 Interoperability	139
Direct3D 11 Interoperability [DEPRECATED]	143
VDPAU Interoperability	145
Graphics Interoperability	148
Texture Reference Management	153
Surface Reference Management	159
Texture Object Management	161
Surface Object Management	167
Version Management	169
C++ API Routines	170
Interactions with the CUDA Driver API	187
Profiler Control	189
Data types used by CUDA Runtime	191
Data types used by CUDA driver	214
Initialization	242
Version Management	243
Device Management	244
Device Management [DEPRECATED]	250

Context Management	252
Context Management [DEPRECATED]	263
Module Management	265
Memory Management	272
Unified Addressing	329
Stream Management	333
Event Management	337
Execution Control	341
Execution Control [DEPRECATED]	346
Texture Reference Management	353
Texture Reference Management [DEPRECATED]	366
Surface Reference Management	368
Texture Object Management	370
Surface Object Management	376
Peer Context Memory Access	378
Graphics Interoperability	380
Profiler Control	385
OpenGL Interoperability	387
OpenGL Interoperability [DEPRECATED]	392
Direct3D 9 Interoperability	399
Direct3D 9 Interoperability [DEPRECATED]	405
Direct3D 10 Interoperability	414
Direct3D 10 Interoperability [DEPRECATED]	418
Direct3D 11 Interoperability	429
Direct3D 11 Interoperability [DEPRECATED]	433
VDPAU Interoperability	435
Mathematical Functions	439
Single Precision Mathematical Functions	440
Double Precision Mathematical Functions	467
Single Precision Intrinsic	493
Double Precision Intrinsic	505
Integer Intrinsic	513
Type Casting Intrinsic	518

Chapter 4

Data Structure Index

4.1 Data Structures

Here are the data structures with brief descriptions:

CUDA_ARRAY3D_DESCRIPTOR_st	535
CUDA_ARRAY_DESCRIPTOR_st	537
CUDA_MEMCPY2D_st	538
CUDA_MEMCPY3D_PEER_st	540
CUDA_MEMCPY3D_st	543
CUDA_POINTER_ATTRIBUTE_P2P_TOKENS_st	546
CUDA_RESOURCE_DESC_st	547
CUDA_RESOURCE_VIEW_DESC_st	549
CUDA_TEXTURE_DESC_st	551
cudaChannelFormatDesc	553
cudaDeviceProp	554
cudaExtent	560
cudaFuncAttributes	561
cudaIpcEventHandle_st	563
cudaIpcMemHandle_st	564
cudaMemcpy3DParms	565
cudaMemcpy3DPeerParms	567
cudaPitchedPtr	569
cudaPointerAttributes	570
cudaPos	571
cudaResourceDesc	572
cudaResourceViewDesc	574
cudaTextureDesc	576
CUdevprop_st	578
CUipcEventHandle_st	580
CUipcMemHandle_st	581
surfaceReference	582
textureReference	583

Chapter 5

Module Documentation

5.1 Device Management

Functions

- [cudaError_t cudaChooseDevice](#) (int *device, const struct [cudaDeviceProp](#) *prop)
Select compute-device which best matches criteria.
- [__cuda_builtin__ cudaError_t cudaDeviceGetAttribute](#) (int *value, enum [cudaDeviceAttr](#) attr, int device)
Returns information about the device.
- [cudaError_t cudaDeviceGetByPCIBusId](#) (int *device, char *pciBusId)
Returns a handle to a compute device.
- [__cuda_builtin__ cudaError_t cudaDeviceGetCacheConfig](#) (enum [cudaFuncCache](#) *pCacheConfig)
Returns the preferred cache configuration for the current device.
- [__cuda_builtin__ cudaError_t cudaDeviceGetLimit](#) (size_t *pValue, enum [cudaLimit](#) limit)
Returns resource limits.
- [cudaError_t cudaDeviceGetPCIBusId](#) (char *pciBusId, int len, int device)
Returns a PCI Bus Id string for the device.
- [__cuda_builtin__ cudaError_t cudaDeviceGetSharedMemConfig](#) (enum [cudaSharedMemConfig](#) *pConfig)
Returns the shared memory configuration for the current device.
- [cudaError_t cudaDeviceReset](#) (void)
Destroy all allocations and reset all state on the current device in the current process.
- [cudaError_t cudaDeviceSetCacheConfig](#) (enum [cudaFuncCache](#) cacheConfig)
Sets the preferred cache configuration for the current device.
- [cudaError_t cudaDeviceSetLimit](#) (enum [cudaLimit](#) limit, size_t value)
Set resource limits.
- [cudaError_t cudaDeviceSetSharedMemConfig](#) (enum [cudaSharedMemConfig](#) config)

Sets the shared memory configuration for the current device.

- `__cuda_builtin__ cudaError_t cudaDeviceSynchronize` (void)
Wait for compute device to finish.
- `__cuda_builtin__ cudaError_t cudaGetDevice` (int *device)
Returns which device is currently being used.
- `__cuda_builtin__ cudaError_t cudaGetDeviceCount` (int *count)
Returns the number of compute-capable devices.
- `__cuda_builtin__ cudaError_t cudaGetDeviceProperties` (struct `cudaDeviceProp` *prop, int device)
Returns information about the compute-device.
- `cudaError_t cudaIpcCloseMemHandle` (void *devPtr)
Close memory mapped with `cudaIpcOpenMemHandle`.
- `cudaError_t cudaIpcGetEventHandle` (`cudaIpcEventHandle_t` *handle, `cudaEvent_t` event)
Gets an interprocess handle for a previously allocated event.
- `cudaError_t cudaIpcGetMemHandle` (`cudaIpcMemHandle_t` *handle, void *devPtr)
- `cudaError_t cudaIpcOpenEventHandle` (`cudaEvent_t` *event, `cudaIpcEventHandle_t` handle)
Opens an interprocess event handle for use in the current process.
- `cudaError_t cudaIpcOpenMemHandle` (void **devPtr, `cudaIpcMemHandle_t` handle, unsigned int flags)
- `cudaError_t cudaSetDevice` (int device)
Set device to be used for GPU executions.
- `cudaError_t cudaSetDeviceFlags` (unsigned int flags)
Sets flags to be used for device executions.
- `cudaError_t cudaSetValidDevices` (int *device_arr, int len)
Set a list of devices that can be used for CUDA.

5.1.1 Detailed Description

This section describes the device management functions of the CUDA runtime application programming interface.

5.1.2 Function Documentation

5.1.2.1 `cudaError_t cudaChooseDevice` (int * device, const struct `cudaDeviceProp` * prop)

Returns in *device the device which has properties that best match *prop.

Parameters:

device - Device with best match

prop - Desired device properties

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGetDeviceCount](#), [cudaGetDevice](#), [cudaSetDevice](#), [cudaGetDeviceProperties](#)

5.1.2.2 `__cuda_builtin__ cudaError_t cudaDeviceGetAttribute (int * value, enum cudaDeviceAttr attr, int device)`

Returns in `*value` the integer value of the attribute `attr` on device `device`. The supported attributes are:

- [cudaDevAttrMaxThreadsPerBlock](#): Maximum number of threads per block;
- [cudaDevAttrMaxBlockDimX](#): Maximum x-dimension of a block;
- [cudaDevAttrMaxBlockDimY](#): Maximum y-dimension of a block;
- [cudaDevAttrMaxBlockDimZ](#): Maximum z-dimension of a block;
- [cudaDevAttrMaxGridDimX](#): Maximum x-dimension of a grid;
- [cudaDevAttrMaxGridDimY](#): Maximum y-dimension of a grid;
- [cudaDevAttrMaxGridDimZ](#): Maximum z-dimension of a grid;
- [cudaDevAttrMaxSharedMemoryPerBlock](#): Maximum amount of shared memory available to a thread block in bytes; this amount is shared by all thread blocks simultaneously resident on a multiprocessor;
- [cudaDevAttrTotalConstantMemory](#): Memory available on device for `__constant__` variables in a CUDA C kernel in bytes;
- [cudaDevAttrWarpSize](#): Warp size in threads;
- [cudaDevAttrMaxPitch](#): Maximum pitch in bytes allowed by the memory copy functions that involve memory regions allocated through [cudaMallocPitch\(\)](#);
- [cudaDevAttrMaxTexture1DWidth](#): Maximum 1D texture width;
- [cudaDevAttrMaxTexture1DLinearWidth](#): Maximum width for a 1D texture bound to linear memory;
- [cudaDevAttrMaxTexture1DMipmappedWidth](#): Maximum mipmapped 1D texture width;
- [cudaDevAttrMaxTexture2DWidth](#): Maximum 2D texture width;
- [cudaDevAttrMaxTexture2DHeight](#): Maximum 2D texture height;
- [cudaDevAttrMaxTexture2DLinearWidth](#): Maximum width for a 2D texture bound to linear memory;
- [cudaDevAttrMaxTexture2DLinearHeight](#): Maximum height for a 2D texture bound to linear memory;
- [cudaDevAttrMaxTexture2DLinearPitch](#): Maximum pitch in bytes for a 2D texture bound to linear memory;
- [cudaDevAttrMaxTexture2DMipmappedWidth](#): Maximum mipmapped 2D texture width;
- [cudaDevAttrMaxTexture2DMipmappedHeight](#): Maximum mipmapped 2D texture height;

- `cudaDevAttrMaxTexture3DWidth`: Maximum 3D texture width;
- `cudaDevAttrMaxTexture3DHeight`: Maximum 3D texture height;
- `cudaDevAttrMaxTexture3DDepth`: Maximum 3D texture depth;
- `cudaDevAttrMaxTexture3DWidthAlt`: Alternate maximum 3D texture width, 0 if no alternate maximum 3D texture size is supported;
- `cudaDevAttrMaxTexture3DHeightAlt`: Alternate maximum 3D texture height, 0 if no alternate maximum 3D texture size is supported;
- `cudaDevAttrMaxTexture3DDepthAlt`: Alternate maximum 3D texture depth, 0 if no alternate maximum 3D texture size is supported;
- `cudaDevAttrMaxTextureCubemapWidth`: Maximum cubemap texture width or height;
- `cudaDevAttrMaxTexture1DLayeredWidth`: Maximum 1D layered texture width;
- `cudaDevAttrMaxTexture1DLayeredLayers`: Maximum layers in a 1D layered texture;
- `cudaDevAttrMaxTexture2DLayeredWidth`: Maximum 2D layered texture width;
- `cudaDevAttrMaxTexture2DLayeredHeight`: Maximum 2D layered texture height;
- `cudaDevAttrMaxTexture2DLayeredLayers`: Maximum layers in a 2D layered texture;
- `cudaDevAttrMaxTextureCubemapLayeredWidth`: Maximum cubemap layered texture width or height;
- `cudaDevAttrMaxTextureCubemapLayeredLayers`: Maximum layers in a cubemap layered texture;
- `cudaDevAttrMaxSurface1DWidth`: Maximum 1D surface width;
- `cudaDevAttrMaxSurface2DWidth`: Maximum 2D surface width;
- `cudaDevAttrMaxSurface2DHeight`: Maximum 2D surface height;
- `cudaDevAttrMaxSurface3DWidth`: Maximum 3D surface width;
- `cudaDevAttrMaxSurface3DHeight`: Maximum 3D surface height;
- `cudaDevAttrMaxSurface3DDepth`: Maximum 3D surface depth;
- `cudaDevAttrMaxSurface1DLayeredWidth`: Maximum 1D layered surface width;
- `cudaDevAttrMaxSurface1DLayeredLayers`: Maximum layers in a 1D layered surface;
- `cudaDevAttrMaxSurface2DLayeredWidth`: Maximum 2D layered surface width;
- `cudaDevAttrMaxSurface2DLayeredHeight`: Maximum 2D layered surface height;
- `cudaDevAttrMaxSurface2DLayeredLayers`: Maximum layers in a 2D layered surface;
- `cudaDevAttrMaxSurfaceCubemapWidth`: Maximum cubemap surface width;
- `cudaDevAttrMaxSurfaceCubemapLayeredWidth`: Maximum cubemap layered surface width;
- `cudaDevAttrMaxSurfaceCubemapLayeredLayers`: Maximum layers in a cubemap layered surface;
- `cudaDevAttrMaxRegistersPerBlock`: Maximum number of 32-bit registers available to a thread block; this number is shared by all thread blocks simultaneously resident on a multiprocessor;
- `cudaDevAttrClockRate`: Peak clock frequency in kilohertz;

- `cudaDevAttrTextureAlignment`: Alignment requirement; texture base addresses aligned to `textureAlign` bytes do not need an offset applied to texture fetches;
- `cudaDevAttrTexturePitchAlignment`: Pitch alignment requirement for 2D texture references bound to pitched memory;
- `cudaDevAttrGpuOverlap`: 1 if the device can concurrently copy memory between host and device while executing a kernel, or 0 if not;
- `cudaDevAttrMultiProcessorCount`: Number of multiprocessors on the device;
- `cudaDevAttrKernelExecTimeout`: 1 if there is a run time limit for kernels executed on the device, or 0 if not;
- `cudaDevAttrIntegrated`: 1 if the device is integrated with the memory subsystem, or 0 if not;
- `cudaDevAttrCanMapHostMemory`: 1 if the device can map host memory into the CUDA address space, or 0 if not;
- `cudaDevAttrComputeMode`: Compute mode is the compute mode that the device is currently in. Available modes are as follows:
 - `cudaComputeModeDefault`: Default mode - Device is not restricted and multiple threads can use `cudaSetDevice()` with this device.
 - `cudaComputeModeExclusive`: Compute-exclusive mode - Only one thread will be able to use `cudaSetDevice()` with this device.
 - `cudaComputeModeProhibited`: Compute-prohibited mode - No threads can use `cudaSetDevice()` with this device.
 - `cudaComputeModeExclusiveProcess`: Compute-exclusive-process mode - Many threads in one process will be able to use `cudaSetDevice()` with this device.
- `cudaDevAttrConcurrentKernels`: 1 if the device supports executing multiple kernels within the same context simultaneously, or 0 if not. It is not guaranteed that multiple kernels will be resident on the device concurrently so this feature should not be relied upon for correctness;
- `cudaDevAttrEccEnabled`: 1 if error correction is enabled on the device, 0 if error correction is disabled or not supported by the device;
- `cudaDevAttrPciBusId`: PCI bus identifier of the device;
- `cudaDevAttrPciDeviceId`: PCI device (also known as slot) identifier of the device;
- `cudaDevAttrTccDriver`: 1 if the device is using a TCC driver. TCC is only available on Tesla hardware running Windows Vista or later;
- `cudaDevAttrMemoryClockRate`: Peak memory clock frequency in kilohertz;
- `cudaDevAttrGlobalMemoryBusWidth`: Global memory bus width in bits;
- `cudaDevAttrL2CacheSize`: Size of L2 cache in bytes. 0 if the device doesn't have L2 cache;
- `cudaDevAttrMaxThreadsPerMultiProcessor`: Maximum resident threads per multiprocessor;
- `cudaDevAttrUnifiedAddressing`: 1 if the device shares a unified address space with the host, or 0 if not;
- `cudaDevAttrComputeCapabilityMajor`: Major compute capability version number;
- `cudaDevAttrComputeCapabilityMinor`: Minor compute capability version number;

Parameters:

value - Returned device attribute value

attr - Device attribute to query
device - Device number to query

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGetDeviceCount](#), [cudaGetDevice](#), [cudaSetDevice](#), [cudaChooseDevice](#), [cudaGetDeviceProperties](#)

5.1.2.3 `cudaError_t cudaDeviceGetByPCIBusId (int * device, char * pciBusId)`

Returns in **device* a device ordinal given a PCI bus ID string.

Parameters:

device - Returned device ordinal
pciBusId - String in one of the following forms: [domain]:[bus]:[device].[function] [domain]:[bus]:[device]
 [bus]:[device].[function] where *domain*, *bus*, *device*, and *function* are all hexadecimal values

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceGetPCIBusId](#)

5.1.2.4 `__cuda_builtin__ cudaError_t cudaDeviceGetCacheConfig (enum cudaFuncCache * pCacheConfig)`

On devices where the L1 cache and shared memory use the same hardware resources, this returns through *pCacheConfig* the preferred cache configuration for the current device. This is only a preference. The runtime will use the requested configuration if possible, but it is free to choose a different configuration if required to execute functions.

This will return a *pCacheConfig* of [cudaFuncCachePreferNone](#) on devices where the size of the L1 cache and shared memory are fixed.

The supported cache configurations are:

- [cudaFuncCachePreferNone](#): no preference for shared memory or L1 (default)
- [cudaFuncCachePreferShared](#): prefer larger shared memory and smaller L1 cache
- [cudaFuncCachePreferL1](#): prefer larger L1 cache and smaller shared memory

Parameters:

pCacheConfig - Returned cache configuration

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceSetCacheConfig](#), [cudaFuncSetCacheConfig \(C API\)](#), [cudaFuncSetCacheConfig \(C++ API\)](#)

5.1.2.5 `__cuda_builtin__ cudaError_t cudaDeviceGetLimit (size_t * pValue, enum cudaLimit limit)`

Returns in *pValue the current size of limit. The supported [cudaLimit](#) values are:

- [cudaLimitStackSize](#): stack size in bytes of each GPU thread;
- [cudaLimitPrintfFifoSize](#): size in bytes of the shared FIFO used by the printf() and fprintf() device system calls.
- [cudaLimitMallocHeapSize](#): size in bytes of the heap used by the malloc() and free() device system calls;
- [cudaLimitDevRuntimeSyncDepth](#): maximum grid depth at which a thread can issue the device runtime call [cudaDeviceSynchronize\(\)](#) to wait on child grid launches to complete.
- [cudaLimitDevRuntimePendingLaunchCount](#): maximum number of outstanding device runtime launches.

Parameters:

limit - Limit to query

pValue - Returned size of the limit

Returns:

[cudaSuccess](#), [cudaErrorUnsupportedLimit](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceSetLimit](#)

5.1.2.6 `cudaError_t cudaDeviceGetPCIBusId (char * pciBusId, int len, int device)`

Returns an ASCII string identifying the device dev in the NULL-terminated string pointed to by pciBusId. len specifies the maximum length of the string that may be returned.

Parameters:

pciBusId - Returned identifier string for the device in the following format [domain]:[bus]:[device].[function] where domain, bus, device, and function are all hexadecimal values. pciBusId should be large enough to store 13 characters including the NULL-terminator.

len - Maximum length of string to store in *name*

device - Device to get identifier string for

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceGetByPCIBusId](#)

5.1.2.7 `__cudart_builtin__ cudaError_t cudaDeviceGetSharedMemConfig (enum cudaSharedMemConfig * pConfig)`

This function will return in `pConfig` the current size of shared memory banks on the current device. On devices with configurable shared memory banks, [cudaDeviceSetSharedMemConfig](#) can be used to change this setting, so that all subsequent kernel launches will by default use the new bank size. When [cudaDeviceGetSharedMemConfig](#) is called on devices without configurable shared memory, it will return the fixed bank size of the hardware.

The returned bank configurations can be either:

- `cudaSharedMemBankSizeFourByte` - shared memory bank width is four bytes.
- `cudaSharedMemBankSizeEightByte` - shared memory bank width is eight bytes.

Parameters:

pConfig - Returned cache configuration

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceSetCacheConfig](#), [cudaDeviceGetCacheConfig](#), [cudaDeviceSetSharedMemConfig](#), [cudaFuncSetCacheConfig](#)

5.1.2.8 `cudaError_t cudaDeviceReset (void)`

Explicitly destroys and cleans up all resources associated with the current device in the current process. Any subsequent API call to this device will reinitialize the device.

Note that this function will reset the device immediately. It is the caller's responsibility to ensure that the device is not being accessed by any other host threads from the process when this function is called.

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceSynchronize](#)

5.1.2.9 `cudaError_t cudaDeviceSetCacheConfig (enum cudaFuncCache cacheConfig)`

On devices where the L1 cache and shared memory use the same hardware resources, this sets through `cacheConfig` the preferred cache configuration for the current device. This is only a preference. The runtime will use the requested configuration if possible, but it is free to choose a different configuration if required to execute the function. Any function preference set via [`cudaFuncSetCacheConfig` \(C API\)](#) or [`cudaFuncSetCacheConfig` \(C++ API\)](#) will be preferred over this device-wide setting. Setting the device-wide cache configuration to [`cudaFuncCachePreferNone`](#) will cause subsequent kernel launches to prefer to not change the cache configuration unless required to launch the kernel.

This setting does nothing on devices where the size of the L1 cache and shared memory are fixed.

Launching a kernel with a different preference than the most recent preference setting may insert a device-side synchronization point.

The supported cache configurations are:

- [`cudaFuncCachePreferNone`](#): no preference for shared memory or L1 (default)
- [`cudaFuncCachePreferShared`](#): prefer larger shared memory and smaller L1 cache
- [`cudaFuncCachePreferL1`](#): prefer larger L1 cache and smaller shared memory

Parameters:

cacheConfig - Requested cache configuration

Returns:

[`cudaSuccess`](#), [`cudaErrorInitializationError`](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[`cudaDeviceGetCacheConfig`](#), [`cudaFuncSetCacheConfig` \(C API\)](#), [`cudaFuncSetCacheConfig` \(C++ API\)](#)

5.1.2.10 `cudaError_t cudaDeviceSetLimit (enum cudaLimit limit, size_t value)`

Setting `limit` to `value` is a request by the application to update the current limit maintained by the device. The driver is free to modify the requested value to meet h/w requirements (this could be clamping to minimum or maximum values, rounding up to nearest element size, etc). The application can use [`cudaDeviceGetLimit\(\)`](#) to find out exactly what the limit has been set to.

Setting each [`cudaLimit`](#) has its own specific restrictions, so each is discussed here.

- [`cudaLimitStackSize`](#) controls the stack size in bytes of each GPU thread. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [`cudaErrorUnsupportedLimit`](#) being returned.

- [cudaLimitPrintfFifoSize](#) controls the size in bytes of the shared FIFO used by the `printf()` and `fprintf()` device system calls. Setting [cudaLimitPrintfFifoSize](#) must be performed before launching any kernel that uses the `printf()` or `fprintf()` device system calls, otherwise [cudaErrorInvalidValue](#) will be returned. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [cudaErrorUnsupportedLimit](#) being returned.
- [cudaLimitMallocHeapSize](#) controls the size in bytes of the heap used by the `malloc()` and `free()` device system calls. Setting [cudaLimitMallocHeapSize](#) must be performed before launching any kernel that uses the `malloc()` or `free()` device system calls, otherwise [cudaErrorInvalidValue](#) will be returned. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [cudaErrorUnsupportedLimit](#) being returned.
- [cudaLimitDevRuntimeSyncDepth](#) controls the maximum nesting depth of a grid at which a thread can safely call [cudaDeviceSynchronize\(\)](#). Setting this limit must be performed before any launch of a kernel that uses the device runtime and calls [cudaDeviceSynchronize\(\)](#) above the default sync depth, two levels of grids. Calls to [cudaDeviceSynchronize\(\)](#) will fail with error code [cudaErrorSyncDepthExceeded](#) if the limitation is violated. This limit can be set smaller than the default or up the maximum launch depth of 24. When setting this limit, keep in mind that additional levels of sync depth require the runtime to reserve large amounts of device memory which can no longer be used for user allocations. If these reservations of device memory fail, [cudaDeviceSetLimit](#) will return [cudaErrorMemoryAllocation](#), and the limit can be reset to a lower value. This limit is only applicable to devices of compute capability 3.5 and higher. Attempting to set this limit on devices of compute capability less than 3.5 will result in the error [cudaErrorUnsupportedLimit](#) being returned.
- [cudaLimitDevRuntimePendingLaunchCount](#) controls the maximum number of outstanding device runtime launches that can be made from the current device. A grid is outstanding from the point of launch up until the grid is known to have been completed. Device runtime launches which violate this limitation fail and return [cudaErrorLaunchPendingCountExceeded](#) when [cudaGetLastError\(\)](#) is called after launch. If more pending launches than the default (2048 launches) are needed for a module using the device runtime, this limit can be increased. Keep in mind that being able to sustain additional pending launches will require the runtime to reserve larger amounts of device memory upfront which can no longer be used for allocations. If these reservations fail, [cudaDeviceSetLimit](#) will return [cudaErrorMemoryAllocation](#), and the limit can be reset to a lower value. This limit is only applicable to devices of compute capability 3.5 and higher. Attempting to set this limit on devices of compute capability less than 3.5 will result in the error [cudaErrorUnsupportedLimit](#) being returned.

Parameters:

limit - Limit to set
value - Size of limit

Returns:

[cudaSuccess](#), [cudaErrorUnsupportedLimit](#), [cudaErrorInvalidValue](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceGetLimit](#)

5.1.2.11 `cudaError_t cudaDeviceSetSharedMemConfig` (enum `cudaSharedMemConfig config`)

On devices with configurable shared memory banks, this function will set the shared memory bank size which is used for all subsequent kernel launches. Any per-function setting of shared memory set via [cudaFuncSetSharedMemConfig](#) will override the device wide setting.

Changing the shared memory configuration between launches may introduce a device side synchronization point.

Changing the shared memory bank size will not increase shared memory usage or affect occupancy of kernels, but may have major effects on performance. Larger bank sizes will allow for greater potential bandwidth to shared memory, but will change what kinds of accesses to shared memory will result in bank conflicts.

This function will do nothing on devices with fixed shared memory bank size.

The supported bank configurations are:

- `cudaSharedMemBankSizeDefault`: set bank width the device default (currently, four bytes)
- `cudaSharedMemBankSizeFourByte`: set shared memory bank width to be four bytes natively.
- `cudaSharedMemBankSizeEightByte`: set shared memory bank width to be eight bytes natively.

Parameters:

config - Requested cache configuration

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceSetCacheConfig](#), [cudaDeviceGetCacheConfig](#), [cudaDeviceGetSharedMemConfig](#), [cudaFuncSetCacheConfig](#)

5.1.2.12 `__cuda_builtin__ cudaError_t cudaDeviceSynchronize (void)`

Blocks until the device has completed all preceding requested tasks. `cudaDeviceSynchronize()` returns an error if one of the preceding tasks has failed. If the `cudaDeviceScheduleBlockingSync` flag was set for this device, the host thread will block until the device has finished its work.

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceReset](#)

5.1.2.13 `__cuda_builtin__ cudaError_t cudaGetDevice (int * device)`

Returns in `*device` the current device for the calling host thread.

Parameters:

device - Returns the device on which the active host thread executes the device code.

Returns:[cudaSuccess](#)**Note:**

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGetDeviceCount](#), [cudaSetDevice](#), [cudaGetDeviceProperties](#), [cudaChooseDevice](#)

5.1.2.14 `__cuda_builtin__ cudaError_t cudaGetDeviceCount (int * count)`

Returns in `*count` the number of devices with compute capability greater or equal to 1.0 that are available for execution. If there is no such device then `cudaGetDeviceCount()` will return `cudaErrorNoDevice`. If no driver can be loaded to determine if any such devices exist then `cudaGetDeviceCount()` will return `cudaErrorInsufficientDriver`.

Parameters:

`count` - Returns the number of devices with compute capability greater or equal to 1.0

Returns:

[cudaSuccess](#), [cudaErrorNoDevice](#), [cudaErrorInsufficientDriver](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGetDevice](#), [cudaSetDevice](#), [cudaGetDeviceProperties](#), [cudaChooseDevice](#)

5.1.2.15 `__cuda_builtin__ cudaError_t cudaGetDeviceProperties (struct cudaDeviceProp * prop, int device)`

Returns in `*prop` the properties of device `dev`. The `cudaDeviceProp` structure is defined as:

```
struct cudaDeviceProp {
    char name[256];
    size_t totalGlobalMem;
    size_t sharedMemPerBlock;
    int regsPerBlock;
    int warpSize;
    size_t memPitch;
    int maxThreadsPerBlock;
    int maxThreadsDim[3];
    int maxGridSize[3];
    int clockRate;
    size_t totalConstMem;
    int major;
    int minor;
    size_t textureAlignment;
    size_t texturePitchAlignment;
    int deviceOverlap;
    int multiProcessorCount;
    int kernelExecTimeoutEnabled;
```

```

int integrated;
int canMapHostMemory;
int computeMode;
int maxTexture1D;
int maxTexture1DMipmap;
int maxTexture1DLinear;
int maxTexture2D[2];
int maxTexture2DMipmap[2];
int maxTexture2DLinear[3];
int maxTexture2DGather[2];
int maxTexture3D[3];
int maxTextureCubemap;
int maxTexture1DLayered[2];
int maxTexture2DLayered[3];
int maxTextureCubemapLayered[2];
int maxSurface1D;
int maxSurface2D[2];
int maxSurface3D[3];
int maxSurface1DLayered[2];
int maxSurface2DLayered[3];
int maxSurfaceCubemap;
int maxSurfaceCubemapLayered[2];
size_t surfaceAlignment;
int concurrentKernels;
int ECCEnabled;
int pciBusID;
int pciDeviceID;
int pciDomainID;
int tccDriver;
int asyncEngineCount;
int unifiedAddressing;
int memoryClockRate;
int memoryBusWidth;
int l2CacheSize;
int maxThreadsPerMultiProcessor;
}

```

where:

- `name[256]` is an ASCII string identifying the device;
- `totalGlobalMem` is the total amount of global memory available on the device in bytes;
- `sharedMemPerBlock` is the maximum amount of shared memory available to a thread block in bytes; this amount is shared by all thread blocks simultaneously resident on a multiprocessor;
- `regsPerBlock` is the maximum number of 32-bit registers available to a thread block; this number is shared by all thread blocks simultaneously resident on a multiprocessor;
- `warpSize` is the warp size in threads;
- `memPitch` is the maximum pitch in bytes allowed by the memory copy functions that involve memory regions allocated through `cudaMallocPitch()`;
- `maxThreadsPerBlock` is the maximum number of threads per block;
- `maxThreadsDim[3]` contains the maximum size of each dimension of a block;
- `maxGridSize[3]` contains the maximum size of each dimension of a grid;
- `clockRate` is the clock frequency in kilohertz;
- `totalConstMem` is the total amount of constant memory available on the device in bytes;
- `major`, `minor` are the major and minor revision numbers defining the device's compute capability;

- [textureAlignment](#) is the alignment requirement; texture base addresses that are aligned to [textureAlignment](#) bytes do not need an offset applied to texture fetches;
- [texturePitchAlignment](#) is the pitch alignment requirement for 2D texture references that are bound to pitched memory;
- [deviceOverlap](#) is 1 if the device can concurrently copy memory between host and device while executing a kernel, or 0 if not. Deprecated, use instead [asyncEngineCount](#).
- [multiProcessorCount](#) is the number of multiprocessors on the device;
- [kernelExecTimeoutEnabled](#) is 1 if there is a run time limit for kernels executed on the device, or 0 if not.
- [integrated](#) is 1 if the device is an integrated (motherboard) GPU and 0 if it is a discrete (card) component.
- [canMapHostMemory](#) is 1 if the device can map host memory into the CUDA address space for use with [cudaHostAlloc\(\)/cudaHostGetDevicePointer\(\)](#), or 0 if not;
- [computeMode](#) is the compute mode that the device is currently in. Available modes are as follows:
 - [cudaComputeModeDefault](#): Default mode - Device is not restricted and multiple threads can use [cudaSetDevice\(\)](#) with this device.
 - [cudaComputeModeExclusive](#): Compute-exclusive mode - Only one thread will be able to use [cudaSetDevice\(\)](#) with this device.
 - [cudaComputeModeProhibited](#): Compute-prohibited mode - No threads can use [cudaSetDevice\(\)](#) with this device.
 - [cudaComputeModeExclusiveProcess](#): Compute-exclusive-process mode - Many threads in one process will be able to use [cudaSetDevice\(\)](#) with this device.

If [cudaSetDevice\(\)](#) is called on an already occupied device with computeMode [cudaComputeModeExclusive](#), [cudaErrorDeviceAlreadyInUse](#) will be immediately returned indicating the device cannot be used. When an occupied exclusive mode device is chosen with [cudaSetDevice](#), all subsequent non-device management runtime functions will return [cudaErrorDevicesUnavailable](#).
- [maxTexture1D](#) is the maximum 1D texture size.
- [maxTexture1DMipmap](#) is the maximum 1D mipmapped texture texture size.
- [maxTexture1DLinear](#) is the maximum 1D texture size for textures bound to linear memory.
- [maxTexture2D\[2\]](#) contains the maximum 2D texture dimensions.
- [maxTexture2DMipmap\[2\]](#) contains the maximum 2D mipmapped texture dimensions.
- [maxTexture2DLinear\[3\]](#) contains the maximum 2D texture dimensions for 2D textures bound to pitch linear memory.
- [maxTexture2DGather\[2\]](#) contains the maximum 2D texture dimensions if texture gather operations have to be performed.
- [maxTexture3D\[3\]](#) contains the maximum 3D texture dimensions.
- [maxTextureCubemap](#) is the maximum cubemap texture width or height.
- [maxTexture1DLayered\[2\]](#) contains the maximum 1D layered texture dimensions.
- [maxTexture2DLayered\[3\]](#) contains the maximum 2D layered texture dimensions.
- [maxTextureCubemapLayered\[2\]](#) contains the maximum cubemap layered texture dimensions.
- [maxSurface1D](#) is the maximum 1D surface size.

- [maxSurface2D\[2\]](#) contains the maximum 2D surface dimensions.
- [maxSurface3D\[3\]](#) contains the maximum 3D surface dimensions.
- [maxSurface1DLayered\[2\]](#) contains the maximum 1D layered surface dimensions.
- [maxSurface2DLayered\[3\]](#) contains the maximum 2D layered surface dimensions.
- [maxSurfaceCubemap](#) is the maximum cubemap surface width or height.
- [maxSurfaceCubemapLayered\[2\]](#) contains the maximum cubemap layered surface dimensions.
- [surfaceAlignment](#) specifies the alignment requirements for surfaces.
- [concurrentKernels](#) is 1 if the device supports executing multiple kernels within the same context simultaneously, or 0 if not. It is not guaranteed that multiple kernels will be resident on the device concurrently so this feature should not be relied upon for correctness;
- [ECCEnabled](#) is 1 if the device has ECC support turned on, or 0 if not.
- [pciBusID](#) is the PCI bus identifier of the device.
- [pciDeviceID](#) is the PCI device (sometimes called slot) identifier of the device.
- [pciDomainID](#) is the PCI domain identifier of the device.
- [tccDriver](#) is 1 if the device is using a TCC driver or 0 if not.
- [asyncEngineCount](#) is 1 when the device can concurrently copy memory between host and device while executing a kernel. It is 2 when the device can concurrently copy memory between host and device in both directions and execute a kernel at the same time. It is 0 if neither of these is supported.
- [unifiedAddressing](#) is 1 if the device shares a unified address space with the host and 0 otherwise.
- [memoryClockRate](#) is the peak memory clock frequency in kilohertz.
- [memoryBusWidth](#) is the memory bus width in bits.
- [l2CacheSize](#) is L2 cache size in bytes.
- [maxThreadsPerMultiProcessor](#) is the number of maximum resident threads per multiprocessor.

Parameters:

prop - Properties for the specified device

device - Device number to get properties for

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#)

See also:

[cudaGetDeviceCount](#), [cudaGetDevice](#), [cudaSetDevice](#), [cudaChooseDevice](#), [cudaDeviceGetAttribute](#)

5.1.2.16 `cudaError_t cudaIpcCloseMemHandle (void * devPtr)`

Unmaps memory returned by `cudaIpcOpenMemHandle`. The original allocation in the exporting process as well as imported mappings in other processes will be unaffected.

Any resources used to enable peer access will be freed if this is the last mapping using them.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

devPtr - Device pointer returned by `cudaIpcOpenMemHandle`

Returns:

`cudaSuccess`, `cudaErrorMapBufferObjectFailed`, `cudaErrorInvalidResourceHandle`,

See also:

`cudaMalloc`, `cudaFree`, `cudaIpcGetEventHandle`, `cudaIpcOpenEventHandle`, `cudaIpcGetMemHandle`, `cudaIpcOpenMemHandle`,

5.1.2.17 `cudaError_t cudaIpcGetEventHandle (cudaIpcEventHandle_t * handle, cudaEvent_t event)`

Takes as input a previously allocated event. This event must have been created with the `cudaEventInterprocess` and `cudaEventDisableTiming` flags set. This opaque handle may be copied into other processes and opened with `cudaIpcOpenEventHandle` to allow efficient hardware synchronization between GPU work in different processes.

After the event has been opened in the importing process, `cudaEventRecord`, `cudaEventSynchronize`, `cudaStreamWaitEvent` and `cudaEventQuery` may be used in either process. Performing operations on the imported event after the exported event has been freed with `cudaEventDestroy` will result in undefined behavior.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

handle - Pointer to a user allocated `cudaIpcEventHandle` in which to return the opaque event handle

event - Event allocated with `cudaEventInterprocess` and `cudaEventDisableTiming` flags.

Returns:

`cudaSuccess`, `cudaErrorInvalidResourceHandle`, `cudaErrorMemoryAllocation`, `cudaErrorMapBufferObjectFailed`

See also:

`cudaEventCreate`, `cudaEventDestroy`, `cudaEventSynchronize`, `cudaEventQuery`, `cudaStreamWaitEvent`, `cudaIpcOpenEventHandle`, `cudaIpcGetMemHandle`, `cudaIpcOpenMemHandle`, `cudaIpcCloseMemHandle`

5.1.2.18 `cudaError_t cudaIpcGetMemHandle (cudaIpcMemHandle_t * handle, void * devPtr)`

/brief Gets an interprocess memory handle for an existing device memory allocation

Takes a pointer to the base of an existing device memory allocation created with `cudaMalloc` and exports it for use in another process. This is a lightweight operation and may be called multiple times on an allocation without adverse effects.

If a region of memory is freed with [cudaFree](#) and a subsequent call to [cudaMalloc](#) returns memory with the same device address, [cudaIpcGetMemHandle](#) will return a unique handle for the new memory.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

handle - Pointer to user allocated [cudaIpcMemHandle](#) to return the handle in.

devPtr - Base pointer to previously allocated device memory

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorMemoryAllocation](#), [cudaErrorMapBufferObjectFailed](#),

See also:

[cudaMalloc](#), [cudaFree](#), [cudaIpcGetEventHandle](#), [cudaIpcOpenEventHandle](#), [cudaIpcOpenMemHandle](#), [cudaIpcCloseMemHandle](#)

5.1.2.19 [cudaError_t cudaIpcOpenEventHandle \(cudaEvent_t * event, cudaIpcEventHandle_t handle\)](#)

Opens an interprocess event handle exported from another process with [cudaIpcGetEventHandle](#). This function returns a [cudaEvent_t](#) that behaves like a locally created event with the [cudaEventDisableTiming](#) flag specified. This event must be freed with [cudaEventDestroy](#).

Performing operations on the imported event after the exported event has been freed with [cudaEventDestroy](#) will result in undefined behavior.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

event - Returns the imported event

handle - Interprocess handle to open

Returns:

[cudaSuccess](#), [cudaErrorMapBufferObjectFailed](#), [cudaErrorInvalidResourceHandle](#)

See also:

[cudaEventCreate](#), [cudaEventDestroy](#), [cudaEventSynchronize](#), [cudaEventQuery](#), [cudaStreamWaitEvent](#), [cudaIpcGetEventHandle](#), [cudaIpcGetMemHandle](#), [cudaIpcOpenMemHandle](#), [cudaIpcCloseMemHandle](#)

5.1.2.20 [cudaError_t cudaIpcOpenMemHandle \(void ** devPtr, cudaIpcMemHandle_t handle, unsigned int flags\)](#)

/brief Opens an interprocess memory handle exported from another process and returns a device pointer usable in the local process.

Maps memory exported from another process with [cudaIpcGetMemHandle](#) into the current device address space. For contexts on different devices [cudaIpcOpenMemHandle](#) can attempt to enable peer access between the devices as if the user called [cudaDeviceEnablePeerAccess](#). This behavior is controlled by the [cudaIpcMemLazyEnablePeerAccess](#) flag. [cudaDeviceCanAccessPeer](#) can determine if a mapping is possible.

Contexts that may open `cudaIpcMemHandles` are restricted in the following way. `cudaIpcMemHandles` from each device in a given process may only be opened by one context per device per other process.

Memory returned from `cudaIpcOpenMemHandle` must be freed with `cudaIpcCloseMemHandle`.

Calling `cudaFree` on an exported memory region before calling `cudaIpcCloseMemHandle` in the importing context will result in undefined behavior.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

devPtr - Returned device pointer

handle - `cudaIpcMemHandle` to open

flags - Flags for this operation. Must be specified as `cudaIpcMemLazyEnablePeerAccess`

Returns:

`cudaSuccess`, `cudaErrorMapBufferObjectFailed`, `cudaErrorInvalidResourceHandle`, `cudaErrorTooManyPeers`

See also:

`cudaMalloc`, `cudaFree`, `cudaIpcGetEventHandle`, `cudaIpcOpenEventHandle`, `cudaIpcGetMemHandle`, `cudaIpcCloseMemHandle`, `cudaDeviceEnablePeerAccess`, `cudaDeviceCanAccessPeer`,

5.1.2.21 `cudaError_t cudaSetDevice (int device)`

Sets `device` as the current device for the calling host thread.

Any device memory subsequently allocated from this host thread using `cudaMalloc()`, `cudaMallocPitch()` or `cudaMallocArray()` will be physically resident on `device`. Any host memory allocated from this host thread using `cudaMallocHost()` or `cudaHostAlloc()` or `cudaHostRegister()` will have its lifetime associated with `device`. Any streams or events created from this host thread will be associated with `device`. Any kernels launched from this host thread using the `<<<<>>>` operator or `cudaLaunch()` will be executed on `device`.

This call may be made from any host thread, to any device, and at any time. This function will do no synchronization with the previous or new device, and should be considered a very low overhead call.

Parameters:

device - Device on which the active host thread should execute the device code.

Returns:

`cudaSuccess`, `cudaErrorInvalidDevice`, `cudaErrorDeviceAlreadyInUse`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaGetDeviceCount`, `cudaGetDevice`, `cudaGetDeviceProperties`, `cudaChooseDevice`

5.1.2.22 `cudaError_t cudaSetDeviceFlags (unsigned int flags)`

Records `flags` as the flags to use when initializing the current device. If no device has been made current to the calling thread then `flags` will be applied to the initialization of any device initialized by the calling host thread, unless that device has had its initialization flags set explicitly by this or any host thread.

If the current device has been set and that device has already been initialized then this call will fail with the error `cudaErrorSetOnActiveProcess`. In this case it is necessary to reset `device` using `cudaDeviceReset()` before the device's initialization flags may be set.

The two LSBs of the `flags` parameter can be used to control how the CPU thread interacts with the OS scheduler when waiting for results from the device.

- `cudaDeviceScheduleAuto`: The default value if the `flags` parameter is zero, uses a heuristic based on the number of active CUDA contexts in the process C and the number of logical processors in the system P . If $C > P$, then CUDA will yield to other OS threads when waiting for the device, otherwise CUDA will not yield while waiting for results and actively spin on the processor.
- `cudaDeviceScheduleSpin`: Instruct CUDA to actively spin when waiting for results from the device. This can decrease latency when waiting for the device, but may lower the performance of CPU threads if they are performing work in parallel with the CUDA thread.
- `cudaDeviceScheduleYield`: Instruct CUDA to yield its thread when waiting for results from the device. This can increase latency when waiting for the device, but can increase the performance of CPU threads performing work in parallel with the device.
- `cudaDeviceScheduleBlockingSync`: Instruct CUDA to block the CPU thread on a synchronization primitive when waiting for the device to finish work.
- `cudaDeviceBlockingSync`: Instruct CUDA to block the CPU thread on a synchronization primitive when waiting for the device to finish work.
Deprecated: This flag was deprecated as of CUDA 4.0 and replaced with `cudaDeviceScheduleBlockingSync`.
- `cudaDeviceMapHost`: This flag must be set in order to allocate pinned host memory that is accessible to the device. If this flag is not set, `cudaHostGetDevicePointer()` will always return a failure code.
- `cudaDeviceLmemResizeToMax`: Instruct CUDA to not reduce local memory after resizing local memory for a kernel. This can prevent thrashing by local memory allocations when launching many kernels with high local memory usage at the cost of potentially increased memory usage.

Parameters:

`flags` - Parameters for device operation

Returns:

`cudaSuccess`, `cudaErrorInvalidDevice`, `cudaErrorSetOnActiveProcess`

See also:

`cudaGetDeviceCount`, `cudaGetDevice`, `cudaGetDeviceProperties`, `cudaSetDevice`, `cudaSetValidDevices`, `cudaChooseDevice`

5.1.2.23 `cudaError_t cudaSetValidDevices (int * device_arr, int len)`

Sets a list of devices for CUDA execution in priority order using `device_arr`. The parameter `len` specifies the number of elements in the list. CUDA will try devices from the list sequentially until it finds one that works. If this function is not called, or if it is called with a `len` of 0, then CUDA will go back to its default behavior of trying devices sequentially from a default list containing all of the available CUDA devices in the system. If a specified device ID in the list does not exist, this function will return `cudaErrorInvalidDevice`. If `len` is not 0 and `device_arr` is NULL or if `len` exceeds the number of devices in the system, then `cudaErrorInvalidValue` is returned.

Parameters:

device_arr - List of devices to try

len - Number of devices in specified list

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGetDeviceCount](#), [cudaSetDevice](#), [cudaGetDeviceProperties](#), [cudaSetDeviceFlags](#), [cudaChooseDevice](#)

5.2 Thread Management [DEPRECATED]

Functions

- `cudaError_t cudaThreadExit` (void)
Exit and clean up from CUDA launches.
- `cudaError_t cudaThreadGetCacheConfig` (enum `cudaFuncCache` *pCacheConfig)
Returns the preferred cache configuration for the current device.
- `cudaError_t cudaThreadGetLimit` (size_t *pValue, enum `cudaLimit` limit)
Returns resource limits.
- `cudaError_t cudaThreadSetCacheConfig` (enum `cudaFuncCache` cacheConfig)
Sets the preferred cache configuration for the current device.
- `cudaError_t cudaThreadSetLimit` (enum `cudaLimit` limit, size_t value)
Set resource limits.
- `cudaError_t cudaThreadSynchronize` (void)
Wait for compute device to finish.

5.2.1 Detailed Description

This section describes deprecated thread management functions of the CUDA runtime application programming interface.

5.2.2 Function Documentation

5.2.2.1 `cudaError_t cudaThreadExit` (void)

Deprecated

Note that this function is deprecated because its name does not reflect its behavior. Its functionality is identical to the non-deprecated function `cudaDeviceReset()`, which should be used instead.

Explicitly destroys all cleans up all resources associated with the current device in the current process. Any subsequent API call to this device will reinitialize the device.

Note that this function will reset the device immediately. It is the caller's responsibility to ensure that the device is not being accessed by any other host threads from the process when this function is called.

Returns:

`cudaSuccess`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceReset](#)

5.2.2.2 `cudaError_t cudaThreadGetCacheConfig (enum cudaFuncCache * pCacheConfig)`

Deprecated

Note that this function is deprecated because its name does not reflect its behavior. Its functionality is identical to the non-deprecated function [cudaDeviceGetCacheConfig\(\)](#), which should be used instead.

On devices where the L1 cache and shared memory use the same hardware resources, this returns through `pCacheConfig` the preferred cache configuration for the current device. This is only a preference. The runtime will use the requested configuration if possible, but it is free to choose a different configuration if required to execute functions.

This will return a `pCacheConfig` of [cudaFuncCachePreferNone](#) on devices where the size of the L1 cache and shared memory are fixed.

The supported cache configurations are:

- [cudaFuncCachePreferNone](#): no preference for shared memory or L1 (default)
- [cudaFuncCachePreferShared](#): prefer larger shared memory and smaller L1 cache
- [cudaFuncCachePreferL1](#): prefer larger L1 cache and smaller shared memory

Parameters:

pCacheConfig - Returned cache configuration

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceGetCacheConfig](#)

5.2.2.3 `cudaError_t cudaThreadGetLimit (size_t * pValue, enum cudaLimit limit)`

Deprecated

Note that this function is deprecated because its name does not reflect its behavior. Its functionality is identical to the non-deprecated function [cudaDeviceGetLimit\(\)](#), which should be used instead.

Returns in `*pValue` the current size of `limit`. The supported [cudaLimit](#) values are:

- [cudaLimitStackSize](#): stack size of each GPU thread;

- [cudaLimitPrintfFifoSize](#): size of the shared FIFO used by the `printf()` and `fprintf()` device system calls.
- [cudaLimitMallocHeapSize](#): size of the heap used by the `malloc()` and `free()` device system calls;

Parameters:

limit - Limit to query

pValue - Returned size in bytes of limit

Returns:

[cudaSuccess](#), [cudaErrorUnsupportedLimit](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceGetLimit](#)

5.2.2.4 `cudaError_t cudaThreadSetCacheConfig (enum cudaFuncCache cacheConfig)`

Deprecated

Note that this function is deprecated because its name does not reflect its behavior. Its functionality is identical to the non-deprecated function [cudaDeviceSetCacheConfig\(\)](#), which should be used instead.

On devices where the L1 cache and shared memory use the same hardware resources, this sets through `cacheConfig` the preferred cache configuration for the current device. This is only a preference. The runtime will use the requested configuration if possible, but it is free to choose a different configuration if required to execute the function. Any function preference set via [cudaFuncSetCacheConfig \(C API\)](#) or [cudaFuncSetCacheConfig \(C++ API\)](#) will be preferred over this device-wide setting. Setting the device-wide cache configuration to [cudaFuncCachePreferNone](#) will cause subsequent kernel launches to prefer to not change the cache configuration unless required to launch the kernel.

This setting does nothing on devices where the size of the L1 cache and shared memory are fixed.

Launching a kernel with a different preference than the most recent preference setting may insert a device-side synchronization point.

The supported cache configurations are:

- [cudaFuncCachePreferNone](#): no preference for shared memory or L1 (default)
- [cudaFuncCachePreferShared](#): prefer larger shared memory and smaller L1 cache
- [cudaFuncCachePreferL1](#): prefer larger L1 cache and smaller shared memory

Parameters:

cacheConfig - Requested cache configuration

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceSetCacheConfig](#)

5.2.2.5 `cudaError_t cudaThreadSetLimit (enum cudaLimit limit, size_t value)`**Deprecated**

Note that this function is deprecated because its name does not reflect its behavior. Its functionality is identical to the non-deprecated function [cudaDeviceSetLimit\(\)](#), which should be used instead.

Setting `limit` to `value` is a request by the application to update the current limit maintained by the device. The driver is free to modify the requested value to meet h/w requirements (this could be clamping to minimum or maximum values, rounding up to nearest element size, etc). The application can use [cudaThreadGetLimit\(\)](#) to find out exactly what the limit has been set to.

Setting each [cudaLimit](#) has its own specific restrictions, so each is discussed here.

- [cudaLimitStackSize](#) controls the stack size of each GPU thread. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [cudaErrorUnsupportedLimit](#) being returned.
- [cudaLimitPrintfFifoSize](#) controls the size of the shared FIFO used by the `printf()` and `fprintf()` device system calls. Setting [cudaLimitPrintfFifoSize](#) must be performed before launching any kernel that uses the `printf()` or `fprintf()` device system calls, otherwise [cudaErrorInvalidValue](#) will be returned. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [cudaErrorUnsupportedLimit](#) being returned.
- [cudaLimitMallocHeapSize](#) controls the size of the heap used by the `malloc()` and `free()` device system calls. Setting [cudaLimitMallocHeapSize](#) must be performed before launching any kernel that uses the `malloc()` or `free()` device system calls, otherwise [cudaErrorInvalidValue](#) will be returned. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [cudaErrorUnsupportedLimit](#) being returned.

Parameters:

limit - Limit to set

value - Size in bytes of limit

Returns:

[cudaSuccess](#), [cudaErrorUnsupportedLimit](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceSetLimit](#)

5.2.2.6 `cudaError_t cudaThreadSynchronize (void)`

Deprecated

Note that this function is deprecated because its name does not reflect its behavior. Its functionality is similar to the non-deprecated function [cudaDeviceSynchronize\(\)](#), which should be used instead.

Blocks until the device has completed all preceding requested tasks. [cudaThreadSynchronize\(\)](#) returns an error if one of the preceding tasks has failed. If the [cudaDeviceScheduleBlockingSync](#) flag was set for this device, the host thread will block until the device has finished its work.

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceSynchronize](#)

5.3 Error Handling

Functions

- `__cuda_builtin__ const char * cudaGetErrorString (cudaError_t error)`
Returns the message string from an error code.
- `__cuda_builtin__ cudaError_t cudaGetLastError (void)`
Returns the last error from a runtime call.
- `__cuda_builtin__ cudaError_t cudaPeekAtLastError (void)`
Returns the last error from a runtime call.

5.3.1 Detailed Description

This section describes the error handling functions of the CUDA runtime application programming interface.

5.3.2 Function Documentation

5.3.2.1 `__cuda_builtin__ const char* cudaGetErrorString (cudaError_t error)`

Returns the message string from an error code.

Parameters:

error - Error code to convert to string

Returns:

char* pointer to a NULL-terminated string

See also:

[cudaGetLastError](#), [cudaPeekAtLastError](#), [cudaError](#)

5.3.2.2 `__cuda_builtin__ cudaError_t cudaGetLastError (void)`

Returns the last error that has been produced by any of the runtime calls in the same host thread and resets it to [cudaSuccess](#).

Returns:

[cudaSuccess](#), [cudaErrorMissingConfiguration](#), [cudaErrorMemoryAllocation](#), [cudaErrorInitializationError](#), [cudaErrorLaunchFailure](#), [cudaErrorLaunchTimeout](#), [cudaErrorLaunchOutOfResources](#), [cudaErrorInvalidDeviceFunction](#), [cudaErrorInvalidConfiguration](#), [cudaErrorInvalidDevice](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidSymbol](#), [cudaErrorUnmapBufferObjectFailed](#), [cudaErrorInvalidHostPointer](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#), [cudaErrorInvalidTextureBinding](#), [cudaErrorInvalidChannelDescriptor](#), [cudaErrorInvalidMemcpyDirection](#), [cudaErrorInvalidFilterSetting](#), [cudaErrorInvalidNormSetting](#), [cudaErrorUnknown](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorInsufficientDriver](#), [cudaErrorSetOnActiveProcess](#), [cudaErrorStartupFailure](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaPeekAtLastError](#), [cudaGetErrorString](#), [cudaError](#)

5.3.2.3 `__cuda_builtin__ cudaError_t cudaPeekAtLastError (void)`

Returns the last error that has been produced by any of the runtime calls in the same host thread. Note that this call does not reset the error to [cudaSuccess](#) like [cudaGetLastError\(\)](#).

Returns:

[cudaSuccess](#), [cudaErrorMissingConfiguration](#), [cudaErrorMemoryAllocation](#), [cudaErrorInitializationError](#), [cudaErrorLaunchFailure](#), [cudaErrorLaunchTimeout](#), [cudaErrorLaunchOutOfResources](#), [cudaErrorInvalidDeviceFunction](#), [cudaErrorInvalidConfiguration](#), [cudaErrorInvalidDevice](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidSymbol](#), [cudaErrorUnmapBufferObjectFailed](#), [cudaErrorInvalidHostPointer](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#), [cudaErrorInvalidTextureBinding](#), [cudaErrorInvalidChannelDescriptor](#), [cudaErrorInvalidMemcpyDirection](#), [cudaErrorInvalidFilterSetting](#), [cudaErrorInvalidNormSetting](#), [cudaErrorUnknown](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorInsufficientDriver](#), [cudaErrorSetOnActiveProcess](#), [cudaErrorStartupFailure](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGetLastError](#), [cudaGetErrorString](#), [cudaError](#)

5.4 Stream Management

Typedefs

- typedef void(CUDART_CB * [cudaStreamCallback_t](#))([cudaStream_t](#) stream, [cudaError_t](#) status, void *userData)

Functions

- [cudaError_t cudaStreamAddCallback](#) ([cudaStream_t](#) stream, [cudaStreamCallback_t](#) callback, void *userData, unsigned int flags)
Add a callback to a compute stream.
- [cudaError_t cudaStreamCreate](#) ([cudaStream_t](#) *pStream)
Create an asynchronous stream.
- [__cuda_builtin__ cudaError_t cudaStreamCreateWithFlags](#) ([cudaStream_t](#) *pStream, unsigned int flags)
Create an asynchronous stream.
- [__cuda_builtin__ cudaError_t cudaStreamDestroy](#) ([cudaStream_t](#) stream)
Destroys and cleans up an asynchronous stream.
- [cudaError_t cudaStreamQuery](#) ([cudaStream_t](#) stream)
Queries an asynchronous stream for completion status.
- [cudaError_t cudaStreamSynchronize](#) ([cudaStream_t](#) stream)
Waits for stream tasks to complete.
- [__cuda_builtin__ cudaError_t cudaStreamWaitEvent](#) ([cudaStream_t](#) stream, [cudaEvent_t](#) event, unsigned int flags)
Make a compute stream wait on an event.

5.4.1 Detailed Description

This section describes the stream management functions of the CUDA runtime application programming interface.

5.4.2 Typedef Documentation

5.4.2.1 typedef void(CUDART_CB * [cudaStreamCallback_t](#))([cudaStream_t](#) stream, [cudaError_t](#) status, void *userData)

Type of stream callback functions.

Parameters:

- stream* The stream as passed to [cudaStreamAddCallback](#), may be NULL.
- status* [cudaSuccess](#) or any persistent error on the stream.
- userData* User parameter provided at registration.

5.4.3 Function Documentation

5.4.3.1 `cudaError_t cudaStreamAddCallback(cudaStream_t stream, cudaStreamCallback_t callback, void * userData, unsigned int flags)`

Adds a callback to be called on the host after all currently enqueued items in the stream have completed. For each `cudaStreamAddCallback` call, a callback will be executed exactly once. The callback will block later work in the stream until it is finished.

The callback may be passed `cudaSuccess` or an error code. In the event of a device error, all subsequently executed callbacks will receive an appropriate `cudaError_t`.

Callbacks must not make any CUDA API calls. Attempting to use CUDA APIs will result in `cudaErrorNotPermitted`. Callbacks must not perform any synchronization that may depend on outstanding device work or other callbacks that are not mandated to run earlier. Callbacks without a mandated order (in independent streams) execute in undefined order and may be serialized.

This API requires compute capability 1.1 or greater. See `cudaDeviceGetAttribute` or `cudaGetDeviceProperties` to query compute capability. Calling this API with an earlier compute version will return `cudaErrorNotSupported`.

Parameters:

stream - Stream to add callback to

callback - The function to call once preceding stream operations are complete

userData - User specified data to be passed to the callback function

flags - Reserved for future use, must be 0

Returns:

`cudaSuccess`, `cudaErrorInvalidResourceHandle`, `cudaErrorNotSupported`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaStreamCreate`, `cudaStreamCreateWithFlags`, `cudaStreamQuery`, `cudaStreamSynchronize`, `cudaStreamWaitEvent`, `cudaStreamDestroy`

5.4.3.2 `cudaError_t cudaStreamCreate(cudaStream_t * pStream)`

Creates a new asynchronous stream.

Parameters:

pStream - Pointer to new stream identifier

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaStreamCreate](#), [cudaStreamCreateWithFlags](#), [cudaStreamQuery](#), [cudaStreamSynchronize](#), [cudaStreamWaitEvent](#), [cudaStreamAddCallback](#), [cudaStreamDestroy](#)

5.4.3.3 `__cuda_builtin__ cudaError_t cudaStreamCreateWithFlags (cudaStream_t * pStream, unsigned int flags)`

Creates a new asynchronous stream. The `flags` argument determines the behaviors of the stream. Valid values for `flags` are

- [cudaStreamDefault](#): Default stream creation flag.
- [cudaStreamNonBlocking](#): Specifies that work running in the created stream may run concurrently with work in stream 0 (the NULL stream), and that the created stream should perform no implicit synchronization with stream 0.

Parameters:

pStream - Pointer to new stream identifier

flags - Parameters for stream creation

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaStreamCreate](#), [cudaStreamQuery](#), [cudaStreamSynchronize](#), [cudaStreamWaitEvent](#), [cudaStreamAddCallback](#), [cudaStreamDestroy](#)

5.4.3.4 `__cuda_builtin__ cudaError_t cudaStreamDestroy (cudaStream_t stream)`

Destroys and cleans up the asynchronous stream specified by `stream`.

In case the device is still doing work in the stream `stream` when [cudaStreamDestroy\(\)](#) is called, the function will return immediately and the resources associated with `stream` will be released automatically once the device has completed all work in `stream`.

Parameters:

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaStreamCreate](#), [cudaStreamCreateWithFlags](#), [cudaStreamQuery](#), [cudaStreamWaitEvent](#), [cudaStreamSynchronize](#), [cudaStreamAddCallback](#)

5.4.3.5 `cudaError_t cudaStreamQuery (cudaStream_t stream)`

Returns [cudaSuccess](#) if all operations in `stream` have completed, or [cudaErrorNotReady](#) if not.

Parameters:

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorNotReady](#), [cudaErrorInvalidResourceHandle](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaStreamCreate](#), [cudaStreamCreateWithFlags](#), [cudaStreamWaitEvent](#), [cudaStreamSynchronize](#), [cudaStreamAddCallback](#), [cudaStreamDestroy](#)

5.4.3.6 `cudaError_t cudaStreamSynchronize (cudaStream_t stream)`

Blocks until `stream` has completed all operations. If the [cudaDeviceScheduleBlockingSync](#) flag was set for this device, the host thread will block until the stream is finished with all of its tasks.

Parameters:

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaStreamCreate](#), [cudaStreamCreateWithFlags](#), [cudaStreamQuery](#), [cudaStreamWaitEvent](#), [cudaStreamAddCallback](#), [cudaStreamDestroy](#)

5.4.3.7 `__cuda_builtin__ cudaError_t cudaStreamWaitEvent (cudaStream_t stream, cudaEvent_t event, unsigned int flags)`

Makes all future work submitted to `stream` wait until `event` reports completion before beginning execution. This synchronization will be performed efficiently on the device. The event `event` may be from a different context than `stream`, in which case this function will perform cross-device synchronization.

The stream `stream` will wait only for the completion of the most recent host call to [cudaEventRecord\(\)](#) on `event`. Once this call has returned, any functions (including [cudaEventRecord\(\)](#) and [cudaEventDestroy\(\)](#)) may be called on `event` again, and the subsequent calls will not have any effect on `stream`.

If `stream` is NULL, any future work submitted in any stream will wait for `event` to complete before beginning execution. This effectively creates a barrier for all future work submitted to the device on this thread.

If [cudaEventRecord\(\)](#) has not been called on `event`, this call acts as if the record has already completed, and so is a functional no-op.

Parameters:

stream - Stream to wait

event - Event to wait on

flags - Parameters for the operation (must be 0)

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaStreamCreate](#), [cudaStreamCreateWithFlags](#), [cudaStreamQuery](#), [cudaStreamSynchronize](#), [cudaStreamAddCallback](#), [cudaStreamDestroy](#)

5.5 Event Management

Functions

- [cudaError_t cudaEventCreate \(cudaEvent_t *event\)](#)
Creates an event object.
- [__cuda_builtin__ cudaError_t cudaEventCreateWithFlags \(cudaEvent_t *event, unsigned int flags\)](#)
Creates an event object with the specified flags.
- [__cuda_builtin__ cudaError_t cudaEventDestroy \(cudaEvent_t event\)](#)
Destroys an event object.
- [cudaError_t cudaEventElapsedTime \(float *ms, cudaEvent_t start, cudaEvent_t end\)](#)
Computes the elapsed time between events.
- [cudaError_t cudaEventQuery \(cudaEvent_t event\)](#)
Queries an event's status.
- [__cuda_builtin__ cudaError_t cudaEventRecord \(cudaEvent_t event, cudaStream_t stream=0\)](#)
Records an event.
- [cudaError_t cudaEventSynchronize \(cudaEvent_t event\)](#)
Waits for an event to complete.

5.5.1 Detailed Description

This section describes the event management functions of the CUDA runtime application programming interface.

5.5.2 Function Documentation

5.5.2.1 [cudaError_t cudaEventCreate \(cudaEvent_t * event\)](#)

Creates an event object using [cudaEventDefault](#).

Parameters:

event - Newly created event

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidValue](#), [cudaErrorLaunchFailure](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaEventCreate \(C++ API\)](#), [cudaEventCreateWithFlags](#), [cudaEventRecord](#), [cudaEventQuery](#), [cudaEventSynchronize](#), [cudaEventDestroy](#), [cudaEventElapsedTime](#), [cudaStreamWaitEvent](#)

5.5.2.2 `__cuda_builtin__ cudaError_t cudaEventCreateWithFlags (cudaEvent_t * event, unsigned int flags)`

Creates an event object with the specified flags. Valid flags include:

- `cudaEventDefault`: Default event creation flag.
- `cudaEventBlockingSync`: Specifies that event should use blocking synchronization. A host thread that uses `cudaEventSynchronize()` to wait on an event created with this flag will block until the event actually completes.
- `cudaEventDisableTiming`: Specifies that the created event does not need to record timing data. Events created with this flag specified and the `cudaEventBlockingSync` flag not specified will provide the best performance when used with `cudaStreamWaitEvent()` and `cudaEventQuery()`.
- `cudaEventInterprocess`: Specifies that the created event may be used as an interprocess event by `cudaIpcGetEventHandle()`. `cudaEventInterprocess` must be specified along with `cudaEventDisableTiming`.

Parameters:

event - Newly created event
flags - Flags for new event

Returns:

`cudaSuccess`, `cudaErrorInitializationError`, `cudaErrorInvalidValue`, `cudaErrorLaunchFailure`, `cudaErrorMemoryAllocation`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaEventCreate \(C API\)](#), [cudaEventSynchronize](#), [cudaEventDestroy](#), [cudaEventElapsedTime](#), [cudaStreamWaitEvent](#)

5.5.2.3 `__cuda_builtin__ cudaError_t cudaEventDestroy (cudaEvent_t event)`

Destroys the event specified by *event*.

In case *event* has been recorded but has not yet been completed when `cudaEventDestroy()` is called, the function will return immediately and the resources associated with *event* will be released automatically once the device has completed *event*.

Parameters:

event - Event to destroy

Returns:

`cudaSuccess`, `cudaErrorInitializationError`, `cudaErrorInvalidValue`, `cudaErrorLaunchFailure`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaEventCreate \(C API\)](#), [cudaEventCreateWithFlags](#), [cudaEventQuery](#), [cudaEventSynchronize](#), [cudaEventRecord](#), [cudaEventElapsedTime](#)

5.5.2.4 `cudaError_t cudaEventElapsedTime (float * ms, cudaEvent_t start, cudaEvent_t end)`

Computes the elapsed time between two events (in milliseconds with a resolution of around 0.5 microseconds).

If either event was last recorded in a non-NULL stream, the resulting time may be greater than expected (even if both used the same stream handle). This happens because the `cudaEventRecord()` operation takes place asynchronously and there is no guarantee that the measured latency is actually just between the two events. Any number of other different stream operations could execute in between the two measured events, thus altering the timing in a significant way.

If `cudaEventRecord()` has not been called on either event, then `cudaErrorInvalidResourceHandle` is returned. If `cudaEventRecord()` has been called on both events but one or both of them has not yet been completed (that is, `cudaEventQuery()` would return `cudaErrorNotReady` on at least one of the events), `cudaErrorNotReady` is returned. If either event was created with the `cudaEventDisableTiming` flag, then this function will return `cudaErrorInvalidResourceHandle`.

Parameters:

ms - Time between `start` and `end` in ms
start - Starting event
end - Ending event

Returns:

`cudaSuccess`, `cudaErrorNotReady`, `cudaErrorInvalidValue`, `cudaErrorInitializationError`, `cudaErrorInvalidResourceHandle`, `cudaErrorLaunchFailure`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaEventCreate` (C API), `cudaEventCreateWithFlags`, `cudaEventQuery`, `cudaEventSynchronize`, `cudaEventDestroy`, `cudaEventRecord`

5.5.2.5 `cudaError_t cudaEventQuery (cudaEvent_t event)`

Query the status of all device work preceding the most recent call to `cudaEventRecord()` (in the appropriate compute streams, as specified by the arguments to `cudaEventRecord()`).

If this work has successfully been completed by the device, or if `cudaEventRecord()` has not been called on `event`, then `cudaSuccess` is returned. If this work has not yet been completed by the device then `cudaErrorNotReady` is returned.

Parameters:

event - Event to query

Returns:

`cudaSuccess`, `cudaErrorNotReady`, `cudaErrorInitializationError`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorLaunchFailure`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaEventCreate` (C API), `cudaEventCreateWithFlags`, `cudaEventRecord`, `cudaEventSynchronize`, `cudaEventDestroy`, `cudaEventElapsedTime`

5.5.2.6 `__cudart_builtin__ cudaError_t cudaEventRecord (cudaEvent_t event, cudaStream_t stream = 0)`

Records an event. If `stream` is non-zero, the event is recorded after all preceding operations in `stream` have been completed; otherwise, it is recorded after all preceding operations in the CUDA context have been completed. Since operation is asynchronous, [cudaEventQuery\(\)](#) and/or [cudaEventSynchronize\(\)](#) must be used to determine when the event has actually been recorded.

If [cudaEventRecord\(\)](#) has previously been called on `event`, then this call will overwrite any existing state in `event`. Any subsequent calls which examine the status of `event` will only examine the completion of this most recent call to [cudaEventRecord\(\)](#).

Parameters:

- `event` - Event to record
- `stream` - Stream in which to record event

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInitializationError](#), [cudaErrorInvalidResourceHandle](#), [cudaError-LaunchFailure](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaEventCreate \(C API\)](#), [cudaEventCreateWithFlags](#), [cudaEventQuery](#), [cudaEventSynchronize](#), [cudaEventDestroy](#), [cudaEventElapsedTime](#), [cudaStreamWaitEvent](#)

5.5.2.7 `cudaError_t cudaEventSynchronize (cudaEvent_t event)`

Wait until the completion of all device work preceding the most recent call to [cudaEventRecord\(\)](#) (in the appropriate compute streams, as specified by the arguments to [cudaEventRecord\(\)](#)).

If [cudaEventRecord\(\)](#) has not been called on `event`, [cudaSuccess](#) is returned immediately.

Waiting for an event that was created with the [cudaEventBlockingSync](#) flag will cause the calling CPU thread to block until the event has been completed by the device. If the [cudaEventBlockingSync](#) flag has not been set, then the CPU thread will busy-wait until the event has been completed by the device.

Parameters:

- `event` - Event to wait for

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaError-LaunchFailure](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaEventCreate \(C API\)](#), [cudaEventCreateWithFlags](#), [cudaEventRecord](#), [cudaEventQuery](#), [cudaEventDestroy](#), [cudaEventElapsedTime](#)

5.6 Execution Control

Functions

- `cudaError_t cudaConfigureCall` (`dim3 gridDim`, `dim3 blockDim`, `size_t sharedMem=0`, `cudaStream_t stream=0`)
Configure a device-launch.
- `__cuda_builtin__ cudaError_t cudaFuncGetAttributes` (`struct cudaFuncAttributes *attr`, `const void *func`)
Find out attributes for a given function.
- `cudaError_t cudaFuncSetCacheConfig` (`const void *func`, `enum cudaFuncCache cacheConfig`)
Sets the preferred cache configuration for a device function.
- `cudaError_t cudaFuncSetSharedMemConfig` (`const void *func`, `enum cudaSharedMemConfig config`)
Sets the shared memory configuration for a device function.
- `cudaError_t cudaLaunch` (`const void *func`)
Launches a device function.
- `cudaError_t cudaSetDoubleForDevice` (`double *d`)
Converts a double argument to be executed on a device.
- `cudaError_t cudaSetDoubleForHost` (`double *d`)
Converts a double argument after execution on a device.
- `cudaError_t cudaSetupArgument` (`const void *arg`, `size_t size`, `size_t offset`)
Configure a device launch.

5.6.1 Detailed Description

This section describes the execution control functions of the CUDA runtime application programming interface. Some functions have overloaded C++ API template versions documented separately in the [C++ API Routines](#) module.

5.6.2 Function Documentation

5.6.2.1 `cudaError_t cudaConfigureCall` (`dim3 gridDim`, `dim3 blockDim`, `size_t sharedMem = 0`, `cudaStream_t stream = 0`)

Specifies the grid and block dimensions for the device call to be executed similar to the execution configuration syntax. `cudaConfigureCall()` is stack based. Each call pushes data on top of an execution stack. This data contains the dimension for the grid and thread blocks, together with any arguments for the call.

Parameters:

- gridDim* - Grid dimensions
- blockDim* - Block dimensions
- sharedMem* - Shared memory

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidConfiguration](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaFuncSetCacheConfig \(C API\)](#), [cudaFuncGetAttributes \(C API\)](#), [cudaLaunch \(C API\)](#), [cudaSetDoubleForDevice](#), [cudaSetDoubleForHost](#), [cudaSetupArgument \(C API\)](#),

5.6.2.2 `__cudart_builtin__ cudaError_t cudaFuncGetAttributes (struct cudaFuncAttributes * attr, const void * func)`

This function obtains the attributes of a function specified via `func`. `func` is a device function symbol and must be declared as a `__global__` function. The fetched attributes are placed in `attr`. If the specified function does not exist, then [cudaErrorInvalidDeviceFunction](#) is returned.

Note that some function attributes such as [maxThreadsPerBlock](#) may vary based on the device that is currently being used.

Parameters:

attr - Return pointer to function's attributes

func - Device function symbol

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidDeviceFunction](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

Use of a string naming a function as the `func` parameter was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaConfigureCall](#), [cudaFuncSetCacheConfig \(C API\)](#), [cudaFuncGetAttributes \(C++ API\)](#), [cudaLaunch \(C API\)](#), [cudaSetDoubleForDevice](#), [cudaSetDoubleForHost](#), [cudaSetupArgument \(C API\)](#)

5.6.2.3 `cudaError_t cudaFuncSetCacheConfig (const void * func, enum cudaFuncCache cacheConfig)`

On devices where the L1 cache and shared memory use the same hardware resources, this sets through `cacheConfig` the preferred cache configuration for the function specified via `func`. This is only a preference. The runtime will use the requested configuration if possible, but it is free to choose a different configuration if required to execute `func`.

`func` is a device function symbol and must be declared as a `__global__` function. If the specified function does not exist, then [cudaErrorInvalidDeviceFunction](#) is returned.

This setting does nothing on devices where the size of the L1 cache and shared memory are fixed.

Launching a kernel with a different preference than the most recent preference setting may insert a device-side synchronization point.

The supported cache configurations are:

- [cudaFuncCachePreferNone](#): no preference for shared memory or L1 (default)
- [cudaFuncCachePreferShared](#): prefer larger shared memory and smaller L1 cache
- [cudaFuncCachePreferL1](#): prefer larger L1 cache and smaller shared memory

Parameters:

func - Device function symbol

cacheConfig - Requested cache configuration

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidDeviceFunction](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

Use of a string naming a function as the `func` parameter was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaConfigureCall](#), [cudaFuncSetCacheConfig](#) (C++ API), [cudaFuncGetAttributes](#) (C API), [cudaLaunch](#) (C API), [cudaSetDoubleForDevice](#), [cudaSetDoubleForHost](#), [cudaSetupArgument](#) (C API), [cudaThreadGetCacheConfig](#), [cudaThreadSetCacheConfig](#)

5.6.2.4 `cudaError_t cudaFuncSetSharedMemConfig (const void *func, enum cudaSharedMemConfig config)`

On devices with configurable shared memory banks, this function will force all subsequent launches of the specified device function to have the given shared memory bank size configuration. On any given launch of the function, the shared memory configuration of the device will be temporarily changed if needed to suit the function's preferred configuration. Changes in shared memory configuration between subsequent launches of functions, may introduce a device side synchronization point.

Any per-function setting of shared memory bank size set via [cudaFuncSetSharedMemConfig](#) will override the device wide setting set by [cudaDeviceSetSharedMemConfig](#).

Changing the shared memory bank size will not increase shared memory usage or affect occupancy of kernels, but may have major effects on performance. Larger bank sizes will allow for greater potential bandwidth to shared memory, but will change what kinds of accesses to shared memory will result in bank conflicts.

This function will do nothing on devices with fixed shared memory bank size.

The supported bank configurations are:

- [cudaSharedMemBankSizeDefault](#): use the device's shared memory configuration when launching this function.
- [cudaSharedMemBankSizeFourByte](#): set shared memory bank width to be four bytes natively when launching this function.
- [cudaSharedMemBankSizeEightByte](#): set shared memory bank width to be eight bytes natively when launching this function.

Parameters:

func - Device function symbol
config - Requested shared memory configuration

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidDeviceFunction](#), [cudaErrorInvalidValue](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.
 Use of a string naming a function as the *func* parameter was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaConfigureCall](#), [cudaDeviceSetSharedMemConfig](#), [cudaDeviceGetSharedMemConfig](#), [cudaDeviceSetCacheConfig](#), [cudaDeviceGetCacheConfig](#), [cudaFuncSetCacheConfig](#)

5.6.2.5 `cudaError_t cudaLaunch (const void *func)`

Launches the function *func* on the device. The parameter *func* must be a device function symbol. The parameter specified by *func* must be declared as a `__global__` function. [cudaLaunch\(\)](#) must be preceded by a call to [cudaConfigureCall\(\)](#) since it pops the data that was pushed by [cudaConfigureCall\(\)](#) from the execution stack.

Parameters:

func - Device function symbol

Returns:

[cudaSuccess](#), [cudaErrorInvalidDeviceFunction](#), [cudaErrorInvalidConfiguration](#), [cudaErrorLaunchFailure](#), [cudaErrorLaunchTimeout](#), [cudaErrorLaunchOutOfResources](#), [cudaErrorSharedObjectInitFailed](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
 Use of a string naming a variable as the *symbol* parameter was removed in CUDA 5.0.

See also:

[cudaConfigureCall](#), [cudaFuncSetCacheConfig \(C API\)](#), [cudaFuncGetAttributes \(C API\)](#), [cudaLaunch \(C++ API\)](#), [cudaSetDoubleForDevice](#), [cudaSetDoubleForHost](#), [cudaSetupArgument \(C API\)](#), [cudaThreadGetCacheConfig](#), [cudaThreadSetCacheConfig](#)

5.6.2.6 `cudaError_t cudaSetDoubleForDevice (double *d)`**Parameters:**

d - Double to convert

Converts the double value of *d* to an internal float representation if the device does not support double arithmetic. If the device does natively support doubles, then this function does nothing.

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaConfigureCall](#), [cudaFuncSetCacheConfig](#) (C API), [cudaFuncGetAttributes](#) (C API), [cudaLaunch](#) (C API), [cudaSetDoubleForHost](#), [cudaSetupArgument](#) (C API)

5.6.2.7 `cudaError_t cudaSetDoubleForHost (double * d)`

Converts the double value of `d` from a potentially internal float representation if the device does not support double arithmetic. If the device does natively support doubles, then this function does nothing.

Parameters:

d - Double to convert

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaConfigureCall](#), [cudaFuncSetCacheConfig](#) (C API), [cudaFuncGetAttributes](#) (C API), [cudaLaunch](#) (C API), [cudaSetDoubleForDevice](#), [cudaSetupArgument](#) (C API)

5.6.2.8 `cudaError_t cudaSetupArgument (const void * arg, size_t size, size_t offset)`

Pushes `size` bytes of the argument pointed to by `arg` at `offset` bytes from the start of the parameter passing area, which starts at offset 0. The arguments are stored in the top of the execution stack. [cudaSetupArgument\(\)](#) must be preceded by a call to [cudaConfigureCall\(\)](#).

Parameters:

arg - Argument to push for a kernel launch

size - Size of argument

offset - Offset in argument stack to push new arg

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaConfigureCall](#), [cudaFuncSetCacheConfig](#) (C API), [cudaFuncGetAttributes](#) (C API), [cudaLaunch](#) (C API), [cudaSetDoubleForDevice](#), [cudaSetDoubleForHost](#), [cudaSetupArgument](#) (C++ API),

5.7 Memory Management

Functions

- [cudaError_t cudaArrayGetInfo](#) (struct [cudaChannelFormatDesc](#) *desc, struct [cudaExtent](#) *extent, unsigned int *flags, [cudaArray_t](#) array)
Gets info about the specified cudaArray.
- [__cuda_builtin__ cudaError_t cudaFree](#) (void *devPtr)
Frees memory on the device.
- [cudaError_t cudaFreeArray](#) ([cudaArray_t](#) array)
Frees an array on the device.
- [cudaError_t cudaFreeHost](#) (void *ptr)
Frees page-locked memory.
- [cudaError_t cudaFreeMipmappedArray](#) ([cudaMipmappedArray_t](#) mipmappedArray)
Frees a mipmapped array on the device.
- [cudaError_t cudaGetMipmappedArrayLevel](#) ([cudaArray_t](#) *levelArray, [cudaMipmappedArray_const_t](#) mipmappedArray, unsigned int level)
Gets a mipmap level of a CUDA mipmapped array.
- [cudaError_t cudaGetSymbolAddress](#) (void **devPtr, const void *symbol)
Finds the address associated with a CUDA symbol.
- [cudaError_t cudaGetSymbolSize](#) (size_t *size, const void *symbol)
Finds the size of the object associated with a CUDA symbol.
- [cudaError_t cudaHostAlloc](#) (void **pHost, size_t size, unsigned int flags)
Allocates page-locked memory on the host.
- [cudaError_t cudaHostGetDevicePointer](#) (void **pDevice, void *pHost, unsigned int flags)
Passes back device pointer of mapped host memory allocated by cudaHostAlloc or registered by cudaHostRegister.
- [cudaError_t cudaHostGetFlags](#) (unsigned int *pFlags, void *pHost)
Passes back flags used to allocate pinned host memory allocated by cudaHostAlloc.
- [cudaError_t cudaHostRegister](#) (void *ptr, size_t size, unsigned int flags)
Registers an existing host memory range for use by CUDA.
- [cudaError_t cudaHostUnregister](#) (void *ptr)
Unregisters a memory range that was registered with cudaHostRegister.
- [__cuda_builtin__ cudaError_t cudaMalloc](#) (void **devPtr, size_t size)
Allocate memory on the device.
- [cudaError_t cudaMalloc3D](#) (struct [cudaPitchedPtr](#) *pitchedDevPtr, struct [cudaExtent](#) extent)
Allocates logical 1D, 2D, or 3D memory objects on the device.

- `cudaError_t cudaMalloc3DArray` (`cudaArray_t *array`, `const struct cudaChannelFormatDesc *desc`, `struct cudaExtent extent`, `unsigned int flags=0`)
Allocate an array on the device.
- `cudaError_t cudaMallocArray` (`cudaArray_t *array`, `const struct cudaChannelFormatDesc *desc`, `size_t width`, `size_t height=0`, `unsigned int flags=0`)
Allocate an array on the device.
- `cudaError_t cudaMallocHost` (`void **ptr`, `size_t size`)
Allocates page-locked memory on the host.
- `cudaError_t cudaMallocMipmappedArray` (`cudaMipmappedArray_t *mipmappedArray`, `const struct cudaChannelFormatDesc *desc`, `struct cudaExtent extent`, `unsigned int numLevels`, `unsigned int flags=0`)
Allocate a mipmapped array on the device.
- `cudaError_t cudaMallocPitch` (`void **devPtr`, `size_t *pitch`, `size_t width`, `size_t height`)
Allocates pitched memory on the device.
- `cudaError_t cudaMemcpy` (`void *dst`, `const void *src`, `size_t count`, `enum cudaMemcpyKind kind`)
Copies data between host and device.
- `cudaError_t cudaMemcpy2D` (`void *dst`, `size_t dpitch`, `const void *src`, `size_t spitch`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind`)
Copies data between host and device.
- `cudaError_t cudaMemcpy2DFromArray` (`cudaArray_t dst`, `size_t wOffsetDst`, `size_t hOffsetDst`, `cudaArray_const_t src`, `size_t wOffsetSrc`, `size_t hOffsetSrc`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind=cudaMemcpyDeviceToDevice`)
Copies data between host and device.
- `__cuda_builtin__ cudaError_t cudaMemcpy2DAsync` (`void *dst`, `size_t dpitch`, `const void *src`, `size_t spitch`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind`, `cudaStream_t stream=0`)
Copies data between host and device.
- `cudaError_t cudaMemcpy2DFromArray` (`void *dst`, `size_t dpitch`, `cudaArray_const_t src`, `size_t wOffset`, `size_t hOffset`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind`)
Copies data between host and device.
- `cudaError_t cudaMemcpy2DFromArrayAsync` (`void *dst`, `size_t dpitch`, `cudaArray_const_t src`, `size_t wOffset`, `size_t hOffset`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind`, `cudaStream_t stream=0`)
Copies data between host and device.
- `cudaError_t cudaMemcpy2DToArray` (`cudaArray_t dst`, `size_t wOffset`, `size_t hOffset`, `const void *src`, `size_t spitch`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind`)
Copies data between host and device.
- `cudaError_t cudaMemcpy2DToArrayAsync` (`cudaArray_t dst`, `size_t wOffset`, `size_t hOffset`, `const void *src`, `size_t spitch`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind`, `cudaStream_t stream=0`)
Copies data between host and device.

- [cudaError_t cudaMemcpy3D](#) (const struct [cudaMemcpy3DParms](#) *p)
Copies data between 3D objects.
- [__cuda_builtin__ cudaError_t cudaMemcpy3DAsync](#) (const struct [cudaMemcpy3DParms](#) *p, [cudaStream_t](#) stream=0)
Copies data between 3D objects.
- [cudaError_t cudaMemcpy3DPeer](#) (const struct [cudaMemcpy3DPeerParms](#) *p)
Copies memory between devices.
- [cudaError_t cudaMemcpy3DPeerAsync](#) (const struct [cudaMemcpy3DPeerParms](#) *p, [cudaStream_t](#) stream=0)
Copies memory between devices asynchronously.
- [cudaError_t cudaMemcpyArrayToArray](#) ([cudaArray_t](#) dst, size_t wOffsetDst, size_t hOffsetDst, [cudaArray_const_t](#) src, size_t wOffsetSrc, size_t hOffsetSrc, size_t count, enum [cudaMemcpyKind](#) kind=cudaMemcpyDeviceToDevice)
Copies data between host and device.
- [__cuda_builtin__ cudaError_t cudaMemcpyAsync](#) (void *dst, const void *src, size_t count, enum [cudaMemcpyKind](#) kind, [cudaStream_t](#) stream=0)
Copies data between host and device.
- [cudaError_t cudaMemcpyFromArray](#) (void *dst, [cudaArray_const_t](#) src, size_t wOffset, size_t hOffset, size_t count, enum [cudaMemcpyKind](#) kind)
Copies data between host and device.
- [cudaError_t cudaMemcpyFromArrayAsync](#) (void *dst, [cudaArray_const_t](#) src, size_t wOffset, size_t hOffset, size_t count, enum [cudaMemcpyKind](#) kind, [cudaStream_t](#) stream=0)
Copies data between host and device.
- [cudaError_t cudaMemcpyFromSymbol](#) (void *dst, const void *symbol, size_t count, size_t offset=0, enum [cudaMemcpyKind](#) kind=cudaMemcpyDeviceToHost)
Copies data from the given symbol on the device.
- [cudaError_t cudaMemcpyFromSymbolAsync](#) (void *dst, const void *symbol, size_t count, size_t offset, enum [cudaMemcpyKind](#) kind, [cudaStream_t](#) stream=0)
Copies data from the given symbol on the device.
- [cudaError_t cudaMemcpyPeer](#) (void *dst, int dstDevice, const void *src, int srcDevice, size_t count)
Copies memory between two devices.
- [cudaError_t cudaMemcpyPeerAsync](#) (void *dst, int dstDevice, const void *src, int srcDevice, size_t count, [cudaStream_t](#) stream=0)
Copies memory between two devices asynchronously.
- [cudaError_t cudaMemcpyToArray](#) ([cudaArray_t](#) dst, size_t wOffset, size_t hOffset, const void *src, size_t count, enum [cudaMemcpyKind](#) kind)
Copies data between host and device.
- [cudaError_t cudaMemcpyToArrayAsync](#) ([cudaArray_t](#) dst, size_t wOffset, size_t hOffset, const void *src, size_t count, enum [cudaMemcpyKind](#) kind, [cudaStream_t](#) stream=0)

Copies data between host and device.

- [cudaError_t cudaMemcpyToSymbol](#) (const void *symbol, const void *src, size_t count, size_t offset=0, enum [cudaMemcpyKind](#) kind=cudaMemcpyHostToDevice)

Copies data to the given symbol on the device.

- [cudaError_t cudaMemcpyToSymbolAsync](#) (const void *symbol, const void *src, size_t count, size_t offset, enum [cudaMemcpyKind](#) kind, [cudaStream_t](#) stream=0)

Copies data to the given symbol on the device.

- [cudaError_t cudaMemGetInfo](#) (size_t *free, size_t *total)

Gets free and total device memory.

- [cudaError_t cudaMemset](#) (void *devPtr, int value, size_t count)

Initializes or sets device memory to a value.

- [cudaError_t cudaMemset2D](#) (void *devPtr, size_t pitch, int value, size_t width, size_t height)

Initializes or sets device memory to a value.

- [__cuda_builtin__ cudaError_t cudaMemset2DAsync](#) (void *devPtr, size_t pitch, int value, size_t width, size_t height, [cudaStream_t](#) stream=0)

Initializes or sets device memory to a value.

- [cudaError_t cudaMemset3D](#) (struct [cudaPitchedPtr](#) pitchedDevPtr, int value, struct [cudaExtent](#) extent)

Initializes or sets device memory to a value.

- [__cuda_builtin__ cudaError_t cudaMemset3DAsync](#) (struct [cudaPitchedPtr](#) pitchedDevPtr, int value, struct [cudaExtent](#) extent, [cudaStream_t](#) stream=0)

Initializes or sets device memory to a value.

- [__cuda_builtin__ cudaError_t cudaMemsetAsync](#) (void *devPtr, int value, size_t count, [cudaStream_t](#) stream=0)

Initializes or sets device memory to a value.

- struct [cudaExtent](#) [make_cudaExtent](#) (size_t w, size_t h, size_t d)

Returns a [cudaExtent](#) based on input parameters.

- struct [cudaPitchedPtr](#) [make_cudaPitchedPtr](#) (void *d, size_t p, size_t xsz, size_t ysz)

Returns a [cudaPitchedPtr](#) based on input parameters.

- struct [cudaPos](#) [make_cudaPos](#) (size_t x, size_t y, size_t z)

Returns a [cudaPos](#) based on input parameters.

5.7.1 Detailed Description

This section describes the memory management functions of the CUDA runtime application programming interface. Some functions have overloaded C++ API template versions documented separately in the [C++ API Routines](#) module.

5.7.2 Function Documentation

5.7.2.1 `cudaError_t cudaArrayGetInfo (struct cudaChannelFormatDesc * desc, struct cudaExtent * extent, unsigned int * flags, cudaArray_t array)`

Returns in `*desc`, `*extent` and `*flags` respectively, the type, shape and flags of `array`.

Any of `*desc`, `*extent` and `*flags` may be specified as `NULL`.

Parameters:

desc - Returned array type

extent - Returned array shape. 2D arrays will have depth of zero

flags - Returned array flags

array - The `cudaArray` to get info for

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

5.7.2.2 `__cuda_builtin__ cudaError_t cudaFree (void * devPtr)`

Frees the memory space pointed to by `devPtr`, which must have been returned by a previous call to [cudaMalloc\(\)](#) or [cudaMallocPitch\(\)](#). Otherwise, or if [cudaFree\(devPtr\)](#) has already been called before, an error is returned. If `devPtr` is 0, no operation is performed. [cudaFree\(\)](#) returns [cudaErrorInvalidDevicePointer](#) in case of failure.

Parameters:

devPtr - Device pointer to memory to free

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc](#), [cudaMallocPitch](#), [cudaMallocArray](#), [cudaFreeArray](#), [cudaMallocHost \(C API\)](#), [cudaFreeHost](#), [cudaMalloc3D](#), [cudaMalloc3DArray](#), [cudaHostAlloc](#)

5.7.2.3 `cudaError_t cudaFreeArray (cudaArray_t array)`

Frees the CUDA array `array`, which must have been * returned by a previous call to [cudaMallocArray\(\)](#). If [cudaFreeArray\(array\)](#) has already been called before, [cudaErrorInvalidValue](#) is returned. If `devPtr` is 0, no operation is performed.

Parameters:

array - Pointer to array to free

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc](#), [cudaMallocPitch](#), [cudaFree](#), [cudaMallocArray](#), [cudaMallocHost \(C API\)](#), [cudaFreeHost](#), [cudaHostAlloc](#)

5.7.2.4 `cudaError_t cudaFreeHost (void * ptr)`

Frees the memory space pointed to by `hostPtr`, which must have been returned by a previous call to [cudaMallocHost\(\)](#) or [cudaHostAlloc\(\)](#).

Parameters:

ptr - Pointer to memory to free

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc](#), [cudaMallocPitch](#), [cudaFree](#), [cudaMallocArray](#), [cudaFreeArray](#), [cudaMallocHost \(C API\)](#), [cudaMalloc3D](#), [cudaMalloc3DArray](#), [cudaHostAlloc](#)

5.7.2.5 `cudaError_t cudaFreeMipmappedArray (cudaMipmappedArray_t mipmappedArray)`

Frees the CUDA mipmapped array `mipmappedArray`, which must have been returned by a previous call to [cudaMallocMipmappedArray\(\)](#). If [cudaFreeMipmappedArray\(mipmappedArray\)](#) has already been called before, [cudaErrorInvalidValue](#) is returned.

Parameters:

mipmappedArray - Pointer to mipmapped array to free

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc](#), [cudaMallocPitch](#), [cudaFree](#), [cudaMallocArray](#), [cudaMallocHost \(C API\)](#), [cudaFreeHost](#), [cudaHostAlloc](#)

5.7.2.6 `cudaError_t cudaGetMipmappedArrayLevel (cudaArray_t * levelArray, cudaMipmappedArray_const_t mipmappedArray, unsigned int level)`

Returns in `*levelArray` a CUDA array that represents a single mipmap level of the CUDA mipmapped array `mipmappedArray`.

If `level` is greater than the maximum number of levels in this mipmapped array, `cudaErrorInvalidValue` is returned.

Parameters:

levelArray - Returned mipmap level CUDA array

mipmappedArray - CUDA mipmapped array

level - Mipmap level

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaMalloc3D`, `cudaMalloc`, `cudaMallocPitch`, `cudaFree`, `cudaFreeArray`, `cudaMallocHost (C API)`, `cudaFreeHost`, `cudaHostAlloc`, `make_cudaExtent`

5.7.2.7 `cudaError_t cudaGetSymbolAddress (void ** devPtr, const void * symbol)`

Returns in `*devPtr` the address of symbol `symbol` on the device. `symbol` is a variable that resides in global or constant memory space. If `symbol` cannot be found, or if `symbol` is not declared in the global or constant memory space, `*devPtr` is unchanged and the error `cudaErrorInvalidSymbol` is returned.

Parameters:

devPtr - Return device pointer associated with symbol

symbol - Device symbol address

Returns:

`cudaSuccess`, `cudaErrorInvalidSymbol`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

Use of a string naming a variable as the `symbol` paramater was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

`cudaGetSymbolAddress (C++ API)`, `cudaGetSymbolSize (C API)`

5.7.2.8 `cudaError_t cudaGetSymbolSize (size_t * size, const void * symbol)`

Returns in `*size` the size of symbol `symbol`. `symbol` is a variable that resides in global or constant memory space. If `symbol` cannot be found, or if `symbol` is not declared in global or constant memory space, `*size` is unchanged and the error `cudaErrorInvalidSymbol` is returned.

Parameters:

`size` - Size of object associated with symbol

`symbol` - Device symbol address

Returns:

`cudaSuccess`, `cudaErrorInvalidSymbol`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

Use of a string naming a variable as the `symbol` parameter was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

`cudaGetSymbolAddress` (C API), `cudaGetSymbolSize` (C++ API)

5.7.2.9 `cudaError_t cudaHostAlloc (void ** pHost, size_t size, unsigned int flags)`

Allocates `size` bytes of host memory that is page-locked and accessible to the device. The driver tracks the virtual memory ranges allocated with this function and automatically accelerates calls to functions such as `cudaMemcpy()`. Since the memory can be accessed directly by the device, it can be read or written with much higher bandwidth than pageable memory obtained with functions such as `malloc()`. Allocating excessive amounts of pinned memory may degrade system performance, since it reduces the amount of memory available to the system for paging. As a result, this function is best used sparingly to allocate staging areas for data exchange between host and device.

The `flags` parameter enables different options to be specified that affect the allocation, as follows.

- `cudaHostAllocDefault`: This flag's value is defined to be 0 and causes `cudaHostAlloc()` to emulate `cudaMallocHost()`.
- `cudaHostAllocPortable`: The memory returned by this call will be considered as pinned memory by all CUDA contexts, not just the one that performed the allocation.
- `cudaHostAllocMapped`: Maps the allocation into the CUDA address space. The device pointer to the memory may be obtained by calling `cudaHostGetDevicePointer()`.
- `cudaHostAllocWriteCombined`: Allocates the memory as write-combined (WC). WC memory can be transferred across the PCI Express bus more quickly on some system configurations, but cannot be read efficiently by most CPUs. WC memory is a good option for buffers that will be written by the CPU and read by the device via mapped pinned memory or host->device transfers.

All of these flags are orthogonal to one another: a developer may allocate memory that is portable, mapped and/or write-combined with no restrictions.

`cudaSetDeviceFlags()` must have been called with the `cudaDeviceMapHost` flag in order for the `cudaHostAllocMapped` flag to have any effect.

The [cudaHostAllocMapped](#) flag may be specified on CUDA contexts for devices that do not support mapped pinned memory. The failure is deferred to [cudaHostGetDevicePointer\(\)](#) because the memory may be mapped into other CUDA contexts via the [cudaHostAllocPortable](#) flag.

Memory allocated by this function must be freed with [cudaFreeHost\(\)](#).

Parameters:

- pHost* - Device pointer to allocated memory
- size* - Requested allocation size in bytes
- flags* - Requested properties of allocated memory

Returns:

[cudaSuccess](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaSetDeviceFlags](#), [cudaMallocHost \(C API\)](#), [cudaFreeHost](#)

5.7.2.10 `cudaError_t cudaHostGetDevicePointer (void ** pDevice, void * pHost, unsigned int flags)`

Passes back the device pointer corresponding to the mapped, pinned host buffer allocated by [cudaHostAlloc\(\)](#) or registered by [cudaHostRegister\(\)](#).

[cudaHostGetDevicePointer\(\)](#) will fail if the [cudaDeviceMapHost](#) flag was not specified before deferred context creation occurred, or if called on a device that does not support mapped, pinned memory.

`flags` provides for future releases. For now, it must be set to 0.

Parameters:

- pDevice* - Returned device pointer for mapped memory
- pHost* - Requested host pointer mapping
- flags* - Flags for extensions (must be 0 for now)

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaSetDeviceFlags](#), [cudaHostAlloc](#)

5.7.2.11 `cudaError_t cudaHostGetFlags (unsigned int * pFlags, void * pHost)`

`cudaHostGetFlags()` will fail if the input pointer does not reside in an address range allocated by `cudaHostAlloc()`.

Parameters:

pFlags - Returned flags word

pHost - Host pointer

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaHostAlloc`

5.7.2.12 `cudaError_t cudaHostRegister (void * ptr, size_t size, unsigned int flags)`

Page-locks the memory range specified by `ptr` and `size` and maps it for the device(s) as specified by `flags`. This memory range also is added to the same tracking mechanism as `cudaHostAlloc()` to automatically accelerate calls to functions such as `cudaMemcpy()`. Since the memory can be accessed directly by the device, it can be read or written with much higher bandwidth than pageable memory that has not been registered. Page-locking excessive amounts of memory may degrade system performance, since it reduces the amount of memory available to the system for paging. As a result, this function is best used sparingly to register staging areas for data exchange between host and device.

The `flags` parameter enables different options to be specified that affect the allocation, as follows.

- `cudaHostRegisterPortable`: The memory returned by this call will be considered as pinned memory by all CUDA contexts, not just the one that performed the allocation.
- `cudaHostRegisterMapped`: Maps the allocation into the CUDA address space. The device pointer to the memory may be obtained by calling `cudaHostGetDevicePointer()`. This feature is available only on GPUs with compute capability greater than or equal to 1.1.

All of these flags are orthogonal to one another: a developer may page-lock memory that is portable or mapped with no restrictions.

The CUDA context must have been created with the `cudaMapHost` flag in order for the `cudaHostRegisterMapped` flag to have any effect.

The `cudaHostRegisterMapped` flag may be specified on CUDA contexts for devices that do not support mapped pinned memory. The failure is deferred to `cudaHostGetDevicePointer()` because the memory may be mapped into other CUDA contexts via the `cudaHostRegisterPortable` flag.

The memory page-locked by this function must be unregistered with `cudaHostUnregister()`.

Parameters:

ptr - Host pointer to memory to page-lock

size - Size in bytes of the address range to page-lock in bytes

flags - Flags for allocation request

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaHostUnregister](#), [cudaHostGetFlags](#), [cudaHostGetDevicePointer](#)

5.7.2.13 `cudaError_t cudaHostUnregister (void *ptr)`

Unmaps the memory range whose base address is specified by `ptr`, and makes it pageable again.

The base address must be the same one specified to [cudaHostRegister\(\)](#).

Parameters:

ptr - Host pointer to memory to unregister

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaHostUnregister](#)

5.7.2.14 `__cuda_builtin__ cudaError_t cudaMalloc (void **devPtr, size_t size)`

Allocates `size` bytes of linear memory on the device and returns in `*devPtr` a pointer to the allocated memory. The allocated memory is suitably aligned for any kind of variable. The memory is not cleared. [cudaMalloc\(\)](#) returns [cudaErrorMemoryAllocation](#) in case of failure.

Parameters:

devPtr - Pointer to allocated device memory

size - Requested allocation size in bytes

Returns:

[cudaSuccess](#), [cudaErrorMemoryAllocation](#)

See also:

[cudaMallocPitch](#), [cudaFree](#), [cudaMallocArray](#), [cudaFreeArray](#), [cudaMalloc3D](#), [cudaMalloc3DArray](#), [cudaMallocHost](#) (C API), [cudaFreeHost](#), [cudaHostAlloc](#)

5.7.2.15 `cudaError_t cudaMalloc3D (struct cudaPitchedPtr * pitchedDevPtr, struct cudaExtent extent)`

Allocates at least `width * height * depth` bytes of linear memory on the device and returns a `cudaPitchedPtr` in which `ptr` is a pointer to the allocated memory. The function may pad the allocation to ensure hardware alignment requirements are met. The pitch returned in the `pitch` field of `pitchedDevPtr` is the width in bytes of the allocation.

The returned `cudaPitchedPtr` contains additional fields `xsize` and `ysize`, the logical width and height of the allocation, which are equivalent to the `width` and `height` extent parameters provided by the programmer during allocation.

For allocations of 2D and 3D objects, it is highly recommended that programmers perform allocations using `cudaMalloc3D()` or `cudaMallocPitch()`. Due to alignment restrictions in the hardware, this is especially true if the application will be performing memory copies involving 2D or 3D objects (whether linear memory or CUDA arrays).

Parameters:

- pitchedDevPtr* - Pointer to allocated pitched device memory
- extent* - Requested allocation size (`width` field in bytes)

Returns:

`cudaSuccess`, `cudaErrorMemoryAllocation`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaMallocPitch`, `cudaFree`, `cudaMemcpy3D`, `cudaMemset3D`, `cudaMalloc3DArray`, `cudaMallocArray`, `cudaFreeArray`, `cudaMallocHost` (C API), `cudaFreeHost`, `cudaHostAlloc`, `make_cudaPitchedPtr`, `make_cudaExtent`

5.7.2.16 `cudaError_t cudaMalloc3DArray (cudaArray_t * array, const struct cudaChannelFormatDesc * desc, struct cudaExtent extent, unsigned int flags = 0)`

Allocates a CUDA array according to the `cudaChannelFormatDesc` structure `desc` and returns a handle to the new CUDA array in `*array`.

The `cudaChannelFormatDesc` is defined as:

```
struct cudaChannelFormatDesc {
    int x, y, z, w;
    enum cudaChannelFormatKind f;
};
```

where `cudaChannelFormatKind` is one of `cudaChannelFormatKindSigned`, `cudaChannelFormatKindUnsigned`, or `cudaChannelFormatKindFloat`.

`cudaMalloc3DArray()` can allocate the following:

- A 1D array is allocated if the height and depth extents are both zero.
- A 2D array is allocated if only the depth extent is zero.
- A 3D array is allocated if all three extents are non-zero.

- A 1D layered CUDA array is allocated if only the height extent is zero and the `cudaArrayLayered` flag is set. Each layer is a 1D array. The number of layers is determined by the depth extent.
- A 2D layered CUDA array is allocated if all three extents are non-zero and the `cudaArrayLayered` flag is set. Each layer is a 2D array. The number of layers is determined by the depth extent.
- A cubemap CUDA array is allocated if all three extents are non-zero and the `cudaArrayCubemap` flag is set. Width must be equal to height, and depth must be six. A cubemap is a special type of 2D layered CUDA array, where the six layers represent the six faces of a cube. The order of the six layers in memory is the same as that listed in [cudaGraphicsCubeFace](#).
- A cubemap layered CUDA array is allocated if all three extents are non-zero, and both, `cudaArrayCubemap` and `cudaArrayLayered` flags are set. Width must be equal to height, and depth must be a multiple of six. A cubemap layered CUDA array is a special type of 2D layered CUDA array that consists of a collection of cubemaps. The first six layers represent the first cubemap, the next six layers form the second cubemap, and so on.

The `flags` parameter enables different options to be specified that affect the allocation, as follows.

- `cudaArrayDefault`: This flag's value is defined to be 0 and provides default array allocation
- `cudaArrayLayered`: Allocates a layered CUDA array, with the depth extent indicating the number of layers
- `cudaArrayCubemap`: Allocates a cubemap CUDA array. Width must be equal to height, and depth must be six. If the `cudaArrayLayered` flag is also set, depth must be a multiple of six.
- `cudaArraySurfaceLoadStore`: Allocates a CUDA array that could be read from or written to using a surface reference.
- `cudaArrayTextureGather`: This flag indicates that texture gather operations will be performed on the CUDA array. Texture gather can only be performed on 2D CUDA arrays.

The width, height and depth extents must meet certain size requirements as listed in the following table. All values are specified in elements.

Note that 2D CUDA arrays have different size requirements if the `cudaArrayTextureGather` flag is set. In that case, the valid range for (width, height, depth) is ((1,maxTexture2DGather[0]), (1,maxTexture2DGather[1]), 0).

CUDA array type	Valid extents that must always be met {(width range in elements), (height range), (depth range)}	Valid extents with <code>cudaArraySurfaceLoadStore</code> set {(width range in elements), (height range), (depth range)}
1D	{ (1,maxTexture1D), 0, 0 }	{ (1,maxSurface1D), 0, 0 }
2D	{ (1,maxTexture2D[0]), (1,maxTexture2D[1]), 0 }	{ (1,maxSurface2D[0]), (1,maxSurface2D[1]), 0 }
3D	{ (1,maxTexture3D[0]), (1,maxTexture3D[1]), (1,maxTexture3D[2]) }	{ (1,maxSurface3D[0]), (1,maxSurface3D[1]), (1,maxSurface3D[2]) }
1D Layered	{ (1,maxTexture1DLayered[0]), 0, (1,maxTexture1DLayered[1]) }	{ (1,maxSurface1DLayered[0]), 0, (1,maxSurface1DLayered[1]) }
2D Layered	{ (1,maxTexture2DLayered[0]), (1,maxTexture2DLayered[1]), (1,maxTexture2DLayered[2]) }	{ (1,maxSurface2DLayered[0]), (1,maxSurface2DLayered[1]), (1,maxSurface2DLayered[2]) }
Cubemap	{ (1,maxTextureCubemap), (1,maxTextureCubemap), 6 }	{ (1,maxSurfaceCubemap), (1,maxSurfaceCubemap), 6 }
Cubemap Layered	{ (1,maxTextureCubemapLayered[0]), (1,maxTextureCubemapLayered[0]), (1,maxTextureCubemapLayered[1]) }	{ (1,maxSurfaceCubemapLayered[0]), (1,maxSurfaceCubemapLayered[0]), (1,maxSurfaceCubemapLayered[1]) }

Parameters:

- array* - Pointer to allocated array in device memory
- desc* - Requested channel format
- extent* - Requested allocation size (*width* field in elements)
- flags* - Flags for extensions

Returns:

[cudaSuccess](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc3D](#), [cudaMalloc](#), [cudaMallocPitch](#), [cudaFree](#), [cudaFreeArray](#), [cudaMallocHost](#) (C API), [cudaFreeHost](#), [cudaHostAlloc](#), [make_cudaExtent](#)

5.7.2.17 `cudaError_t cudaMallocArray(cudaArray_t *array, const struct cudaChannelFormatDesc *desc, size_t width, size_t height = 0, unsigned int flags = 0)`

Allocates a CUDA array according to the [cudaChannelFormatDesc](#) structure `desc` and returns a handle to the new CUDA array in `*array`.

The [cudaChannelFormatDesc](#) is defined as:

```
struct cudaChannelFormatDesc {
    int x, y, z, w;
    enum cudaChannelFormatKind f;
};
```

where [cudaChannelFormatKind](#) is one of [cudaChannelFormatKindSigned](#), [cudaChannelFormatKindUnsigned](#), or [cudaChannelFormatKindFloat](#).

The `flags` parameter enables different options to be specified that affect the allocation, as follows.

- [cudaArrayDefault](#): This flag's value is defined to be 0 and provides default array allocation
- [cudaArraySurfaceLoadStore](#): Allocates an array that can be read from or written to using a surface reference
- [cudaArrayTextureGather](#): This flag indicates that texture gather operations will be performed on the array.

`width` and `height` must meet certain size requirements. See [cudaMalloc3DArray\(\)](#) for more details.

Parameters:

- array* - Pointer to allocated array in device memory
- desc* - Requested channel format
- width* - Requested array allocation width
- height* - Requested array allocation height
- flags* - Requested properties of allocated array

Returns:

[cudaSuccess](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc](#), [cudaMallocPitch](#), [cudaFree](#), [cudaFreeArray](#), [cudaMallocHost \(C API\)](#), [cudaFreeHost](#), [cudaMalloc3D](#), [cudaMalloc3DArray](#), [cudaHostAlloc](#)

5.7.2.18 `cudaError_t cudaMallocHost (void **ptr, size_t size)`

Allocates `size` bytes of host memory that is page-locked and accessible to the device. The driver tracks the virtual memory ranges allocated with this function and automatically accelerates calls to functions such as [cudaMemcpy*\(\)](#). Since the memory can be accessed directly by the device, it can be read or written with much higher bandwidth than pageable memory obtained with functions such as `malloc()`. Allocating excessive amounts of memory with [cudaMallocHost\(\)](#) may degrade system performance, since it reduces the amount of memory available to the system for paging. As a result, this function is best used sparingly to allocate staging areas for data exchange between host and device.

Parameters:

ptr - Pointer to allocated host memory
size - Requested allocation size in bytes

Returns:

[cudaSuccess](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc](#), [cudaMallocPitch](#), [cudaMallocArray](#), [cudaMalloc3D](#), [cudaMalloc3DArray](#), [cudaHostAlloc](#), [cudaFree](#), [cudaFreeArray](#), [cudaMallocHost \(C++ API\)](#), [cudaFreeHost](#), [cudaHostAlloc](#)

5.7.2.19 `cudaError_t cudaMallocMipmappedArray (cudaMipmappedArray_t *mipmappedArray, const struct cudaChannelFormatDesc *desc, struct cudaExtent extent, unsigned int numLevels, unsigned int flags = 0)`

Allocates a CUDA mipmapped array according to the [cudaChannelFormatDesc](#) structure `desc` and returns a handle to the new CUDA mipmapped array in `*mipmappedArray`. `numLevels` specifies the number of mipmap levels to be allocated. This value is clamped to the range $[1, 1 + \text{floor}(\log_2(\max(\text{width}, \text{height}, \text{depth})))]$.

The [cudaChannelFormatDesc](#) is defined as:

```
struct cudaChannelFormatDesc {
    int x, y, z, w;
    enum cudaChannelFormatKind f;
};
```

where `cudaChannelFormatKind` is one of `cudaChannelFormatKindSigned`, `cudaChannelFormatKindUnsigned`, or `cudaChannelFormatKindFloat`.

`cudaMallocMipmappedArray()` can allocate the following:

- A 1D mipmapped array is allocated if the height and depth extents are both zero.
- A 2D mipmapped array is allocated if only the depth extent is zero.
- A 3D mipmapped array is allocated if all three extents are non-zero.
- A 1D layered CUDA mipmapped array is allocated if only the height extent is zero and the `cudaArrayLayered` flag is set. Each layer is a 1D mipmapped array. The number of layers is determined by the depth extent.
- A 2D layered CUDA mipmapped array is allocated if all three extents are non-zero and the `cudaArrayLayered` flag is set. Each layer is a 2D mipmapped array. The number of layers is determined by the depth extent.
- A cubemap CUDA mipmapped array is allocated if all three extents are non-zero and the `cudaArrayCubemap` flag is set. Width must be equal to height, and depth must be six. The order of the six layers in memory is the same as that listed in `cudaGraphicsCubeFace`.
- A cubemap layered CUDA mipmapped array is allocated if all three extents are non-zero, and both, `cudaArrayCubemap` and `cudaArrayLayered` flags are set. Width must be equal to height, and depth must be a multiple of six. A cubemap layered CUDA mipmapped array is a special type of 2D layered CUDA mipmapped array that consists of a collection of cubemap mipmapped arrays. The first six layers represent the first cubemap mipmapped array, the next six layers form the second cubemap mipmapped array, and so on.

The `flags` parameter enables different options to be specified that affect the allocation, as follows.

- `cudaArrayDefault`: This flag's value is defined to be 0 and provides default mipmapped array allocation
- `cudaArrayLayered`: Allocates a layered CUDA mipmapped array, with the depth extent indicating the number of layers
- `cudaArrayCubemap`: Allocates a cubemap CUDA mipmapped array. Width must be equal to height, and depth must be six. If the `cudaArrayLayered` flag is also set, depth must be a multiple of six.
- `cudaArraySurfaceLoadStore`: This flag indicates that individual mipmap levels of the CUDA mipmapped array will be read from or written to using a surface reference.
- `cudaArrayTextureGather`: This flag indicates that texture gather operations will be performed on the CUDA array. Texture gather can only be performed on 2D CUDA mipmapped arrays, and the gather operations are performed only on the most detailed mipmap level.

The width, height and depth extents must meet certain size requirements as listed in the following table. All values are specified in elements.

CUDA array type	Valid extents {(width range in elements), (height range), (depth range)}
1D	{ (1,maxTexture1DMipmap), 0, 0 }
2D	{ (1,maxTexture2DMipmap[0]), (1,maxTexture2DMipmap[1]), 0 }
3D	{ (1,maxTexture3D[0]), (1,maxTexture3D[1]), (1,maxTexture3D[2]) }
1D Layered	{ (1,maxTexture1DLayered[0]), 0, (1,maxTexture1DLayered[1]) }
2D Layered	{ (1,maxTexture2DLayered[0]), (1,maxTexture2DLayered[1]), (1,maxTexture2DLayered[2]) }
Cubemap	{ (1,maxTextureCubemap), (1,maxTextureCubemap), 6 }
Cubemap Layered	{ (1,maxTextureCubemapLayered[0]), (1,maxTextureCubemapLayered[0]), (1,maxTextureCubemapLayered[1]) }

Parameters:

mipmappedArray - Pointer to allocated mipmapped array in device memory
desc - Requested channel format
extent - Requested allocation size (*width* field in elements)
numLevels - Number of mipmap levels to allocate
flags - Flags for extensions

Returns:

[cudaSuccess](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc3D](#), [cudaMalloc](#), [cudaMallocPitch](#), [cudaFree](#), [cudaFreeArray](#), [cudaMallocHost \(C API\)](#), [cudaFreeHost](#), [cudaHostAlloc](#), [make_cudaExtent](#)

5.7.2.20 `cudaError_t cudaMallocPitch (void ** devPtr, size_t * pitch, size_t width, size_t height)`

Allocates at least *width* (in bytes) * *height* bytes of linear memory on the device and returns in **devPtr* a pointer to the allocated memory. The function may pad the allocation to ensure that corresponding pointers in any given row will continue to meet the alignment requirements for coalescing as the address is updated from row to row. The pitch returned in **pitch* by [cudaMallocPitch\(\)](#) is the width in bytes of the allocation. The intended usage of *pitch* is as a separate parameter of the allocation, used to compute addresses within the 2D array. Given the row and column of an array element of type *T*, the address is computed as:

```
T* pElement = (T*)((char*)BaseAddress + Row * pitch) + Column;
```

For allocations of 2D arrays, it is recommended that programmers consider performing pitch allocations using [cudaMallocPitch\(\)](#). Due to pitch alignment restrictions in the hardware, this is especially true if the application will be performing 2D memory copies between different regions of device memory (whether linear memory or CUDA arrays).

Parameters:

devPtr - Pointer to allocated pitched device memory
pitch - Pitch for allocation
width - Requested pitched allocation width (in bytes)
height - Requested pitched allocation height

Returns:

[cudaSuccess](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMalloc](#), [cudaFree](#), [cudaMallocArray](#), [cudaFreeArray](#), [cudaMallocHost \(C API\)](#), [cudaFreeHost](#), [cudaMalloc3D](#), [cudaMalloc3DArray](#), [cudaHostAlloc](#)

5.7.2.21 `cudaError_t cudaMemcpy (void * dst, const void * src, size_t count, enum cudaMemcpyKind kind)`

Copies `count` bytes from the memory area pointed to by `src` to the memory area pointed to by `dst`, where `kind` is one of `cudaMemcpyHostToHost`, `cudaMemcpyHostToDevice`, `cudaMemcpyDeviceToHost`, or `cudaMemcpyDeviceToDevice`, and specifies the direction of the copy. The memory areas may not overlap. Calling `cudaMemcpy()` with `dst` and `src` pointers that do not match the direction of the copy results in an undefined behavior.

Parameters:

dst - Destination memory address

src - Source memory address

count - Size in bytes to copy

kind - Type of transfer

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidDevicePointer`, `cudaErrorInvalidMemcpyDirection`

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits `synchronous` behavior for most use cases.

See also:

`cudaMemcpy2D`, `cudaMemcpyToArray`, `cudaMemcpy2DToArray`, `cudaMemcpyFromArray`, `cudaMemcpy2DFromArray`, `cudaMemcpyFromArrayToArray`, `cudaMemcpy2DFromArrayToArray`, `cudaMemcpyToSymbol`, `cudaMemcpyFromSymbol`, `cudaMemcpyAsync`, `cudaMemcpy2DAsync`, `cudaMemcpyToArrayAsync`, `cudaMemcpy2DToArrayAsync`, `cudaMemcpyFromArrayAsync`, `cudaMemcpy2DFromArrayAsync`, `cudaMemcpyToSymbolAsync`, `cudaMemcpyFromSymbolAsync`

5.7.2.22 `cudaError_t cudaMemcpy2D (void * dst, size_t dpitch, const void * src, size_t spitch, size_t width, size_t height, enum cudaMemcpyKind kind)`

Copies a matrix (`height` rows of `width` bytes each) from the memory area pointed to by `src` to the memory area pointed to by `dst`, where `kind` is one of `cudaMemcpyHostToHost`, `cudaMemcpyHostToDevice`, `cudaMemcpyDeviceToHost`, or `cudaMemcpyDeviceToDevice`, and specifies the direction of the copy. `dpitch` and `spitch` are the widths in memory in bytes of the 2D arrays pointed to by `dst` and `src`, including any padding added to the end of each row. The memory areas may not overlap. `width` must not exceed either `dpitch` or `spitch`. Calling `cudaMemcpy2D()` with `dst` and `src` pointers that do not match the direction of the copy results in an undefined behavior. `cudaMemcpy2D()` returns an error if `dpitch` or `spitch` exceeds the maximum allowed.

Parameters:

dst - Destination memory address

dpitch - Pitch of destination memory

src - Source memory address

spitch - Pitch of source memory

width - Width of matrix transfer (columns in bytes)

height - Height of matrix transfer (rows)

kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMemcpy](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyFromArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.23 `cudaError_t cudaMemcpy2DArrayToArray(cudaArray_t dst, size_t wOffsetDst, size_t hOffsetDst, cudaArray_const_t src, size_t wOffsetSrc, size_t hOffsetSrc, size_t width, size_t height, enum cudaMemcpyKind kind = cudaMemcpyDeviceToDevice)`

Copies a matrix (height rows of width bytes each) from the CUDA array `srcArray` starting at the upper left corner (`wOffsetSrc`, `hOffsetSrc`) to the CUDA array `dst` starting at the upper left corner (`wOffsetDst`, `hOffsetDst`), where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy. `wOffsetDst + width` must not exceed the width of the CUDA array `dst`. `wOffsetSrc + width` must not exceed the width of the CUDA array `src`.

Parameters:

dst - Destination memory address
wOffsetDst - Destination starting X offset
hOffsetDst - Destination starting Y offset
src - Source memory address
wOffsetSrc - Source starting X offset
hOffsetSrc - Source starting Y offset
width - Width of matrix transfer (columns in bytes)
height - Height of matrix transfer (rows)
kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [synchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyFromArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.24 `__cuda_builtin__ cudaError_t cudaMemcpy2DAsync` (`void *dst`, `size_t dpitch`, `const void *src`, `size_t spitch`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind`, `cudaStream_t stream = 0`)

Copies a matrix (`height` rows of `width` bytes each) from the memory area pointed to by `src` to the memory area pointed to by `dst`, where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy. `dpitch` and `spitch` are the widths in memory in bytes of the 2D arrays pointed to by `dst` and `src`, including any padding added to the end of each row. The memory areas may not overlap. `width` must not exceed either `dpitch` or `spitch`. Calling [cudaMemcpy2DAsync\(\)](#) with `dst` and `src` pointers that do not match the direction of the copy results in an undefined behavior. [cudaMemcpy2DAsync\(\)](#) returns an error if `dpitch` or `spitch` is greater than the maximum allowed.

[cudaMemcpy2DAsync\(\)](#) is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is [cudaMemcpyHostToDevice](#) or [cudaMemcpyDeviceToHost](#) and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

dst - Destination memory address
dpitch - Pitch of destination memory
src - Source memory address
spitch - Pitch of source memory
width - Width of matrix transfer (columns in bytes)
height - Height of matrix transfer (rows)
kind - Type of transfer
stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [asynchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyFromArrayToArray](#), [cudaMemcpy2DFromArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.25 `cudaError_t cudaMemcpy2DFromArray` (`void *dst`, `size_t dpitch`, `cudaArray_const_t src`, `size_t wOffset`, `size_t hOffset`, `size_t width`, `size_t height`, `enum cudaMemcpyKind kind`)

Copies a matrix (`height` rows of `width` bytes each) from the CUDA array `srcArray` starting at the upper left corner (`wOffset`, `hOffset`) to the memory area pointed to by `dst`, where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy. `dpitch` is the width in memory in bytes of the 2D array pointed to by `dst`, including any padding added to the end of each row. `wOffset + width` must not exceed the width of the CUDA array `src`. `width` must not exceed `dpitch`. [cudaMemcpy2DFromArray\(\)](#) returns an error if `dpitch` exceeds the maximum allowed.

Parameters:

dst - Destination memory address
dpitch - Pitch of destination memory
src - Source memory address
wOffset - Source starting X offset
hOffset - Source starting Y offset
width - Width of matrix transfer (columns in bytes)
height - Height of matrix transfer (rows)
kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [synchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.26 `cudaError_t cudaMemcpy2DFromArrayAsync(void *dst, size_t dpitch, cudaArray_const_t src, size_t wOffset, size_t hOffset, size_t width, size_t height, enum cudaMemcpyKind kind, cudaStream_t stream = 0)`

Copies a matrix (*height* rows of *width* bytes each) from the CUDA array *srcArray* starting at the upper left corner (*wOffset*, *hOffset*) to the memory area pointed to by *dst*, where *kind* is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy. *dpitch* is the width in memory in bytes of the 2D array pointed to by *dst*, including any padding added to the end of each row. *wOffset* + *width* must not exceed the width of the CUDA array *src*. *width* must not exceed *dpitch*. [cudaMemcpy2DFromArrayAsync\(\)](#) returns an error if *dpitch* exceeds the maximum allowed.

[cudaMemcpy2DFromArrayAsync\(\)](#) is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero *stream* argument. If *kind* is [cudaMemcpyHostToDevice](#) or [cudaMemcpyDeviceToHost](#) and *stream* is non-zero, the copy may overlap with operations in other streams.

Parameters:

dst - Destination memory address
dpitch - Pitch of destination memory
src - Source memory address
wOffset - Source starting X offset
hOffset - Source starting Y offset

width - Width of matrix transfer (columns in bytes)

height - Height of matrix transfer (rows)

kind - Type of transfer

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [asynchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.27 `cudaError_t cudaMemcpy2DToArray(cudaArray_t dst, size_t wOffset, size_t hOffset, const void * src, size_t spitch, size_t width, size_t height, enum cudaMemcpyKind kind)`

Copies a matrix (*height* rows of *width* bytes each) from the memory area pointed to by *src* to the CUDA array *dst* starting at the upper left corner (*wOffset*, *hOffset*) where *kind* is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy. *spitch* is the width in memory in bytes of the 2D array pointed to by *src*, including any padding added to the end of each row. *wOffset* + *width* must not exceed the width of the CUDA array *dst*. *width* must not exceed *spitch*. [cudaMemcpy2DToArray\(\)](#) returns an error if *spitch* exceeds the maximum allowed.

Parameters:

dst - Destination memory address

wOffset - Destination starting X offset

hOffset - Destination starting Y offset

src - Source memory address

spitch - Pitch of source memory

width - Width of matrix transfer (columns in bytes)

height - Height of matrix transfer (rows)

kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [synchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.28 `cudaError_t cudaMemcpy2DToArrayAsync(cudaArray_t dst, size_t wOffset, size_t hOffset, const void * src, size_t spitch, size_t width, size_t height, enum cudaMemcpyKind kind, cudaStream_t stream = 0)`

Copies a matrix (height rows of width bytes each) from the memory area pointed to by `src` to the CUDA array `dst` starting at the upper left corner (`wOffset`, `hOffset`) where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy. `spitch` is the width in memory in bytes of the 2D array pointed to by `src`, including any padding added to the end of each row. `wOffset + width` must not exceed the width of the CUDA array `dst`. `width` must not exceed `spitch`. [cudaMemcpy2DToArrayAsync\(\)](#) returns an error if `spitch` exceeds the maximum allowed.

[cudaMemcpy2DToArrayAsync\(\)](#) is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is [cudaMemcpyHostToDevice](#) or [cudaMemcpyDeviceToHost](#) and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

dst - Destination memory address
wOffset - Destination starting X offset
hOffset - Destination starting Y offset
src - Source memory address
spitch - Pitch of source memory
width - Width of matrix transfer (columns in bytes)
height - Height of matrix transfer (rows)
kind - Type of transfer
stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [asynchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.29 `cudaError_t cudaMemcpy3D (const struct cudaMemcpy3DParms * p)`

```

struct cudaExtent {
    size_t width;
    size_t height;
    size_t depth;
};
struct cudaExtent make_cudaExtent(size_t w, size_t h, size_t d);

struct cudaPos {
    size_t x;
    size_t y;
    size_t z;
};
struct cudaPos make_cudaPos(size_t x, size_t y, size_t z);

struct cudaMemcpy3DParms {
    cudaArray_t      srcArray;
    struct cudaPos   srcPos;
    struct cudaPitchedPtr srcPtr;
    cudaArray_t      dstArray;
    struct cudaPos   dstPos;
    struct cudaPitchedPtr dstPtr;
    struct cudaExtent extent;
    enum cudaMemcpyKind kind;
};

```

`cudaMemcpy3D()` copies data between two 3D objects. The source and destination objects may be in either host memory, device memory, or a CUDA array. The source, destination, extent, and kind of copy performed is specified by the `cudaMemcpy3DParms` struct which should be initialized to zero before use:

```
cudaMemcpy3DParms myParms = {0};
```

The struct passed to `cudaMemcpy3D()` must specify one of `srcArray` or `srcPtr` and one of `dstArray` or `dstPtr`. Passing more than one non-zero source or destination will cause `cudaMemcpy3D()` to return an error.

The `srcPos` and `dstPos` fields are optional offsets into the source and destination objects and are defined in units of each object's elements. The element for a host or device pointer is assumed to be **unsigned char**. For CUDA arrays, positions must be in the range [0, 2048) for any dimension.

The `extent` field defines the dimensions of the transferred area in elements. If a CUDA array is participating in the copy, the extent is defined in terms of that array's elements. If no CUDA array is participating in the copy then the extents are defined in elements of **unsigned char**.

The `kind` field defines the direction of the copy. It must be one of `cudaMemcpyHostToHost`, `cudaMemcpyHostToDevice`, `cudaMemcpyDeviceToHost`, or `cudaMemcpyDeviceToDevice`.

If the source and destination are both arrays, `cudaMemcpy3D()` will return an error if they do not have the same element size.

The source and destination object may not overlap. If overlapping source and destination objects are specified, undefined behavior will result.

The source object must lie entirely within the region defined by `srcPos` and `extent`. The destination object must lie entirely within the region defined by `dstPos` and `extent`.

`cudaMemcpy3D()` returns an error if the pitch of `srcPtr` or `dstPtr` exceeds the maximum allowed. The pitch of a `cudaPitchedPtr` allocated with `cudaMalloc3D()` will always be valid.

Parameters:

p - 3D memory copy parameters

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidPitchValue](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [synchronous](#) behavior for most use cases.

See also:

[cudaMalloc3D](#), [cudaMalloc3DArray](#), [cudaMemset3D](#), [cudaMemcpy3DAsync](#), [cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#), [make_cudaExtent](#), [make_cudaPos](#)

5.7.2.30 `__cuda_builtin__ cudaError_t cudaMemcpy3DAsync(const struct cudaMemcpy3DParms *p, cudaStream_t stream = 0)`

```
struct cudaExtent {
    size_t width;
    size_t height;
    size_t depth;
};
struct cudaExtent make_cudaExtent(size_t w, size_t h, size_t d);

struct cudaPos {
    size_t x;
    size_t y;
    size_t z;
};
struct cudaPos make_cudaPos(size_t x, size_t y, size_t z);

struct cudaMemcpy3DParms {
    cudaArray_t      srcArray;
    struct cudaPos   srcPos;
    struct cudaPitchedPtr srcPtr;
    cudaArray_t      dstArray;
    struct cudaPos   dstPos;
    struct cudaPitchedPtr dstPtr;
    struct cudaExtent extent;
    enum cudaMemcpyKind kind;
};
```

[cudaMemcpy3DAsync\(\)](#) copies data between two 3D objects. The source and destination objects may be in either host memory, device memory, or a CUDA array. The source, destination, extent, and kind of copy performed is specified by the [cudaMemcpy3DParms](#) struct which should be initialized to zero before use:

```
cudaMemcpy3DParms myParms = {0};
```

The struct passed to [cudaMemcpy3DAsync\(\)](#) must specify one of `srcArray` or `srcPtr` and one of `dstArray` or `dstPtr`. Passing more than one non-zero source or destination will cause [cudaMemcpy3DAsync\(\)](#) to return an error.

The `srcPos` and `dstPos` fields are optional offsets into the source and destination objects and are defined in units of each object's elements. The element for a host or device pointer is assumed to be **unsigned char**. For CUDA arrays, positions must be in the range [0, 2048) for any dimension.

The `extent` field defines the dimensions of the transferred area in elements. If a CUDA array is participating in the copy, the extent is defined in terms of that array's elements. If no CUDA array is participating in the copy then the extents are defined in elements of **unsigned char**.

The `kind` field defines the direction of the copy. It must be one of `cudaMemcpyHostToHost`, `cudaMemcpyHostToDevice`, `cudaMemcpyDeviceToHost`, or `cudaMemcpyDeviceToDevice`.

If the source and destination are both arrays, `cudaMemcpy3DAsync()` will return an error if they do not have the same element size.

The source and destination object may not overlap. If overlapping source and destination objects are specified, undefined behavior will result.

The source object must lie entirely within the region defined by `srcPos` and `extent`. The destination object must lie entirely within the region defined by `dstPos` and `extent`.

`cudaMemcpy3DAsync()` returns an error if the pitch of `srcPtr` or `dstPtr` exceeds the maximum allowed. The pitch of a `cudaPitchedPtr` allocated with `cudaMalloc3D()` will always be valid.

`cudaMemcpy3DAsync()` is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is `cudaMemcpyHostToDevice` or `cudaMemcpyDeviceToHost` and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

p - 3D memory copy parameters

stream - Stream identifier

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidDevicePointer`, `cudaErrorInvalidPitchValue`, `cudaErrorInvalidMemcpyDirection`

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits **asynchronous** behavior for most use cases.

See also:

`cudaMalloc3D`, `cudaMalloc3DArray`, `cudaMemset3D`, `cudaMemcpy3D`, `cudaMemcpy`, `cudaMemcpy2D`, `cudaMemcpyToArray`, `cudaMemcpy2DToArray`, `cudaMemcpyFromArray`, `cudaMemcpy2DFromArray`, `cudaMemcpyFromArrayToArray`, `cudaMemcpy2DFromArrayToArray`, `cudaMemcpyToSymbol`, `cudaMemcpyFromSymbol`, `cudaMemcpyAsync`, `cudaMemcpy2DAsync`, `cudaMemcpyToArrayAsync`, `cudaMemcpy2DToArrayAsync`, `cudaMemcpyFromArrayAsync`, `cudaMemcpy2DFromArrayAsync`, `cudaMemcpyToSymbolAsync`, `cudaMemcpyFromSymbolAsync`, `make_cudaExtent`, `make_cudaPos`

5.7.2.31 `cudaError_t cudaMemcpy3DPeer (const struct cudaMemcpy3DPeerParms * p)`

Perform a 3D memory copy according to the parameters specified in `p`. See the definition of the `cudaMemcpy3DPeerParms` structure for documentation of its parameters.

Note that this function is synchronous with respect to the host only if the source or destination of the transfer is host memory. Note also that this copy is serialized with respect to all pending and future asynchronous work in to the current device, the copy's source device, and the copy's destination device (use `cudaMemcpy3DPeerAsync` to avoid this synchronization).

Parameters:

p - Parameters for the memory copy

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [synchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpyPeer](#), [cudaMemcpyAsync](#), [cudaMemcpyPeerAsync](#), [cudaMemcpy3DPeerAsync](#)

5.7.2.32 `cudaError_t cudaMemcpy3DPeerAsync (const struct cudaMemcpy3DPeerParms *p, cudaStream_t stream = 0)`

Perform a 3D memory copy according to the parameters specified in *p*. See the definition of the [cudaMemcpy3DPeerParms](#) structure for documentation of its parameters.

Parameters:

p - Parameters for the memory copy

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [asynchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpyPeer](#), [cudaMemcpyAsync](#), [cudaMemcpyPeerAsync](#), [cudaMemcpy3DPeerAsync](#)

5.7.2.33 `cudaError_t cudaMemcpyArrayToArray (cudaArray_t dst, size_t wOffsetDst, size_t hOffsetDst, cudaArray_const_t src, size_t wOffsetSrc, size_t hOffsetSrc, size_t count, enum cudaMemcpyKind kind = cudaMemcpyDeviceToDevice)`

Copies *count* bytes from the CUDA array *src* starting at the upper left corner (*wOffsetSrc*, *hOffsetSrc*) to the CUDA array *dst* starting at the upper left corner (*wOffsetDst*, *hOffsetDst*) where *kind* is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy.

Parameters:

dst - Destination memory address

wOffsetDst - Destination starting X offset

hOffsetDst - Destination starting Y offset
src - Source memory address
wOffsetSrc - Source starting X offset
hOffsetSrc - Source starting Y offset
count - Size in bytes to copy
kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.34 `__cuda_builtin__ cudaError_t cudaMemcpyAsync(void * dst, const void * src, size_t count, enum cudaMemcpyKind kind, cudaStream_t stream = 0)`

Copies `count` bytes from the memory area pointed to by `src` to the memory area pointed to by `dst`, where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy. The memory areas may not overlap. Calling [cudaMemcpyAsync\(\)](#) with `dst` and `src` pointers that do not match the direction of the copy results in an undefined behavior.

[cudaMemcpyAsync\(\)](#) is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is [cudaMemcpyHostToDevice](#) or [cudaMemcpyDeviceToHost](#) and the `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

dst - Destination memory address
src - Source memory address
count - Size in bytes to copy
kind - Type of transfer
stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
This function exhibits [asynchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.35 `cudaError_t cudaMemcpyFromArray (void * dst, cudaArray_const_t src, size_t wOffset, size_t hOffset, size_t count, enum cudaMemcpyKind kind)`

Copies `count` bytes from the CUDA array `src` starting at the upper left corner (`wOffset`, `hOffset`) to the memory area pointed to by `dst`, where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy.

Parameters:

dst - Destination memory address
src - Source memory address
wOffset - Source starting X offset
hOffset - Source starting Y offset
count - Size in bytes to copy
kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [synchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.36 `cudaError_t cudaMemcpyFromArrayAsync (void * dst, cudaArray_const_t src, size_t wOffset, size_t hOffset, size_t count, enum cudaMemcpyKind kind, cudaStream_t stream = 0)`

Copies `count` bytes from the CUDA array `src` starting at the upper left corner (`wOffset`, `hOffset`) to the memory area pointed to by `dst`, where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy.

[cudaMemcpyFromArrayAsync\(\)](#) is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is [cudaMemcpyHostToDevice](#) or [cudaMemcpyDeviceToHost](#) and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

dst - Destination memory address

src - Source memory address

wOffset - Source starting X offset

hOffset - Source starting Y offset

count - Size in bytes to copy

kind - Type of transfer

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [asynchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.37 `cudaError_t cudaMemcpyFromSymbol (void *dst, const void *symbol, size_t count, size_t offset = 0, enum cudaMemcpyKind kind = cudaMemcpyDeviceToHost)`

Copies `count` bytes from the memory area pointed to by `offset` bytes from the start of symbol `symbol` to the memory area pointed to by `dst`. The memory areas may not overlap. `symbol` is a variable that resides in global or constant memory space. `kind` can be either [cudaMemcpyDeviceToHost](#) or [cudaMemcpyDeviceToDevice](#).

Parameters:

dst - Destination memory address

symbol - Device symbol address

count - Size in bytes to copy

offset - Offset from start of symbol in bytes

kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidSymbol](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

This function exhibits [synchronous](#) behavior for most use cases.

Use of a string naming a variable as the `symbol` paramater was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.38 `cudaError_t cudaMemcpyFromSymbolAsync(void *dst, const void *symbol, size_t count, size_t offset, enum cudaMemcpyKind kind, cudaStream_t stream = 0)`

Copies `count` bytes from the memory area pointed to by `offset` bytes from the start of symbol `symbol` to the memory area pointed to by `dst`. The memory areas may not overlap. `symbol` is a variable that resides in global or constant memory space. `kind` can be either [cudaMemcpyDeviceToHost](#) or [cudaMemcpyDeviceToDevice](#).

[cudaMemcpyFromSymbolAsync\(\)](#) is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is [cudaMemcpyDeviceToHost](#) and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

dst - Destination memory address
symbol - Device symbol address
count - Size in bytes to copy
offset - Offset from start of symbol in bytes
kind - Type of transfer
stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidSymbol](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
 This function exhibits [asynchronous](#) behavior for most use cases.
 Use of a string naming a variable as the `symbol` paramater was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#)

5.7.2.39 `cudaError_t cudaMemcpyPeer(void *dst, int dstDevice, const void *src, int srcDevice, size_t count)`

Copies memory from one device to memory on another device. `dst` is the base device pointer of the destination memory and `dstDevice` is the destination device. `src` is the base device pointer of the source memory and `srcDevice` is the source device. `count` specifies the number of bytes to copy.

Note that this function is asynchronous with respect to the host, but serialized with respect all pending and future asynchronous work in to the current device, `srcDevice`, and `dstDevice` (use [cudaMemcpyPeerAsync](#) to avoid this synchronization).

Parameters:

dst - Destination device pointer
dstDevice - Destination device
src - Source device pointer
srcDevice - Source device
count - Size of memory copy in bytes

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [synchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpyAsync](#), [cudaMemcpyPeerAsync](#), [cudaMemcpy3DPeerAsync](#)

5.7.2.40 `cudaError_t cudaMemcpyPeerAsync (void * dst, int dstDevice, const void * src, int srcDevice, size_t count, cudaStream_t stream = 0)`

Copies memory from one device to memory on another device. `dst` is the base device pointer of the destination memory and `dstDevice` is the destination device. `src` is the base device pointer of the source memory and `srcDevice` is the source device. `count` specifies the number of bytes to copy.

Note that this function is asynchronous with respect to the host and all work in other streams and other devices.

Parameters:

dst - Destination device pointer
dstDevice - Destination device
src - Source device pointer
srcDevice - Source device
count - Size of memory copy in bytes
stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [asynchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpyPeer](#), [cudaMemcpyAsync](#), [cudaMemcpy3DPeerAsync](#)

5.7.2.41 `cudaError_t cudaMemcpyToArray(cudaArray_t dst, size_t wOffset, size_t hOffset, const void * src, size_t count, enum cudaMemcpyKind kind)`

Copies `count` bytes from the memory area pointed to by `src` to the CUDA array `dst` starting at the upper left corner (`wOffset`, `hOffset`), where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy.

Parameters:

dst - Destination memory address
wOffset - Destination starting X offset
hOffset - Destination starting Y offset
src - Source memory address
count - Size in bytes to copy
kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [synchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.42 `cudaError_t cudaMemcpyToArrayAsync(cudaArray_t dst, size_t wOffset, size_t hOffset, const void * src, size_t count, enum cudaMemcpyKind kind, cudaStream_t stream = 0)`

Copies `count` bytes from the memory area pointed to by `src` to the CUDA array `dst` starting at the upper left corner (`wOffset`, `hOffset`), where `kind` is one of [cudaMemcpyHostToHost](#), [cudaMemcpyHostToDevice](#), [cudaMemcpyDeviceToHost](#), or [cudaMemcpyDeviceToDevice](#), and specifies the direction of the copy.

[cudaMemcpyToArrayAsync\(\)](#) is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is [cudaMemcpyHostToDevice](#) or [cudaMemcpyDeviceToHost](#) and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

dst - Destination memory address
wOffset - Destination starting X offset
hOffset - Destination starting Y offset
src - Source memory address
count - Size in bytes to copy

kind - Type of transfer

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits [asynchronous](#) behavior for most use cases.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.43 `cudaError_t cudaMemcpyToSymbol (const void * symbol, const void * src, size_t count, size_t offset = 0, enum cudaMemcpyKind kind = cudaMemcpyHostToDevice)`

Copies `count` bytes from the memory area pointed to by `src` to the memory area pointed to by `offset` bytes from the start of symbol `symbol`. The memory areas may not overlap. `symbol` is a variable that resides in global or constant memory space. `kind` can be either [cudaMemcpyHostToDevice](#) or [cudaMemcpyDeviceToDevice](#).

Parameters:

symbol - Device symbol address

src - Source memory address

count - Size in bytes to copy

offset - Offset from start of symbol in bytes

kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidSymbol](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

This function exhibits [synchronous](#) behavior for most use cases.

Use of a string naming a variable as the `symbol` paramater was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.44 `cudaError_t cudaMemcpyToSymbolAsync(const void *symbol, const void *src, size_t count, size_t offset, enum cudaMemcpyKind kind, cudaStream_t stream = 0)`

Copies `count` bytes from the memory area pointed to by `src` to the memory area pointed to by `offset` bytes from the start of symbol `symbol`. The memory areas may not overlap. `symbol` is a variable that resides in global or constant memory space. `kind` can be either `cudaMemcpyHostToDevice` or `cudaMemcpyDeviceToDevice`.

`cudaMemcpyToSymbolAsync()` is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is `cudaMemcpyHostToDevice` and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

symbol - Device symbol address
src - Source memory address
count - Size in bytes to copy
offset - Offset from start of symbol in bytes
kind - Type of transfer
stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidSymbol](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
 This function exhibits [asynchronous](#) behavior for most use cases.
 Use of a string naming a variable as the `symbol` paramater was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.7.2.45 `cudaError_t cudaMemGetInfo(size_t *free, size_t *total)`

Returns in `*free` and `*total` respectively, the free and total amount of memory available for allocation by the device in bytes.

Parameters:

free - Returned free memory in bytes
total - Returned total memory in bytes

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidValue](#), [cudaErrorLaunchFailure](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

5.7.2.46 `cudaError_t cudaMemset (void * devPtr, int value, size_t count)`

Fills the first `count` bytes of the memory area pointed to by `devPtr` with the constant byte value `value`.

Note that this function is asynchronous with respect to the host unless `devPtr` refers to pinned host memory.

Parameters:

devPtr - Pointer to device memory

value - Value to set for each byte of specified memory

count - Size in bytes to set

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
See also [memset synchronization details](#).

See also:

[cudaMemset2D](#), [cudaMemset3D](#), [cudaMemsetAsync](#), [cudaMemset2DAsync](#), [cudaMemset3DAsync](#)

5.7.2.47 `cudaError_t cudaMemset2D (void * devPtr, size_t pitch, int value, size_t width, size_t height)`

Sets to the specified value `value` a matrix (`height` rows of `width` bytes each) pointed to by `dstPtr`. `pitch` is the width in bytes of the 2D array pointed to by `dstPtr`, including any padding added to the end of each row. This function performs fastest when the pitch is one that has been passed back by [cudaMallocPitch\(\)](#).

Note that this function is asynchronous with respect to the host unless `devPtr` refers to pinned host memory.

Parameters:

devPtr - Pointer to 2D device memory

pitch - Pitch in bytes of 2D device memory

value - Value to set for each byte of specified memory

width - Width of matrix set (columns in bytes)

height - Height of matrix set (rows)

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
See also [memset synchronization details](#).

See also:

[cudaMemset](#), [cudaMemset3D](#), [cudaMemsetAsync](#), [cudaMemset2DAsync](#), [cudaMemset3DAsync](#)

5.7.2.48 `__cuda_builtin__ cudaError_t cudaMemset2DAsync (void * devPtr, size_t pitch, int value, size_t width, size_t height, cudaStream_t stream = 0)`

Sets to the specified value `value` a matrix (`height` rows of `width` bytes each) pointed to by `dstPtr`. `pitch` is the width in bytes of the 2D array pointed to by `dstPtr`, including any padding added to the end of each row. This function performs fastest when the pitch is one that has been passed back by `cudaMallocPitch()`.

`cudaMemset2DAsync()` is asynchronous with respect to the host, so the call may return before the memset is complete. The operation can optionally be associated to a stream by passing a non-zero `stream` argument. If `stream` is non-zero, the operation may overlap with operations in other streams.

Parameters:

- devPtr* - Pointer to 2D device memory
- pitch* - Pitch in bytes of 2D device memory
- value* - Value to set for each byte of specified memory
- width* - Width of matrix set (columns in bytes)
- height* - Height of matrix set (rows)
- stream* - Stream identifier

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidDevicePointer`

Note:

Note that this function may also return error codes from previous, asynchronous launches. See also [memset synchronization details](#).

See also:

`cudaMemset`, `cudaMemset2D`, `cudaMemset3D`, `cudaMemsetAsync`, `cudaMemset3DAsync`

5.7.2.49 `cudaError_t cudaMemset3D (struct cudaPitchedPtr pitchedDevPtr, int value, struct cudaExtent extent)`

Initializes each element of a 3D array to the specified value `value`. The object to initialize is defined by `pitchedDevPtr`. The `pitch` field of `pitchedDevPtr` is the width in memory in bytes of the 3D array pointed to by `pitchedDevPtr`, including any padding added to the end of each row. The `xsize` field specifies the logical width of each row in bytes, while the `ysize` field specifies the height of each 2D slice in rows.

The extents of the initialized region are specified as a `width` in bytes, a `height` in rows, and a `depth` in slices.

Extents with `width` greater than or equal to the `xsize` of `pitchedDevPtr` may perform significantly faster than extents narrower than the `xsize`. Secondly, extents with `height` equal to the `ysize` of `pitchedDevPtr` will perform faster than when the `height` is shorter than the `ysize`.

This function performs fastest when the `pitchedDevPtr` has been allocated by `cudaMalloc3D()`.

Note that this function is asynchronous with respect to the host unless `pitchedDevPtr` refers to pinned host memory.

Parameters:

- pitchedDevPtr* - Pointer to pitched device memory
- value* - Value to set for each byte of specified memory

extent - Size parameters for where to set device memory (*width* field in bytes)

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
See also [memset synchronization details](#).

See also:

[cudaMemset](#), [cudaMemset2D](#), [cudaMemsetAsync](#), [cudaMemset2DAsync](#), [cudaMemset3DAsync](#), [cudaMalloc3D](#), [make_cudaPitchedPtr](#), [make_cudaExtent](#)

5.7.2.50 `__cuda_builtin__ cudaError_t cudaMemset3DAsync(struct cudaPitchedPtr pitchedDevPtr, int value, struct cudaExtent extent, cudaStream_t stream = 0)`

Initializes each element of a 3D array to the specified value *value*. The object to initialize is defined by *pitchedDevPtr*. The *pitch* field of *pitchedDevPtr* is the width in memory in bytes of the 3D array pointed to by *pitchedDevPtr*, including any padding added to the end of each row. The *xsize* field specifies the logical width of each row in bytes, while the *ysize* field specifies the height of each 2D slice in rows.

The extents of the initialized region are specified as a *width* in bytes, a *height* in rows, and a *depth* in slices.

Extents with *width* greater than or equal to the *xsize* of *pitchedDevPtr* may perform significantly faster than extents narrower than the *xsize*. Secondly, extents with *height* equal to the *ysize* of *pitchedDevPtr* will perform faster than when the *height* is shorter than the *ysize*.

This function performs fastest when the *pitchedDevPtr* has been allocated by [cudaMalloc3D\(\)](#).

[cudaMemset3DAsync\(\)](#) is asynchronous with respect to the host, so the call may return before the memset is complete. The operation can optionally be associated to a stream by passing a non-zero *stream* argument. If *stream* is non-zero, the operation may overlap with operations in other streams.

Parameters:

pitchedDevPtr - Pointer to pitched device memory

value - Value to set for each byte of specified memory

extent - Size parameters for where to set device memory (*width* field in bytes)

stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
See also [memset synchronization details](#).

See also:

[cudaMemset](#), [cudaMemset2D](#), [cudaMemset3D](#), [cudaMemsetAsync](#), [cudaMemset2DAsync](#), [cudaMalloc3D](#), [make_cudaPitchedPtr](#), [make_cudaExtent](#)

5.7.2.51 `__cuda_builtin__ cudaError_t cudaMemsetAsync (void * devPtr, int value, size_t count, cudaStream_t stream = 0)`

Fills the first `count` bytes of the memory area pointed to by `devPtr` with the constant byte value `value`.

`cudaMemsetAsync()` is asynchronous with respect to the host, so the call may return before the memset is complete. The operation can optionally be associated to a stream by passing a non-zero `stream` argument. If `stream` is non-zero, the operation may overlap with operations in other streams.

Parameters:

- `devPtr` - Pointer to device memory
- `value` - Value to set for each byte of specified memory
- `count` - Size in bytes to set
- `stream` - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches. See also [memset synchronization details](#).

See also:

[cudaMemset](#), [cudaMemset2D](#), [cudaMemset3D](#), [cudaMemset2DAsync](#), [cudaMemset3DAsync](#)

5.7.2.52 `struct cudaExtent make_cudaExtent (size_t w, size_t h, size_t d) [read]`

Returns a [cudaExtent](#) based on the specified input parameters `w`, `h`, and `d`.

Parameters:

- `w` - Width in bytes
- `h` - Height in elements
- `d` - Depth in elements

Returns:

[cudaExtent](#) specified by `w`, `h`, and `d`

See also:

[make_cudaPitchedPtr](#), [make_cudaPos](#)

5.7.2.53 `struct cudaPitchedPtr make_cudaPitchedPtr (void * d, size_t p, size_t xsz, size_t ysz) [read]`

Returns a [cudaPitchedPtr](#) based on the specified input parameters `d`, `p`, `xsz`, and `ysz`.

Parameters:

- `d` - Pointer to allocated memory

p - Pitch of allocated memory in bytes
xsz - Logical width of allocation in elements
ysz - Logical height of allocation in elements

Returns:

[cudaPitchedPtr](#) specified by *d*, *p*, *xsz*, and *ysz*

See also:

[make_cudaExtent](#), [make_cudaPos](#)

5.7.2.54 struct cudaPos make_cudaPos (size_t x, size_t y, size_t z) [read]

Returns a [cudaPos](#) based on the specified input parameters *x*, *y*, and *z*.

Parameters:

x - X position
y - Y position
z - Z position

Returns:

[cudaPos](#) specified by *x*, *y*, and *z*

See also:

[make_cudaExtent](#), [make_cudaPitchedPtr](#)

5.8 Unified Addressing

Functions

- [cudaError_t cudaPointerGetAttributes](#) (struct [cudaPointerAttributes](#) *attributes, const void *ptr)

Returns attributes about a specified pointer.

5.8.1 Detailed Description

This section describes the unified addressing functions of the CUDA runtime application programming interface.

5.8.2 Overview

CUDA devices can share a unified address space with the host. For these devices there is no distinction between a device pointer and a host pointer – the same pointer value may be used to access memory from the host program and from a kernel running on the device (with exceptions enumerated below).

5.8.3 Supported Platforms

Whether or not a device supports unified addressing may be queried by calling [cudaGetDeviceProperties\(\)](#) with the device property [cudaDeviceProp::unifiedAddressing](#).

Unified addressing is automatically enabled in 64-bit processes on devices with compute capability greater than or equal to 2.0.

Unified addressing is not yet supported on Windows Vista or Windows 7 for devices that do not use the TCC driver model.

5.8.4 Looking Up Information from Pointer Values

It is possible to look up information about the memory which backs a pointer value. For instance, one may want to know if a pointer points to host or device memory. As another example, in the case of device memory, one may want to know on which CUDA device the memory resides. These properties may be queried using the function [cudaPointerGetAttributes\(\)](#)

Since pointers are unique, it is not necessary to specify information about the pointers specified to [cudaMemcpy\(\)](#) and other copy functions. The copy direction [cudaMemcpyDefault](#) may be used to specify that the CUDA runtime should infer the location of the pointer from its value.

5.8.5 Automatic Mapping of Host Allocated Host Memory

All host memory allocated through all devices using [cudaMallocHost\(\)](#) and [cudaHostAlloc\(\)](#) is always directly accessible from all devices that support unified addressing. This is the case regardless of whether or not the flags [cudaHostAllocPortable](#) and [cudaHostAllocMapped](#) are specified.

The pointer value through which allocated host memory may be accessed in kernels on all devices that support unified addressing is the same as the pointer value through which that memory is accessed on the host. It is not necessary to call [cudaHostGetDevicePointer\(\)](#) to get the device pointer for these allocations.

Note that this is not the case for memory allocated using the flag [cudaHostAllocWriteCombined](#), as discussed below.

5.8.6 Direct Access of Peer Memory

Upon enabling direct access from a device that supports unified addressing to another peer device that supports unified addressing using `cudaDeviceEnablePeerAccess()` all memory allocated in the peer device using `cudaMalloc()` and `cudaMallocPitch()` will immediately be accessible by the current device. The device pointer value through which any peer's memory may be accessed in the current device is the same pointer value through which that memory may be accessed from the peer device.

5.8.7 Exceptions, Disjoint Addressing

Not all memory may be accessed on devices through the same pointer value through which they are accessed on the host. These exceptions are host memory registered using `cudaHostRegister()` and host memory allocated using the flag `cudaHostAllocWriteCombined`. For these exceptions, there exists a distinct host and device address for the memory. The device address is guaranteed to not overlap any valid host pointer range and is guaranteed to have the same value across all devices that support unified addressing.

This device address may be queried using `cudaHostGetDevicePointer()` when a device using unified addressing is current. Either the host or the unified device pointer value may be used to refer to this memory in `cudaMemcpy()` and similar functions using the `cudaMemcpyDefault` memory direction.

5.8.8 Function Documentation

5.8.8.1 `cudaError_t cudaPointerGetAttributes (struct cudaPointerAttributes * attributes, const void * ptr)`

Returns in `*attributes` the attributes of the pointer `ptr`.

The `cudaPointerAttributes` structure is defined as:

```
struct cudaPointerAttributes {
    enum cudaMemoryType memoryType;
    int device;
    void *devicePointer;
    void *hostPointer;
}
```

In this structure, the individual fields mean

- `memoryType` identifies the physical location of the memory associated with pointer `ptr`. It can be `cudaMemoryTypeHost` for host memory or `cudaMemoryTypeDevice` for device memory.
- `device` is the device against which `ptr` was allocated. If `ptr` has memory type `cudaMemoryTypeDevice` then this identifies the device on which the memory referred to by `ptr` physically resides. If `ptr` has memory type `cudaMemoryTypeHost` then this identifies the device which was current when the allocation was made (and if that device is deinitialized then this allocation will vanish with that device's state).
- `devicePointer` is the device pointer alias through which the memory referred to by `ptr` may be accessed on the current device. If the memory referred to by `ptr` cannot be accessed directly by the current device then this is `NULL`.
- `hostPointer` is the host pointer alias through which the memory referred to by `ptr` may be accessed on the host. If the memory referred to by `ptr` cannot be accessed directly by the host then this is `NULL`.

Parameters:

`attributes` - Attributes for the specified pointer

ptr - Pointer to get attributes for

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#)

See also:

[cudaGetDeviceCount](#), [cudaGetDevice](#), [cudaSetDevice](#), [cudaChooseDevice](#)

5.9 Peer Device Memory Access

Functions

- [cudaError_t cudaDeviceCanAccessPeer](#) (int *canAccessPeer, int device, int peerDevice)
Queries if a device may directly access a peer device's memory.
- [cudaError_t cudaDeviceDisablePeerAccess](#) (int peerDevice)
Disables direct access to memory allocations on a peer device.
- [cudaError_t cudaDeviceEnablePeerAccess](#) (int peerDevice, unsigned int flags)
Enables direct access to memory allocations on a peer device.

5.9.1 Detailed Description

This section describes the peer device memory access functions of the CUDA runtime application programming interface.

5.9.2 Function Documentation

5.9.2.1 [cudaError_t cudaDeviceCanAccessPeer](#) (int * *canAccessPeer*, int *device*, int *peerDevice*)

Returns in *canAccessPeer a value of 1 if device device is capable of directly accessing memory from peerDevice and 0 otherwise. If direct access of peerDevice from device is possible, then access may be enabled by calling [cudaDeviceEnablePeerAccess\(\)](#).

Parameters:

canAccessPeer - Returned access capability

device - Device from which allocations on peerDevice are to be directly accessed.

peerDevice - Device on which the allocations to be directly accessed by device reside.

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceEnablePeerAccess](#), [cudaDeviceDisablePeerAccess](#)

5.9.2.2 [cudaError_t cudaDeviceDisablePeerAccess](#) (int *peerDevice*)

Returns [cudaErrorPeerAccessNotEnabled](#) if direct access to memory on peerDevice has not yet been enabled from the current device.

Parameters:

peerDevice - Peer device to disable direct access to

Returns:

[cudaSuccess](#), [cudaErrorPeerAccessNotEnabled](#), [cudaErrorInvalidDevice](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceCanAccessPeer](#), [cudaDeviceEnablePeerAccess](#)

5.9.2.3 `cudaError_t cudaDeviceEnablePeerAccess (int peerDevice, unsigned int flags)`

On success, all allocations from `peerDevice` will immediately be accessible by the current device. They will remain accessible until access is explicitly disabled using [cudaDeviceDisablePeerAccess\(\)](#) or either device is reset using [cudaDeviceReset\(\)](#).

Note that access granted by this call is unidirectional and that in order to access memory on the current device from `peerDevice`, a separate symmetric call to [cudaDeviceEnablePeerAccess\(\)](#) is required.

Peer access is not supported in 32 bit applications.

Returns [cudaErrorInvalidDevice](#) if [cudaDeviceCanAccessPeer\(\)](#) indicates that the current device cannot directly access memory from `peerDevice`.

Returns [cudaErrorPeerAccessAlreadyEnabled](#) if direct access of `peerDevice` from the current device has already been enabled.

Returns [cudaErrorInvalidValue](#) if `flags` is not 0.

Parameters:

peerDevice - Peer device to enable direct access to from the current device

flags - Reserved for future use and must be set to 0

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#), [cudaErrorPeerAccessAlreadyEnabled](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaDeviceCanAccessPeer](#), [cudaDeviceDisablePeerAccess](#)

5.10 OpenGL Interoperability

Enumerations

- enum `cudaGLDeviceList` {
`cudaGLDeviceListAll` = 1,
`cudaGLDeviceListCurrentFrame` = 2,
`cudaGLDeviceListNextFrame` = 3 }

Functions

- `cudaError_t cudaGLGetDevices` (unsigned int *pCudaDeviceCount, int *pCudaDevices, unsigned int cudaDeviceCount, enum `cudaGLDeviceList` deviceList)
Gets the CUDA devices associated with the current OpenGL context.
- `cudaError_t cudaGraphicsGLRegisterBuffer` (struct `cudaGraphicsResource` **resource, GLuint buffer, unsigned int flags)
Registers an OpenGL buffer object.
- `cudaError_t cudaGraphicsGLRegisterImage` (struct `cudaGraphicsResource` **resource, GLuint image, GLenum target, unsigned int flags)
Register an OpenGL texture or renderbuffer object.
- `cudaError_t cudaWGLGetDevice` (int *device, HGPUNV hGpu)
Gets the CUDA device associated with hGpu.

5.10.1 Detailed Description

This section describes the OpenGL interoperability functions of the CUDA runtime application programming interface. Note that mapping of OpenGL resources is performed with the graphics API agnostic, resource mapping interface described in [Graphics Interopability](#).

5.10.2 Enumeration Type Documentation

5.10.2.1 enum `cudaGLDeviceList`

CUDA devices corresponding to the current OpenGL context

Enumerator:

cudaGLDeviceListAll The CUDA devices for all GPUs used by the current OpenGL context

cudaGLDeviceListCurrentFrame The CUDA devices for the GPUs used by the current OpenGL context in its currently rendering frame

cudaGLDeviceListNextFrame The CUDA devices for the GPUs to be used by the current OpenGL context in the next frame

5.10.3 Function Documentation

5.10.3.1 `cudaError_t cudaGLGetDevices (unsigned int *pCudaDeviceCount, int *pCudaDevices, unsigned int cudaDeviceCount, enum cudaGLDeviceList deviceList)`

Returns in `*pCudaDeviceCount` the number of CUDA-compatible devices corresponding to the current OpenGL context. Also returns in `*pCudaDevices` at most `cudaDeviceCount` of the CUDA-compatible devices corresponding to the current OpenGL context. If any of the GPUs being used by the current OpenGL context are not CUDA capable then the call will return `cudaErrorNoDevice`.

Parameters:

pCudaDeviceCount - Returned number of CUDA devices corresponding to the current OpenGL context

pCudaDevices - Returned CUDA devices corresponding to the current OpenGL context

cudaDeviceCount - The size of the output device array `pCudaDevices`

deviceList - The set of devices to return. This set may be `cudaGLDeviceListAll` for all devices, `cudaGLDeviceListCurrentFrame` for the devices used to render the current frame (in SLI), or `cudaGLDeviceListNextFrame` for the devices used to render the next frame (in SLI).

Returns:

`cudaSuccess`, `cudaErrorNoDevice`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaGraphicsUnregisterResource`, `cudaGraphicsMapResources`, `cudaGraphicsSubResourceGetMappedArray`, `cudaGraphicsResourceGetMappedPointer`

5.10.3.2 `cudaError_t cudaGraphicsGLRegisterBuffer (struct cudaGraphicsResource **resource, GLuint buffer, unsigned int flags)`

Registers the buffer object specified by `buffer` for access by CUDA. A handle to the registered object is returned as `resource`. The register flags `flags` specify the intended usage, as follows:

- `cudaGraphicsRegisterFlagsNone`: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA. This is the default value.
- `cudaGraphicsRegisterFlagsReadOnly`: Specifies that CUDA will not write to this resource.
- `cudaGraphicsRegisterFlagsWriteDiscard`: Specifies that CUDA will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

Parameters:

resource - Pointer to the returned object handle

buffer - name of buffer object to be registered

flags - Register flags

Returns:

`cudaSuccess`, `cudaErrorInvalidDevice`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnregisterResource](#), [cudaGraphicsMapResources](#), [cudaGraphicsResourceGetMappedPointer](#)

5.10.3.3 `cudaError_t cudaGraphicsGLRegisterImage` (struct `cudaGraphicsResource ** resource`, `GLuint image`, `GLenum target`, unsigned int `flags`)

Registers the texture or renderbuffer object specified by `image` for access by CUDA. A handle to the registered object is returned as `resource`.

`target` must match the type of the object, and must be one of `GL_TEXTURE_2D`, `GL_TEXTURE_RECTANGLE`, `GL_TEXTURE_CUBE_MAP`, `GL_TEXTURE_3D`, `GL_TEXTURE_2D_ARRAY`, or `GL_RENDERBUFFER`.

The register flags `flags` specify the intended usage, as follows:

- [cudaGraphicsRegisterFlagsNone](#): Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA. This is the default value.
- [cudaGraphicsRegisterFlagsReadOnly](#): Specifies that CUDA will not write to this resource.
- [cudaGraphicsRegisterFlagsWriteDiscard](#): Specifies that CUDA will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.
- [cudaGraphicsRegisterFlagsSurfaceLoadStore](#): Specifies that CUDA will bind this resource to a surface reference.
- [cudaGraphicsRegisterFlagsTextureGather](#): Specifies that CUDA will perform texture gather operations on this resource.

The following image formats are supported. For brevity's sake, the list is abbreviated. For ex., `{GL_R, GL_RG} X {8, 16}` would expand to the following 4 formats `{GL_R8, GL_R16, GL_RG8, GL_RG16}` :

- `GL_RED, GL_RG, GL_RGBA, GL_LUMINANCE, GL_ALPHA, GL_LUMINANCE_ALPHA, GL_INTENSITY`
- `{GL_R, GL_RG, GL_RGBA} X {8, 16, 16F, 32F, 8UI, 16UI, 32UI, 8I, 16I, 32I}`
- `{GL_LUMINANCE, GL_ALPHA, GL_LUMINANCE_ALPHA, GL_INTENSITY} X {8, 16, 16F_ARB, 32F_ARB, 8UI_EXT, 16UI_EXT, 32UI_EXT, 8I_EXT, 16I_EXT, 32I_EXT}`

The following image classes are currently disallowed:

- Textures with borders
- Multisampled renderbuffers

Parameters:

resource - Pointer to the returned object handle

image - name of texture or renderbuffer object to be registered

target - Identifies the type of object specified by *image*

flags - Register flags

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnregisterResource](#), [cudaGraphicsMapResources](#), [cudaGraphicsSubResourceGetMappedArray](#)

5.10.3.4 `cudaError_t cudaWGLGetDevice (int * device, HGPUNV hGpu)`

Returns the CUDA device associated with a hGpu, if applicable.

Parameters:

device - Returns the device associated with hGpu, or -1 if hGpu is not a compute device.

hGpu - Handle to a GPU, as queried via WGL_NV_gpu_affinity

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

WGL_NV_gpu_affinity

5.11 OpenGL Interoperability [DEPRECATED]

Enumerations

- enum `cudaGLMapFlags` {
 `cudaGLMapFlagsNone` = 0,
 `cudaGLMapFlagsReadOnly` = 1,
 `cudaGLMapFlagsWriteDiscard` = 2 }

Functions

- `cudaError_t cudaGLMapBufferObject` (void **devPtr, GLuint bufObj)
Maps a buffer object for access by CUDA.
- `cudaError_t cudaGLMapBufferObjectAsync` (void **devPtr, GLuint bufObj, `cudaStream_t` stream)
Maps a buffer object for access by CUDA.
- `cudaError_t cudaGLRegisterBufferObject` (GLuint bufObj)
Registers a buffer object for access by CUDA.
- `cudaError_t cudaGLSetBufferObjectMapFlags` (GLuint bufObj, unsigned int flags)
Set usage flags for mapping an OpenGL buffer.
- `cudaError_t cudaGLSetGLDevice` (int device)
Sets a CUDA device to use OpenGL interoperability.
- `cudaError_t cudaGLUnmapBufferObject` (GLuint bufObj)
Unmaps a buffer object for access by CUDA.
- `cudaError_t cudaGLUnmapBufferObjectAsync` (GLuint bufObj, `cudaStream_t` stream)
Unmaps a buffer object for access by CUDA.
- `cudaError_t cudaGLUnregisterBufferObject` (GLuint bufObj)
Unregisters a buffer object for access by CUDA.

5.11.1 Detailed Description

This section describes deprecated OpenGL interoperability functionality.

5.11.2 Enumeration Type Documentation

5.11.2.1 enum `cudaGLMapFlags`

CUDA GL Map Flags

Enumerator:

`cudaGLMapFlagsNone` Default; Assume resource can be read/written

cudaGLMapFlagsReadOnly CUDA kernels will not write to this resource

cudaGLMapFlagsWriteDiscard CUDA kernels will only write to and will not read from this resource

5.11.3 Function Documentation

5.11.3.1 `cudaError_t cudaGLMapBufferObject (void ** devPtr, GLuint bufObj)`

Deprecated

This function is deprecated as of CUDA 3.0.

Maps the buffer object of ID `bufObj` into the address space of CUDA and returns in `*devPtr` the base pointer of the resulting mapping. The buffer must have previously been registered by calling `cudaGLRegisterBufferObject()`. While a buffer is mapped by CUDA, any OpenGL operation which references the buffer will result in undefined behavior. The OpenGL context used to create the buffer, or another context from the same share group, must be bound to the current thread when this is called.

All streams in the current thread are synchronized with the current GL context.

Parameters:

devPtr - Returned device pointer to CUDA object

bufObj - Buffer object ID to map

Returns:

`cudaSuccess`, `cudaErrorMapBufferObjectFailed`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsMapResources](#)

5.11.3.2 `cudaError_t cudaGLMapBufferObjectAsync (void ** devPtr, GLuint bufObj, cudaStream_t stream)`

Deprecated

This function is deprecated as of CUDA 3.0.

Maps the buffer object of ID `bufObj` into the address space of CUDA and returns in `*devPtr` the base pointer of the resulting mapping. The buffer must have previously been registered by calling `cudaGLRegisterBufferObject()`. While a buffer is mapped by CUDA, any OpenGL operation which references the buffer will result in undefined behavior. The OpenGL context used to create the buffer, or another context from the same share group, must be bound to the current thread when this is called.

Stream `/p stream` is synchronized with the current GL context.

Parameters:

devPtr - Returned device pointer to CUDA object

bufObj - Buffer object ID to map

stream - Stream to synchronize

Returns:

[cudaSuccess](#), [cudaErrorMapBufferObjectFailed](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsMapResources](#)

5.11.3.3 `cudaError_t cudaGLRegisterBufferObject (GLuint bufObj)`

Deprecated

This function is deprecated as of CUDA 3.0.

Registers the buffer object of ID `bufObj` for access by CUDA. This function must be called before CUDA can map the buffer object. The OpenGL context used to create the buffer, or another context from the same share group, must be bound to the current thread when this is called.

Parameters:

bufObj - Buffer object ID to register

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsGLRegisterBuffer](#)

5.11.3.4 `cudaError_t cudaGLSetBufferObjectMapFlags (GLuint bufObj, unsigned int flags)`

Deprecated

This function is deprecated as of CUDA 3.0.

Set flags for mapping the OpenGL buffer `bufObj`

Changes to flags will take effect the next time `bufObj` is mapped. The `flags` argument may be any of the following:

- [cudaGLMapFlagsNone](#): Specifies no hints about how this buffer will be used. It is therefore assumed that this buffer will be read from and written to by CUDA kernels. This is the default value.
- [cudaGLMapFlagsReadOnly](#): Specifies that CUDA kernels which access this buffer will not write to the buffer.

- [cudaGLMapFlagsWriteDiscard](#): Specifies that CUDA kernels which access this buffer will not read from the buffer and will write over the entire contents of the buffer, so none of the data previously stored in the buffer will be preserved.

If `bufObj` has not been registered for use with CUDA, then [cudaErrorInvalidResourceHandle](#) is returned. If `bufObj` is presently mapped for access by CUDA, then [cudaErrorUnknown](#) is returned.

Parameters:

bufObj - Registered buffer object to set flags for

flags - Parameters for buffer mapping

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceSetMapFlags](#)

5.11.3.5 `cudaError_t cudaGLSetGLDevice (int device)`

Deprecated

This function is deprecated as of CUDA 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA device with an OpenGL context in order to achieve maximum interoperability performance.

Parameters:

device - Device to use for OpenGL interoperability

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#), [cudaErrorSetOnActiveProcess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsGLRegisterBuffer](#), [cudaGraphicsGLRegisterImage](#)

5.11.3.6 `cudaError_t cudaGLUnmapBufferObject (GLuint bufObj)`

Deprecated

This function is deprecated as of CUDA 3.0.

Unmaps the buffer object of ID `bufObj` for access by CUDA. When a buffer is unmapped, the base address returned by [cudaGLMapBufferObject\(\)](#) is invalid and subsequent references to the address result in undefined behavior. The OpenGL context used to create the buffer, or another context from the same share group, must be bound to the current thread when this is called.

All streams in the current thread are synchronized with the current GL context.

Parameters:

bufObj - Buffer object to unmap

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorUnmapBufferObjectFailed](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnmapResources](#)

5.11.3.7 `cudaError_t cudaGLUnmapBufferObjectAsync (GLuint bufObj, cudaStream_t stream)`

Deprecated

This function is deprecated as of CUDA 3.0.

Unmaps the buffer object of ID `bufObj` for access by CUDA. When a buffer is unmapped, the base address returned by [cudaGLMapBufferObject\(\)](#) is invalid and subsequent references to the address result in undefined behavior. The OpenGL context used to create the buffer, or another context from the same share group, must be bound to the current thread when this is called.

Stream `/p stream` is synchronized with the current GL context.

Parameters:

bufObj - Buffer object to unmap

stream - Stream to synchronize

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorUnmapBufferObjectFailed](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnmapResources](#)

5.11.3.8 `cudaError_t cudaGLUnregisterBufferObject (GLuint bufObj)`

Deprecated

This function is deprecated as of CUDA 3.0.

Unregisters the buffer object of ID `bufObj` for access by CUDA and releases any CUDA resources associated with the buffer. Once a buffer is unregistered, it may no longer be mapped by CUDA. The GL context used to create the buffer, or another context from the same share group, must be bound to the current thread when this is called.

Parameters:

bufObj - Buffer object to unregister

Returns:

`cudaSuccess`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnregisterResource](#)

5.12 Direct3D 9 Interoperability

Enumerations

- enum `cudaD3D9DeviceList` {
`cudaD3D9DeviceListAll` = 1,
`cudaD3D9DeviceListCurrentFrame` = 2,
`cudaD3D9DeviceListNextFrame` = 3 }

Functions

- `cudaError_t cudaD3D9GetDevice` (int *device, const char *pszAdapterName)
Gets the device number for an adapter.
- `cudaError_t cudaD3D9GetDevices` (unsigned int *pCudaDeviceCount, int *pCudaDevices, unsigned int cudaDeviceCount, IDirect3DDevice9 *pD3D9Device, enum `cudaD3D9DeviceList` deviceList)
Gets the CUDA devices corresponding to a Direct3D 9 device.
- `cudaError_t cudaD3D9GetDirect3DDevice` (IDirect3DDevice9 **ppD3D9Device)
Gets the Direct3D device against which the current CUDA context was created.
- `cudaError_t cudaD3D9SetDirect3DDevice` (IDirect3DDevice9 *pD3D9Device, int device=-1)
Sets the Direct3D 9 device to use for interoperability with a CUDA device.
- `cudaError_t cudaGraphicsD3D9RegisterResource` (struct `cudaGraphicsResource` **resource, IDirect3DResource9 *pD3DResource, unsigned int flags)
Register a Direct3D 9 resource for access by CUDA.

5.12.1 Detailed Description

This section describes the Direct3D 9 interoperability functions of the CUDA runtime application programming interface. Note that mapping of Direct3D 9 resources is performed with the graphics API agnostic, resource mapping interface described in [Graphics Interoperability](#).

5.12.2 Enumeration Type Documentation

5.12.2.1 enum `cudaD3D9DeviceList`

CUDA devices corresponding to a D3D9 device

Enumerator:

`cudaD3D9DeviceListAll` The CUDA devices for all GPUs used by a D3D9 device

`cudaD3D9DeviceListCurrentFrame` The CUDA devices for the GPUs used by a D3D9 device in its currently rendering frame

`cudaD3D9DeviceListNextFrame` The CUDA devices for the GPUs to be used by a D3D9 device in the next frame

5.12.3 Function Documentation

5.12.3.1 `cudaError_t cudaD3D9GetDevice (int * device, const char * pszAdapterName)`

Returns in *device* the CUDA-compatible device corresponding to the adapter name *pszAdapterName* obtained from `EnumDisplayDevices` or `IDirect3D9::GetAdapterIdentifier()`. If no device on the adapter with name *pszAdapterName* is CUDA-compatible then the call will fail.

Parameters:

device - Returns the device corresponding to *pszAdapterName*
pszAdapterName - D3D9 adapter to get device for

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaD3D9SetDirect3DDevice](#), [cudaGraphicsD3D9RegisterResource](#),

5.12.3.2 `cudaError_t cudaD3D9GetDevices (unsigned int * pCudaDeviceCount, int * pCudaDevices, unsigned int cudaDeviceCount, IDirect3DDevice9 * pD3D9Device, enum cudaD3D9DeviceList deviceList)`

Returns in *pCudaDeviceCount* the number of CUDA-compatible devices corresponding to the Direct3D 9 device *pD3D9Device*. Also returns in *pCudaDevices* at most *cudaDeviceCount* of the the CUDA-compatible devices corresponding to the Direct3D 9 device *pD3D9Device*.

If any of the GPUs being used to render *pDevice* are not CUDA capable then the call will return [cudaErrorNoDevice](#).

Parameters:

pCudaDeviceCount - Returned number of CUDA devices corresponding to *pD3D9Device*
pCudaDevices - Returned CUDA devices corresponding to *pD3D9Device*
cudaDeviceCount - The size of the output device array *pCudaDevices*
pD3D9Device - Direct3D 9 device to query for CUDA devices
deviceList - The set of devices to return. This set may be [cudaD3D9DeviceListAll](#) for all devices, [cudaD3D9DeviceListCurrentFrame](#) for the devices used to render the current frame (in SLI), or [cudaD3D9DeviceListNextFrame](#) for the devices used to render the next frame (in SLI).

Returns:

[cudaSuccess](#), [cudaErrorNoDevice](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnregisterResource](#), [cudaGraphicsMapResources](#), [cudaGraphicsSubResourceGetMappedArray](#), [cudaGraphicsResourceGetMappedPointer](#)

5.12.3.3 `cudaError_t cudaD3D9GetDirect3DDevice (IDirect3DDevice9 ** ppD3D9Device)`

Returns in `*ppD3D9Device` the Direct3D device against which this CUDA context was created in `cudaD3D9SetDirect3DDevice()`.

Parameters:

`ppD3D9Device` - Returns the Direct3D device for this thread

Returns:

`cudaSuccess`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaD3D9SetDirect3DDevice`

5.12.3.4 `cudaError_t cudaD3D9SetDirect3DDevice (IDirect3DDevice9 * pD3D9Device, int device = -1)`

Records `pD3D9Device` as the Direct3D 9 device to use for Direct3D 9 interoperability with the CUDA device `device` and sets `device` as the current device for the calling host thread.

If `device` has already been initialized then this call will fail with the error `cudaErrorSetOnActiveProcess`. In this case it is necessary to reset `device` using `cudaDeviceReset()` before Direct3D 9 interoperability on `device` may be enabled.

Successfully initializing CUDA interoperability with `pD3D9Device` will increase the internal reference count on `pD3D9Device`. This reference count will be decremented when `device` is reset using `cudaDeviceReset()`.

Note that this function is never required for correct functionality. Use of this function will result in accelerated interoperability only when the operating system is Windows Vista or Windows 7, and the device `pD3DDdevice` is not an `IDirect3DDevice9Ex`. In all other circumstances, this function is not necessary.

Parameters:

`pD3D9Device` - Direct3D device to use for this thread

`device` - The CUDA device to use. This device must be among the devices returned when querying `cudaD3D9DeviceListAll` from `cudaD3D9GetDevices`, may be set to -1 to automatically select an appropriate CUDA device.

Returns:

`cudaSuccess`, `cudaErrorInitializationError`, `cudaErrorInvalidValue`, `cudaErrorSetOnActiveProcess`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaD3D9GetDevice`, `cudaGraphicsD3D9RegisterResource`, `cudaDeviceReset`

5.12.3.5 `cudaError_t cudaGraphicsD3D9RegisterResource (struct cudaGraphicsResource ** resource, IDirect3DResource9 * pD3DResource, unsigned int flags)`

Registers the Direct3D 9 resource `pD3DResource` for access by CUDA.

If this call is successful then the application will be able to map and unmap this resource until it is unregistered through `cudaGraphicsUnregisterResource()`. Also on success, this call will increase the internal reference count on `pD3DResource`. This reference count will be decremented when this resource is unregistered through `cudaGraphicsUnregisterResource()`.

This call potentially has a high-overhead and should not be called every frame in interactive applications.

The type of `pD3DResource` must be one of the following.

- `IDirect3DVertexBuffer9`: may be accessed through a device pointer
- `IDirect3DIndexBuffer9`: may be accessed through a device pointer
- `IDirect3DSurface9`: may be accessed through an array. Only stand-alone objects of type `IDirect3DSurface9` may be explicitly shared. In particular, individual mipmap levels and faces of cube maps may not be registered directly. To access individual surfaces associated with a texture, one must register the base texture object.
- `IDirect3DBaseTexture9`: individual surfaces on this texture may be accessed through an array.

The `flags` argument may be used to specify additional parameters at register time. The valid values for this parameter are

- `cudaGraphicsRegisterFlagsNone`: Specifies no hints about how this resource will be used.
- `cudaGraphicsRegisterFlagsSurfaceLoadStore`: Specifies that CUDA will bind this resource to a surface reference.
- `cudaGraphicsRegisterFlagsTextureGather`: Specifies that CUDA will perform texture gather operations on this resource.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

A complete list of supported formats is as follows:

- `D3DFMT_L8`
- `D3DFMT_L16`
- `D3DFMT_A8R8G8B8`
- `D3DFMT_X8R8G8B8`
- `D3DFMT_G16R16`

- D3DFMT_A8B8G8R8
- D3DFMT_A8
- D3DFMT_A8L8
- D3DFMT_Q8W8V8U8
- D3DFMT_V16U16
- D3DFMT_A16B16G16R16F
- D3DFMT_A16B16G16R16
- D3DFMT_R32F
- D3DFMT_G16R16F
- D3DFMT_A32B32G32R32F
- D3DFMT_G32R32F
- D3DFMT_R16F

If `pD3DResource` is of incorrect type or is already registered, then [cudaErrorInvalidResourceHandle](#) is returned. If `pD3DResource` cannot be registered, then [cudaErrorUnknown](#) is returned.

Parameters:

- resource* - Pointer to returned resource handle
- pD3DResource* - Direct3D resource to register
- flags* - Parameters for resource registration

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaD3D9SetDirect3DDevice](#), [cudaGraphicsUnregisterResource](#), [cudaGraphicsMapResources](#), [cudaGraphicsSubResourceGetMappedArray](#), [cudaGraphicsResourceGetMappedPointer](#)

5.13 Direct3D 9 Interoperability [DEPRECATED]

Enumerations

- enum `cudaD3D9MapFlags` {
`cudaD3D9MapFlagsNone` = 0,
`cudaD3D9MapFlagsReadOnly` = 1,
`cudaD3D9MapFlagsWriteDiscard` = 2 }
- enum `cudaD3D9RegisterFlags` {
`cudaD3D9RegisterFlagsNone` = 0,
`cudaD3D9RegisterFlagsArray` = 1 }

Functions

- `cudaError_t cudaD3D9MapResources` (int count, IDirect3DResource9 **ppResources)
Map Direct3D resources for access by CUDA.
- `cudaError_t cudaD3D9RegisterResource` (IDirect3DResource9 *pResource, unsigned int flags)
Registers a Direct3D resource for access by CUDA.
- `cudaError_t cudaD3D9ResourceGetMappedArray` (cudaArray **ppArray, IDirect3DResource9 *pResource, unsigned int face, unsigned int level)
Get an array through which to access a subresource of a Direct3D resource which has been mapped for access by CUDA.
- `cudaError_t cudaD3D9ResourceGetMappedPitch` (size_t *pPitch, size_t *pPitchSlice, IDirect3DResource9 *pResource, unsigned int face, unsigned int level)
Get the pitch of a subresource of a Direct3D resource which has been mapped for access by CUDA.
- `cudaError_t cudaD3D9ResourceGetMappedPointer` (void **pPointer, IDirect3DResource9 *pResource, unsigned int face, unsigned int level)
Get a pointer through which to access a subresource of a Direct3D resource which has been mapped for access by CUDA.
- `cudaError_t cudaD3D9ResourceGetMappedSize` (size_t *pSize, IDirect3DResource9 *pResource, unsigned int face, unsigned int level)
Get the size of a subresource of a Direct3D resource which has been mapped for access by CUDA.
- `cudaError_t cudaD3D9ResourceGetSurfaceDimensions` (size_t *pWidth, size_t *pHeight, size_t *pDepth, IDirect3DResource9 *pResource, unsigned int face, unsigned int level)
Get the dimensions of a registered Direct3D surface.
- `cudaError_t cudaD3D9ResourceSetMapFlags` (IDirect3DResource9 *pResource, unsigned int flags)
Set usage flags for mapping a Direct3D resource.
- `cudaError_t cudaD3D9UnmapResources` (int count, IDirect3DResource9 **ppResources)
Unmap Direct3D resources for access by CUDA.
- `cudaError_t cudaD3D9UnregisterResource` (IDirect3DResource9 *pResource)
Unregisters a Direct3D resource for access by CUDA.

5.13.1 Detailed Description

This section describes deprecated Direct3D 9 interoperability functions.

5.13.2 Enumeration Type Documentation

5.13.2.1 enum cudaD3D9MapFlags

CUDA D3D9 Map Flags

Enumerator:

cudaD3D9MapFlagsNone Default; Assume resource can be read/written

cudaD3D9MapFlagsReadOnly CUDA kernels will not write to this resource

cudaD3D9MapFlagsWriteDiscard CUDA kernels will only write to and will not read from this resource

5.13.2.2 enum cudaD3D9RegisterFlags

CUDA D3D9 Register Flags

Enumerator:

cudaD3D9RegisterFlagsNone Default; Resource can be accessed through a void*

cudaD3D9RegisterFlagsArray Resource can be accessed through a CUarray*

5.13.3 Function Documentation

5.13.3.1 cudaError_t cudaD3D9MapResources (int count, IDirect3DResource9 ** ppResources)

Deprecated

This function is deprecated as of CUDA 3.0.

Maps the `count` Direct3D resources in `ppResources` for access by CUDA.

The resources in `ppResources` may be accessed in CUDA kernels until they are unmapped. Direct3D should not access any resources while they are mapped by CUDA. If an application does so, the results are undefined.

This function provides the synchronization guarantee that any Direct3D calls issued before `cudaD3D9MapResources()` will complete before any CUDA kernels issued after `cudaD3D9MapResources()` begin.

If any of `ppResources` have not been registered for use with CUDA or if `ppResources` contains any duplicate entries then `cudaErrorInvalidResourceHandle` is returned. If any of `ppResources` are presently mapped for access by CUDA then `cudaErrorUnknown` is returned.

Parameters:

count - Number of resources to map for CUDA

ppResources - Resources to map for CUDA

Returns:

`cudaSuccess`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsMapResources](#)

5.13.3.2 `cudaError_t cudaD3D9RegisterResource (IDirect3DResource9 * pResource, unsigned int flags)`**Deprecated**

This function is deprecated as of CUDA 3.0.

Registers the Direct3D resource `pResource` for access by CUDA.

If this call is successful, then the application will be able to map and unmap this resource until it is unregistered through [cudaD3D9UnregisterResource\(\)](#). Also on success, this call will increase the internal reference count on `pResource`. This reference count will be decremented when this resource is unregistered through [cudaD3D9UnregisterResource\(\)](#).

This call potentially has a high-overhead and should not be called every frame in interactive applications.

The type of `pResource` must be one of the following.

- `IDirect3DVertexBuffer9`: No notes.
- `IDirect3DIndexBuffer9`: No notes.
- `IDirect3DSurface9`: Only stand-alone objects of type `IDirect3DSurface9` may be explicitly shared. In particular, individual mipmap levels and faces of cube maps may not be registered directly. To access individual surfaces associated with a texture, one must register the base texture object.
- `IDirect3DBaseTexture9`: When a texture is registered, all surfaces associated with all mipmap levels of all faces of the texture will be accessible to CUDA.

The `flags` argument specifies the mechanism through which CUDA will access the Direct3D resource. The following value is allowed:

- `cudaD3D9RegisterFlagsNone`: Specifies that CUDA will access this resource through a `void*`. The pointer, size, and pitch for each subresource of this resource may be queried through [cudaD3D9ResourceGetMappedPointer\(\)](#), [cudaD3D9ResourceGetMappedSize\(\)](#), and [cudaD3D9ResourceGetMappedPitch\(\)](#) respectively. This option is valid for all resource types.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations:

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Any resources allocated in `D3DPOOL_SYSTEMMEM` or `D3DPOOL_MANAGED` may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

If Direct3D interoperability is not initialized on this context, then [cudaErrorInvalidDevice](#) is returned. If `pResource` is of incorrect type (e.g, is a non-stand-alone `IDirect3DSurface9`) or is already registered, then [cudaErrorInvalidResourceHandle](#) is returned. If `pResource` cannot be registered then [cudaErrorUnknown](#) is returned.

Parameters:

pResource - Resource to register
flags - Parameters for resource registration

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsD3D9RegisterResource](#)

5.13.3.3 `cudaError_t cudaD3D9ResourceGetMappedArray (cudaArray **ppArray, IDirect3DResource9 *pResource, unsigned int face, unsigned int level)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in `*ppArray` an array through which the subresource of the mapped Direct3D resource `pResource`, which corresponds to `face` and `level` may be accessed. The value set in `pArray` may change every time that `pResource` is mapped.

If `pResource` is not registered then [cudaErrorInvalidResourceHandle](#) is returned. If `pResource` was not registered with usage flags [cudaD3D9RegisterFlagsArray](#), then [cudaErrorInvalidResourceHandle](#) is returned. If `pResource` is not mapped, then [cudaErrorUnknown](#) is returned.

For usage requirements of `face` and `level` parameters, see [cudaD3D9ResourceGetMappedPointer\(\)](#).

Parameters:

ppArray - Returned array corresponding to subresource
pResource - Mapped resource to access
face - Face of resource to access
level - Level of resource to access

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsSubResourceGetMappedArray](#)

5.13.3.4 `cudaError_t cudaD3D9ResourceGetMappedPitch (size_t * pPitch, size_t * pPitchSlice, IDirect3DResource9 * pResource, unsigned int face, unsigned int level)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in `*pPitch` and `*pPitchSlice` the pitch and Z-slice pitch of the subresource of the mapped Direct3D resource `pResource`, which corresponds to `face` and `level`. The values set in `pPitch` and `pPitchSlice` may change every time that `pResource` is mapped.

The pitch and Z-slice pitch values may be used to compute the location of a sample on a surface as follows.

For a 2D surface, the byte offset of the sample at position `x`, `y` from the base pointer of the surface is:

$$y * \text{pitch} + (\text{bytes per pixel}) * x$$

For a 3D surface, the byte offset of the sample at position `x`, `y`, `z` from the base pointer of the surface is:

$$z * \text{slicePitch} + y * \text{pitch} + (\text{bytes per pixel}) * x$$

Both parameters `pPitch` and `pPitchSlice` are optional and may be set to `NULL`.

If `pResource` is not of type `IDirect3DBaseTexture9` or one of its sub-types or if `pResource` has not been registered for use with CUDA, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` was not registered with usage flags `cudaD3D9RegisterFlagsNone`, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` is not mapped for access by CUDA then `cudaErrorUnknown` is returned.

For usage requirements of `face` and `level` parameters, see `cudaD3D9ResourceGetMappedPointer()`.

Parameters:

- pPitch* - Returned pitch of subresource
- pPitchSlice* - Returned Z-slice pitch of subresource
- pResource* - Mapped resource to access
- face* - Face of resource to access
- level* - Level of resource to access

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceGetMappedPointer](#)

5.13.3.5 `cudaError_t cudaD3D9ResourceGetMappedPointer (void ** pPointer, IDirect3DResource9 * pResource, unsigned int face, unsigned int level)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in *pPointer* the base pointer of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *face* and *level*. The value set in *pPointer* may change every time that *pResource* is mapped.

If *pResource* is not registered, then [cudaErrorInvalidResourceHandle](#) is returned. If *pResource* was not registered with usage flags [cudaD3D9RegisterFlagsNone](#), then [cudaErrorInvalidResourceHandle](#) is returned. If *pResource* is not mapped, then [cudaErrorUnknown](#) is returned.

If *pResource* is of type `IDirect3DCubeTexture9`, then *face* must one of the values enumerated by type `D3DCUBEMAP_FACES`. For all other types, *face* must be 0. If *face* is invalid, then [cudaErrorInvalidValue](#) is returned.

If *pResource* is of type `IDirect3DBaseTexture9`, then *level* must correspond to a valid mipmap level. Only mipmap level 0 is supported for now. For all other types *level* must be 0. If *level* is invalid, then [cudaErrorInvalidValue](#) is returned.

Parameters:

pPointer - Returned pointer corresponding to subresource

pResource - Mapped resource to access

face - Face of resource to access

level - Level of resource to access

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceGetMappedPointer](#)

5.13.3.6 `cudaError_t cudaD3D9ResourceGetMappedSize (size_t * pSize, IDirect3DResource9 * pResource, unsigned int face, unsigned int level)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in *pSize* the size of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *face* and *level*. The value set in *pSize* may change every time that *pResource* is mapped.

If *pResource* has not been registered for use with CUDA then [cudaErrorInvalidResourceHandle](#) is returned. If *pResource* was not registered with usage flags [cudaD3D9RegisterFlagsNone](#), then [cudaErrorInvalidResourceHandle](#) is returned. If *pResource* is not mapped for access by CUDA then [cudaErrorUnknown](#) is returned.

For usage requirements of *face* and *level* parameters, see [cudaD3D9ResourceGetMappedPointer\(\)](#).

Parameters:

pSize - Returned size of subresource

pResource - Mapped resource to access

face - Face of resource to access

level - Level of resource to access

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceGetMappedPointer](#)

5.13.3.7 `cudaError_t cudaD3D9ResourceGetSurfaceDimensions (size_t * pWidth, size_t * pHeight, size_t * pDepth, IDirect3DResource9 * pResource, unsigned int face, unsigned int level)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in *pWidth*, *pHeight*, and *pDepth* the dimensions of the subresource of the mapped Direct3D resource *pResource* which corresponds to *face* and *level*.

Since anti-aliased surfaces may have multiple samples per pixel, it is possible that the dimensions of a resource will be an integer factor larger than the dimensions reported by the Direct3D runtime.

The parameters *pWidth*, *pHeight*, and *pDepth* are optional. For 2D surfaces, the value returned in *pDepth* will be 0.

If *pResource* is not of type `IDirect3DBaseTexture9` or `IDirect3DSurface9` or if *pResource* has not been registered for use with CUDA, then [cudaErrorInvalidResourceHandle](#) is returned.

For usage requirements of *face* and *level* parameters, see [cudaD3D9ResourceGetMappedPointer](#).

Parameters:

pWidth - Returned width of surface

pHeight - Returned height of surface

pDepth - Returned depth of surface

pResource - Registered resource to access

face - Face of resource to access

level - Level of resource to access

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsSubResourceGetMappedArray](#)

5.13.3.8 `cudaError_t cudaD3D9ResourceSetMapFlags (IDirect3DResource9 * pResource, unsigned int flags)`

Deprecated

This function is deprecated as of CUDA 3.0.

Set flags for mapping the Direct3D resource `pResource`.

Changes to flags will take effect the next time `pResource` is mapped. The `flags` argument may be any of the following:

- `cudaD3D9MapFlagsNone`: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA kernels. This is the default value.
- `cudaD3D9MapFlagsReadOnly`: Specifies that CUDA kernels which access this resource will not write to this resource.
- `cudaD3D9MapFlagsWriteDiscard`: Specifies that CUDA kernels which access this resource will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

If `pResource` has not been registered for use with CUDA, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` is presently mapped for access by CUDA, then `cudaErrorUnknown` is returned.

Parameters:

pResource - Registered resource to set flags for

flags - Parameters for resource mapping

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaInteropResourceSetMapFlags`

5.13.3.9 `cudaError_t cudaD3D9UnmapResources (int count, IDirect3DResource9 ** ppResources)`

Deprecated

This function is deprecated as of CUDA 3.0.

Unmaps the `count` Direct3D resources in `ppResources`.

This function provides the synchronization guarantee that any CUDA kernels issued before `cudaD3D9UnmapResources()` will complete before any Direct3D calls issued after `cudaD3D9UnmapResources()` begin.

If any of `ppResources` have not been registered for use with CUDA or if `ppResources` contains any duplicate entries, then `cudaErrorInvalidResourceHandle` is returned. If any of `ppResources` are not presently mapped for access by CUDA then `cudaErrorUnknown` is returned.

Parameters:

count - Number of resources to unmap for CUDA

ppResources - Resources to unmap for CUDA

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnmapResources](#)

5.13.3.10 `cudaError_t cudaD3D9UnregisterResource (IDirect3DResource9 * pResource)`**Deprecated**

This function is deprecated as of CUDA 3.0.

Unregisters the Direct3D resource `pResource` so it is not accessible by CUDA unless registered again.

If `pResource` is not registered, then [cudaErrorInvalidResourceHandle](#) is returned.

Parameters:

pResource - Resource to unregister

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnregisterResource](#)

5.14 Direct3D 10 Interoperability

Enumerations

- enum `cudaD3D10DeviceList` {
`cudaD3D10DeviceListAll` = 1,
`cudaD3D10DeviceListCurrentFrame` = 2,
`cudaD3D10DeviceListNextFrame` = 3 }

Functions

- `cudaError_t cudaD3D10GetDevice` (int *device, IDXGIAdapter *pAdapter)
Gets the device number for an adapter.
- `cudaError_t cudaD3D10GetDevices` (unsigned int *pCudaDeviceCount, int *pCudaDevices, unsigned int cudaDeviceCount, ID3D10Device *pD3D10Device, enum `cudaD3D10DeviceList` deviceList)
Gets the CUDA devices corresponding to a Direct3D 10 device.
- `cudaError_t cudaGraphicsD3D10RegisterResource` (struct `cudaGraphicsResource` **resource, ID3D10Resource *pD3DResource, unsigned int flags)
Registers a Direct3D 10 resource for access by CUDA.

5.14.1 Detailed Description

This section describes the Direct3D 10 interoperability functions of the CUDA runtime application programming interface. Note that mapping of Direct3D 10 resources is performed with the graphics API agnostic, resource mapping interface described in [Graphics Interopability](#).

5.14.2 Enumeration Type Documentation

5.14.2.1 enum `cudaD3D10DeviceList`

CUDA devices corresponding to a D3D10 device

Enumerator:

`cudaD3D10DeviceListAll` The CUDA devices for all GPUs used by a D3D10 device

`cudaD3D10DeviceListCurrentFrame` The CUDA devices for the GPUs used by a D3D10 device in its currently rendering frame

`cudaD3D10DeviceListNextFrame` The CUDA devices for the GPUs to be used by a D3D10 device in the next frame

5.14.3 Function Documentation

5.14.3.1 `cudaError_t cudaD3D10GetDevice` (int * device, IDXGIAdapter * pAdapter)

Returns in *device the CUDA-compatible device corresponding to the adapter pAdapter obtained from IDXGI-Factory::EnumAdapters. This call will succeed only if a device on adapter pAdapter is CUDA-compatible.

Parameters:

device - Returns the device corresponding to pAdapter

pAdapter - D3D10 adapter to get device for

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsD3D10RegisterResource](#),

5.14.3.2 `cudaError_t cudaD3D10GetDevices (unsigned int * pCudaDeviceCount, int * pCudaDevices, unsigned int cudaDeviceCount, ID3D10Device * pD3D10Device, enum cudaD3D10DeviceList deviceList)`

Returns in *pCudaDeviceCount the number of CUDA-compatible devices corresponding to the Direct3D 10 device pD3D10Device. Also returns in *pCudaDevices at most cudaDeviceCount of the the CUDA-compatible devices corresponding to the Direct3D 10 device pD3D10Device.

If any of the GPUs being used to render pDevice are not CUDA capable then the call will return [cudaErrorNoDevice](#).

Parameters:

pCudaDeviceCount - Returned number of CUDA devices corresponding to pD3D10Device

pCudaDevices - Returned CUDA devices corresponding to pD3D10Device

cudaDeviceCount - The size of the output device array pCudaDevices

pD3D10Device - Direct3D 10 device to query for CUDA devices

deviceList - The set of devices to return. This set may be [cudaD3D10DeviceListAll](#) for all devices, [cudaD3D10DeviceListCurrentFrame](#) for the devices used to render the current frame (in SLI), or [cudaD3D10DeviceListNextFrame](#) for the devices used to render the next frame (in SLI).

Returns:

[cudaSuccess](#), [cudaErrorNoDevice](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnregisterResource](#), [cudaGraphicsMapResources](#), [cudaGraphicsSubResourceGetMappedArray](#), [cudaGraphicsResourceGetMappedPointer](#)

5.14.3.3 `cudaError_t cudaGraphicsD3D10RegisterResource` (struct `cudaGraphicsResource` ** *resource*, `ID3D10Resource` * *pD3DResource*, unsigned int *flags*)

Registers the Direct3D 10 resource `pD3DResource` for access by CUDA.

If this call is successful, then the application will be able to map and unmap this resource until it is unregistered through `cudaGraphicsUnregisterResource()`. Also on success, this call will increase the internal reference count on `pD3DResource`. This reference count will be decremented when this resource is unregistered through `cudaGraphicsUnregisterResource()`.

This call potentially has a high-overhead and should not be called every frame in interactive applications.

The type of `pD3DResource` must be one of the following.

- `ID3D10Buffer`: may be accessed via a device pointer
- `ID3D10Texture1D`: individual subresources of the texture may be accessed via arrays
- `ID3D10Texture2D`: individual subresources of the texture may be accessed via arrays
- `ID3D10Texture3D`: individual subresources of the texture may be accessed via arrays

The `flags` argument may be used to specify additional parameters at register time. The valid values for this parameter are

- `cudaGraphicsRegisterFlagsNone`: Specifies no hints about how this resource will be used.
- `cudaGraphicsRegisterFlagsSurfaceLoadStore`: Specifies that CUDA will bind this resource to a surface reference.
- `cudaGraphicsRegisterFlagsTextureGather`: Specifies that CUDA will perform texture gather operations on this resource.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

A complete list of supported DXGI formats is as follows. For compactness the notation `A_{B,C,D}` represents `A_B`, `A_C`, and `A_D`.

- `DXGI_FORMAT_A8_UNORM`
- `DXGI_FORMAT_B8G8R8A8_UNORM`
- `DXGI_FORMAT_B8G8R8X8_UNORM`
- `DXGI_FORMAT_R16_FLOAT`
- `DXGI_FORMAT_R16G16B16A16_{FLOAT,SINT,SNORM,UINT,UNORM}`
- `DXGI_FORMAT_R16G16_{FLOAT,SINT,SNORM,UINT,UNORM}`

- `DXGI_FORMAT_R16_{SINT,SNORM,UINT,UNORM}`
- `DXGI_FORMAT_R32_FLOAT`
- `DXGI_FORMAT_R32G32B32A32_{FLOAT,SINT,UINT}`
- `DXGI_FORMAT_R32G32_{FLOAT,SINT,UINT}`
- `DXGI_FORMAT_R32_{SINT,UINT}`
- `DXGI_FORMAT_R8G8B8A8_{SINT,SNORM,UINT,UNORM,UNORM_SRGB}`
- `DXGI_FORMAT_R8G8_{SINT,SNORM,UINT,UNORM}`
- `DXGI_FORMAT_R8_{SINT,SNORM,UINT,UNORM}`

If `pD3DResource` is of incorrect type or is already registered, then `cudaErrorInvalidResourceHandle` is returned. If `pD3DResource` cannot be registered, then `cudaErrorUnknown` is returned.

Parameters:

resource - Pointer to returned resource handle

pD3DResource - Direct3D resource to register

flags - Parameters for resource registration

Returns:

`cudaSuccess`, `cudaErrorInvalidDevice`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaGraphicsUnregisterResource`, `cudaGraphicsMapResources`, `cudaGraphicsSubResourceGetMappedArray`, `cudaGraphicsResourceGetMappedPointer`

5.15 Direct3D 10 Interoperability [DEPRECATED]

Enumerations

- enum `cudaD3D10MapFlags` {
`cudaD3D10MapFlagsNone` = 0,
`cudaD3D10MapFlagsReadOnly` = 1,
`cudaD3D10MapFlagsWriteDiscard` = 2 }
- enum `cudaD3D10RegisterFlags` {
`cudaD3D10RegisterFlagsNone` = 0,
`cudaD3D10RegisterFlagsArray` = 1 }

Functions

- `cudaError_t cudaD3D10GetDirect3DDevice` (`ID3D10Device **ppD3D10Device`)
Gets the Direct3D device against which the current CUDA context was created.
- `cudaError_t cudaD3D10MapResources` (`int count`, `ID3D10Resource **ppResources`)
Maps Direct3D Resources for access by CUDA.
- `cudaError_t cudaD3D10RegisterResource` (`ID3D10Resource *pResource`, `unsigned int flags`)
Registers a Direct3D 10 resource for access by CUDA.
- `cudaError_t cudaD3D10ResourceGetMappedArray` (`cudaArray **ppArray`, `ID3D10Resource *pResource`, `unsigned int subResource`)
Gets an array through which to access a subresource of a Direct3D resource which has been mapped for access by CUDA.
- `cudaError_t cudaD3D10ResourceGetMappedPitch` (`size_t *pPitch`, `size_t *pPitchSlice`, `ID3D10Resource *pResource`, `unsigned int subResource`)
Gets the pitch of a subresource of a Direct3D resource which has been mapped for access by CUDA.
- `cudaError_t cudaD3D10ResourceGetMappedPointer` (`void **ppPointer`, `ID3D10Resource *pResource`, `unsigned int subResource`)
Gets a pointer through which to access a subresource of a Direct3D resource which has been mapped for access by CUDA.
- `cudaError_t cudaD3D10ResourceGetMappedSize` (`size_t *pSize`, `ID3D10Resource *pResource`, `unsigned int subResource`)
Gets the size of a subresource of a Direct3D resource which has been mapped for access by CUDA.
- `cudaError_t cudaD3D10ResourceGetSurfaceDimensions` (`size_t *pWidth`, `size_t *pHeight`, `size_t *pDepth`, `ID3D10Resource *pResource`, `unsigned int subResource`)
Gets the dimensions of a registered Direct3D surface.
- `cudaError_t cudaD3D10ResourceSetMapFlags` (`ID3D10Resource *pResource`, `unsigned int flags`)
Set usage flags for mapping a Direct3D resource.
- `cudaError_t cudaD3D10SetDirect3DDevice` (`ID3D10Device *pD3D10Device`, `int device=-1`)

Sets the Direct3D 10 device to use for interoperability with a CUDA device.

- [cudaError_t cudaD3D10UnmapResources](#) (int count, ID3D10Resource **ppResources)
Unmaps Direct3D resources.
- [cudaError_t cudaD3D10UnregisterResource](#) (ID3D10Resource *pResource)
Unregisters a Direct3D resource.

5.15.1 Detailed Description

This section describes deprecated Direct3D 10 interoperability functions.

5.15.2 Enumeration Type Documentation

5.15.2.1 enum cudaD3D10MapFlags

CUDA D3D10 Map Flags

Enumerator:

cudaD3D10MapFlagsNone Default; Assume resource can be read/written
cudaD3D10MapFlagsReadOnly CUDA kernels will not write to this resource
cudaD3D10MapFlagsWriteDiscard CUDA kernels will only write to and will not read from this resource

5.15.2.2 enum cudaD3D10RegisterFlags

CUDA D3D10 Register Flags

Enumerator:

cudaD3D10RegisterFlagsNone Default; Resource can be accessed through a void*
cudaD3D10RegisterFlagsArray Resource can be accessed through a CUarray*

5.15.3 Function Documentation

5.15.3.1 cudaError_t cudaD3D10GetDirect3DDevice (ID3D10Device ** ppD3D10Device)

Deprecated

This function is deprecated as of CUDA 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA device with a D3D10 device in order to achieve maximum interoperability performance.

Parameters:

ppD3D10Device - Returns the Direct3D device for this thread

Returns:

[cudaSuccess](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaD3D10SetDirect3DDevice](#)

5.15.3.2 `cudaError_t cudaD3D10MapResources (int count, ID3D10Resource ** ppResources)`**Deprecated**

This function is deprecated as of CUDA 3.0.

Maps the `count` Direct3D resources in `ppResources` for access by CUDA.

The resources in `ppResources` may be accessed in CUDA kernels until they are unmapped. Direct3D should not access any resources while they are mapped by CUDA. If an application does so, the results are undefined.

This function provides the synchronization guarantee that any Direct3D calls issued before [cudaD3D10MapResources\(\)](#) will complete before any CUDA kernels issued after [cudaD3D10MapResources\(\)](#) begin.

If any of `ppResources` have not been registered for use with CUDA or if `ppResources` contains any duplicate entries then [cudaErrorInvalidResourceHandle](#) is returned. If any of `ppResources` are presently mapped for access by CUDA then [cudaErrorUnknown](#) is returned.

Parameters:

count - Number of resources to map for CUDA

ppResources - Resources to map for CUDA

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsMapResources](#)

5.15.3.3 `cudaError_t cudaD3D10RegisterResource (ID3D10Resource * pResource, unsigned int flags)`**Deprecated**

This function is deprecated as of CUDA 3.0.

Registers the Direct3D resource `pResource` for access by CUDA.

If this call is successful, then the application will be able to map and unmap this resource until it is unregistered through [cudaD3D10UnregisterResource\(\)](#). Also on success, this call will increase the internal reference count on `pResource`. This reference count will be decremented when this resource is unregistered through [cudaD3D10UnregisterResource\(\)](#).

This call potentially has a high-overhead and should not be called every frame in interactive applications.

The type of `pResource` must be one of the following:

- `ID3D10Buffer`: Cannot be used with `flags` set to `cudaD3D10RegisterFlagsArray`.
- `ID3D10Texture1D`: No restrictions.
- `ID3D10Texture2D`: No restrictions.
- `ID3D10Texture3D`: No restrictions.

The `flags` argument specifies the mechanism through which CUDA will access the Direct3D resource. The following values are allowed.

- `cudaD3D10RegisterFlagsNone`: Specifies that CUDA will access this resource through a `void*`. The pointer, size, and pitch for each subresource of this resource may be queried through `cudaD3D10ResourceGetMappedPointer()`, `cudaD3D10ResourceGetMappedSize()`, and `cudaD3D10ResourceGetMappedPitch()` respectively. This option is valid for all resource types.
- `cudaD3D10RegisterFlagsArray`: Specifies that CUDA will access this resource through a `CUarray` queried on a sub-resource basis through `cudaD3D10ResourceGetMappedArray()`. This option is only valid for resources of type `ID3D10Texture1D`, `ID3D10Texture2D`, and `ID3D10Texture3D`.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

If Direct3D interoperability is not initialized on this context then `cudaErrorInvalidDevice` is returned. If `pResource` is of incorrect type or is already registered then `cudaErrorInvalidResourceHandle` is returned. If `pResource` cannot be registered then `cudaErrorUnknown` is returned.

Parameters:

pResource - Resource to register

flags - Parameters for resource registration

Returns:

`cudaSuccess`, `cudaErrorInvalidDevice`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsD3D10RegisterResource](#)

5.15.3.4 `cudaError_t cudaD3D10ResourceGetMappedArray (cudaArray ** ppArray, ID3D10Resource * pResource, unsigned int subResource)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in `*ppArray` an array through which the subresource of the mapped Direct3D resource `pResource` which corresponds to `subResource` may be accessed. The value set in `ppArray` may change every time that `pResource` is mapped.

If `pResource` is not registered, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` was not registered with usage flags `cudaD3D10RegisterFlagsArray`, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` is not mapped then `cudaErrorUnknown` is returned.

For usage requirements of the `subResource` parameter, see `cudaD3D10ResourceGetMappedPointer()`.

Parameters:

- `ppArray` - Returned array corresponding to subresource
- `pResource` - Mapped resource to access
- `subResource` - Subresource of `pResource` to access

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsSubResourceGetMappedArray](#)

5.15.3.5 `cudaError_t cudaD3D10ResourceGetMappedPitch (size_t * pPitch, size_t * pPitchSlice, ID3D10Resource * pResource, unsigned int subResource)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in `*pPitch` and `*pPitchSlice` the pitch and Z-slice pitch of the subresource of the mapped Direct3D resource `pResource`, which corresponds to `subResource`. The values set in `pPitch` and `pPitchSlice` may change every time that `pResource` is mapped.

The pitch and Z-slice pitch values may be used to compute the location of a sample on a surface as follows.

For a 2D surface, the byte offset of the sample at position `x`, `y` from the base pointer of the surface is:

$$y * \text{pitch} + (\text{bytes per pixel}) * x$$

For a 3D surface, the byte offset of the sample at position `x`, `y`, `z` from the base pointer of the surface is:

$$z * \text{slicePitch} + y * \text{pitch} + (\text{bytes per pixel}) * x$$

Both parameters `pPitch` and `pPitchSlice` are optional and may be set to NULL.

If `pResource` is not of type `ID3D10Texture1D`, `ID3D10Texture2D`, or `ID3D10Texture3D`, or if `pResource` has not been registered for use with CUDA, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` was

not registered with usage flags `cudaD3D10RegisterFlagsNone`, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` is not mapped for access by CUDA then `cudaErrorUnknown` is returned.

For usage requirements of the `subResource` parameter see `cudaD3D10ResourceGetMappedPointer()`.

Parameters:

- pPitch* - Returned pitch of subresource
- pPitchSlice* - Returned Z-slice pitch of subresource
- pResource* - Mapped resource to access
- subResource* - Subresource of `pResource` to access

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsSubResourceGetMappedArray](#)

5.15.3.6 `cudaError_t cudaD3D10ResourceGetMappedPointer (void ** pPointer, ID3D10Resource * pResource, unsigned int subResource)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in `*pPointer` the base pointer of the subresource of the mapped Direct3D resource `pResource` which corresponds to `subResource`. The value set in `pPointer` may change every time that `pResource` is mapped.

If `pResource` is not registered, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` was not registered with usage flags `cudaD3D9RegisterFlagsNone`, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` is not mapped then `cudaErrorUnknown` is returned.

If `pResource` is of type `ID3D10Buffer` then `subResource` must be 0. If `pResource` is of any other type, then the value of `subResource` must come from the subresource calculation in `D3D10CalcSubResource()`.

Parameters:

- pPointer* - Returned pointer corresponding to subresource
- pResource* - Mapped resource to access
- subResource* - Subresource of `pResource` to access

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceGetMappedPointer](#)

5.15.3.7 `cudaError_t cudaD3D10ResourceGetMappedSize (size_t * pSize, ID3D10Resource * pResource, unsigned int subResource)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in `*pSize` the size of the subresource of the mapped Direct3D resource `pResource` which corresponds to `subResource`. The value set in `pSize` may change every time that `pResource` is mapped.

If `pResource` has not been registered for use with CUDA then `cudaErrorInvalidHandle` is returned. If `pResource` was not registered with usage flags `cudaD3D10RegisterFlagsNone`, then `cudaErrorInvalidResourceHandle` is returned. If `pResource` is not mapped for access by CUDA then `cudaErrorUnknown` is returned.

For usage requirements of the `subResource` parameter see `cudaD3D10ResourceGetMappedPointer()`.

Parameters:

- `pSize` - Returned size of subresource
- `pResource` - Mapped resource to access
- `subResource` - Subresource of `pResource` to access

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaGraphicsResourceGetMappedPointer`

5.15.3.8 `cudaError_t cudaD3D10ResourceGetSurfaceDimensions (size_t * pWidth, size_t * pHeight, size_t * pDepth, ID3D10Resource * pResource, unsigned int subResource)`

Deprecated

This function is deprecated as of CUDA 3.0.

Returns in `*pWidth`, `*pHeight`, and `*pDepth` the dimensions of the subresource of the mapped Direct3D resource `pResource` which corresponds to `subResource`.

Since anti-aliased surfaces may have multiple samples per pixel, it is possible that the dimensions of a resource will be an integer factor larger than the dimensions reported by the Direct3D runtime.

The parameters `pWidth`, `pHeight`, and `pDepth` are optional. For 2D surfaces, the value returned in `*pDepth` will be 0.

If `pResource` is not of type `ID3D10Texture1D`, `ID3D10Texture2D`, or `ID3D10Texture3D`, or if `pResource` has not been registered for use with CUDA, then `cudaErrorInvalidHandle` is returned.

For usage requirements of `subResource` parameters see `cudaD3D10ResourceGetMappedPointer()`.

Parameters:

- `pWidth` - Returned width of surface

pHeight - Returned height of surface
pDepth - Returned depth of surface
pResource - Registered resource to access
subResource - Subresource of pResource to access

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsSubResourceGetMappedArray](#)

5.15.3.9 `cudaError_t cudaD3D10ResourceSetMapFlags (ID3D10Resource * pResource, unsigned int flags)`

Deprecated

This function is deprecated as of CUDA 3.0.

Set usage flags for mapping the Direct3D resource pResource.

Changes to flags will take effect the next time pResource is mapped. The flags argument may be any of the following:

- [cudaD3D10MapFlagsNone](#): Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA kernels. This is the default value.
- [cudaD3D10MapFlagsReadOnly](#): Specifies that CUDA kernels which access this resource will not write to this resource.
- [cudaD3D10MapFlagsWriteDiscard](#): Specifies that CUDA kernels which access this resource will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

If pResource has not been registered for use with CUDA then [cudaErrorInvalidHandle](#) is returned. If pResource is presently mapped for access by CUDA then [cudaErrorUnknown](#) is returned.

Parameters:

pResource - Registered resource to set flags for
flags - Parameters for resource mapping

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceSetMapFlags](#)

5.15.3.10 `cudaError_t cudaD3D10SetDirect3DDevice (ID3D10Device * pD3D10Device, int device = -1)`**Deprecated**

This function is deprecated as of CUDA 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA device with a D3D10 device in order to achieve maximum interoperability performance.

Parameters:

pD3D10Device - Direct3D device to use for interoperability

device - The CUDA device to use. This device must be among the devices returned when querying `cudaD3D10DeviceListAll` from `cudaD3D10GetDevices`, may be set to -1 to automatically select an appropriate CUDA device.

Returns:

`cudaSuccess`, `cudaErrorInitializationError`, `cudaErrorInvalidValue`, `cudaErrorSetOnActiveProcess`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaD3D10GetDevice`, `cudaGraphicsD3D10RegisterResource`, `cudaDeviceReset`

5.15.3.11 `cudaError_t cudaD3D10UnmapResources (int count, ID3D10Resource ** ppResources)`**Deprecated**

This function is deprecated as of CUDA 3.0.

Unmaps the `count` Direct3D resource in `ppResources`.

This function provides the synchronization guarantee that any CUDA kernels issued before `cudaD3D10UnmapResources()` will complete before any Direct3D calls issued after `cudaD3D10UnmapResources()` begin.

If any of `ppResources` have not been registered for use with CUDA or if `ppResources` contains any duplicate entries, then `cudaErrorInvalidResourceHandle` is returned. If any of `ppResources` are not presently mapped for access by CUDA then `cudaErrorUnknown` is returned.

Parameters:

count - Number of resources to unmap for CUDA

ppResources - Resources to unmap for CUDA

Returns:

`cudaSuccess`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaGraphicsUnmapResources`

5.15.3.12 `cudaError_t cudaD3D10UnregisterResource (ID3D10Resource * pResource)`

Deprecated

This function is deprecated as of CUDA 3.0.

Unregisters the Direct3D resource `resource` so it is not accessible by CUDA unless registered again.

If `pResource` is not registered, then [`cudaErrorInvalidResourceHandle`](#) is returned.

Parameters:

pResource - Resource to unregister

Returns:

[`cudaSuccess`](#), [`cudaErrorInvalidResourceHandle`](#), [`cudaErrorUnknown`](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[`cudaGraphicsUnregisterResource`](#)

5.16 Direct3D 11 Interoperability

Enumerations

- enum `cudaD3D11DeviceList` {
`cudaD3D11DeviceListAll` = 1,
`cudaD3D11DeviceListCurrentFrame` = 2,
`cudaD3D11DeviceListNextFrame` = 3 }

Functions

- `cudaError_t cudaD3D11GetDevice` (int *device, IDXGIAdapter *pAdapter)
Gets the device number for an adapter.
- `cudaError_t cudaD3D11GetDevices` (unsigned int *pCudaDeviceCount, int *pCudaDevices, unsigned int cudaDeviceCount, ID3D11Device *pD3D11Device, enum `cudaD3D11DeviceList` deviceList)
Gets the CUDA devices corresponding to a Direct3D 11 device.
- `cudaError_t cudaGraphicsD3D11RegisterResource` (struct `cudaGraphicsResource` **resource, ID3D11Resource *pD3DResource, unsigned int flags)
Register a Direct3D 11 resource for access by CUDA.

5.16.1 Detailed Description

This section describes the Direct3D 11 interoperability functions of the CUDA runtime application programming interface. Note that mapping of Direct3D 11 resources is performed with the graphics API agnostic, resource mapping interface described in [Graphics Interopability](#).

5.16.2 Enumeration Type Documentation

5.16.2.1 enum `cudaD3D11DeviceList`

CUDA devices corresponding to a D3D11 device

Enumerator:

`cudaD3D11DeviceListAll` The CUDA devices for all GPUs used by a D3D11 device

`cudaD3D11DeviceListCurrentFrame` The CUDA devices for the GPUs used by a D3D11 device in its currently rendering frame

`cudaD3D11DeviceListNextFrame` The CUDA devices for the GPUs to be used by a D3D11 device in the next frame

5.16.3 Function Documentation

5.16.3.1 `cudaError_t cudaD3D11GetDevice` (int * device, IDXGIAdapter * pAdapter)

Returns in *device the CUDA-compatible device corresponding to the adapter pAdapter obtained from IDXGI-Factory::EnumAdapters. This call will succeed only if a device on adapter pAdapter is CUDA-compatible.

Parameters:

device - Returns the device corresponding to pAdapter

pAdapter - D3D11 adapter to get device for

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnregisterResource](#), [cudaGraphicsMapResources](#), [cudaGraphicsSubResourceGetMappedArray](#), [cudaGraphicsResourceGetMappedPointer](#)

5.16.3.2 `cudaError_t cudaD3D11GetDevices (unsigned int * pCudaDeviceCount, int * pCudaDevices, unsigned int cudaDeviceCount, ID3D11Device * pD3D11Device, enum cudaD3D11DeviceList deviceList)`

Returns in *pCudaDeviceCount the number of CUDA-compatible devices corresponding to the Direct3D 11 device pD3D11Device. Also returns in *pCudaDevices at most cudaDeviceCount of the the CUDA-compatible devices corresponding to the Direct3D 11 device pD3D11Device.

If any of the GPUs being used to render pDevice are not CUDA capable then the call will return [cudaErrorNoDevice](#).

Parameters:

pCudaDeviceCount - Returned number of CUDA devices corresponding to pD3D11Device

pCudaDevices - Returned CUDA devices corresponding to pD3D11Device

cudaDeviceCount - The size of the output device array pCudaDevices

pD3D11Device - Direct3D 11 device to query for CUDA devices

deviceList - The set of devices to return. This set may be [cudaD3D11DeviceListAll](#) for all devices, [cudaD3D11DeviceListCurrentFrame](#) for the devices used to render the current frame (in SLI), or [cudaD3D11DeviceListNextFrame](#) for the devices used to render the next frame (in SLI).

Returns:

[cudaSuccess](#), [cudaErrorNoDevice](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsUnregisterResource](#), [cudaGraphicsMapResources](#), [cudaGraphicsSubResourceGetMappedArray](#), [cudaGraphicsResourceGetMappedPointer](#)

5.16.3.3 `cudaError_t cudaGraphicsD3D11RegisterResource` (struct `cudaGraphicsResource` ** *resource*, `ID3D11Resource` * *pD3DResource*, unsigned int *flags*)

Registers the Direct3D 11 resource `pD3DResource` for access by CUDA.

If this call is successful, then the application will be able to map and unmap this resource until it is unregistered through `cudaGraphicsUnregisterResource()`. Also on success, this call will increase the internal reference count on `pD3DResource`. This reference count will be decremented when this resource is unregistered through `cudaGraphicsUnregisterResource()`.

This call potentially has a high-overhead and should not be called every frame in interactive applications.

The type of `pD3DResource` must be one of the following.

- `ID3D11Buffer`: may be accessed via a device pointer
- `ID3D11Texture1D`: individual subresources of the texture may be accessed via arrays
- `ID3D11Texture2D`: individual subresources of the texture may be accessed via arrays
- `ID3D11Texture3D`: individual subresources of the texture may be accessed via arrays

The `flags` argument may be used to specify additional parameters at register time. The valid values for this parameter are

- `cudaGraphicsRegisterFlagsNone`: Specifies no hints about how this resource will be used.
- `cudaGraphicsRegisterFlagsSurfaceLoadStore`: Specifies that CUDA will bind this resource to a surface reference.
- `cudaGraphicsRegisterFlagsTextureGather`: Specifies that CUDA will perform texture gather operations on this resource.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

A complete list of supported DXGI formats is as follows. For compactness the notation `A_{B,C,D}` represents `A_B`, `A_C`, and `A_D`.

- `DXGI_FORMAT_A8_UNORM`
- `DXGI_FORMAT_B8G8R8A8_UNORM`
- `DXGI_FORMAT_B8G8R8X8_UNORM`
- `DXGI_FORMAT_R16_FLOAT`
- `DXGI_FORMAT_R16G16B16A16_{FLOAT,SINT,SNORM,UINT,UNORM}`
- `DXGI_FORMAT_R16G16_{FLOAT,SINT,SNORM,UINT,UNORM}`

- `DXGI_FORMAT_R16_{SINT,SNORM,UINT,UNORM}`
- `DXGI_FORMAT_R32_FLOAT`
- `DXGI_FORMAT_R32G32B32A32_{FLOAT,SINT,UINT}`
- `DXGI_FORMAT_R32G32_{FLOAT,SINT,UINT}`
- `DXGI_FORMAT_R32_{SINT,UINT}`
- `DXGI_FORMAT_R8G8B8A8_{SINT,SNORM,UINT,UNORM,UNORM_SRGB}`
- `DXGI_FORMAT_R8G8_{SINT,SNORM,UINT,UNORM}`
- `DXGI_FORMAT_R8_{SINT,SNORM,UINT,UNORM}`

If `pD3DResource` is of incorrect type or is already registered, then `cudaErrorInvalidResourceHandle` is returned. If `pD3DResource` cannot be registered, then `cudaErrorUnknown` is returned.

Parameters:

resource - Pointer to returned resource handle

pD3DResource - Direct3D resource to register

flags - Parameters for resource registration

Returns:

`cudaSuccess`, `cudaErrorInvalidDevice`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaGraphicsUnregisterResource`, `cudaGraphicsMapResources`, `cudaGraphicsSubResourceGetMappedArray`, `cudaGraphicsResourceGetMappedPointer`

5.17 Direct3D 11 Interoperability [DEPRECATED]

Functions

- [cudaError_t cudaD3D11GetDirect3DDevice](#) (ID3D11Device **ppD3D11Device)
Gets the Direct3D device against which the current CUDA context was created.
- [cudaError_t cudaD3D11SetDirect3DDevice](#) (ID3D11Device *pD3D11Device, int device=-1)
Sets the Direct3D 11 device to use for interoperability with a CUDA device.

5.17.1 Detailed Description

This section describes deprecated Direct3D 11 interoperability functions.

5.17.2 Function Documentation

5.17.2.1 [cudaError_t cudaD3D11GetDirect3DDevice](#) (ID3D11Device ** *ppD3D11Device*)

Deprecated

This function is deprecated as of CUDA 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA device with a D3D11 device in order to achieve maximum interoperability performance.

Parameters:

ppD3D11Device - Returns the Direct3D device for this thread

Returns:

[cudaSuccess](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaD3D11SetDirect3DDevice](#)

5.17.2.2 [cudaError_t cudaD3D11SetDirect3DDevice](#) (ID3D11Device * *pD3D11Device*, int *device* = -1)

Deprecated

This function is deprecated as of CUDA 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA device with a D3D11 device in order to achieve maximum interoperability performance.

Parameters:

pD3D11Device - Direct3D device to use for interoperability

device - The CUDA device to use. This device must be among the devices returned when querying [cudaD3D11DeviceListAll](#) from [cudaD3D11GetDevices](#), may be set to -1 to automatically select an appropriate CUDA device.

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidValue](#), [cudaErrorSetOnActiveProcess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaD3D11GetDevice](#), [cudaGraphicsD3D11RegisterResource](#), [cudaDeviceReset](#)

5.18 VDPAU Interoperability

Functions

- [cudaError_t cudaGraphicsVDPAURegisterOutputSurface](#) (struct cudaGraphicsResource **resource, VdpOutputSurface vdpSurface, unsigned int flags)
Register a VdpOutputSurface object.
- [cudaError_t cudaGraphicsVDPAURegisterVideoSurface](#) (struct cudaGraphicsResource **resource, VdpVideoSurface vdpSurface, unsigned int flags)
Register a VdpVideoSurface object.
- [cudaError_t cudaVDPAUGetDevice](#) (int *device, VdpDevice vdpDevice, VdpGetProcAddress *vdpGetProcAddress)
Gets the CUDA device associated with a VdpDevice.
- [cudaError_t cudaVDPAUSetVDPAUDevice](#) (int device, VdpDevice vdpDevice, VdpGetProcAddress *vdpGetProcAddress)
Sets a CUDA device to use VDPAU interoperability.

5.18.1 Detailed Description

This section describes the VDPAU interoperability functions of the CUDA runtime application programming interface.

5.18.2 Function Documentation

5.18.2.1 [cudaError_t cudaGraphicsVDPAURegisterOutputSurface](#) (struct cudaGraphicsResource **resource, VdpOutputSurface vdpSurface, unsigned int flags)

Registers the VdpOutputSurface specified by `vdpSurface` for access by CUDA. A handle to the registered object is returned as `resource`. The surface's intended usage is specified using `flags`, as follows:

- [cudaGraphicsMapFlagsNone](#): Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA. This is the default value.
- [cudaGraphicsMapFlagsReadOnly](#): Specifies that CUDA will not write to this resource.
- [cudaGraphicsMapFlagsWriteDiscard](#): Specifies that CUDA will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

Parameters:

resource - Pointer to the returned object handle
vdpSurface - VDPAU object to be registered
flags - Map flags

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaVDPAUSetVDPAUDevice](#), [cudaGraphicsUnregisterResource](#), [cudaGraphicsSubResourceGetMappedArray](#)

5.18.2.2 `cudaError_t cudaGraphicsVDPAURegisterVideoSurface (struct cudaGraphicsResource ** resource, VdpVideoSurface vdpSurface, unsigned int flags)`

Registers the `VdpVideoSurface` specified by `vdpSurface` for access by CUDA. A handle to the registered object is returned as `resource`. The surface's intended usage is specified using `flags`, as follows:

- `cudaGraphicsMapFlagsNone`: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA. This is the default value.
- `cudaGraphicsMapFlagsReadOnly`: Specifies that CUDA will not write to this resource.
- `cudaGraphicsMapFlagsWriteDiscard`: Specifies that CUDA will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

Parameters:

resource - Pointer to the returned object handle

vdpSurface - VDPAU object to be registered

flags - Map flags

Returns:

`cudaSuccess`, `cudaErrorInvalidDevice`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaVDPAUSetVDPAUDevice](#), [cudaGraphicsUnregisterResource](#), [cudaGraphicsSubResourceGetMappedArray](#)

5.18.2.3 `cudaError_t cudaVDPAUGetDevice (int * device, VdpDevice vdpDevice, VdpGetProcAddress * vdpGetProcAddress)`

Returns the CUDA device associated with a `VdpDevice`, if applicable.

Parameters:

device - Returns the device associated with `vdpDevice`, or -1 if the device associated with `vdpDevice` is not a compute device.

vdpDevice - A `VdpDevice` handle

vdpGetProcAddress - VDPAU's `VdpGetProcAddress` function pointer

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaVDPAUSetVDPAUDevice](#)

5.18.2.4 `cudaError_t cudaVDPAUSetVDPAUDevice (int device, VdpDevice vdpDevice, VdpGetProcAddress * vdpGetProcAddress)`

Records `vdpDevice` as the `VdpDevice` for VDPAU interoperability with the CUDA device `device` and sets `device` as the current device for the calling host thread.

If `device` has already been initialized then this call will fail with the error [cudaErrorSetOnActiveProcess](#). In this case it is necessary to reset `device` using [cudaDeviceReset\(\)](#) before VDPAU interoperability on `device` may be enabled.

Parameters:

device - Device to use for VDPAU interoperability

vdpDevice - The `VdpDevice` to interoperate with

vdpGetProcAddress - VDPAU's `VdpGetProcAddress` function pointer

Returns:

[cudaSuccess](#), [cudaErrorInvalidDevice](#), [cudaErrorSetOnActiveProcess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsVDPAURegisterVideoSurface](#), [cudaGraphicsVDPAURegisterOutputSurface](#), [cudaDeviceReset](#)

5.19 Graphics Interoperability

Functions

- `cudaError_t cudaGraphicsMapResources` (int count, `cudaGraphicsResource_t` *resources, `cudaStream_t` stream=0)
Map graphics resources for access by CUDA.
- `cudaError_t cudaGraphicsResourceGetMappedMipmappedArray` (`cudaMipmappedArray_t` *mipmappedArray, `cudaGraphicsResource_t` resource)
Get a mipmapped array through which to access a mapped graphics resource.
- `cudaError_t cudaGraphicsResourceGetMappedPointer` (void **devPtr, size_t *size, `cudaGraphicsResource_t` resource)
Get an device pointer through which to access a mapped graphics resource.
- `cudaError_t cudaGraphicsResourceSetMapFlags` (`cudaGraphicsResource_t` resource, unsigned int flags)
Set usage flags for mapping a graphics resource.
- `cudaError_t cudaGraphicsSubResourceGetMappedArray` (`cudaArray_t` *array, `cudaGraphicsResource_t` resource, unsigned int arrayIndex, unsigned int mipLevel)
Get an array through which to access a subresource of a mapped graphics resource.
- `cudaError_t cudaGraphicsUnmapResources` (int count, `cudaGraphicsResource_t` *resources, `cudaStream_t` stream=0)
Unmap graphics resources.
- `cudaError_t cudaGraphicsUnregisterResource` (`cudaGraphicsResource_t` resource)
Unregisters a graphics resource for access by CUDA.

5.19.1 Detailed Description

This section describes the graphics interoperability functions of the CUDA runtime application programming interface.

5.19.2 Function Documentation

5.19.2.1 `cudaError_t cudaGraphicsMapResources` (int count, `cudaGraphicsResource_t` *resources, `cudaStream_t` stream = 0)

Maps the count graphics resources in resources for access by CUDA.

The resources in resources may be accessed by CUDA until they are unmapped. The graphics API from which resources were registered should not access any resources while they are mapped by CUDA. If an application does so, the results are undefined.

This function provides the synchronization guarantee that any graphics calls issued before `cudaGraphicsMapResources()` will complete before any subsequent CUDA work issued in stream begins.

If resources contains any duplicate entries then `cudaErrorInvalidResourceHandle` is returned. If any of resources are presently mapped for access by CUDA then `cudaErrorUnknown` is returned.

Parameters:

count - Number of resources to map
resources - Resources to map for CUDA
stream - Stream for synchronization

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceGetMappedPointer](#), [cudaGraphicsSubResourceGetMappedArray](#), [cudaGraphicsUnmapResources](#)

5.19.2.2 `cudaError_t cudaGraphicsResourceGetMappedMipmappedArray (cudaMipmappedArray_t * mipmappedArray, cudaGraphicsResource_t resource)`

Returns in **mipmappedArray* a mipmapped array through which the mapped graphics resource *resource* may be accessed. The value set in *mipmappedArray* may change every time that *resource* is mapped.

If *resource* is not a texture then it cannot be accessed via an array and [cudaErrorUnknown](#) is returned. If *resource* is not mapped then [cudaErrorUnknown](#) is returned.

Parameters:

mipmappedArray - Returned mipmapped array through which *resource* may be accessed
resource - Mapped resource to access

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceGetMappedPointer](#)

5.19.2.3 `cudaError_t cudaGraphicsResourceGetMappedPointer (void ** devPtr, size_t * size, cudaGraphicsResource_t resource)`

Returns in **devPtr* a pointer through which the mapped graphics resource *resource* may be accessed. Returns in **size* the size of the memory in bytes which may be accessed from that pointer. The value set in *devPtr* may change every time that *resource* is mapped.

If *resource* is not a buffer then it cannot be accessed via a pointer and [cudaErrorUnknown](#) is returned. If *resource* is not mapped then [cudaErrorUnknown](#) is returned. *

Parameters:

devPtr - Returned pointer through which `resource` may be accessed

size - Returned size of the buffer accessible starting at `*devPtr`

resource - Mapped resource to access

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsMapResources](#), [cudaGraphicsSubResourceGetMappedArray](#)

5.19.2.4 `cudaError_t cudaGraphicsResourceSetMapFlags (cudaGraphicsResource_t resource, unsigned int flags)`

Set `flags` for mapping the graphics resource `resource`.

Changes to `flags` will take effect the next time `resource` is mapped. The `flags` argument may be any of the following:

- [cudaGraphicsMapFlagsNone](#): Specifies no hints about how `resource` will be used. It is therefore assumed that CUDA may read from or write to `resource`.
- [cudaGraphicsMapFlagsReadOnly](#): Specifies that CUDA will not write to `resource`.
- [cudaGraphicsMapFlagsWriteDiscard](#): Specifies CUDA will not read from `resource` and will write over the entire contents of `resource`, so none of the data previously stored in `resource` will be preserved.

If `resource` is presently mapped for access by CUDA then [cudaErrorUnknown](#) is returned. If `flags` is not one of the above values then [cudaErrorInvalidValue](#) is returned.

Parameters:

resource - Registered resource to set flags for

flags - Parameters for resource mapping

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsMapResources](#)

5.19.2.5 `cudaError_t cudaGraphicsSubResourceGetMappedArray (cudaArray_t * array, cudaGraphicsResource_t resource, unsigned int arrayIndex, unsigned int mipLevel)`

Returns in `*array` an array through which the subresource of the mapped graphics resource `resource` which corresponds to array index `arrayIndex` and mipmap level `mipLevel` may be accessed. The value set in `array` may change every time that `resource` is mapped.

If `resource` is not a texture then it cannot be accessed via an array and `cudaErrorUnknown` is returned. If `arrayIndex` is not a valid array index for `resource` then `cudaErrorInvalidValue` is returned. If `mipLevel` is not a valid mipmap level for `resource` then `cudaErrorInvalidValue` is returned. If `resource` is not mapped then `cudaErrorUnknown` is returned.

Parameters:

array - Returned array through which a subresource of `resource` may be accessed

resource - Mapped resource to access

arrayIndex - Array index for array textures or cubemap face index as defined by `cudaGraphicsCubeFace` for cubemap textures for the subresource to access

mipLevel - Mipmap level for the subresource to access

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsResourceGetMappedPointer](#)

5.19.2.6 `cudaError_t cudaGraphicsUnmapResources (int count, cudaGraphicsResource_t * resources, cudaStream_t stream = 0)`

Unmaps the `count` graphics resources in `resources`.

Once unmapped, the resources in `resources` may not be accessed by CUDA until they are mapped again.

This function provides the synchronization guarantee that any CUDA work issued in `stream` before `cudaGraphicsUnmapResources()` will complete before any subsequently issued graphics work begins.

If `resources` contains any duplicate entries then `cudaErrorInvalidResourceHandle` is returned. If any of `resources` are not presently mapped for access by CUDA then `cudaErrorUnknown` is returned.

Parameters:

count - Number of resources to unmap

resources - Resources to unmap

stream - Stream for synchronization

Returns:

`cudaSuccess`, `cudaErrorInvalidResourceHandle`, `cudaErrorUnknown`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsMapResources](#)

5.19.2.7 `cudaError_t cudaGraphicsUnregisterResource (cudaGraphicsResource_t resource)`

Unregisters the graphics resource `resource` so it is not accessible by CUDA unless registered again.

If `resource` is invalid then [cudaErrorInvalidResourceHandle](#) is returned.

Parameters:

resource - Resource to unregister

Returns:

[cudaSuccess](#), [cudaErrorInvalidResourceHandle](#), [cudaErrorUnknown](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGraphicsD3D9RegisterResource](#), [cudaGraphicsD3D10RegisterResource](#), [cudaGraphicsD3D11RegisterResource](#), [cudaGraphicsGLRegisterBuffer](#), [cudaGraphicsGLRegisterImage](#)

5.20 Texture Reference Management

Functions

- `cudaError_t cudaBindTexture` (`size_t *offset`, `const struct textureReference *texref`, `const void *devPtr`, `const struct cudaChannelFormatDesc *desc`, `size_t size=UINT_MAX`)
Binds a memory area to a texture.
- `cudaError_t cudaBindTexture2D` (`size_t *offset`, `const struct textureReference *texref`, `const void *devPtr`, `const struct cudaChannelFormatDesc *desc`, `size_t width`, `size_t height`, `size_t pitch`)
Binds a 2D memory area to a texture.
- `cudaError_t cudaBindTextureToArray` (`const struct textureReference *texref`, `cudaArray_const_t array`, `const struct cudaChannelFormatDesc *desc`)
Binds an array to a texture.
- `cudaError_t cudaBindTextureToMipmappedArray` (`const struct textureReference *texref`, `cudaMipmappedArray_const_t mipmappedArray`, `const struct cudaChannelFormatDesc *desc`)
Binds a mipmapped array to a texture.
- `struct cudaChannelFormatDesc cudaCreateChannelDesc` (`int x`, `int y`, `int z`, `int w`, `enum cudaChannelFormatKind f`)
Returns a channel descriptor using the specified format.
- `cudaError_t cudaGetChannelDesc` (`struct cudaChannelFormatDesc *desc`, `cudaArray_const_t array`)
Get the channel descriptor of an array.
- `cudaError_t cudaGetTextureAlignmentOffset` (`size_t *offset`, `const struct textureReference *texref`)
Get the alignment offset of a texture.
- `cudaError_t cudaGetTextureReference` (`const struct textureReference **texref`, `const void *symbol`)
Get the texture reference associated with a symbol.
- `cudaError_t cudaUnbindTexture` (`const struct textureReference *texref`)
Unbinds a texture.

5.20.1 Detailed Description

This section describes the low level texture reference management functions of the CUDA runtime application programming interface.

Some functions have overloaded C++ API template versions documented separately in the [C++ API Routines](#) module.

5.20.2 Function Documentation

5.20.2.1 `cudaError_t cudaBindTexture` (`size_t *offset`, `const struct textureReference *texref`, `const void *devPtr`, `const struct cudaChannelFormatDesc *desc`, `size_t size = UINT_MAX`)

Binds `size` bytes of the memory area pointed to by `devPtr` to the texture reference `texref`. `desc` describes how the memory is interpreted when fetching values from the texture. Any memory previously bound to `texref` is unbound.

Since the hardware enforces an alignment requirement on texture base addresses, `cudaBindTexture()` returns in `*offset` a byte offset that must be applied to texture fetches in order to read from the desired memory. This offset must be divided by the texel size and passed to kernels that read from the texture so they can be applied to the `tex1Dfetch()` function. If the device memory pointer was returned from `cudaMalloc()`, the offset is guaranteed to be 0 and NULL may be passed as the `offset` parameter.

The total number of elements (or texels) in the linear address range cannot exceed `cudaDeviceProp::maxTexture1DLinear[0]`. The number of elements is computed as $(size / elementSize)$, where `elementSize` is determined from `desc`.

Parameters:

offset - Offset in bytes
texref - Texture to bind
devPtr - Memory area on device
desc - Channel format
size - Size of the memory area pointed to by *devPtr*

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidDevicePointer`, `cudaErrorInvalidTexture`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaCreateChannelDesc` (C API), `cudaGetChannelDesc`, `cudaGetTextureReference`, `cudaBindTexture` (C++ API), `cudaBindTexture2D` (C API), `cudaBindTextureToArray` (C API), `cudaUnbindTexture` (C API), `cudaGetTextureAlignmentOffset` (C API)

5.20.2.2 `cudaError_t cudaBindTexture2D (size_t * offset, const struct textureReference * texref, const void * devPtr, const struct cudaChannelFormatDesc * desc, size_t width, size_t height, size_t pitch)`

Binds the 2D memory area pointed to by `devPtr` to the texture reference `texref`. The size of the area is constrained by `width` in texel units, `height` in texel units, and `pitch` in byte units. `desc` describes how the memory is interpreted when fetching values from the texture. Any memory previously bound to `texref` is unbound.

Since the hardware enforces an alignment requirement on texture base addresses, `cudaBindTexture2D()` returns in `*offset` a byte offset that must be applied to texture fetches in order to read from the desired memory. This offset must be divided by the texel size and passed to kernels that read from the texture so they can be applied to the `tex2D()` function. If the device memory pointer was returned from `cudaMalloc()`, the offset is guaranteed to be 0 and NULL may be passed as the `offset` parameter.

`width` and `height`, which are specified in elements (or texels), cannot exceed `cudaDeviceProp::maxTexture2DLinear[0]` and `cudaDeviceProp::maxTexture2DLinear[1]` respectively. `pitch`, which is specified in bytes, cannot exceed `cudaDeviceProp::maxTexture2DLinear[2]`.

The driver returns `cudaErrorInvalidValue` if `pitch` is not a multiple of `cudaDeviceProp::texturePitchAlignment`.

Parameters:

offset - Offset in bytes
texref - Texture reference to bind

devPtr - 2D memory area on device

desc - Channel format

width - Width in texel units

height - Height in texel units

pitch - Pitch in bytes

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C API), [cudaBindTexture2D](#) (C++ API), [cudaBindTexture2D](#) (C++ API, inherited channel descriptor), [cudaBindTextureToArray](#) (C API), [cudaBindTextureToArray](#) (C API), [cudaGetTextureAlignmentOffset](#) (C API)

5.20.2.3 `cudaError_t cudaBindTextureToArray` (`const struct textureReference * texref`, `cudaArray_const_t array`, `const struct cudaChannelFormatDesc * desc`)

Binds the CUDA array `array` to the texture reference `texref`. `desc` describes how the memory is interpreted when fetching values from the texture. Any CUDA array previously bound to `texref` is unbound.

Parameters:

texref - Texture to bind

array - Memory array on device

desc - Channel format

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C API), [cudaBindTexture2D](#) (C API), [cudaBindTextureToArray](#) (C++ API), [cudaUnbindTexture](#) (C API), [cudaGetTextureAlignmentOffset](#) (C API)

5.20.2.4 `cudaError_t cudaBindTextureToMipmappedArray` (`const struct textureReference * texref`, `cudaMipmappedArray_const_t mipmappedArray`, `const struct cudaChannelFormatDesc * desc`)

Binds the CUDA mipmapped array `mipmappedArray` to the texture reference `texref`. `desc` describes how the memory is interpreted when fetching values from the texture. Any CUDA mipmapped array previously bound to `texref` is unbound.

Parameters:

texref - Texture to bind
mipmappedArray - Memory mipmapped array on device
desc - Channel format

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C API), [cudaBindTexture2D](#) (C API), [cudaBindTextureToArray](#) (C++ API), [cudaUnbindTexture](#) (C API), [cudaGetTextureAlignmentOffset](#) (C API)

5.20.2.5 `struct cudaChannelFormatDesc cudaCreateChannelDesc (int x, int y, int z, int w, enum cudaChannelFormatKind f)` [read]

Returns a channel descriptor with format *f* and number of bits of each component *x*, *y*, *z*, and *w*. The [cudaChannelFormatDesc](#) is defined as:

```
struct cudaChannelFormatDesc {
    int x, y, z, w;
    enum cudaChannelFormatKind f;
};
```

where [cudaChannelFormatKind](#) is one of [cudaChannelFormatKindSigned](#), [cudaChannelFormatKindUnsigned](#), or [cudaChannelFormatKindFloat](#).

Parameters:

x - X component
y - Y component
z - Z component
w - W component
f - Channel format

Returns:

Channel descriptor with format *f*

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C API), [cudaBindTexture2D](#) (C API), [cudaBindTextureToArray](#) (C API), [cudaUnbindTexture](#) (C API), [cudaGetTextureAlignmentOffset](#) (C API)

5.20.2.6 `cudaError_t cudaGetChannelDesc (struct cudaChannelFormatDesc * desc, cudaArray_const_t array)`

Returns in `*desc` the channel descriptor of the CUDA array `array`.

Parameters:

`desc` - Channel format
`array` - Memory array on device

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc \(C API\)](#), [cudaGetTextureReference](#), [cudaBindTexture \(C API\)](#), [cudaBindTexture2D \(C API\)](#), [cudaBindTextureToArray \(C API\)](#), [cudaUnbindTexture \(C API\)](#), [cudaGetTextureAlignmentOffset \(C API\)](#)

5.20.2.7 `cudaError_t cudaGetTextureAlignmentOffset (size_t * offset, const struct textureReference * texref)`

Returns in `*offset` the offset that was returned when texture reference `texref` was bound.

Parameters:

`offset` - Offset of texture reference in bytes
`texref` - Texture to get offset of

Returns:

[cudaSuccess](#), [cudaErrorInvalidTexture](#), [cudaErrorInvalidTextureBinding](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc \(C API\)](#), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture \(C API\)](#), [cudaBindTexture2D \(C API\)](#), [cudaBindTextureToArray \(C API\)](#), [cudaUnbindTexture \(C API\)](#), [cudaGetTextureAlignmentOffset \(C++ API\)](#)

5.20.2.8 `cudaError_t cudaGetTextureReference (const struct textureReference ** texref, const void * symbol)`

Returns in `*texref` the structure associated to the texture reference defined by symbol `symbol`.

Parameters:

`texref` - Texture reference associated with symbol

symbol - Texture to get reference for

Returns:

[cudaSuccess](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
Use of a string naming a variable as the `symbol` parameter was removed in CUDA 5.0.

See also:

[cudaCreateChannelDesc](#) (C API), [cudaGetChannelDesc](#), [cudaGetTextureAlignmentOffset](#) (C API), [cudaBindTexture](#) (C API), [cudaBindTexture2D](#) (C API), [cudaBindTextureToArray](#) (C API), [cudaUnbindTexture](#) (C API)

5.20.2.9 `cudaError_t cudaUnbindTexture (const struct textureReference * texref)`

Unbinds the texture bound to `texref`.

Parameters:

texref - Texture to unbind

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C API), [cudaBindTexture2D](#) (C API), [cudaBindTextureToArray](#) (C API), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C API)

5.21 Surface Reference Management

Functions

- [cudaError_t cudaBindSurfaceToArray](#) (const struct [surfaceReference](#) *surfref, [cudaArray_const_t](#) array, const struct [cudaChannelFormatDesc](#) *desc)
Binds an array to a surface.
- [cudaError_t cudaGetSurfaceReference](#) (const struct [surfaceReference](#) **surfref, const void *symbol)
Get the surface reference associated with a symbol.

5.21.1 Detailed Description

This section describes the low level surface reference management functions of the CUDA runtime application programming interface.

Some functions have overloaded C++ API template versions documented separately in the [C++ API Routines](#) module.

5.21.2 Function Documentation

5.21.2.1 [cudaError_t cudaBindSurfaceToArray](#) (const struct [surfaceReference](#) * *surfref*, [cudaArray_const_t](#) *array*, const struct [cudaChannelFormatDesc](#) * *desc*)

Binds the CUDA array *array* to the surface reference *surfref*. *desc* describes how the memory is interpreted when fetching values from the surface. Any CUDA array previously bound to *surfref* is unbound.

Parameters:

- surfref* - Surface to bind
- array* - Memory array on device
- desc* - Channel format

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidSurface](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaBindSurfaceToArray](#) (C++ API), [cudaBindSurfaceToArray](#) (C++ API, inherited channel descriptor), [cudaGetSurfaceReference](#)

5.21.2.2 [cudaError_t cudaGetSurfaceReference](#) (const struct [surfaceReference](#) ** *surfref*, const void * *symbol*)

Returns in **surfref* the structure associated to the surface reference defined by symbol *symbol*.

Parameters:

surfref - Surface reference associated with symbol

symbol - Surface to get reference for

Returns:

[cudaSuccess](#), [cudaErrorInvalidSurface](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
Use of a string naming a variable as the `symbol` parameter was removed in CUDA 5.0.

See also:

[cudaBindSurfaceToArray \(C API\)](#)

5.22 Texture Object Management

Functions

- `cudaError_t cudaCreateTextureObject (cudaTextureObject_t *pTexObject, const struct cudaResourceDesc *pResDesc, const struct cudaTextureDesc *pTexDesc, const struct cudaResourceViewDesc *pResViewDesc)`
Creates a texture object.
- `cudaError_t cudaDestroyTextureObject (cudaTextureObject_t texObject)`
Destroys a texture object.
- `cudaError_t cudaGetTextureObjectResourceDesc (struct cudaResourceDesc *pResDesc, cudaTextureObject_t texObject)`
Returns a texture object's resource descriptor.
- `cudaError_t cudaGetTextureObjectResourceViewDesc (struct cudaResourceViewDesc *pResViewDesc, cudaTextureObject_t texObject)`
Returns a texture object's resource view descriptor.
- `cudaError_t cudaGetTextureObjectTextureDesc (struct cudaTextureDesc *pTexDesc, cudaTextureObject_t texObject)`
Returns a texture object's texture descriptor.

5.22.1 Detailed Description

This section describes the low level texture object management functions of the CUDA runtime application programming interface. The texture object API is only supported on devices of compute capability 3.0 or higher.

5.22.2 Function Documentation

5.22.2.1 `cudaError_t cudaCreateTextureObject (cudaTextureObject_t *pTexObject, const struct cudaResourceDesc *pResDesc, const struct cudaTextureDesc *pTexDesc, const struct cudaResourceViewDesc *pResViewDesc)`

Creates a texture object and returns it in `pTexObject`. `pResDesc` describes the data to texture from. `pTexDesc` describes how the data should be sampled. `pResViewDesc` is an optional argument that specifies an alternate format for the data described by `pResDesc`, and also describes the subresource region to restrict access to when texturing. `pResViewDesc` can only be specified if the type of resource is a CUDA array or a CUDA mipmapped array.

Texture objects are only supported on devices of compute capability 3.0 or higher.

The `cudaResourceDesc` structure is defined as:

```
struct cudaResourceDesc {
    enum cudaResourceType resType;

    union {
        struct {
            cudaArray_t array;
        } array;
        struct {
            cudaMipmappedArray_t mipmap;
        } mipmap;
    }
};
```

```

    struct {
        void *devPtr;
        struct cudaChannelFormatDesc desc;
        size_t sizeInBytes;
    } linear;
    struct {
        void *devPtr;
        struct cudaChannelFormatDesc desc;
        size_t width;
        size_t height;
        size_t pitchInBytes;
    } pitch2D;
} res;
};

```

where:

- `cudaResourceDesc::resType` specifies the type of resource to texture from. `CUresourceType` is defined as:

```

enum cudaResourceType {
    cudaResourceTypeArray          = 0x00,
    cudaResourceTypeMipmappedArray = 0x01,
    cudaResourceTypeLinear         = 0x02,
    cudaResourceTypePitch2D       = 0x03
};

```

If `cudaResourceDesc::resType` is set to `cudaResourceTypeArray`, `cudaResourceDesc::res::array::array` must be set to a valid CUDA array handle.

If `cudaResourceDesc::resType` is set to `cudaResourceTypeMipmappedArray`, `cudaResourceDesc::res::mipmap::mipmap` must be set to a valid CUDA mipmapped array handle.

If `cudaResourceDesc::resType` is set to `cudaResourceTypeLinear`, `cudaResourceDesc::res::linear::devPtr` must be set to a valid device pointer, that is aligned to `cudaDeviceProp::textureAlignment`. `cudaResourceDesc::res::linear::desc` describes the format and the number of components per array element. `cudaResourceDesc::res::linear::sizeInBytes` specifies the size of the array in bytes. The total number of elements in the linear address range cannot exceed `cudaDeviceProp::maxTexture1DLinear`. The number of elements is computed as $(\text{sizeInBytes} / \text{sizeof}(\text{desc}))$.

If `cudaResourceDesc::resType` is set to `cudaResourceTypePitch2D`, `cudaResourceDesc::res::pitch2D::devPtr` must be set to a valid device pointer, that is aligned to `cudaDeviceProp::textureAlignment`. `cudaResourceDesc::res::pitch2D::desc` describes the format and the number of components per array element. `cudaResourceDesc::res::pitch2D::width` and `cudaResourceDesc::res::pitch2D::height` specify the width and height of the array in elements, and cannot exceed `cudaDeviceProp::maxTexture2DLinear[0]` and `cudaDeviceProp::maxTexture2DLinear[1]` respectively. `cudaResourceDesc::res::pitch2D::pitchInBytes` specifies the pitch between two rows in bytes and has to be aligned to `cudaDeviceProp::texturePitchAlignment`. Pitch cannot exceed `cudaDeviceProp::maxTexture2DLinear[2]`.

The `cudaTextureDesc` struct is defined as

```

struct cudaTextureDesc {
    enum cudaTextureAddressMode addressMode[3];
    enum cudaTextureFilterMode  filterMode;
    enum cudaTextureReadMode    readMode;
};

```

```

int                sRGB;
int                normalizedCoords;
unsigned int       maxAnisotropy;
enum cudaTextureFilterMode mipmapFilterMode;
float              mipmapLevelBias;
float              minMipmapLevelClamp;
float              maxMipmapLevelClamp;
};

```

where

- `cudaTextureDesc::addressMode` specifies the addressing mode for each dimension of the texture data. `cudaTextureAddressMode` is defined as:

```

enum cudaTextureAddressMode {
    cudaAddressModeWrap    = 0,
    cudaAddressModeClamp   = 1,
    cudaAddressModeMirror  = 2,
    cudaAddressModeBorder  = 3
};

```

This is ignored if `cudaResourceDesc::resType` is `cudaResourceTypeLinear`. Also, if `cudaTextureDesc::normalizedCoords` is set to zero, the only supported address mode is `cudaAddressModeClamp`.

- `cudaTextureDesc::filterMode` specifies the filtering mode to be used when fetching from the texture. `cudaTextureFilterMode` is defined as:

```

enum cudaTextureFilterMode {
    cudaFilterModePoint    = 0,
    cudaFilterModeLinear   = 1
};

```

This is ignored if `cudaResourceDesc::resType` is `cudaResourceTypeLinear`.

- `cudaTextureDesc::readMode` specifies whether integer data should be converted to floating point or not. `cudaTextureReadMode` is defined as:

```

enum cudaTextureReadMode {
    cudaReadModeElementType    = 0,
    cudaReadModeNormalizedFloat = 1
};

```

Note that this applies only to 8-bit and 16-bit integer formats. 32-bit integer format would not be promoted, regardless of whether or not this `cudaTextureDesc::readMode` is set `cudaReadModeNormalizedFloat` is specified.

- `cudaTextureDesc::sRGB` specifies whether sRGB to linear conversion should be performed during texture fetch.
- `cudaTextureDesc::normalizedCoords` specifies whether the texture coordinates will be normalized or not.
- `cudaTextureDesc::maxAnisotropy` specifies the maximum anisotropy ratio to be used when doing anisotropic filtering. This value will be clamped to the range [1,16].
- `cudaTextureDesc::mipmapFilterMode` specifies the filter mode when the calculated mipmap level lies between two defined mipmap levels.
- `cudaTextureDesc::mipmapLevelBias` specifies the offset to be applied to the calculated mipmap level.

- [cudaTextureDesc::minMipmapLevelClamp](#) specifies the lower end of the mipmap level range to clamp access to.
- [cudaTextureDesc::maxMipmapLevelClamp](#) specifies the upper end of the mipmap level range to clamp access to.

The [cudaResourceViewDesc](#) struct is defined as

```
struct cudaResourceViewDesc {
    enum cudaResourceViewFormat format;
    size_t width;
    size_t height;
    size_t depth;
    unsigned int firstMipmapLevel;
    unsigned int lastMipmapLevel;
    unsigned int firstLayer;
    unsigned int lastLayer;
};
```

where:

- [cudaResourceViewDesc::format](#) specifies how the data contained in the CUDA array or CUDA mipmapped array should be interpreted. Note that this can incur a change in size of the texture data. If the resource view format is a block compressed format, then the underlying CUDA array or CUDA mipmapped array has to have a 32-bit unsigned integer format with 2 or 4 channels, depending on the block compressed format. For ex., BC1 and BC4 require the underlying CUDA array to have a 32-bit unsigned int with 2 channels. The other BC formats require the underlying resource to have the same 32-bit unsigned int format but with 4 channels.
- [cudaResourceViewDesc::width](#) specifies the new width of the texture data. If the resource view format is a block compressed format, this value has to be 4 times the original width of the resource. For non block compressed formats, this value has to be equal to that of the original resource.
- [cudaResourceViewDesc::height](#) specifies the new height of the texture data. If the resource view format is a block compressed format, this value has to be 4 times the original height of the resource. For non block compressed formats, this value has to be equal to that of the original resource.
- [cudaResourceViewDesc::depth](#) specifies the new depth of the texture data. This value has to be equal to that of the original resource.
- [cudaResourceViewDesc::firstMipmapLevel](#) specifies the most detailed mipmap level. This will be the new mipmap level zero. For non-mipmapped resources, this value has to be zero. [cudaTextureDesc::minMipmapLevelClamp](#) and [cudaTextureDesc::maxMipmapLevelClamp](#) will be relative to this value. For ex., if the firstMipmapLevel is set to 2, and a minMipmapLevelClamp of 1.2 is specified, then the actual minimum mipmap level clamp will be 3.2.
- [cudaResourceViewDesc::lastMipmapLevel](#) specifies the least detailed mipmap level. For non-mipmapped resources, this value has to be zero.
- [cudaResourceViewDesc::firstLayer](#) specifies the first layer index for layered textures. This will be the new layer zero. For non-layered resources, this value has to be zero.
- [cudaResourceViewDesc::lastLayer](#) specifies the last layer index for layered textures. For non-layered resources, this value has to be zero.

Parameters:

pTexObject - Texture object to create
pResDesc - Resource descriptor
pTexDesc - Texture descriptor
pResViewDesc - Resource view descriptor

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

See also:

[cudaDestroyTextureObject](#)

5.22.2.2 `cudaError_t cudaDestroyTextureObject (cudaTextureObject_t texObject)`

Destroys the texture object specified by `texObject`.

Parameters:

texObject - Texture object to destroy

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

See also:

[cudaCreateTextureObject](#)

5.22.2.3 `cudaError_t cudaGetTextureObjectResourceDesc (struct cudaResourceDesc * pResDesc, cudaTextureObject_t texObject)`

Returns the resource descriptor for the texture object specified by `texObject`.

Parameters:

pResDesc - Resource descriptor
texObject - Texture object

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

See also:

[cudaCreateTextureObject](#)

5.22.2.4 `cudaError_t cudaGetTextureObjectResourceViewDesc (struct cudaResourceViewDesc * pResViewDesc, cudaTextureObject_t texObject)`

Returns the resource view descriptor for the texture object specified by `texObject`. If no resource view was specified, `cudaErrorInvalidValue` is returned.

Parameters:

pResViewDesc - Resource view descriptor
texObject - Texture object

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`

See also:

`cudaCreateTextureObject`

5.22.2.5 `cudaError_t cudaGetTextureObjectTextureDesc (struct cudaTextureDesc * pTexDesc, cudaTextureObject_t texObject)`

Returns the texture descriptor for the texture object specified by `texObject`.

Parameters:

pTexDesc - Texture descriptor
texObject - Texture object

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`

See also:

`cudaCreateTextureObject`

5.23 Surface Object Management

Functions

- `cudaError_t cudaCreateSurfaceObject (cudaSurfaceObject_t *pSurfObject, const struct cudaResourceDesc *pResDesc)`
Creates a surface object.
- `cudaError_t cudaDestroySurfaceObject (cudaSurfaceObject_t surfObject)`
Destroys a surface object.
- `cudaError_t cudaGetSurfaceObjectResourceDesc (struct cudaResourceDesc *pResDesc, cudaSurfaceObject_t surfObject)`
Returns a surface object's resource descriptor Returns the resource descriptor for the surface object specified by surfObject.

5.23.1 Detailed Description

This section describes the low level texture object management functions of the CUDA runtime application programming interface. The surface object API is only supported on devices of compute capability 3.0 or higher.

5.23.2 Function Documentation

5.23.2.1 `cudaError_t cudaCreateSurfaceObject (cudaSurfaceObject_t * pSurfObject, const struct cudaResourceDesc * pResDesc)`

Creates a surface object and returns it in `pSurfObject`. `pResDesc` describes the data to perform surface load/stores on. `cudaResourceDesc::resType` must be `cudaResourceTypeArray` and `cudaResourceDesc::res::array::array` must be set to a valid CUDA array handle.

Surface objects are only supported on devices of compute capability 3.0 or higher.

Parameters:

- pSurfObject* - Surface object to create
- pResDesc* - Resource descriptor

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`

See also:

`cudaDestroySurfaceObject`

5.23.2.2 `cudaError_t cudaDestroySurfaceObject (cudaSurfaceObject_t surfObject)`

Destroys the surface object specified by `surfObject`.

Parameters:

- surfObject* - Surface object to destroy

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

See also:

[cudaCreateSurfaceObject](#)

**5.23.2.3 `cudaError_t cudaGetSurfaceObjectResourceDesc (struct cudaResourceDesc * pResDesc,
cudaSurfaceObject_t surfObject)`****Parameters:**

pResDesc - Resource descriptor

surfObject - Surface object

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

See also:

[cudaCreateSurfaceObject](#)

5.24 Version Management

Functions

- [cudaError_t cudaDriverGetVersion](#) (int *driverVersion)
Returns the CUDA driver version.
- [__cuda_builtin__ cudaError_t cudaRuntimeGetVersion](#) (int *runtimeVersion)
Returns the CUDA Runtime version.

5.24.1 Function Documentation

5.24.1.1 [cudaError_t cudaDriverGetVersion](#) (int * *driverVersion*)

Returns in **driverVersion* the version number of the installed CUDA driver. If no driver is installed, then 0 is returned as the driver version (via *driverVersion*). This function automatically returns [cudaErrorInvalidValue](#) if the *driverVersion* argument is NULL.

Parameters:

driverVersion - Returns the CUDA driver version.

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaRuntimeGetVersion](#)

5.24.1.2 [__cuda_builtin__ cudaError_t cudaRuntimeGetVersion](#) (int * *runtimeVersion*)

Returns in **runtimeVersion* the version number of the installed CUDA Runtime. This function automatically returns [cudaErrorInvalidValue](#) if the *runtimeVersion* argument is NULL.

Parameters:

runtimeVersion - Returns the CUDA Runtime version.

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#)

See also:

[cudaDriverGetVersion](#)

5.25 C++ API Routines

C++-style interface built on top of CUDA runtime API.

Functions

- `template<class T, int dim>`
`cudaError_t cudaBindSurfaceToArray` (const struct surface< T, dim > &surf, `cudaArray_const_t` array)
[C++ API] Binds an array to a surface
- `template<class T, int dim>`
`cudaError_t cudaBindSurfaceToArray` (const struct surface< T, dim > &surf, `cudaArray_const_t` array, const struct `cudaChannelFormatDesc` &desc)
[C++ API] Binds an array to a surface
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaBindTexture` (size_t *offset, const struct texture< T, dim, readMode > &tex, const void *devPtr, size_t size=UINT_MAX)
[C++ API] Binds a memory area to a texture
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaBindTexture` (size_t *offset, const struct texture< T, dim, readMode > &tex, const void *devPtr, const struct `cudaChannelFormatDesc` &desc, size_t size=UINT_MAX)
[C++ API] Binds a memory area to a texture
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaBindTexture2D` (size_t *offset, const struct texture< T, dim, readMode > &tex, const void *devPtr, size_t width, size_t height, size_t pitch)
[C++ API] Binds a 2D memory area to a texture
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaBindTexture2D` (size_t *offset, const struct texture< T, dim, readMode > &tex, const void *devPtr, const struct `cudaChannelFormatDesc` &desc, size_t width, size_t height, size_t pitch)
[C++ API] Binds a 2D memory area to a texture
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaBindTextureToArray` (const struct texture< T, dim, readMode > &tex, `cudaArray_const_t` array)
[C++ API] Binds an array to a texture
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaBindTextureToArray` (const struct texture< T, dim, readMode > &tex, `cudaArray_const_t` array, const struct `cudaChannelFormatDesc` &desc)
[C++ API] Binds an array to a texture
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaBindTextureToMipmappedArray` (const struct texture< T, dim, readMode > &tex, `cudaMipmappedArray_const_t` mipmappedArray)
[C++ API] Binds a mipmapped array to a texture

- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaBindTextureToMipmappedArray` (const struct texture< T, dim, readMode > &tex, `cudaMipmappedArray_const_t` mipmappedArray, const struct `cudaChannelFormatDesc` &desc)
[C++ API] Binds a mipmapped array to a texture
- `template<class T >`
`cudaChannelFormatDesc cudaCreateChannelDesc` (void)
[C++ API] Returns a channel descriptor using the specified format
- `cudaError_t cudaEventCreate` (`cudaEvent_t` *event, unsigned int flags)
[C++ API] Creates an event object with the specified flags
- `template<class T >`
`cudaError_t cudaFuncGetAttributes` (struct `cudaFuncAttributes` *attr, T *entry)
[C++ API] Find out attributes for a given function
- `template<class T >`
`cudaError_t cudaFuncSetCacheConfig` (T *func, enum `cudaFuncCache` cacheConfig)
[C++ API] Sets the preferred cache configuration for a device function
- `template<class T >`
`cudaError_t cudaGetSymbolAddress` (void **devPtr, const T &symbol)
[C++ API] Finds the address associated with a CUDA symbol
- `template<class T >`
`cudaError_t cudaGetSymbolSize` (size_t *size, const T &symbol)
[C++ API] Finds the size of the object associated with a CUDA symbol
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaGetTextureAlignmentOffset` (size_t *offset, const struct texture< T, dim, readMode > &tex)
[C++ API] Get the alignment offset of a texture
- `template<class T >`
`cudaError_t cudaLaunch` (T *func)
[C++ API] Launches a device function
- `cudaError_t cudaMallocHost` (void **ptr, size_t size, unsigned int flags)
[C++ API] Allocates page-locked memory on the host
- `template<class T >`
`cudaError_t cudaMemcpyFromSymbol` (void *dst, const T &symbol, size_t count, size_t offset=0, enum `cudaMemcpyKind` kind=`cudaMemcpyDeviceToHost`)
[C++ API] Copies data from the given symbol on the device
- `template<class T >`
`cudaError_t cudaMemcpyFromSymbolAsync` (void *dst, const T &symbol, size_t count, size_t offset=0, enum `cudaMemcpyKind` kind=`cudaMemcpyDeviceToHost`, `cudaStream_t` stream=0)
[C++ API] Copies data from the given symbol on the device
- `template<class T >`
`cudaError_t cudaMemcpyToSymbol` (const T &symbol, const void *src, size_t count, size_t offset=0, enum `cudaMemcpyKind` kind=`cudaMemcpyHostToDevice`)

[C++ API] Copies data to the given symbol on the device

- `template<class T >`
`cudaError_t cudaMemcpyToSymbolAsync` (`const T &symbol`, `const void *src`, `size_t count`, `size_t offset=0`,
`enum cudaMemcpyKind kind=cudaMemcpyHostToDevice`, `cudaStream_t stream=0`)
[C++ API] Copies data to the given symbol on the device
- `template<class T >`
`cudaError_t cudaSetupArgument` (`T arg`, `size_t offset`)
[C++ API] Configure a device launch
- `template<class T, int dim, enum cudaTextureReadMode readMode>`
`cudaError_t cudaUnbindTexture` (`const struct texture< T, dim, readMode > &tex`)
[C++ API] Unbinds a texture

5.25.1 Detailed Description

This section describes the C++ high level API functions of the CUDA runtime application programming interface. To use these functions, your application needs to be compiled with the `nvcc` compiler.

5.25.2 Function Documentation

5.25.2.1 `template<class T, int dim> cudaError_t cudaBindSurfaceToArray (const struct surface< T, dim > & surf, cudaArray_const_t array)`

Binds the CUDA array `array` to the surface reference `surf`. The channel descriptor is inherited from the CUDA array. Any CUDA array previously bound to `surf` is unbound.

Parameters:

- `surf` - Surface to bind
- `array` - Memory array on device

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidSurface`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaBindSurfaceToArray (C API)`, `cudaBindSurfaceToArray (C++ API)`

5.25.2.2 `template<class T, int dim> cudaError_t cudaBindSurfaceToArray (const struct surface< T, dim > & surf, cudaArray_const_t array, const struct cudaChannelFormatDesc & desc)`

Binds the CUDA array `array` to the surface reference `surf`. `desc` describes how the memory is interpreted when dealing with the surface. Any CUDA array previously bound to `surf` is unbound.

Parameters:

surf - Surface to bind
array - Memory array on device
desc - Channel format

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidSurface](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaBindSurfaceToArray \(C API\)](#), [cudaBindSurfaceToArray \(C++ API, inherited channel descriptor\)](#)

5.25.2.3 `template<class T , int dim, enum cudaTextureReadMode readMode> cudaError_t
 cudaBindTexture (size_t * offset, const struct texture< T, dim, readMode > & tex, const void *
 devPtr, size_t size = UINT_MAX)`

Binds *size* bytes of the memory area pointed to by *devPtr* to texture reference *tex*. The channel descriptor is inherited from the texture reference type. The *offset* parameter is an optional byte offset as with the low-level [cudaBindTexture\(size_t*, const struct textureReference*, const void*, const struct cudaChannelFormatDesc*, size_t\)](#) function. Any memory previously bound to *tex* is unbound.

Parameters:

offset - Offset in bytes
tex - Texture to bind
devPtr - Memory area on device
size - Size of the memory area pointed to by *devPtr*

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc \(C++ API\)](#), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture \(C API\)](#), [cudaBindTexture \(C++ API\)](#), [cudaBindTexture2D \(C++ API\)](#), [cudaBindTexture2D \(C++ API, inherited channel descriptor\)](#), [cudaBindTextureToArray \(C++ API\)](#), [cudaBindTextureToArray \(C++ API, inherited channel descriptor\)](#), [cudaUnbindTexture \(C++ API\)](#), [cudaGetTextureAlignmentOffset \(C++ API\)](#)

5.25.2.4 `template<class T , int dim, enum cudaTextureReadMode readMode> cudaError_t
 cudaBindTexture (size_t * offset, const struct texture< T, dim, readMode > & tex, const void *
 devPtr, const struct cudaChannelFormatDesc & desc, size_t size = UINT_MAX)`

Binds *size* bytes of the memory area pointed to by *devPtr* to texture reference *tex*. *desc* describes how the memory is interpreted when fetching values from the texture. The *offset* parameter is an optional byte offset as with the low-level [cudaBindTexture\(\)](#) function. Any memory previously bound to *tex* is unbound.

Parameters:

offset - Offset in bytes
tex - Texture to bind
devPtr - Memory area on device
desc - Channel format
size - Size of the memory area pointed to by devPtr

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C API), [cudaBindTexture](#) (C++ API, inherited channel descriptor), [cudaBindTexture2D](#) (C++ API), [cudaBindTexture2D](#) (C++ API, inherited channel descriptor), [cudaBindTextureToArray](#) (C++ API), [cudaBindTextureToArray](#) (C++ API, inherited channel descriptor), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C++ API)

5.25.2.5 `template<class T , int dim, enum cudaTextureReadMode readMode> cudaError_t
 cudaBindTexture2D (size_t * offset, const struct texture< T, dim, readMode > & tex, const void *
 devPtr, size_t width, size_t height, size_t pitch)`

Binds the 2D memory area pointed to by `devPtr` to the texture reference `tex`. The size of the area is constrained by `width` in texel units, `height` in texel units, and `pitch` in byte units. The channel descriptor is inherited from the texture reference type. Any memory previously bound to `tex` is unbound.

Since the hardware enforces an alignment requirement on texture base addresses, `cudaBindTexture2D()` returns in `*offset` a byte offset that must be applied to texture fetches in order to read from the desired memory. This offset must be divided by the texel size and passed to kernels that read from the texture so they can be applied to the `tex2D()` function. If the device memory pointer was returned from `cudaMalloc()`, the offset is guaranteed to be 0 and NULL may be passed as the `offset` parameter.

Parameters:

offset - Offset in bytes
tex - Texture reference to bind
devPtr - 2D memory area on device
width - Width in texel units
height - Height in texel units
pitch - Pitch in bytes

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C++ API), [cudaBindTexture](#) (C++ API, inherited channel descriptor), [cudaBindTexture2D](#) (C API), [cudaBindTexture2D](#) (C++ API), [cudaBindTextureToArray](#) (C++ API), [cudaBindTextureToArray](#) (C++ API, inherited channel descriptor), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C++ API)

5.25.2.6 `template<class T , int dim, enum cudaTextureReadMode readMode> cudaError_t
 cudaBindTexture2D (size_t *offset, const struct texture< T, dim, readMode > & tex, const void *
 devPtr, const struct cudaChannelFormatDesc & desc, size_t width, size_t height, size_t pitch)`

Binds the 2D memory area pointed to by `devPtr` to the texture reference `tex`. The size of the area is constrained by `width` in texel units, `height` in texel units, and `pitch` in byte units. `desc` describes how the memory is interpreted when fetching values from the texture. Any memory previously bound to `tex` is unbound.

Since the hardware enforces an alignment requirement on texture base addresses, `cudaBindTexture2D()` returns in `*offset` a byte offset that must be applied to texture fetches in order to read from the desired memory. This offset must be divided by the texel size and passed to kernels that read from the texture so they can be applied to the `tex2D()` function. If the device memory pointer was returned from `cudaMalloc()`, the offset is guaranteed to be 0 and NULL may be passed as the `offset` parameter.

Parameters:

offset - Offset in bytes
tex - Texture reference to bind
devPtr - 2D memory area on device
desc - Channel format
width - Width in texel units
height - Height in texel units
pitch - Pitch in bytes

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C++ API), [cudaBindTexture](#) (C++ API, inherited channel descriptor), [cudaBindTexture2D](#) (C API), [cudaBindTexture2D](#) (C++ API, inherited channel descriptor), [cudaBindTextureToArray](#) (C++ API), [cudaBindTextureToArray](#) (C++ API, inherited channel descriptor), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C++ API)

5.25.2.7 `template<class T , int dim, enum cudaTextureReadMode readMode> cudaError_t
 cudaBindTextureToArray (const struct texture< T, dim, readMode > & tex, cudaArray_const_t
 array)`

Binds the CUDA array `array` to the texture reference `tex`. The channel descriptor is inherited from the CUDA array. Any CUDA array previously bound to `tex` is unbound.

Parameters:

tex - Texture to bind
array - Memory array on device

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C++ API), [cudaBindTexture](#) (C++ API, inherited channel descriptor), [cudaBindTexture2D](#) (C++ API), [cudaBindTexture2D](#) (C++ API, inherited channel descriptor), [cudaBindTextureToArray](#) (C API), [cudaBindTextureToArray](#) (C++ API), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C++ API)

5.25.2.8 `template<class T , int dim, enum cudaTextureReadMode readMode> cudaError_t cudaBindTextureToArray (const struct texture< T, dim, readMode > & tex, cudaArray_const_t array, const struct cudaChannelFormatDesc & desc)`

Binds the CUDA array *array* to the texture reference *tex*. *desc* describes how the memory is interpreted when fetching values from the texture. Any CUDA array previously bound to *tex* is unbound.

Parameters:

tex - Texture to bind
array - Memory array on device
desc - Channel format

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C++ API), [cudaBindTexture](#) (C++ API, inherited channel descriptor), [cudaBindTexture2D](#) (C++ API), [cudaBindTexture2D](#) (C++ API, inherited channel descriptor), [cudaBindTextureToArray](#) (C API), [cudaBindTextureToArray](#) (C++ API, inherited channel descriptor), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C++ API)

5.25.2.9 `template<class T , int dim, enum cudaTextureReadMode readMode> cudaError_t cudaBindTextureToMipmappedArray (const struct texture< T, dim, readMode > & tex, cudaMipmappedArray_const_t mipmappedArray)`

Binds the CUDA mipmapped array *mipmappedArray* to the texture reference *tex*. The channel descriptor is inherited from the CUDA array. Any CUDA mipmapped array previously bound to *tex* is unbound.

Parameters:

tex - Texture to bind
mipmappedArray - Memory mipmapped array on device

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C++ API), [cudaBindTexture](#) (C++ API, inherited channel descriptor), [cudaBindTexture2D](#) (C++ API), [cudaBindTexture2D](#) (C++ API, inherited channel descriptor), [cudaBindTextureToArray](#) (C API), [cudaBindTextureToArray](#) (C++ API), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C++ API)

5.25.2.10 `template<class T, int dim, enum cudaTextureReadMode readMode> cudaError_t
 cudaBindTextureToMipmappedArray (const struct texture< T, dim, readMode > & tex,
 cudaMipmappedArray_const_t mipmappedArray, const struct cudaChannelFormatDesc & desc)`

Binds the CUDA mipmapped array *mipmappedArray* to the texture reference *tex*. *desc* describes how the memory is interpreted when fetching values from the texture. Any CUDA mipmapped array previously bound to *tex* is unbound.

Parameters:

tex - Texture to bind
mipmappedArray - Memory mipmapped array on device
desc - Channel format

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidTexture](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C++ API), [cudaBindTexture](#) (C++ API, inherited channel descriptor), [cudaBindTexture2D](#) (C++ API), [cudaBindTexture2D](#) (C++ API, inherited channel descriptor), [cudaBindTextureToArray](#) (C API), [cudaBindTextureToArray](#) (C++ API, inherited channel descriptor), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C++ API)

5.25.2.11 `template<class T > cudaChannelFormatDesc cudaCreateChannelDesc (void)`

Returns a channel descriptor with format *f* and number of bits of each component *x*, *y*, *z*, and *w*. The [cudaChannelFormatDesc](#) is defined as:

```
struct cudaChannelFormatDesc {
    int x, y, z, w;
    enum cudaChannelFormatKind f;
};
```

where `cudaChannelFormatKind` is one of `cudaChannelFormatKindSigned`, `cudaChannelFormatKindUnsigned`, or `cudaChannelFormatKindFloat`.

Returns:

Channel descriptor with format `f`

See also:

[cudaCreateChannelDesc](#) (Low level), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (High level), [cudaBindTexture](#) (High level, inherited channel descriptor), [cudaBindTexture2D](#) (High level), [cudaBindTextureToArray](#) (High level), [cudaBindTextureToArray](#) (High level, inherited channel descriptor), [cudaUnbindTexture](#) (High level), [cudaGetTextureAlignmentOffset](#) (High level)

5.25.2.12 `cudaError_t cudaEventCreate (cudaEvent_t * event, unsigned int flags)`

Creates an event object with the specified flags. Valid flags include:

- [cudaEventDefault](#): Default event creation flag.
- [cudaEventBlockingSync](#): Specifies that event should use blocking synchronization. A host thread that uses [cudaEventSynchronize\(\)](#) to wait on an event created with this flag will block until the event actually completes.
- [cudaEventDisableTiming](#): Specifies that the created event does not need to record timing data. Events created with this flag specified and the [cudaEventBlockingSync](#) flag not specified will provide the best performance when used with [cudaStreamWaitEvent\(\)](#) and [cudaEventQuery\(\)](#).

Parameters:

event - Newly created event

flags - Flags for new event

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidValue](#), [cudaErrorLaunchFailure](#), [cudaErrorMemoryAllocation](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaEventCreate](#) (C API), [cudaEventCreateWithFlags](#), [cudaEventRecord](#), [cudaEventQuery](#), [cudaEventSynchronize](#), [cudaEventDestroy](#), [cudaEventElapsedTime](#), [cudaStreamWaitEvent](#)

5.25.2.13 `template<class T > cudaError_t cudaFuncGetAttributes (struct cudaFuncAttributes * attr, T * entry)`

This function obtains the attributes of a function specified via `entry`. The parameter `entry` must be a pointer to a function that executes on the device. The parameter specified by `entry` must be declared as a `__global__` function. The fetched attributes are placed in `attr`. If the specified function does not exist, then `cudaErrorInvalidDeviceFunction` is returned.

Note that some function attributes such as `maxThreadsPerBlock` may vary based on the device that is currently being used.

Parameters:

attr - Return pointer to function's attributes

entry - Function to get attributes of

Returns:

`cudaSuccess`, `cudaErrorInitializationError`, `cudaErrorInvalidDeviceFunction`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaConfigureCall`, `cudaFuncSetCacheConfig` (C++ API), `cudaFuncGetAttributes` (C API), `cudaLaunch` (C++ API), `cudaSetDoubleForDevice`, `cudaSetDoubleForHost`, `cudaSetupArgument` (C++ API)

5.25.2.14 `template<class T > cudaError_t cudaFuncSetCacheConfig (T * func, enum cudaFuncCache cacheConfig)`

On devices where the L1 cache and shared memory use the same hardware resources, this sets through `cacheConfig` the preferred cache configuration for the function specified via `func`. This is only a preference. The runtime will use the requested configuration if possible, but it is free to choose a different configuration if required to execute `func`.

`func` must be a pointer to a function that executes on the device. The parameter specified by `func` must be declared as a `__global__` function. If the specified function does not exist, then `cudaErrorInvalidDeviceFunction` is returned.

This setting does nothing on devices where the size of the L1 cache and shared memory are fixed.

Launching a kernel with a different preference than the most recent preference setting may insert a device-side synchronization point.

The supported cache configurations are:

- `cudaFuncCachePreferNone`: no preference for shared memory or L1 (default)
- `cudaFuncCachePreferShared`: prefer larger shared memory and smaller L1 cache
- `cudaFuncCachePreferL1`: prefer larger L1 cache and smaller shared memory

Parameters:

func - device function pointer

cacheConfig - Requested cache configuration

Returns:

[cudaSuccess](#), [cudaErrorInitializationError](#), [cudaErrorInvalidDeviceFunction](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaConfigureCall](#), [cudaFuncSetCacheConfig](#) (C API), [cudaFuncGetAttributes](#) (C++ API), [cudaLaunch](#) (C API), [cudaSetDoubleForDevice](#), [cudaSetDoubleForHost](#), [cudaSetupArgument](#) (C++ API), [cudaThreadGetCacheConfig](#), [cudaThreadSetCacheConfig](#)

5.25.2.15 `template<class T > cudaError_t cudaGetSymbolAddress (void ** devPtr, const T & symbol)`

Returns in `*devPtr` the address of symbol `symbol` on the device. `symbol` can either be a variable that resides in global or constant memory space. If `symbol` cannot be found, or if `symbol` is not declared in the global or constant memory space, `*devPtr` is unchanged and the error [cudaErrorInvalidSymbol](#) is returned.

Parameters:

devPtr - Return device pointer associated with symbol
symbol - Device symbol reference

Returns:

[cudaSuccess](#), [cudaErrorInvalidSymbol](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGetSymbolAddress](#) (C API), [cudaGetSymbolSize](#) (C++ API)

5.25.2.16 `template<class T > cudaError_t cudaGetSymbolSize (size_t * size, const T & symbol)`

Returns in `*size` the size of symbol `symbol`. `symbol` must be a variable that resides in global or constant memory space. If `symbol` cannot be found, or if `symbol` is not declared in global or constant memory space, `*size` is unchanged and the error [cudaErrorInvalidSymbol](#) is returned.

Parameters:

size - Size of object associated with symbol
symbol - Device symbol reference

Returns:

[cudaSuccess](#), [cudaErrorInvalidSymbol](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaGetSymbolAddress](#) (C++ API), [cudaGetSymbolSize](#) (C API)

5.25.2.17 `template<class T, int dim, enum cudaTextureReadMode readMode> cudaError_t
cudaGetTextureAlignmentOffset (size_t * offset, const struct texture< T, dim, readMode > & tex)`

Returns in `*offset` the offset that was returned when texture reference `tex` was bound.

Parameters:

`offset` - Offset of texture reference in bytes

`tex` - Texture to get offset of

Returns:

[cudaSuccess](#), [cudaErrorInvalidTexture](#), [cudaErrorInvalidTextureBinding](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc](#) (C++ API), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture](#) (C++ API), [cudaBindTexture](#) (C++ API, inherited channel descriptor), [cudaBindTexture2D](#) (C++ API), [cudaBindTexture2D](#) (C++ API, inherited channel descriptor), [cudaBindTextureToArray](#) (C++ API), [cudaBindTextureToArray](#) (C++ API, inherited channel descriptor), [cudaUnbindTexture](#) (C++ API), [cudaGetTextureAlignmentOffset](#) (C API)

5.25.2.18 `template<class T > cudaError_t cudaLaunch (T * func)`

Launches the function `entry` on the device. The parameter `entry` must be a function that executes on the device. The parameter specified by `entry` must be declared as a `__global__` function. `cudaLaunch()` must be preceded by a call to [cudaConfigureCall\(\)](#) since it pops the data that was pushed by [cudaConfigureCall\(\)](#) from the execution stack.

Parameters:

`entry` - Device function pointer to execute

Returns:

[cudaSuccess](#), [cudaErrorInvalidDeviceFunction](#), [cudaErrorInvalidConfiguration](#), [cudaErrorLaunchFailure](#), [cudaErrorLaunchTimeout](#), [cudaErrorLaunchOutOfResources](#), [cudaErrorSharedObjectSymbolNotFound](#), [cudaErrorSharedObjectInitFailed](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaConfigureCall](#), [cudaFuncSetCacheConfig](#) (C++ API), [cudaFuncGetAttributes](#) (C++ API), [cudaLaunch](#) (C API), [cudaSetDoubleForDevice](#), [cudaSetDoubleForHost](#), [cudaSetupArgument](#) (C++ API), [cudaThreadGetCacheConfig](#), [cudaThreadSetCacheConfig](#)

5.25.2.19 `cudaError_t cudaMallocHost (void **ptr, size_t size, unsigned int flags)`

Allocates `size` bytes of host memory that is page-locked and accessible to the device. The driver tracks the virtual memory ranges allocated with this function and automatically accelerates calls to functions such as `cudaMemcpy()`. Since the memory can be accessed directly by the device, it can be read or written with much higher bandwidth than pageable memory obtained with functions such as `malloc()`. Allocating excessive amounts of pinned memory may degrade system performance, since it reduces the amount of memory available to the system for paging. As a result, this function is best used sparingly to allocate staging areas for data exchange between host and device.

The `flags` parameter enables different options to be specified that affect the allocation, as follows.

- `cudaHostAllocDefault`: This flag's value is defined to be 0.
- `cudaHostAllocPortable`: The memory returned by this call will be considered as pinned memory by all CUDA contexts, not just the one that performed the allocation.
- `cudaHostAllocMapped`: Maps the allocation into the CUDA address space. The device pointer to the memory may be obtained by calling `cudaHostGetDevicePointer()`.
- `cudaHostAllocWriteCombined`: Allocates the memory as write-combined (WC). WC memory can be transferred across the PCI Express bus more quickly on some system configurations, but cannot be read efficiently by most CPUs. WC memory is a good option for buffers that will be written by the CPU and read by the device via mapped pinned memory or host->device transfers.

All of these flags are orthogonal to one another: a developer may allocate memory that is portable, mapped and/or write-combined with no restrictions.

`cudaSetDeviceFlags()` must have been called with the `cudaDeviceMapHost` flag in order for the `cudaHostAllocMapped` flag to have any effect.

The `cudaHostAllocMapped` flag may be specified on CUDA contexts for devices that do not support mapped pinned memory. The failure is deferred to `cudaHostGetDevicePointer()` because the memory may be mapped into other CUDA contexts via the `cudaHostAllocPortable` flag.

Memory allocated by this function must be freed with `cudaFreeHost()`.

Parameters:

- ptr* - Device pointer to allocated memory
- size* - Requested allocation size in bytes
- flags* - Requested properties of allocated memory

Returns:

`cudaSuccess`, `cudaErrorMemoryAllocation`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cudaSetDeviceFlags`, `cudaMallocHost (C API)`, `cudaFreeHost`, `cudaHostAlloc`

5.25.2.20 `template<class T > cudaError_t cudaMemcpyFromSymbol (void * dst, const T & symbol, size_t count, size_t offset = 0, enum cudaMemcpyKind kind = cudaMemcpyDeviceToHost)`

Copies `count` bytes from the memory area `offset` bytes from the start of symbol `symbol` to the memory area pointed to by `dst`. The memory areas may not overlap. `symbol` is a variable that resides in global or constant memory space. `kind` can be either `cudaMemcpyDeviceToHost` or `cudaMemcpyDeviceToDevice`.

Parameters:

dst - Destination memory address
symbol - Device symbol reference
count - Size in bytes to copy
offset - Offset from start of symbol in bytes
kind - Type of transfer

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidSymbol`, `cudaErrorInvalidDevicePointer`, `cudaErrorInvalidMemcpyDirection`

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits `synchronous` behavior for most use cases. Use of a string naming a variable as the `symbol` parameter was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

`cudaMemcpy`, `cudaMemcpy2D`, `cudaMemcpyToArray`, `cudaMemcpy2DToArray`, `cudaMemcpyFromArray`, `cudaMemcpy2DFromArray`, `cudaMemcpyArrayToArray`, `cudaMemcpy2DArrayToArray`, `cudaMemcpyToSymbol`, `cudaMemcpyAsync`, `cudaMemcpy2DAsync`, `cudaMemcpyToArrayAsync`, `cudaMemcpy2DToArrayAsync`, `cudaMemcpyFromArrayAsync`, `cudaMemcpy2DFromArrayAsync`, `cudaMemcpyToSymbolAsync`, `cudaMemcpyFromSymbolAsync`

5.25.2.21 `template<class T > cudaError_t cudaMemcpyFromSymbolAsync (void * dst, const T & symbol, size_t count, size_t offset = 0, enum cudaMemcpyKind kind = cudaMemcpyDeviceToHost, cudaStream_t stream = 0)`

Copies `count` bytes from the memory area `offset` bytes from the start of symbol `symbol` to the memory area pointed to by `dst`. The memory areas may not overlap. `symbol` is a variable that resides in global or constant memory space. `kind` can be either `cudaMemcpyDeviceToHost` or `cudaMemcpyDeviceToDevice`.

`cudaMemcpyFromSymbolAsync()` is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is `cudaMemcpyDeviceToHost` and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

dst - Destination memory address
symbol - Device symbol reference
count - Size in bytes to copy
offset - Offset from start of symbol in bytes

kind - Type of transfer
stream - Stream identifier

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidSymbol](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
 This function exhibits [asynchronous](#) behavior for most use cases.
 Use of a string naming a variable as the `symbol` parameter was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyToSymbol](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#)

5.25.2.22 `template<class T > cudaError_t cudaMemcpyToSymbol (const T & symbol, const void * src, size_t count, size_t offset = 0, enum cudaMemcpyKind kind = cudaMemcpyHostToDevice)`

Copies `count` bytes from the memory area pointed to by `src` to the memory area `offset` bytes from the start of symbol `symbol`. The memory areas may not overlap. `symbol` is a variable that resides in global or constant memory space. `kind` can be either [cudaMemcpyHostToDevice](#) or [cudaMemcpyDeviceToDevice](#).

Parameters:

symbol - Device symbol reference
src - Source memory address
count - Size in bytes to copy
offset - Offset from start of symbol in bytes
kind - Type of transfer

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorInvalidSymbol](#), [cudaErrorInvalidDevicePointer](#), [cudaErrorInvalidMemcpyDirection](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.
 This function exhibits [synchronous](#) behavior for most use cases.
 Use of a string naming a variable as the `symbol` parameter was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

[cudaMemcpy](#), [cudaMemcpy2D](#), [cudaMemcpyToArray](#), [cudaMemcpy2DToArray](#), [cudaMemcpyFromArray](#), [cudaMemcpy2DFromArray](#), [cudaMemcpyArrayToArray](#), [cudaMemcpy2DArrayToArray](#), [cudaMemcpyFromSymbol](#), [cudaMemcpyAsync](#), [cudaMemcpy2DAsync](#), [cudaMemcpyToArrayAsync](#), [cudaMemcpy2DToArrayAsync](#), [cudaMemcpyFromArrayAsync](#), [cudaMemcpy2DFromArrayAsync](#), [cudaMemcpyToSymbolAsync](#), [cudaMemcpyFromSymbolAsync](#)

5.25.2.23 `template<class T > cudaError_t cudaMemcpyToSymbolAsync (const T & symbol, const void * src, size_t count, size_t offset = 0, enum cudaMemcpyKind kind = cudaMemcpyHostToDevice, cudaStream_t stream = 0)`

Copies `count` bytes from the memory area pointed to by `src` to the memory area `offset` bytes from the start of symbol `symbol`. The memory areas may not overlap. `symbol` is a variable that resides in global or constant memory space. `kind` can be either `cudaMemcpyHostToDevice` or `cudaMemcpyDeviceToDevice`.

`cudaMemcpyToSymbolAsync()` is asynchronous with respect to the host, so the call may return before the copy is complete. The copy can optionally be associated to a stream by passing a non-zero `stream` argument. If `kind` is `cudaMemcpyHostToDevice` and `stream` is non-zero, the copy may overlap with operations in other streams.

Parameters:

symbol - Device symbol reference
src - Source memory address
count - Size in bytes to copy
offset - Offset from start of symbol in bytes
kind - Type of transfer
stream - Stream identifier

Returns:

`cudaSuccess`, `cudaErrorInvalidValue`, `cudaErrorInvalidSymbol`, `cudaErrorInvalidDevicePointer`, `cudaErrorInvalidMemcpyDirection`

Note:

Note that this function may also return error codes from previous, asynchronous launches. This function exhibits `asynchronous` behavior for most use cases. Use of a string naming a variable as the `symbol` parameter was deprecated in CUDA 4.1 and removed in CUDA 5.0.

See also:

`cudaMemcpy`, `cudaMemcpy2D`, `cudaMemcpyToArray`, `cudaMemcpy2DToArray`, `cudaMemcpyFromArray`, `cudaMemcpy2DFromArray`, `cudaMemcpyArrayToArray`, `cudaMemcpy2DArrayToArray`, `cudaMemcpyToSymbol`, `cudaMemcpyFromSymbol`, `cudaMemcpyAsync`, `cudaMemcpy2DAsync`, `cudaMemcpyToArrayAsync`, `cudaMemcpy2DToArrayAsync`, `cudaMemcpyFromArrayAsync`, `cudaMemcpy2DFromArrayAsync`, `cudaMemcpyFromSymbolAsync`

5.25.2.24 `template<class T > cudaError_t cudaSetupArgument (T arg, size_t offset)`

Pushes `size` bytes of the argument pointed to by `arg` at `offset` bytes from the start of the parameter passing area, which starts at offset 0. The arguments are stored in the top of the execution stack. `cudaSetupArgument()` must be preceded by a call to `cudaConfigureCall()`.

Parameters:

arg - Argument to push for a kernel launch
offset - Offset in argument stack to push new arg

Returns:

`cudaSuccess`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaConfigureCall](#), [cudaFuncGetAttributes \(C++ API\)](#), [cudaLaunch \(C++ API\)](#), [cudaSetDoubleForDevice](#), [cudaSetDoubleForHost](#), [cudaSetupArgument \(C API\)](#)

5.25.2.25 `template<class T , int dim, enum cudaTextureReadMode readMode> cudaError_t
cudaUnbindTexture (const struct texture< T, dim, readMode > & tex)`

Unbinds the texture bound to `tex`.

Parameters:

`tex` - Texture to unbind

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaCreateChannelDesc \(C++ API\)](#), [cudaGetChannelDesc](#), [cudaGetTextureReference](#), [cudaBindTexture \(C++ API\)](#), [cudaBindTexture \(C++ API, inherited channel descriptor\)](#), [cudaBindTexture2D \(C++ API\)](#), [cudaBindTexture2D \(C++ API, inherited channel descriptor\)](#), [cudaBindTextureToArray \(C++ API\)](#), [cudaBindTextureToArray \(C++ API, inherited channel descriptor\)](#), [cudaUnbindTexture \(C API\)](#), [cudaGetTextureAlignmentOffset \(C++ API\)](#)

5.26 Interactions with the CUDA Driver API

Interactions between the CUDA Driver API and the CUDA Runtime API.

This section describes the interactions between the CUDA Driver API and the CUDA Runtime API

5.26.1 Primary Contexts

There exists a one to one relationship between CUDA devices in the CUDA Runtime API and `CUcontext`s in the CUDA Driver API within a process. The specific context which the CUDA Runtime API uses for a device is called the device's primary context. From the perspective of the CUDA Runtime API, a device and its primary context are synonymous.

5.26.2 Initialization and Tear-Down

CUDA Runtime API calls operate on the CUDA Driver API `CUcontext` which is current to to the calling host thread.

The function `cudaSetDevice()` makes the primary context for the specified device current to the calling thread by calling `cuCtxSetCurrent()`.

The CUDA Runtime API will automatically initialize the primary context for a device at the first CUDA Runtime API call which requires an active context. If no `CUcontext` is current to the calling thread when a CUDA Runtime API call which requires an active context is made, then the primary context for a device will be selected, made current to the calling thread, and initialized.

The context which the CUDA Runtime API initializes will be initialized using the parameters specified by the CUDA Runtime API functions `cudaSetDeviceFlags()`, `cudaD3D9SetDirect3DDevice()`, `cudaD3D10SetDirect3DDevice()`, `cudaD3D11SetDirect3DDevice()`, `cudaGLSetGLDevice()`, and `cudaVDPAUSetVDPAUDevice()`. Note that these functions will fail with `cudaErrorSetOnActiveProcess` if they are called when the primary context for the specified device has already been initialized. (or if the current device has already been initialized, in the case of `cudaSetDeviceFlags()`).

Primary contexts will remain active until they are explicitly deinitialized using `cudaDeviceReset()`. The function `cudaDeviceReset()` will deinitialize the primary context for the calling thread's current device immediately. The context will remain current to all of the threads that it was current to. The next CUDA Runtime API call on any thread which requires an active context will trigger the reinitialization of that device's primary context.

Note that there is no reference counting of the primary context's lifetime. It is recommended that the primary context not be deinitialized except just before exit or to recover from an unspecified launch failure.

5.26.3 Context Interoperability

Note that the use of multiple `CUcontext`s per device within a single process will substantially degrade performance and is strongly discouraged. Instead, it is highly recommended that the implicit one-to-one device-to-context mapping for the process provided by the CUDA Runtime API be used.

If a non-primary `CUcontext` created by the CUDA Driver API is current to a thread then the CUDA Runtime API calls to that thread will operate on that `CUcontext`, with some exceptions listed below. Interoperability between data types is discussed in the following sections.

The function `cudaPointerGetAttributes()` will return the error `cudaErrorIncompatibleDriverContext` if the pointer being queried was allocated by a non-primary context. The function `cudaDeviceEnablePeerAccess()` and the rest of the peer access API may not be called when a non-primary `CUcontext` is current. To use the pointer query and peer access APIs with a context created using the CUDA Driver API, it is necessary that the CUDA Driver API be used to access these features.

All CUDA Runtime API state (e.g, global variables' addresses and values) travels with its underlying [CUcontext](#). In particular, if a [CUcontext](#) is moved from one thread to another then all CUDA Runtime API state will move to that thread as well.

Please note that attaching to legacy contexts (those with a version of 3010 as returned by [cuCtxGetApiVersion\(\)](#)) is not possible. The CUDA Runtime will return [cudaErrorIncompatibleDriverContext](#) in such cases.

5.26.4 Interactions between CUstream and cudaStream_t

The types [CUstream](#) and [cudaStream_t](#) are identical and may be used interchangeably.

5.26.5 Interactions between CUevent and cudaEvent_t

The types [CUevent](#) and [cudaEvent_t](#) are identical and may be used interchangeably.

5.26.6 Interactions between CUarray and cudaArray_t

The types [CUarray](#) and struct `cudaArray *` represent the same data type and may be used interchangeably by casting the two types between each other.

In order to use a [CUarray](#) in a CUDA Runtime API function which takes a struct `cudaArray *`, it is necessary to explicitly cast the [CUarray](#) to a struct `cudaArray *`.

In order to use a struct `cudaArray *` in a CUDA Driver API function which takes a [CUarray](#), it is necessary to explicitly cast the struct `cudaArray *` to a [CUarray](#).

5.26.7 Interactions between CUgraphicsResource and cudaGraphicsResource_t

The types [CUgraphicsResource](#) and [cudaGraphicsResource_t](#) represent the same data type and may be used interchangeably by casting the two types between each other.

In order to use a [CUgraphicsResource](#) in a CUDA Runtime API function which takes a [cudaGraphicsResource_t](#), it is necessary to explicitly cast the [CUgraphicsResource](#) to a [cudaGraphicsResource_t](#).

In order to use a [cudaGraphicsResource_t](#) in a CUDA Driver API function which takes a [CUgraphicsResource](#), it is necessary to explicitly cast the [cudaGraphicsResource_t](#) to a [CUgraphicsResource](#).

5.27 Profiler Control

Functions

- `cudaError_t cudaProfilerInitialize` (const char *configFile, const char *outputFile, `cudaOutputMode_t` outputMode)
Initialize the CUDA profiler.
- `cudaError_t cudaProfilerStart` (void)
Enable profiling.
- `cudaError_t cudaProfilerStop` (void)
Disable profiling.

5.27.1 Detailed Description

This section describes the profiler control functions of the CUDA runtime application programming interface.

5.27.2 Function Documentation

5.27.2.1 `cudaError_t cudaProfilerInitialize` (const char * configFile, const char * outputFile, `cudaOutputMode_t` outputMode)

Using this API user can initialize the CUDA profiler by specifying the configuration file, output file and output file format. This API is generally used to profile different set of counters by looping the kernel launch. The `configFile` parameter can be used to select profiling options including profiler counters. Refer to the "Compute Command Line Profiler User Guide" for supported profiler options and counters.

Limitation: The CUDA profiler cannot be initialized with this API if another profiling tool is already active, as indicated by the `cudaErrorProfilerDisabled` return code.

Typical usage of the profiling APIs is as follows:

for each set of counters/options

```
{  
cudaProfilerInitialize(); //Initialize profiling,set the counters/options in the config file  
...  
cudaProfilerStart();  
// code to be profiled  
cudaProfilerStop();  
...  
cudaProfilerStart();  
// code to be profiled  
cudaProfilerStop();  
...  
}
```

Parameters:

configFile - Name of the config file that lists the counters/options for profiling.

outputFile - Name of the outputFile where the profiling results will be stored.

outputMode - outputMode, can be [cudaKeyValuePair](#) OR [cudaCSV](#).

Returns:

[cudaSuccess](#), [cudaErrorInvalidValue](#), [cudaErrorProfilerDisabled](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaProfilerStart](#), [cudaProfilerStop](#)

5.27.2.2 `cudaError_t cudaProfilerStart (void)`

Enables profile collection by the active profiling tool. If profiling is already enabled, then [cudaProfilerStart\(\)](#) has no effect.

[cudaProfilerStart](#) and [cudaProfilerStop](#) APIs are used to programmatically control the profiling granularity by allowing profiling to be done only on selective pieces of code.

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaProfilerInitialize](#), [cudaProfilerStop](#)

5.27.2.3 `cudaError_t cudaProfilerStop (void)`

Disables profile collection by the active profiling tool. If profiling is already disabled, then [cudaProfilerStop\(\)](#) has no effect.

[cudaProfilerStart](#) and [cudaProfilerStop](#) APIs are used to programmatically control the profiling granularity by allowing profiling to be done only on selective pieces of code.

Returns:

[cudaSuccess](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cudaProfilerInitialize](#), [cudaProfilerStart](#)

5.28 Data types used by CUDA Runtime

Data Structures

- struct [cudaChannelFormatDesc](#)
- struct [cudaDeviceProp](#)
- struct [cudaExtent](#)
- struct [cudaFuncAttributes](#)
- struct [cudaIpcEventHandle_st](#)
- struct [cudaIpcMemHandle_st](#)
- struct [cudaMemcpy3DParms](#)
- struct [cudaMemcpy3DPeerParms](#)
- struct [cudaPitchedPtr](#)
- struct [cudaPointerAttributes](#)
- struct [cudaPos](#)
- struct [cudaResourceDesc](#)
- struct [cudaResourceViewDesc](#)
- struct [cudaTextureDesc](#)
- struct [surfaceReference](#)
- struct [textureReference](#)

Defines

- #define [CUDA_IPC_HANDLE_SIZE](#) 64
- #define [cudaArrayCubemap](#) 0x04
- #define [cudaArrayDefault](#) 0x00
- #define [cudaArrayLayered](#) 0x01
- #define [cudaArraySurfaceLoadStore](#) 0x02
- #define [cudaArrayTextureGather](#) 0x08
- #define [cudaDeviceBlockingSync](#) 0x04
- #define [cudaDeviceLmemResizeToMax](#) 0x10
- #define [cudaDeviceMapHost](#) 0x08
- #define [cudaDeviceMask](#) 0x1f
- #define [cudaDevicePropDontCare](#)
- #define [cudaDeviceScheduleAuto](#) 0x00
- #define [cudaDeviceScheduleBlockingSync](#) 0x04
- #define [cudaDeviceScheduleMask](#) 0x07
- #define [cudaDeviceScheduleSpin](#) 0x01
- #define [cudaDeviceScheduleYield](#) 0x02
- #define [cudaEventBlockingSync](#) 0x01
- #define [cudaEventDefault](#) 0x00
- #define [cudaEventDisableTiming](#) 0x02
- #define [cudaEventInterprocess](#) 0x04
- #define [cudaHostAllocDefault](#) 0x00
- #define [cudaHostAllocMapped](#) 0x02
- #define [cudaHostAllocPortable](#) 0x01
- #define [cudaHostAllocWriteCombined](#) 0x04
- #define [cudaHostRegisterDefault](#) 0x00
- #define [cudaHostRegisterMapped](#) 0x02
- #define [cudaHostRegisterPortable](#) 0x01

- #define `cudaIpcMemLazyEnablePeerAccess` 0x01
- #define `cudaPeerAccessDefault` 0x00
- #define `cudaStreamDefault` 0x00
- #define `cudaStreamNonBlocking` 0x01

Typedefs

- typedef struct `cudaArray` * `cudaArray_const_t`
- typedef struct `cudaArray` * `cudaArray_t`
- typedef enum `cudaError` `cudaError_t`
- typedef struct `CUevent_st` * `cudaEvent_t`
- typedef struct `cudaGraphicsResource` * `cudaGraphicsResource_t`
- typedef struct `cudaIpcEventHandle_st` `cudaIpcEventHandle_t`
- typedef struct `cudaIpcMemHandle_st` `cudaIpcMemHandle_t`
- typedef struct `cudaMipmappedArray` * `cudaMipmappedArray_const_t`
- typedef struct `cudaMipmappedArray` * `cudaMipmappedArray_t`
- typedef enum `cudaOutputMode` `cudaOutputMode_t`
- typedef struct `CUstream_st` * `cudaStream_t`
- typedef unsigned long long `cudaSurfaceObject_t`
- typedef unsigned long long `cudaTextureObject_t`
- typedef struct `CUuuid_st` `cudaUUID_t`

Enumerations

- enum `cudaChannelFormatKind` {
`cudaChannelFormatKindSigned` = 0,
`cudaChannelFormatKindUnsigned` = 1,
`cudaChannelFormatKindFloat` = 2,
`cudaChannelFormatKindNone` = 3 }
- enum `cudaComputeMode` {
`cudaComputeModeDefault` = 0,
`cudaComputeModeExclusive` = 1,
`cudaComputeModeProhibited` = 2,
`cudaComputeModeExclusiveProcess` = 3 }
- enum `cudaDeviceAttr` {
`cudaDevAttrMaxThreadsPerBlock` = 1,
`cudaDevAttrMaxBlockDimX` = 2,
`cudaDevAttrMaxBlockDimY` = 3,
`cudaDevAttrMaxBlockDimZ` = 4,
`cudaDevAttrMaxGridDimX` = 5,
`cudaDevAttrMaxGridDimY` = 6,
`cudaDevAttrMaxGridDimZ` = 7,
`cudaDevAttrMaxSharedMemoryPerBlock` = 8,
`cudaDevAttrTotalConstantMemory` = 9,
`cudaDevAttrWarpSize` = 10,

cudaDevAttrMaxPitch = 11,
cudaDevAttrMaxRegistersPerBlock = 12,
cudaDevAttrClockRate = 13,
cudaDevAttrTextureAlignment = 14,
cudaDevAttrGpuOverlap = 15,
cudaDevAttrMultiProcessorCount = 16,
cudaDevAttrKernelExecTimeout = 17,
cudaDevAttrIntegrated = 18,
cudaDevAttrCanMapHostMemory = 19,
cudaDevAttrComputeMode = 20,
cudaDevAttrMaxTexture1DWidth = 21,
cudaDevAttrMaxTexture2DWidth = 22,
cudaDevAttrMaxTexture2DHeight = 23,
cudaDevAttrMaxTexture3DWidth = 24,
cudaDevAttrMaxTexture3DHeight = 25,
cudaDevAttrMaxTexture3DDepth = 26,
cudaDevAttrMaxTexture2DLayeredWidth = 27,
cudaDevAttrMaxTexture2DLayeredHeight = 28,
cudaDevAttrMaxTexture2DLayeredLayers = 29,
cudaDevAttrSurfaceAlignment = 30,
cudaDevAttrConcurrentKernels = 31,
cudaDevAttrEccEnabled = 32,
cudaDevAttrPciBusId = 33,
cudaDevAttrPciDeviceId = 34,
cudaDevAttrTccDriver = 35,
cudaDevAttrMemoryClockRate = 36,
cudaDevAttrGlobalMemoryBusWidth = 37,
cudaDevAttrL2CacheSize = 38,
cudaDevAttrMaxThreadsPerMultiProcessor = 39,
cudaDevAttrAsyncEngineCount = 40,
cudaDevAttrUnifiedAddressing = 41,
cudaDevAttrMaxTexture1DLayeredWidth = 42,
cudaDevAttrMaxTexture1DLayeredLayers = 43,
cudaDevAttrMaxTexture2DGatherWidth = 45,
cudaDevAttrMaxTexture2DGatherHeight = 46,
cudaDevAttrMaxTexture3DWidthAlt = 47,
cudaDevAttrMaxTexture3DHeightAlt = 48,
cudaDevAttrMaxTexture3DDepthAlt = 49,
cudaDevAttrPciDomainId = 50,
cudaDevAttrTexturePitchAlignment = 51,
cudaDevAttrMaxTextureCubemapWidth = 52,

```
cudaDevAttrMaxTextureCubemapLayeredWidth = 53,  
cudaDevAttrMaxTextureCubemapLayeredLayers = 54,  
cudaDevAttrMaxSurface1DWidth = 55,  
cudaDevAttrMaxSurface2DWidth = 56,  
cudaDevAttrMaxSurface2DHeight = 57,  
cudaDevAttrMaxSurface3DWidth = 58,  
cudaDevAttrMaxSurface3DHeight = 59,  
cudaDevAttrMaxSurface3DDepth = 60,  
cudaDevAttrMaxSurface1DLayeredWidth = 61,  
cudaDevAttrMaxSurface1DLayeredLayers = 62,  
cudaDevAttrMaxSurface2DLayeredWidth = 63,  
cudaDevAttrMaxSurface2DLayeredHeight = 64,  
cudaDevAttrMaxSurface2DLayeredLayers = 65,  
cudaDevAttrMaxSurfaceCubemapWidth = 66,  
cudaDevAttrMaxSurfaceCubemapLayeredWidth = 67,  
cudaDevAttrMaxSurfaceCubemapLayeredLayers = 68,  
cudaDevAttrMaxTexture1DLinearWidth = 69,  
cudaDevAttrMaxTexture2DLinearWidth = 70,  
cudaDevAttrMaxTexture2DLinearHeight = 71,  
cudaDevAttrMaxTexture2DLinearPitch = 72,  
cudaDevAttrMaxTexture2DMipmappedWidth = 73,  
cudaDevAttrMaxTexture2DMipmappedHeight = 74,  
cudaDevAttrComputeCapabilityMajor = 75,  
cudaDevAttrComputeCapabilityMinor = 76,  
cudaDevAttrMaxTexture1DMipmappedWidth = 77 }  
• enum cudaError {  
    cudaSuccess = 0,  
    cudaErrorMissingConfiguration = 1,  
    cudaErrorMemoryAllocation = 2,  
    cudaErrorInitializationError = 3,  
    cudaErrorLaunchFailure = 4,  
    cudaErrorPriorLaunchFailure = 5,  
    cudaErrorLaunchTimeout = 6,  
    cudaErrorLaunchOutOfResources = 7,  
    cudaErrorInvalidDeviceFunction = 8,  
    cudaErrorInvalidConfiguration = 9,  
    cudaErrorInvalidDevice = 10,  
    cudaErrorInvalidValue = 11,  
    cudaErrorInvalidPitchValue = 12,  
    cudaErrorInvalidSymbol = 13,  
    cudaErrorMapBufferObjectFailed = 14,
```

`cudaErrorUnmapBufferObjectFailed` = 15,
`cudaErrorInvalidHostPointer` = 16,
`cudaErrorInvalidDevicePointer` = 17,
`cudaErrorInvalidTexture` = 18,
`cudaErrorInvalidTextureBinding` = 19,
`cudaErrorInvalidChannelDescriptor` = 20,
`cudaErrorInvalidMemcpyDirection` = 21,
`cudaErrorAddressOfConstant` = 22,
`cudaErrorTextureFetchFailed` = 23,
`cudaErrorTextureNotBound` = 24,
`cudaErrorSynchronizationError` = 25,
`cudaErrorInvalidFilterSetting` = 26,
`cudaErrorInvalidNormSetting` = 27,
`cudaErrorMixedDeviceExecution` = 28,
`cudaErrorCudartUnloading` = 29,
`cudaErrorUnknown` = 30,
`cudaErrorNotYetImplemented` = 31,
`cudaErrorMemoryValueTooLarge` = 32,
`cudaErrorInvalidResourceHandle` = 33,
`cudaErrorNotReady` = 34,
`cudaErrorInsufficientDriver` = 35,
`cudaErrorSetOnActiveProcess` = 36,
`cudaErrorInvalidSurface` = 37,
`cudaErrorNoDevice` = 38,
`cudaErrorECCUncorrectable` = 39,
`cudaErrorSharedObjectSymbolNotFound` = 40,
`cudaErrorSharedObjectInitFailed` = 41,
`cudaErrorUnsupportedLimit` = 42,
`cudaErrorDuplicateVariableName` = 43,
`cudaErrorDuplicateTextureName` = 44,
`cudaErrorDuplicateSurfaceName` = 45,
`cudaErrorDevicesUnavailable` = 46,
`cudaErrorInvalidKernelImage` = 47,
`cudaErrorNoKernelImageForDevice` = 48,
`cudaErrorIncompatibleDriverContext` = 49,
`cudaErrorPeerAccessAlreadyEnabled` = 50,
`cudaErrorPeerAccessNotEnabled` = 51,
`cudaErrorDeviceAlreadyInUse` = 54,
`cudaErrorProfilerDisabled` = 55,
`cudaErrorProfilerNotInitialized` = 56,
`cudaErrorProfilerAlreadyStarted` = 57,

```

cudaErrorProfilerAlreadyStopped = 58,
cudaErrorAssert = 59,
cudaErrorTooManyPeers = 60,
cudaErrorHostMemoryAlreadyRegistered = 61,
cudaErrorHostMemoryNotRegistered = 62,
cudaErrorOperatingSystem = 63,
cudaErrorPeerAccessUnsupported = 64,
cudaErrorLaunchMaxDepthExceeded = 65,
cudaErrorLaunchFileScopedTex = 66,
cudaErrorLaunchFileScopedSurf = 67,
cudaErrorSyncDepthExceeded = 68,
cudaErrorLaunchPendingCountExceeded = 69,
cudaErrorNotPermitted = 70,
cudaErrorNotSupported = 71,
cudaErrorStartupFailure = 0x7f,
cudaErrorApiFailureBase = 10000 }
• enum cudaFuncCache {
  cudaFuncCachePreferNone = 0,
  cudaFuncCachePreferShared = 1,
  cudaFuncCachePreferL1 = 2,
  cudaFuncCachePreferEqual = 3 }
• enum cudaGraphicsCubeFace {
  cudaGraphicsCubeFacePositiveX = 0x00,
  cudaGraphicsCubeFaceNegativeX = 0x01,
  cudaGraphicsCubeFacePositiveY = 0x02,
  cudaGraphicsCubeFaceNegativeY = 0x03,
  cudaGraphicsCubeFacePositiveZ = 0x04,
  cudaGraphicsCubeFaceNegativeZ = 0x05 }
• enum cudaGraphicsMapFlags {
  cudaGraphicsMapFlagsNone = 0,
  cudaGraphicsMapFlagsReadOnly = 1,
  cudaGraphicsMapFlagsWriteDiscard = 2 }
• enum cudaGraphicsRegisterFlags {
  cudaGraphicsRegisterFlagsNone = 0,
  cudaGraphicsRegisterFlagsReadOnly = 1,
  cudaGraphicsRegisterFlagsWriteDiscard = 2,
  cudaGraphicsRegisterFlagsSurfaceLoadStore = 4,
  cudaGraphicsRegisterFlagsTextureGather = 8 }
• enum cudaLimit {
  cudaLimitStackSize = 0x00,
  cudaLimitPrintfFifoSize = 0x01,
  cudaLimitMallocHeapSize = 0x02,
  cudaLimitDevRuntimeSyncDepth = 0x03,
  cudaLimitDevRuntimePendingLaunchCount = 0x04 }

```

- enum `cudaMemcpyKind` {
 `cudaMemcpyHostToHost` = 0,
 `cudaMemcpyHostToDevice` = 1,
 `cudaMemcpyDeviceToHost` = 2,
 `cudaMemcpyDeviceToDevice` = 3,
 `cudaMemcpyDefault` = 4 }
- enum `cudaMemoryType` {
 `cudaMemoryTypeHost` = 1,
 `cudaMemoryTypeDevice` = 2 }
- enum `cudaOutputMode` {
 `cudaKeyValuePair` = 0x00,
 `cudaCSV` = 0x01 }
- enum `cudaResourceType` {
 `cudaResourceTypeArray` = 0x00,
 `cudaResourceTypeMipmappedArray` = 0x01,
 `cudaResourceTypeLinear` = 0x02,
 `cudaResourceTypePitch2D` = 0x03 }
- enum `cudaResourceViewFormat` {
 `cudaResViewFormatNone` = 0x00,
 `cudaResViewFormatUnsignedChar1` = 0x01,
 `cudaResViewFormatUnsignedChar2` = 0x02,
 `cudaResViewFormatUnsignedChar4` = 0x03,
 `cudaResViewFormatSignedChar1` = 0x04,
 `cudaResViewFormatSignedChar2` = 0x05,
 `cudaResViewFormatSignedChar4` = 0x06,
 `cudaResViewFormatUnsignedShort1` = 0x07,
 `cudaResViewFormatUnsignedShort2` = 0x08,
 `cudaResViewFormatUnsignedShort4` = 0x09,
 `cudaResViewFormatSignedShort1` = 0x0a,
 `cudaResViewFormatSignedShort2` = 0x0b,
 `cudaResViewFormatSignedShort4` = 0x0c,
 `cudaResViewFormatUnsignedInt1` = 0x0d,
 `cudaResViewFormatUnsignedInt2` = 0x0e,
 `cudaResViewFormatUnsignedInt4` = 0x0f,
 `cudaResViewFormatSignedInt1` = 0x10,
 `cudaResViewFormatSignedInt2` = 0x11,
 `cudaResViewFormatSignedInt4` = 0x12,
 `cudaResViewFormatHalf1` = 0x13,
 `cudaResViewFormatHalf2` = 0x14,
 `cudaResViewFormatHalf4` = 0x15,
 `cudaResViewFormatFloat1` = 0x16,
 `cudaResViewFormatFloat2` = 0x17,

```

cudaResViewFormatFloat4 = 0x18,
cudaResViewFormatUnsignedBlockCompressed1 = 0x19,
cudaResViewFormatUnsignedBlockCompressed2 = 0x1a,
cudaResViewFormatUnsignedBlockCompressed3 = 0x1b,
cudaResViewFormatUnsignedBlockCompressed4 = 0x1c,
cudaResViewFormatSignedBlockCompressed4 = 0x1d,
cudaResViewFormatUnsignedBlockCompressed5 = 0x1e,
cudaResViewFormatSignedBlockCompressed5 = 0x1f,
cudaResViewFormatUnsignedBlockCompressed6H = 0x20,
cudaResViewFormatSignedBlockCompressed6H = 0x21,
cudaResViewFormatUnsignedBlockCompressed7 = 0x22 }
• enum cudaSharedMemConfig
• enum cudaSurfaceBoundaryMode {
    cudaBoundaryModeZero = 0,
    cudaBoundaryModeClamp = 1,
    cudaBoundaryModeTrap = 2 }
• enum cudaSurfaceFormatMode {
    cudaFormatModeForced = 0,
    cudaFormatModeAuto = 1 }
• enum cudaTextureAddressMode {
    cudaAddressModeWrap = 0,
    cudaAddressModeClamp = 1,
    cudaAddressModeMirror = 2,
    cudaAddressModeBorder = 3 }
• enum cudaTextureFilterMode {
    cudaFilterModePoint = 0,
    cudaFilterModeLinear = 1 }
• enum cudaTextureReadMode {
    cudaReadModeElementType = 0,
    cudaReadModeNormalizedFloat = 1 }

```

5.28.1 Define Documentation

5.28.1.1 #define CUDA_IPC_HANDLE_SIZE 64

CUDA IPC Handle Size

5.28.1.2 #define cudaArrayCubemap 0x04

Must be set in cudaMalloc3DArray to create a cubemap CUDA array

5.28.1.3 #define cudaArrayDefault 0x00

Default CUDA array allocation flag

5.28.1.4 #define cudaArrayLayered 0x01

Must be set in `cudaMalloc3DArray` to create a layered CUDA array

5.28.1.5 #define cudaArraySurfaceLoadStore 0x02

Must be set in `cudaMallocArray` or `cudaMalloc3DArray` in order to bind surfaces to the CUDA array

5.28.1.6 #define cudaArrayTextureGather 0x08

Must be set in `cudaMallocArray` or `cudaMalloc3DArray` in order to perform texture gather operations on the CUDA array

5.28.1.7 #define cudaDeviceBlockingSync 0x04

Device flag - Use blocking synchronization

Deprecated

This flag was deprecated as of CUDA 4.0 and replaced with [cudaDeviceScheduleBlockingSync](#).

5.28.1.8 #define cudaDeviceLmemResizeToMax 0x10

Device flag - Keep local memory allocation after launch

5.28.1.9 #define cudaDeviceMapHost 0x08

Device flag - Support mapped pinned allocations

5.28.1.10 #define cudaDeviceMask 0x1f

Device flags mask

5.28.1.11 #define cudaDevicePropDontCare

Empty device properties

5.28.1.12 #define cudaDeviceScheduleAuto 0x00

Device flag - Automatic scheduling

5.28.1.13 #define cudaDeviceScheduleBlockingSync 0x04

Device flag - Use blocking synchronization

5.28.1.14 #define cudaDeviceScheduleMask 0x07

Device schedule flags mask

5.28.1.15 #define cudaDeviceScheduleSpin 0x01

Device flag - Spin default scheduling

5.28.1.16 #define cudaDeviceScheduleYield 0x02

Device flag - Yield default scheduling

5.28.1.17 #define cudaEventBlockingSync 0x01

Event uses blocking synchronization

5.28.1.18 #define cudaEventDefault 0x00

Default event flag

5.28.1.19 #define cudaEventDisableTiming 0x02

Event will not record timing data

5.28.1.20 #define cudaEventInterprocess 0x04

Event is suitable for interprocess use. cudaEventDisableTiming must be set

5.28.1.21 #define cudaHostAllocDefault 0x00

Default page-locked allocation flag

5.28.1.22 #define cudaHostAllocMapped 0x02

Map allocation into device space

5.28.1.23 #define cudaHostAllocPortable 0x01

Pinned memory accessible by all CUDA contexts

5.28.1.24 #define cudaHostAllocWriteCombined 0x04

Write-combined memory

5.28.1.25 #define cudaHostRegisterDefault 0x00

Default host memory registration flag

5.28.1.26 #define cudaHostRegisterMapped 0x02

Map registered memory into device space

5.28.1.27 #define cudaHostRegisterPortable 0x01

Pinned memory accessible by all CUDA contexts

5.28.1.28 #define cudaIpcMemLazyEnablePeerAccess 0x01

Automatically enable peer access between remote devices as needed

5.28.1.29 #define cudaPeerAccessDefault 0x00

Default peer addressing enable flag

5.28.1.30 #define cudaStreamDefault 0x00

Default stream flag

5.28.1.31 #define cudaStreamNonBlocking 0x01

Stream does not synchronize with stream 0 (the NULL stream)

5.28.2 Typedef Documentation**5.28.2.1 typedef struct cudaArray* cudaArray_const_t**

CUDA array (as source copy argument)

5.28.2.2 typedef struct cudaArray* cudaArray_t

CUDA array

5.28.2.3 typedef enum cudaError cudaError_t

CUDA Error types

5.28.2.4 typedef struct CUevent_st* cudaEvent_t

CUDA event types

5.28.2.5 typedef struct cudaGraphicsResource* cudaGraphicsResource_t

CUDA graphics resource types

5.28.2.6 typedef struct cudaIpcEventHandle_st cudaIpcEventHandle_t

CUDA IPC event handle

5.28.2.7 typedef struct cudaIpcMemHandle_st cudaIpcMemHandle_t

CUDA IPC memory handle

5.28.2.8 typedef struct cudaMipmappedArray* cudaMipmappedArray_const_t

CUDA mipmapped array (as source argument)

5.28.2.9 typedef struct cudaMipmappedArray* cudaMipmappedArray_t

CUDA mipmapped array

5.28.2.10 typedef enum cudaOutputMode cudaOutputMode_t

CUDA output file modes

5.28.2.11 typedef struct CUstream_st* cudaStream_t

CUDA stream

5.28.2.12 typedef unsigned long long cudaSurfaceObject_t

CUDA Surface object

5.28.2.13 typedef unsigned long long cudaTextureObject_t

CUDA texture object

5.28.2.14 typedef struct CUuuid_st cudaUUID_t

CUDA UUID types

5.28.3 Enumeration Type Documentation**5.28.3.1 enum cudaChannelFormatKind**

Channel format kind

Enumerator:

- cudaChannelFormatKindSigned* Signed channel format
- cudaChannelFormatKindUnsigned* Unsigned channel format
- cudaChannelFormatKindFloat* Float channel format
- cudaChannelFormatKindNone* No channel format

5.28.3.2 enum cudaComputeMode

CUDA device compute modes

Enumerator:

- cudaComputeModeDefault* Default compute mode (Multiple threads can use [cudaSetDevice\(\)](#) with this device)
- cudaComputeModeExclusive* Compute-exclusive-thread mode (Only one thread in one process will be able to use [cudaSetDevice\(\)](#) with this device)
- cudaComputeModeProhibited* Compute-prohibited mode (No threads can use [cudaSetDevice\(\)](#) with this device)
- cudaComputeModeExclusiveProcess* Compute-exclusive-process mode (Many threads in one process will be able to use [cudaSetDevice\(\)](#) with this device)

5.28.3.3 enum cudaDeviceAttr

CUDA device attributes

Enumerator:

- cudaDevAttrMaxThreadsPerBlock* Maximum number of threads per block
- cudaDevAttrMaxBlockDimX* Maximum block dimension X
- cudaDevAttrMaxBlockDimY* Maximum block dimension Y
- cudaDevAttrMaxBlockDimZ* Maximum block dimension Z
- cudaDevAttrMaxGridDimX* Maximum grid dimension X
- cudaDevAttrMaxGridDimY* Maximum grid dimension Y
- cudaDevAttrMaxGridDimZ* Maximum grid dimension Z
- cudaDevAttrMaxSharedMemoryPerBlock* Maximum shared memory available per block in bytes
- cudaDevAttrTotalConstantMemory* Memory available on device for `__constant__` variables in a CUDA C kernel in bytes
- cudaDevAttrWarpSize* Warp size in threads
- cudaDevAttrMaxPitch* Maximum pitch in bytes allowed by memory copies
- cudaDevAttrMaxRegistersPerBlock* Maximum number of 32-bit registers available per block
- cudaDevAttrClockRate* Peak clock frequency in kilohertz
- cudaDevAttrTextureAlignment* Alignment requirement for textures
- cudaDevAttrGpuOverlap* Device can possibly copy memory and execute a kernel concurrently
- cudaDevAttrMultiProcessorCount* Number of multiprocessors on device
- cudaDevAttrKernelExecTimeout* Specifies whether there is a run time limit on kernels
- cudaDevAttrIntegrated* Device is integrated with host memory
- cudaDevAttrCanMapHostMemory* Device can map host memory into CUDA address space
- cudaDevAttrComputeMode* Compute mode (See [cudaComputeMode](#) for details)
- cudaDevAttrMaxTexture1DWidth* Maximum 1D texture width
- cudaDevAttrMaxTexture2DWidth* Maximum 2D texture width
- cudaDevAttrMaxTexture2DHeight* Maximum 2D texture height
- cudaDevAttrMaxTexture3DWidth* Maximum 3D texture width

cudaDevAttrMaxTexture3DHeight Maximum 3D texture height

cudaDevAttrMaxTexture3DDepth Maximum 3D texture depth

cudaDevAttrMaxTexture2DLayeredWidth Maximum 2D layered texture width

cudaDevAttrMaxTexture2DLayeredHeight Maximum 2D layered texture height

cudaDevAttrMaxTexture2DLayeredLayers Maximum layers in a 2D layered texture

cudaDevAttrSurfaceAlignment Alignment requirement for surfaces

cudaDevAttrConcurrentKernels Device can possibly execute multiple kernels concurrently

cudaDevAttrEccEnabled Device has ECC support enabled

cudaDevAttrPciBusId PCI bus ID of the device

cudaDevAttrPciDeviceId PCI device ID of the device

cudaDevAttrTccDriver Device is using TCC driver model

cudaDevAttrMemoryClockRate Peak memory clock frequency in kilohertz

cudaDevAttrGlobalMemoryBusWidth Global memory bus width in bits

cudaDevAttrL2CacheSize Size of L2 cache in bytes

cudaDevAttrMaxThreadsPerMultiProcessor Maximum resident threads per multiprocessor

cudaDevAttrAsyncEngineCount Number of asynchronous engines

cudaDevAttrUnifiedAddressing Device shares a unified address space with the host

cudaDevAttrMaxTexture1DLayeredWidth Maximum 1D layered texture width

cudaDevAttrMaxTexture1DLayeredLayers Maximum layers in a 1D layered texture

cudaDevAttrMaxTexture2DGatherWidth Maximum 2D texture width if `cudaArrayTextureGather` is set

cudaDevAttrMaxTexture2DGatherHeight Maximum 2D texture height if `cudaArrayTextureGather` is set

cudaDevAttrMaxTexture3DWidthAlt Alternate maximum 3D texture width

cudaDevAttrMaxTexture3DHeightAlt Alternate maximum 3D texture height

cudaDevAttrMaxTexture3DDepthAlt Alternate maximum 3D texture depth

cudaDevAttrPciDomainId PCI domain ID of the device

cudaDevAttrTexturePitchAlignment Pitch alignment requirement for textures

cudaDevAttrMaxTextureCubemapWidth Maximum cubemap texture width/height

cudaDevAttrMaxTextureCubemapLayeredWidth Maximum cubemap layered texture width/height

cudaDevAttrMaxTextureCubemapLayeredLayers Maximum layers in a cubemap layered texture

cudaDevAttrMaxSurface1DWidth Maximum 1D surface width

cudaDevAttrMaxSurface2DWidth Maximum 2D surface width

cudaDevAttrMaxSurface2DHeight Maximum 2D surface height

cudaDevAttrMaxSurface3DWidth Maximum 3D surface width

cudaDevAttrMaxSurface3DHeight Maximum 3D surface height

cudaDevAttrMaxSurface3DDepth Maximum 3D surface depth

cudaDevAttrMaxSurface1DLayeredWidth Maximum 1D layered surface width

cudaDevAttrMaxSurface1DLayeredLayers Maximum layers in a 1D layered surface

cudaDevAttrMaxSurface2DLayeredWidth Maximum 2D layered surface width

cudaDevAttrMaxSurface2DLayeredHeight Maximum 2D layered surface height

cudaDevAttrMaxSurface2DLayeredLayers Maximum layers in a 2D layered surface

cudaDevAttrMaxSurfaceCubemapWidth Maximum cubemap surface width

cudaDevAttrMaxSurfaceCubemapLayeredWidth Maximum cubemap layered surface width
cudaDevAttrMaxSurfaceCubemapLayeredLayers Maximum layers in a cubemap layered surface
cudaDevAttrMaxTexture1DLinearWidth Maximum 1D linear texture width
cudaDevAttrMaxTexture2DLinearWidth Maximum 2D linear texture width
cudaDevAttrMaxTexture2DLinearHeight Maximum 2D linear texture height
cudaDevAttrMaxTexture2DLinearPitch Maximum 2D linear texture pitch in bytes
cudaDevAttrMaxTexture2DMipmappedWidth Maximum mipmapped 2D texture width
cudaDevAttrMaxTexture2DMipmappedHeight Maximum mipmapped 2D texture height
cudaDevAttrComputeCapabilityMajor Major compute capability version number
cudaDevAttrComputeCapabilityMinor Minor compute capability version number
cudaDevAttrMaxTexture1DMipmappedWidth Maximum mipmapped 1D texture width

5.28.3.4 enum cudaError

CUDA error types

Enumerator:

cudaSuccess The API call returned with no errors. In the case of query calls, this can also mean that the operation being queried is complete (see [cudaEventQuery\(\)](#) and [cudaStreamQuery\(\)](#)).

cudaErrorMissingConfiguration The device function being invoked (usually via [cudaLaunch\(\)](#)) was not previously configured via the [cudaConfigureCall\(\)](#) function.

cudaErrorMemoryAllocation The API call failed because it was unable to allocate enough memory to perform the requested operation.

cudaErrorInitializationError The API call failed because the CUDA driver and runtime could not be initialized.

cudaErrorLaunchFailure An exception occurred on the device while executing a kernel. Common causes include dereferencing an invalid device pointer and accessing out of bounds shared memory. The device cannot be used until [cudaThreadExit\(\)](#) is called. All existing device memory allocations are invalid and must be reconstructed if the program is to continue using CUDA.

cudaErrorPriorLaunchFailure This indicated that a previous kernel launch failed. This was previously used for device emulation of kernel launches.

Deprecated

This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

cudaErrorLaunchTimeout This indicates that the device kernel took too long to execute. This can only occur if timeouts are enabled - see the device property [kernelExecTimeoutEnabled](#) for more information. The device cannot be used until [cudaThreadExit\(\)](#) is called. All existing device memory allocations are invalid and must be reconstructed if the program is to continue using CUDA.

cudaErrorLaunchOutOfResources This indicates that a launch did not occur because it did not have appropriate resources. Although this error is similar to [cudaErrorInvalidConfiguration](#), this error usually indicates that the user has attempted to pass too many arguments to the device kernel, or the kernel launch specifies too many threads for the kernel's register count.

cudaErrorInvalidDeviceFunction The requested device function does not exist or is not compiled for the proper device architecture.

cudaErrorInvalidConfiguration This indicates that a kernel launch is requesting resources that can never be satisfied by the current device. Requesting more shared memory per block than the device supports will trigger this error, as will requesting too many threads or blocks. See [cudaDeviceProp](#) for more device limitations.

cudaErrorInvalidDevice This indicates that the device ordinal supplied by the user does not correspond to a valid CUDA device.

cudaErrorInvalidValue This indicates that one or more of the parameters passed to the API call is not within an acceptable range of values.

cudaErrorInvalidPitchValue This indicates that one or more of the pitch-related parameters passed to the API call is not within the acceptable range for pitch.

cudaErrorInvalidSymbol This indicates that the symbol name/identifier passed to the API call is not a valid name or identifier.

cudaErrorMapBufferObjectFailed This indicates that the buffer object could not be mapped.

cudaErrorUnmapBufferObjectFailed This indicates that the buffer object could not be unmapped.

cudaErrorInvalidHostPointer This indicates that at least one host pointer passed to the API call is not a valid host pointer.

cudaErrorInvalidDevicePointer This indicates that at least one device pointer passed to the API call is not a valid device pointer.

cudaErrorInvalidTexture This indicates that the texture passed to the API call is not a valid texture.

cudaErrorInvalidTextureBinding This indicates that the texture binding is not valid. This occurs if you call [cudaGetTextureAlignmentOffset\(\)](#) with an unbound texture.

cudaErrorInvalidChannelDescriptor This indicates that the channel descriptor passed to the API call is not valid. This occurs if the format is not one of the formats specified by [cudaChannelFormatKind](#), or if one of the dimensions is invalid.

cudaErrorInvalidMemcpyDirection This indicates that the direction of the memcpy passed to the API call is not one of the types specified by [cudaMemcpyKind](#).

cudaErrorAddressOfConstant This indicated that the user has taken the address of a constant variable, which was forbidden up until the CUDA 3.1 release.

Deprecated

This error return is deprecated as of CUDA 3.1. Variables in constant memory may now have their address taken by the runtime via [cudaGetSymbolAddress\(\)](#).

cudaErrorTextureFetchFailed This indicated that a texture fetch was not able to be performed. This was previously used for device emulation of texture operations.

Deprecated

This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

cudaErrorTextureNotBound This indicated that a texture was not bound for access. This was previously used for device emulation of texture operations.

Deprecated

This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

cudaErrorSynchronizationError This indicated that a synchronization operation had failed. This was previously used for some device emulation functions.

Deprecated

This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

cudaErrorInvalidFilterSetting This indicates that a non-float texture was being accessed with linear filtering. This is not supported by CUDA.

cudaErrorInvalidNormSetting This indicates that an attempt was made to read a non-float texture as a normalized float. This is not supported by CUDA.

cudaErrorMixedDeviceExecution Mixing of device and device emulation code was not allowed.

Deprecated

This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

cudaErrorCudartUnloading This indicates that a CUDA Runtime API call cannot be executed because it is being called during process shut down, at a point in time after CUDA driver has been unloaded.

cudaErrorUnknown This indicates that an unknown internal error has occurred.

cudaErrorNotYetImplemented This indicates that the API call is not yet implemented. Production releases of CUDA will never return this error.

Deprecated

This error return is deprecated as of CUDA 4.1.

cudaErrorMemoryValueTooLarge This indicated that an emulated device pointer exceeded the 32-bit address range.

Deprecated

This error return is deprecated as of CUDA 3.1. Device emulation mode was removed with the CUDA 3.1 release.

cudaErrorInvalidResourceHandle This indicates that a resource handle passed to the API call was not valid. Resource handles are opaque types like [cudaStream_t](#) and [cudaEvent_t](#).

cudaErrorNotReady This indicates that asynchronous operations issued previously have not completed yet. This result is not actually an error, but must be indicated differently than [cudaSuccess](#) (which indicates completion). Calls that may return this value include [cudaEventQuery\(\)](#) and [cudaStreamQuery\(\)](#).

cudaErrorInsufficientDriver This indicates that the installed NVIDIA CUDA driver is older than the CUDA runtime library. This is not a supported configuration. Users should install an updated NVIDIA display driver to allow the application to run.

cudaErrorSetOnActiveProcess This indicates that the user has called [cudaSetValidDevices\(\)](#), [cudaSetDeviceFlags\(\)](#), [cudaD3D9SetDirect3DDevice\(\)](#), [cudaD3D10SetDirect3DDevice\(\)](#), [cudaD3D11SetDirect3DDevice\(\)](#), or [cudaVDPASetVDPASUDevice\(\)](#) after initializing the CUDA runtime by calling non-device management operations (allocating memory and launching kernels are examples of non-device management operations). This error can also be returned if using runtime/driver interoperability and there is an existing [CUcontext](#) active on the host thread.

cudaErrorInvalidSurface This indicates that the surface passed to the API call is not a valid surface.

cudaErrorNoDevice This indicates that no CUDA-capable devices were detected by the installed CUDA driver.

cudaErrorECCUncorrectable This indicates that an uncorrectable ECC error was detected during execution.

cudaErrorSharedObjectSymbolNotFound This indicates that a link to a shared object failed to resolve.

cudaErrorSharedObjectInitFailed This indicates that initialization of a shared object failed.

cudaErrorUnsupportedLimit This indicates that the [cudaLimit](#) passed to the API call is not supported by the active device.

cudaErrorDuplicateVariableName This indicates that multiple global or constant variables (across separate CUDA source files in the application) share the same string name.

cudaErrorDuplicateTextureName This indicates that multiple textures (across separate CUDA source files in the application) share the same string name.

cudaErrorDuplicateSurfaceName This indicates that multiple surfaces (across separate CUDA source files in the application) share the same string name.

cudaErrorDevicesUnavailable This indicates that all CUDA devices are busy or unavailable at the current time. Devices are often busy/unavailable due to use of [cudaComputeModeExclusive](#), [cudaComputeModeProhibited](#) or when long running CUDA kernels have filled up the GPU and are blocking new work from starting. They can also be unavailable due to memory constraints on a device that already has active CUDA work being performed.

cudaErrorInvalidKernelImage This indicates that the device kernel image is invalid.

cudaErrorNoKernelImageForDevice This indicates that there is no kernel image available that is suitable for the device. This can occur when a user specifies code generation options for a particular CUDA source file that do not include the corresponding device configuration.

cudaErrorIncompatibleDriverContext This indicates that the current context is not compatible with this the CUDA Runtime. This can only occur if you are using CUDA Runtime/Driver interoperability and have created an existing Driver context using the driver API. The Driver context may be incompatible either because the Driver context was created using an older version of the API, because the Runtime API call expects a primary driver context and the Driver context is not primary, or because the Driver context has been destroyed. Please see [Interactions](#) with the CUDA Driver API" for more information.

cudaErrorPeerAccessAlreadyEnabled This error indicates that a call to [cudaDeviceEnablePeerAccess\(\)](#) is trying to re-enable peer addressing on from a context which has already had peer addressing enabled.

cudaErrorPeerAccessNotEnabled This error indicates that [cudaDeviceDisablePeerAccess\(\)](#) is trying to disable peer addressing which has not been enabled yet via [cudaDeviceEnablePeerAccess\(\)](#).

cudaErrorDeviceAlreadyInUse This indicates that a call tried to access an exclusive-thread device that is already in use by a different thread.

cudaErrorProfilerDisabled This indicates profiler is not initialized for this run. This can happen when the application is running with external profiling tools like visual profiler.

Deprecated

cudaErrorProfilerNotInitialized This error return is deprecated as of CUDA 5.0. It is no longer an error to attempt to enable/disable the profiling via [cudaProfilerStart](#) or [cudaProfilerStop](#) without initialization.

Deprecated

cudaErrorProfilerAlreadyStarted This error return is deprecated as of CUDA 5.0. It is no longer an error to call [cudaProfilerStart\(\)](#) when profiling is already enabled.

Deprecated

cudaErrorProfilerAlreadyStopped This error return is deprecated as of CUDA 5.0. It is no longer an error to call [cudaProfilerStop\(\)](#) when profiling is already disabled.

cudaErrorAssert An assert triggered in device code during kernel execution. The device cannot be used again until [cudaThreadExit\(\)](#) is called. All existing allocations are invalid and must be reconstructed if the program is to continue using CUDA.

cudaErrorTooManyPeers This error indicates that the hardware resources required to enable peer access have been exhausted for one or more of the devices passed to [cudaEnablePeerAccess\(\)](#).

cudaErrorHostMemoryAlreadyRegistered This error indicates that the memory range passed to [cudaHostRegister\(\)](#) has already been registered.

cudaErrorHostMemoryNotRegistered This error indicates that the pointer passed to [cudaHostUnregister\(\)](#) does not correspond to any currently registered memory region.

cudaErrorOperatingSystem This error indicates that an OS call failed.

cudaErrorPeerAccessUnsupported This error indicates that P2P access is not supported across the given devices.

cudaErrorLaunchMaxDepthExceeded This error indicates that a device runtime grid launch did not occur because the depth of the child grid would exceed the maximum supported number of nested grid launches.

cudaErrorLaunchFileScopedTex This error indicates that a grid launch did not occur because the kernel uses file-scoped textures which are unsupported by the device runtime. Kernels launched via the device runtime only support textures created with the Texture Object API's.

cudaErrorLaunchFileScopedSurf This error indicates that a grid launch did not occur because the kernel uses file-scoped surfaces which are unsupported by the device runtime. Kernels launched via the device runtime only support surfaces created with the Surface Object API's.

cudaErrorSyncDepthExceeded This error indicates that a call to [cudaDeviceSynchronize](#) made from the device runtime failed because the call was made at grid depth greater than either the default (2 levels of grids) or user specified device limit [cudaLimitDevRuntimeSyncDepth](#). To be able to synchronize on launched grids at a greater depth successfully, the maximum nested depth at which [cudaDeviceSynchronize](#) will be called must be specified with the [cudaLimitDevRuntimeSyncDepth](#) limit to the [cudaDeviceSetLimit](#) api before the host-side launch of a kernel using the device runtime. Keep in mind that additional levels of sync depth require the runtime to reserve large amounts of device memory that cannot be used for user allocations.

cudaErrorLaunchPendingCountExceeded This error indicates that a device runtime grid launch failed because the launch would exceed the limit [cudaLimitDevRuntimePendingLaunchCount](#). For this launch to proceed successfully, [cudaDeviceSetLimit](#) must be called to set the [cudaLimitDevRuntimePendingLaunchCount](#) to be higher than the upper bound of outstanding launches that can be issued to the device runtime. Keep in mind that raising the limit of pending device runtime launches will require the runtime to reserve device memory that cannot be used for user allocations.

cudaErrorNotPermitted This error indicates the attempted operation is not permitted.

cudaErrorNotSupported This error indicates the attempted operation is not supported on the current system or device.

cudaErrorStartupFailure This indicates an internal startup failure in the CUDA runtime.

cudaErrorApiFailureBase Any unhandled CUDA driver error is added to this value and returned via the runtime. Production releases of CUDA should not return such errors.

Deprecated

This error return is deprecated as of CUDA 4.1.

5.28.3.5 enum cudaFuncCache

CUDA function cache configurations

Enumerator:

cudaFuncCachePreferNone Default function cache configuration, no preference

cudaFuncCachePreferShared Prefer larger shared memory and smaller L1 cache

cudaFuncCachePreferL1 Prefer larger L1 cache and smaller shared memory

cudaFuncCachePreferEqual Prefer equal size L1 cache and shared memory

5.28.3.6 enum cudaGraphicsCubeFace

CUDA graphics interop array indices for cube maps

Enumerator:

cudaGraphicsCubeFacePositiveX Positive X face of cubemap

cudaGraphicsCubeFaceNegativeX Negative X face of cubemap

cudaGraphicsCubeFacePositiveY Positive Y face of cubemap

cudaGraphicsCubeFaceNegativeY Negative Y face of cubemap

cudaGraphicsCubeFacePositiveZ Positive Z face of cubemap

cudaGraphicsCubeFaceNegativeZ Negative Z face of cubemap

5.28.3.7 enum cudaGraphicsMapFlags

CUDA graphics interop map flags

Enumerator:

cudaGraphicsMapFlagsNone Default; Assume resource can be read/written

cudaGraphicsMapFlagsReadOnly CUDA will not write to this resource

cudaGraphicsMapFlagsWriteDiscard CUDA will only write to and will not read from this resource

5.28.3.8 enum cudaGraphicsRegisterFlags

CUDA graphics interop register flags

Enumerator:

cudaGraphicsRegisterFlagsNone Default

cudaGraphicsRegisterFlagsReadOnly CUDA will not write to this resource

cudaGraphicsRegisterFlagsWriteDiscard CUDA will only write to and will not read from this resource

cudaGraphicsRegisterFlagsSurfaceLoadStore CUDA will bind this resource to a surface reference

cudaGraphicsRegisterFlagsTextureGather CUDA will perform texture gather operations on this resource

5.28.3.9 enum cudaLimit

CUDA Limits

Enumerator:

cudaLimitStackSize GPU thread stack size

cudaLimitPrintfFifoSize GPU printf/fprintf FIFO size

cudaLimitMallocHeapSize GPU malloc heap size

cudaLimitDevRuntimeSyncDepth GPU device runtime synchronize depth

cudaLimitDevRuntimePendingLaunchCount GPU device runtime pending launch count

5.28.3.10 enum cudaMemcpyKind

CUDA memory copy types

Enumerator:

cudaMemcpyHostToHost Host -> Host

cudaMemcpyHostToDevice Host -> Device

cudaMemcpyDeviceToHost Device -> Host

cudaMemcpyDeviceToDevice Device -> Device

cudaMemcpyDefault Default based unified virtual address space

5.28.3.11 enum cudaMemoryType

CUDA memory types

Enumerator:

cudaMemoryTypeHost Host memory
cudaMemoryTypeDevice Device memory

5.28.3.12 enum cudaOutputMode

CUDA Profiler Output modes

Enumerator:

cudaKeyValuePair Output mode Key-Value pair format.
cudaCSV Output mode Comma separated values format.

5.28.3.13 enum cudaResourceType

CUDA resource types

Enumerator:

cudaResourceTypeArray Array resource
cudaResourceTypeMipmappedArray Mipmapped array resource
cudaResourceTypeLinear Linear resource
cudaResourceTypePitch2D Pitch 2D resource

5.28.3.14 enum cudaResourceViewFormat

CUDA texture resource view formats

Enumerator:

cudaResViewFormatNone No resource view format (use underlying resource format)
cudaResViewFormatUnsignedChar1 1 channel unsigned 8-bit integers
cudaResViewFormatUnsignedChar2 2 channel unsigned 8-bit integers
cudaResViewFormatUnsignedChar4 4 channel unsigned 8-bit integers
cudaResViewFormatSignedChar1 1 channel signed 8-bit integers
cudaResViewFormatSignedChar2 2 channel signed 8-bit integers
cudaResViewFormatSignedChar4 4 channel signed 8-bit integers
cudaResViewFormatUnsignedShort1 1 channel unsigned 16-bit integers
cudaResViewFormatUnsignedShort2 2 channel unsigned 16-bit integers
cudaResViewFormatUnsignedShort4 4 channel unsigned 16-bit integers
cudaResViewFormatSignedShort1 1 channel signed 16-bit integers
cudaResViewFormatSignedShort2 2 channel signed 16-bit integers

cudaResViewFormatSignedShort4 4 channel signed 16-bit integers
cudaResViewFormatUnsignedInt1 1 channel unsigned 32-bit integers
cudaResViewFormatUnsignedInt2 2 channel unsigned 32-bit integers
cudaResViewFormatUnsignedInt4 4 channel unsigned 32-bit integers
cudaResViewFormatSignedInt1 1 channel signed 32-bit integers
cudaResViewFormatSignedInt2 2 channel signed 32-bit integers
cudaResViewFormatSignedInt4 4 channel signed 32-bit integers
cudaResViewFormatHalf1 1 channel 16-bit floating point
cudaResViewFormatHalf2 2 channel 16-bit floating point
cudaResViewFormatHalf4 4 channel 16-bit floating point
cudaResViewFormatFloat1 1 channel 32-bit floating point
cudaResViewFormatFloat2 2 channel 32-bit floating point
cudaResViewFormatFloat4 4 channel 32-bit floating point
cudaResViewFormatUnsignedBlockCompressed1 Block compressed 1
cudaResViewFormatUnsignedBlockCompressed2 Block compressed 2
cudaResViewFormatUnsignedBlockCompressed3 Block compressed 3
cudaResViewFormatUnsignedBlockCompressed4 Block compressed 4 unsigned
cudaResViewFormatSignedBlockCompressed4 Block compressed 4 signed
cudaResViewFormatUnsignedBlockCompressed5 Block compressed 5 unsigned
cudaResViewFormatSignedBlockCompressed5 Block compressed 5 signed
cudaResViewFormatUnsignedBlockCompressed6H Block compressed 6 unsigned half-float
cudaResViewFormatSignedBlockCompressed6H Block compressed 6 signed half-float
cudaResViewFormatUnsignedBlockCompressed7 Block compressed 7

5.28.3.15 enum cudaSharedMemConfig

CUDA shared memory configuration

5.28.3.16 enum cudaSurfaceBoundaryMode

CUDA Surface boundary modes

Enumerator:

cudaBoundaryModeZero Zero boundary mode
cudaBoundaryModeClamp Clamp boundary mode
cudaBoundaryModeTrap Trap boundary mode

5.28.3.17 enum cudaSurfaceFormatMode

CUDA Surface format modes

Enumerator:

cudaFormatModeForced Forced format mode
cudaFormatModeAuto Auto format mode

5.28.3.18 enum cudaTextureAddressMode

CUDA texture address modes

Enumerator:

- cudaAddressModeWrap* Wrapping address mode
- cudaAddressModeClamp* Clamp to edge address mode
- cudaAddressModeMirror* Mirror address mode
- cudaAddressModeBorder* Border address mode

5.28.3.19 enum cudaTextureFilterMode

CUDA texture filter modes

Enumerator:

- cudaFilterModePoint* Point filter mode
- cudaFilterModeLinear* Linear filter mode

5.28.3.20 enum cudaTextureReadMode

CUDA texture read modes

Enumerator:

- cudaReadModeElementType* Read texture as specified element type
- cudaReadModeNormalizedFloat* Read texture as normalized float

5.29 Data types used by CUDA driver

Data Structures

- struct [CUDA_ARRAY3D_DESCRIPTOR_st](#)
- struct [CUDA_ARRAY_DESCRIPTOR_st](#)
- struct [CUDA_MEMCPY2D_st](#)
- struct [CUDA_MEMCPY3D_PEER_st](#)
- struct [CUDA_MEMCPY3D_st](#)
- struct [CUDA_POINTER_ATTRIBUTE_P2P_TOKENS_st](#)
- struct [CUDA_RESOURCE_DESC_st](#)
- struct [CUDA_RESOURCE_VIEW_DESC_st](#)
- struct [CUDA_TEXTURE_DESC_st](#)
- struct [CUdevprop_st](#)
- struct [CUipcEventHandle_st](#)
- struct [CUipcMemHandle_st](#)

Defines

- `#define CU_IPC_HANDLE_SIZE 64`
- `#define CU_LAUNCH_PARAM_BUFFER_POINTER ((void*)0x01)`
- `#define CU_LAUNCH_PARAM_BUFFER_SIZE ((void*)0x02)`
- `#define CU_LAUNCH_PARAM_END ((void*)0x00)`
- `#define CU_MEMHOSTALLOC_DEVICEMAP 0x02`
- `#define CU_MEMHOSTALLOC_PORTABLE 0x01`
- `#define CU_MEMHOSTALLOC_WRITECOMBINED 0x04`
- `#define CU_MEMHOSTREGISTER_DEVICEMAP 0x02`
- `#define CU_MEMHOSTREGISTER_PORTABLE 0x01`
- `#define CU_PARAM_TR_DEFAULT -1`
- `#define CU_TRSA_OVERRIDE_FORMAT 0x01`
- `#define CU_TRSF_NORMALIZED_COORDINATES 0x02`
- `#define CU_TRSF_READ_AS_INTEGER 0x01`
- `#define CU_TRSF_SRGB 0x10`
- `#define CUDA_ARRAY3D_2DARRAY 0x01`
- `#define CUDA_ARRAY3D_CUBEMAP 0x04`
- `#define CUDA_ARRAY3D_LAYERED 0x01`
- `#define CUDA_ARRAY3D_SURFACE_LDST 0x02`
- `#define CUDA_ARRAY3D_TEXTURE_GATHER 0x08`
- `#define CUDA_VERSION 5000`

Typedefs

- typedef enum [CUaddress_mode_enum](#) [CUaddress_mode](#)
- typedef struct [CUarray_st](#) * [CUarray](#)
- typedef enum [CUarray_cubemap_face_enum](#) [CUarray_cubemap_face](#)
- typedef enum [CUarray_format_enum](#) [CUarray_format](#)
- typedef enum [CUcomputemode_enum](#) [CUcomputemode](#)
- typedef struct [CUctx_st](#) * [CUcontext](#)
- typedef enum [CUctx_flags_enum](#) [CUctx_flags](#)

- typedef struct [CUDA_ARRAY3D_DESCRIPTOR_st](#) [CUDA_ARRAY3D_DESCRIPTOR](#)
- typedef struct [CUDA_ARRAY_DESCRIPTOR_st](#) [CUDA_ARRAY_DESCRIPTOR](#)
- typedef struct [CUDA_MEMCPY2D_st](#) [CUDA_MEMCPY2D](#)
- typedef struct [CUDA_MEMCPY3D_st](#) [CUDA_MEMCPY3D](#)
- typedef struct [CUDA_MEMCPY3D_PEER_st](#) [CUDA_MEMCPY3D_PEER](#)
- typedef struct [CUDA_POINTER_ATTRIBUTE_P2P_TOKENS_st](#) [CUDA_POINTER_ATTRIBUTE_P2P_TOKENS](#)
- typedef struct [CUDA_RESOURCE_DESC_st](#) [CUDA_RESOURCE_DESC](#)
- typedef struct [CUDA_RESOURCE_VIEW_DESC_st](#) [CUDA_RESOURCE_VIEW_DESC](#)
- typedef struct [CUDA_TEXTURE_DESC_st](#) [CUDA_TEXTURE_DESC](#)
- typedef int [CUdevice](#)
- typedef enum [CUdevice_attribute_enum](#) [CUdevice_attribute](#)
- typedef unsigned int [CUdeviceptr](#)
- typedef struct [CUdevprop_st](#) [CUdevprop](#)
- typedef struct [CUevent_st](#) * [CUevent](#)
- typedef enum [CUevent_flags_enum](#) [CUevent_flags](#)
- typedef enum [CUfilter_mode_enum](#) [CUfilter_mode](#)
- typedef enum [CUfunc_cache_enum](#) [CUfunc_cache](#)
- typedef struct [CUfunc_st](#) * [CUfunction](#)
- typedef enum [CUfunction_attribute_enum](#) [CUfunction_attribute](#)
- typedef enum [CUgraphicsMapResourceFlags_enum](#) [CUgraphicsMapResourceFlags](#)
- typedef enum [CUgraphicsRegisterFlags_enum](#) [CUgraphicsRegisterFlags](#)
- typedef struct [CUgraphicsResource_st](#) * [CUgraphicsResource](#)
- typedef struct [CUipcEventHandle_st](#) [CUipcEventHandle](#)
- typedef enum [CUipcMem_flags_enum](#) [CUipcMem_flags](#)
- typedef struct [CUipcMemHandle_st](#) [CUipcMemHandle](#)
- typedef enum [CUjit_fallback_enum](#) [CUjit_fallback](#)
- typedef enum [CUjit_option_enum](#) [CUjit_option](#)
- typedef enum [CUjit_target_enum](#) [CUjit_target](#)
- typedef enum [CUlimit_enum](#) [CUlimit](#)
- typedef enum [CUmemorytype_enum](#) [CUmemorytype](#)
- typedef struct [CUmipmappedArray_st](#) * [CUmipmappedArray](#)
- typedef struct [CUmod_st](#) * [CUmodule](#)
- typedef enum [CUpointer_attribute_enum](#) [CUpointer_attribute](#)
- typedef enum [CUresourcetype_enum](#) [CUresourcetype](#)
- typedef enum [CUresourceViewFormat_enum](#) [CUresourceViewFormat](#)
- typedef enum [cudaError_enum](#) [CUresult](#)
- typedef enum [CUsharedconfig_enum](#) [CUsharedconfig](#)
- typedef struct [CUstream_st](#) * [CUstream](#)
- typedef enum [CUstream_flags_enum](#) [CUstream_flags](#)
- typedef void([CUDA_CB](#) * [CUstreamCallback](#))([CUstream](#) hStream, [CUresult](#) status, void *[userData](#))
- typedef unsigned long [CUSurfObject](#)
- typedef struct [CUSurfref_st](#) * [CUSurfref](#)
- typedef unsigned long long [CUTexObject](#)
- typedef struct [CUTexref_st](#) * [CUTexref](#)

Enumerations

- enum CUaddress_mode_enum {
CU_TR_ADDRESS_MODE_WRAP = 0,
CU_TR_ADDRESS_MODE_CLAMP = 1,
CU_TR_ADDRESS_MODE_MIRROR = 2,
CU_TR_ADDRESS_MODE_BORDER = 3 }
- enum CUarray_cubemap_face_enum {
CU_CUBEMAP_FACE_POSITIVE_X = 0x00,
CU_CUBEMAP_FACE_NEGATIVE_X = 0x01,
CU_CUBEMAP_FACE_POSITIVE_Y = 0x02,
CU_CUBEMAP_FACE_NEGATIVE_Y = 0x03,
CU_CUBEMAP_FACE_POSITIVE_Z = 0x04,
CU_CUBEMAP_FACE_NEGATIVE_Z = 0x05 }
- enum CUarray_format_enum {
CU_AD_FORMAT_UNSIGNED_INT8 = 0x01,
CU_AD_FORMAT_UNSIGNED_INT16 = 0x02,
CU_AD_FORMAT_UNSIGNED_INT32 = 0x03,
CU_AD_FORMAT_SIGNED_INT8 = 0x08,
CU_AD_FORMAT_SIGNED_INT16 = 0x09,
CU_AD_FORMAT_SIGNED_INT32 = 0x0a,
CU_AD_FORMAT_HALF = 0x10,
CU_AD_FORMAT_FLOAT = 0x20 }
- enum CUcomputemode_enum {
CU_COMPUTEMODE_DEFAULT = 0,
CU_COMPUTEMODE_EXCLUSIVE = 1,
CU_COMPUTEMODE_PROHIBITED = 2,
CU_COMPUTEMODE_EXCLUSIVE_PROCESS = 3 }
- enum CUctx_flags_enum {
CU_CTX_SCHED_AUTO = 0x00,
CU_CTX_SCHED_SPIN = 0x01,
CU_CTX_SCHED_YIELD = 0x02,
CU_CTX_SCHED_BLOCKING_SYNC = 0x04,
CU_CTX_BLOCKING_SYNC = 0x04 ,
CU_CTX_MAP_HOST = 0x08,
CU_CTX_LMEM_RESIZE_TO_MAX = 0x10 }
- enum cudaError_enum {
CUDA_SUCCESS = 0,
CUDA_ERROR_INVALID_VALUE = 1,
CUDA_ERROR_OUT_OF_MEMORY = 2,
CUDA_ERROR_NOT_INITIALIZED = 3,
CUDA_ERROR_DEINITIALIZED = 4,

CUDA_ERROR_PROFILER_DISABLED = 5,
CUDA_ERROR_PROFILER_NOT_INITIALIZED = 6,
CUDA_ERROR_PROFILER_ALREADY_STARTED = 7,
CUDA_ERROR_PROFILER_ALREADY_STOPPED = 8,
CUDA_ERROR_NO_DEVICE = 100,
CUDA_ERROR_INVALID_DEVICE = 101,
CUDA_ERROR_INVALID_IMAGE = 200,
CUDA_ERROR_INVALID_CONTEXT = 201,
CUDA_ERROR_CONTEXT_ALREADY_CURRENT = 202,
CUDA_ERROR_MAP_FAILED = 205,
CUDA_ERROR_UNMAP_FAILED = 206,
CUDA_ERROR_ARRAY_IS_MAPPED = 207,
CUDA_ERROR_ALREADY_MAPPED = 208,
CUDA_ERROR_NO_BINARY_FOR_GPU = 209,
CUDA_ERROR_ALREADY_ACQUIRED = 210,
CUDA_ERROR_NOT_MAPPED = 211,
CUDA_ERROR_NOT_MAPPED_AS_ARRAY = 212,
CUDA_ERROR_NOT_MAPPED_AS_POINTER = 213,
CUDA_ERROR_ECC_UNCORRECTABLE = 214,
CUDA_ERROR_UNSUPPORTED_LIMIT = 215,
CUDA_ERROR_CONTEXT_ALREADY_IN_USE = 216,
CUDA_ERROR_PEER_ACCESS_UNSUPPORTED = 217,
CUDA_ERROR_INVALID_SOURCE = 300,
CUDA_ERROR_FILE_NOT_FOUND = 301,
CUDA_ERROR_SHARED_OBJECT_SYMBOL_NOT_FOUND = 302,
CUDA_ERROR_SHARED_OBJECT_INIT_FAILED = 303,
CUDA_ERROR_OPERATING_SYSTEM = 304,
CUDA_ERROR_INVALID_HANDLE = 400,
CUDA_ERROR_NOT_FOUND = 500,
CUDA_ERROR_NOT_READY = 600,
CUDA_ERROR_LAUNCH_FAILED = 700,
CUDA_ERROR_LAUNCH_OUT_OF_RESOURCES = 701,
CUDA_ERROR_LAUNCH_TIMEOUT = 702,
CUDA_ERROR_LAUNCH_INCOMPATIBLE_TEXTURING = 703,
CUDA_ERROR_PEER_ACCESS_ALREADY_ENABLED = 704,
CUDA_ERROR_PEER_ACCESS_NOT_ENABLED = 705,
CUDA_ERROR_PRIMARY_CONTEXT_ACTIVE = 708,
CUDA_ERROR_CONTEXT_IS_DESTROYED = 709,
CUDA_ERROR_ASSERT = 710,
CUDA_ERROR_TOO_MANY_PEERS = 711,
CUDA_ERROR_HOST_MEMORY_ALREADY_REGISTERED = 712,

```
CUDA_ERROR_HOST_MEMORY_NOT_REGISTERED = 713,  
CUDA_ERROR_NOT_PERMITTED = 800,  
CUDA_ERROR_NOT_SUPPORTED = 801,  
CUDA_ERROR_UNKNOWN = 999 }  
• enum CUdevice_attribute_enum {  
  CU_DEVICE_ATTRIBUTE_MAX_THREADS_PER_BLOCK = 1,  
  CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_X = 2,  
  CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Y = 3,  
  CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Z = 4,  
  CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_X = 5,  
  CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Y = 6,  
  CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Z = 7,  
  CU_DEVICE_ATTRIBUTE_MAX_SHARED_MEMORY_PER_BLOCK = 8,  
  CU_DEVICE_ATTRIBUTE_SHARED_MEMORY_PER_BLOCK = 8,  
  CU_DEVICE_ATTRIBUTE_TOTAL_CONSTANT_MEMORY = 9,  
  CU_DEVICE_ATTRIBUTE_WARP_SIZE = 10,  
  CU_DEVICE_ATTRIBUTE_MAX_PITCH = 11,  
  CU_DEVICE_ATTRIBUTE_MAX_REGISTERS_PER_BLOCK = 12,  
  CU_DEVICE_ATTRIBUTE_REGISTERS_PER_BLOCK = 12,  
  CU_DEVICE_ATTRIBUTE_CLOCK_RATE = 13,  
  CU_DEVICE_ATTRIBUTE_TEXTURE_ALIGNMENT = 14,  
  CU_DEVICE_ATTRIBUTE_GPU_OVERLAP = 15,  
  CU_DEVICE_ATTRIBUTE_MULTIPROCESSOR_COUNT = 16,  
  CU_DEVICE_ATTRIBUTE_KERNEL_EXEC_TIMEOUT = 17,  
  CU_DEVICE_ATTRIBUTE_INTEGRATED = 18,  
  CU_DEVICE_ATTRIBUTE_CAN_MAP_HOST_MEMORY = 19,  
  CU_DEVICE_ATTRIBUTE_COMPUTE_MODE = 20,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_WIDTH = 21,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_WIDTH = 22,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_HEIGHT = 23,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_WIDTH = 24,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_HEIGHT = 25,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_DEPTH = 26,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_WIDTH = 27,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_HEIGHT = 28,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_LAYERS = 29,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_ARRAY_WIDTH = 27,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_ARRAY_HEIGHT = 28,  
  CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_ARRAY_NUMSLICES = 29,  
  CU_DEVICE_ATTRIBUTE_SURFACE_ALIGNMENT = 30,  
  CU_DEVICE_ATTRIBUTE_CONCURRENT_KERNELS = 31,
```

CU_DEVICE_ATTRIBUTE_ECC_ENABLED = 32,
CU_DEVICE_ATTRIBUTE_PCI_BUS_ID = 33,
CU_DEVICE_ATTRIBUTE_PCI_DEVICE_ID = 34,
CU_DEVICE_ATTRIBUTE_TCC_DRIVER = 35,
CU_DEVICE_ATTRIBUTE_MEMORY_CLOCK_RATE = 36,
CU_DEVICE_ATTRIBUTE_GLOBAL_MEMORY_BUS_WIDTH = 37,
CU_DEVICE_ATTRIBUTE_L2_CACHE_SIZE = 38,
CU_DEVICE_ATTRIBUTE_MAX_THREADS_PER_MULTIPROCESSOR = 39,
CU_DEVICE_ATTRIBUTE_ASYNC_ENGINE_COUNT = 40,
CU_DEVICE_ATTRIBUTE_UNIFIED_ADDRESSING = 41,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LAYERED_WIDTH = 42,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LAYERED_LAYERS = 43,
CU_DEVICE_ATTRIBUTE_CAN_TEX2D_GATHER = 44,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_GATHER_WIDTH = 45,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_GATHER_HEIGHT = 46,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_WIDTH_ALTERNATE = 47,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_HEIGHT_ALTERNATE = 48,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_DEPTH_ALTERNATE = 49,
CU_DEVICE_ATTRIBUTE_PCI_DOMAIN_ID = 50,
CU_DEVICE_ATTRIBUTE_TEXTURE_PITCH_ALIGNMENT = 51,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_WIDTH = 52,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_LAYERED_WIDTH = 53,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_LAYERED_LAYERS = 54,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_WIDTH = 55,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_WIDTH = 56,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_HEIGHT = 57,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_WIDTH = 58,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_HEIGHT = 59,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_DEPTH = 60,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_LAYERED_WIDTH = 61,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_LAYERED_LAYERS = 62,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_WIDTH = 63,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_HEIGHT = 64,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_LAYERS = 65,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_WIDTH = 66,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_LAYERED_WIDTH = 67,
CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_LAYERED_LAYERS = 68,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LINEAR_WIDTH = 69,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_WIDTH = 70,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_HEIGHT = 71,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_PITCH = 72,

```

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_MIPMAPPED_WIDTH = 73,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_MIPMAPPED_HEIGHT = 74,
CU_DEVICE_ATTRIBUTE_COMPUTE_CAPABILITY_MAJOR = 75,
CU_DEVICE_ATTRIBUTE_COMPUTE_CAPABILITY_MINOR = 76,
CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_MIPMAPPED_WIDTH = 77 }
• enum CUevent_flags_enum {
    CU_EVENT_DEFAULT = 0x0,
    CU_EVENT_BLOCKING_SYNC = 0x1,
    CU_EVENT_DISABLE_TIMING = 0x2,
    CU_EVENT_INTERPROCESS = 0x4 }
• enum CUfilter_mode_enum {
    CU_TR_FILTER_MODE_POINT = 0,
    CU_TR_FILTER_MODE_LINEAR = 1 }
• enum CUfunc_cache_enum {
    CU_FUNC_CACHE_PREFER_NONE = 0x00,
    CU_FUNC_CACHE_PREFER_SHARED = 0x01,
    CU_FUNC_CACHE_PREFER_L1 = 0x02,
    CU_FUNC_CACHE_PREFER_EQUAL = 0x03 }
• enum CUfunction_attribute_enum {
    CU_FUNC_ATTRIBUTE_MAX_THREADS_PER_BLOCK = 0,
    CU_FUNC_ATTRIBUTE_SHARED_SIZE_BYTES = 1,
    CU_FUNC_ATTRIBUTE_CONST_SIZE_BYTES = 2,
    CU_FUNC_ATTRIBUTE_LOCAL_SIZE_BYTES = 3,
    CU_FUNC_ATTRIBUTE_NUM_REGS = 4,
    CU_FUNC_ATTRIBUTE_PTX_VERSION = 5,
    CU_FUNC_ATTRIBUTE_BINARY_VERSION = 6 }
• enum CUgraphicsMapResourceFlags_enum
• enum CUgraphicsRegisterFlags_enum
• enum CUipcMem_flags_enum { CU_IPC_MEM_LAZY_ENABLE_PEER_ACCESS = 0x1 }
• enum CUjit_fallback_enum {
    CU_PREFER_PTX = 0,
    CU_PREFER_BINARY }
• enum CUjit_option_enum {
    CU_JIT_MAX_REGISTERS = 0,
    CU_JIT_THREADS_PER_BLOCK,
    CU_JIT_WALL_TIME,
    CU_JIT_INFO_LOG_BUFFER,
    CU_JIT_INFO_LOG_BUFFER_SIZE_BYTES,
    CU_JIT_ERROR_LOG_BUFFER,
    CU_JIT_ERROR_LOG_BUFFER_SIZE_BYTES,
    CU_JIT_OPTIMIZATION_LEVEL,
    CU_JIT_TARGET_FROM_CUCONTEXT,
    CU_JIT_TARGET,
    CU_JIT_FALLBACK_STRATEGY }

```

- enum CUjit_target_enum {
CU_TARGET_COMPUTE_10 = 0,
CU_TARGET_COMPUTE_11,
CU_TARGET_COMPUTE_12,
CU_TARGET_COMPUTE_13,
CU_TARGET_COMPUTE_20,
CU_TARGET_COMPUTE_21,
CU_TARGET_COMPUTE_30,
CU_TARGET_COMPUTE_35 }
- enum CUlimit_enum {
CU_LIMIT_STACK_SIZE = 0x00,
CU_LIMIT_PRINTF_FIFO_SIZE = 0x01,
CU_LIMIT_MALLOC_HEAP_SIZE = 0x02,
CU_LIMIT_DEV_RUNTIME_SYNC_DEPTH = 0x03,
CU_LIMIT_DEV_RUNTIME_PENDING_LAUNCH_COUNT = 0x04 }
- enum CUmemorytype_enum {
CU_MEMORYTYPE_HOST = 0x01,
CU_MEMORYTYPE_DEVICE = 0x02,
CU_MEMORYTYPE_ARRAY = 0x03,
CU_MEMORYTYPE_UNIFIED = 0x04 }
- enum CUpointer_attribute_enum {
CU_POINTER_ATTRIBUTE_CONTEXT = 1,
CU_POINTER_ATTRIBUTE_MEMORY_TYPE = 2,
CU_POINTER_ATTRIBUTE_DEVICE_POINTER = 3,
CU_POINTER_ATTRIBUTE_HOST_POINTER = 4,
CU_POINTER_ATTRIBUTE_P2P_TOKENS = 5 }
- enum CUresourcetype_enum {
CU_RESOURCE_TYPE_ARRAY = 0x00,
CU_RESOURCE_TYPE_MIPMAPPED_ARRAY = 0x01,
CU_RESOURCE_TYPE_LINEAR = 0x02,
CU_RESOURCE_TYPE_PITCH2D = 0x03 }
- enum CUresourceViewFormat_enum {
CU_RES_VIEW_FORMAT_NONE = 0x00,
CU_RES_VIEW_FORMAT_UINT_1X8 = 0x01,
CU_RES_VIEW_FORMAT_UINT_2X8 = 0x02,
CU_RES_VIEW_FORMAT_UINT_4X8 = 0x03,
CU_RES_VIEW_FORMAT_SINT_1X8 = 0x04,
CU_RES_VIEW_FORMAT_SINT_2X8 = 0x05,
CU_RES_VIEW_FORMAT_SINT_4X8 = 0x06,
CU_RES_VIEW_FORMAT_UINT_1X16 = 0x07,
CU_RES_VIEW_FORMAT_UINT_2X16 = 0x08,
CU_RES_VIEW_FORMAT_UINT_4X16 = 0x09,

```

CU_RES_VIEW_FORMAT_SINT_1X16 = 0x0a,
CU_RES_VIEW_FORMAT_SINT_2X16 = 0x0b,
CU_RES_VIEW_FORMAT_SINT_4X16 = 0x0c,
CU_RES_VIEW_FORMAT_UINT_1X32 = 0x0d,
CU_RES_VIEW_FORMAT_UINT_2X32 = 0x0e,
CU_RES_VIEW_FORMAT_UINT_4X32 = 0x0f,
CU_RES_VIEW_FORMAT_SINT_1X32 = 0x10,
CU_RES_VIEW_FORMAT_SINT_2X32 = 0x11,
CU_RES_VIEW_FORMAT_SINT_4X32 = 0x12,
CU_RES_VIEW_FORMAT_FLOAT_1X16 = 0x13,
CU_RES_VIEW_FORMAT_FLOAT_2X16 = 0x14,
CU_RES_VIEW_FORMAT_FLOAT_4X16 = 0x15,
CU_RES_VIEW_FORMAT_FLOAT_1X32 = 0x16,
CU_RES_VIEW_FORMAT_FLOAT_2X32 = 0x17,
CU_RES_VIEW_FORMAT_FLOAT_4X32 = 0x18,
CU_RES_VIEW_FORMAT_UNSIGNED_BC1 = 0x19,
CU_RES_VIEW_FORMAT_UNSIGNED_BC2 = 0x1a,
CU_RES_VIEW_FORMAT_UNSIGNED_BC3 = 0x1b,
CU_RES_VIEW_FORMAT_UNSIGNED_BC4 = 0x1c,
CU_RES_VIEW_FORMAT_SIGNED_BC4 = 0x1d,
CU_RES_VIEW_FORMAT_UNSIGNED_BC5 = 0x1e,
CU_RES_VIEW_FORMAT_SIGNED_BC5 = 0x1f,
CU_RES_VIEW_FORMAT_UNSIGNED_BC6H = 0x20,
CU_RES_VIEW_FORMAT_SIGNED_BC6H = 0x21,
CU_RES_VIEW_FORMAT_UNSIGNED_BC7 = 0x22 }
• enum CUsharedconfig_enum {
    CU_SHARED_MEM_CONFIG_DEFAULT_BANK_SIZE = 0x00,
    CU_SHARED_MEM_CONFIG_FOUR_BYTE_BANK_SIZE = 0x01,
    CU_SHARED_MEM_CONFIG_EIGHT_BYTE_BANK_SIZE = 0x02 }
• enum CUstream_flags_enum {
    CU_STREAM_DEFAULT = 0x0,
    CU_STREAM_NON_BLOCKING = 0x1 }

```

5.29.1 Define Documentation

5.29.1.1 #define CU_IPC_HANDLE_SIZE 64

CUDA IPC handle size

5.29.1.2 #define CU_LAUNCH_PARAM_BUFFER_POINTER ((void*)0x01)

Indicator that the next value in the `extra` parameter to `cuLaunchKernel` will be a pointer to a buffer containing all kernel parameters used for launching kernel `f`. This buffer needs to honor all alignment/padding requirements of the individual parameters. If `CU_LAUNCH_PARAM_BUFFER_SIZE` is not also specified in the `extra` array, then `CU_LAUNCH_PARAM_BUFFER_POINTER` will have no effect.

5.29.1.3 #define CU_LAUNCH_PARAM_BUFFER_SIZE ((void*)0x02)

Indicator that the next value in the `extra` parameter to `cuLaunchKernel` will be a pointer to a `size_t` which contains the size of the buffer specified with `CU_LAUNCH_PARAM_BUFFER_POINTER`. It is required that `CU_LAUNCH_PARAM_BUFFER_POINTER` also be specified in the `extra` array if the value associated with `CU_LAUNCH_PARAM_BUFFER_SIZE` is not zero.

5.29.1.4 #define CU_LAUNCH_PARAM_END ((void*)0x00)

End of array terminator for the `extra` parameter to `cuLaunchKernel`

5.29.1.5 #define CU_MEMHOSTALLOC_DEVICEMAP 0x02

If set, host memory is mapped into CUDA address space and `cuMemHostGetDevicePointer()` may be called on the host pointer. Flag for `cuMemHostAlloc()`

5.29.1.6 #define CU_MEMHOSTALLOC_PORTABLE 0x01

If set, host memory is portable between CUDA contexts. Flag for `cuMemHostAlloc()`

5.29.1.7 #define CU_MEMHOSTALLOC_WRITECOMBINED 0x04

If set, host memory is allocated as write-combined - fast to write, faster to DMA, slow to read except via SSE4 streaming load instruction (MOVNTDQA). Flag for `cuMemHostAlloc()`

5.29.1.8 #define CU_MEMHOSTREGISTER_DEVICEMAP 0x02

If set, host memory is mapped into CUDA address space and `cuMemHostGetDevicePointer()` may be called on the host pointer. Flag for `cuMemHostRegister()`

5.29.1.9 #define CU_MEMHOSTREGISTER_PORTABLE 0x01

If set, host memory is portable between CUDA contexts. Flag for `cuMemHostRegister()`

5.29.1.10 #define CU_PARAM_TR_DEFAULT -1

For texture references loaded into the module, use default texunit from texture reference.

5.29.1.11 #define CU_TRSA_OVERRIDE_FORMAT 0x01

Override the texref format with a format inferred from the array. Flag for `cuTexRefSetArray()`

5.29.1.12 #define CU_TRSF_NORMALIZED_COORDINATES 0x02

Use normalized texture coordinates in the range [0,1) instead of [0,dim). Flag for `cuTexRefSetFlags()`

5.29.1.13 #define CU_TRSF_READ_AS_INTEGER 0x01

Read the texture as integers rather than promoting the values to floats in the range [0,1]. Flag for [cuTexRefSetFlags\(\)](#)

5.29.1.14 #define CU_TRSF_SRGB 0x10

Perform sRGB->linear conversion during texture read. Flag for [cuTexRefSetFlags\(\)](#)

5.29.1.15 #define CUDA_ARRAY3D_2DARRAY 0x01

Deprecated, use [CUDA_ARRAY3D_LAYERED](#)

5.29.1.16 #define CUDA_ARRAY3D_CUBEMAP 0x04

If set, the CUDA array is a collection of six 2D arrays, representing faces of a cube. The width of such a CUDA array must be equal to its height, and Depth must be six. If [CUDA_ARRAY3D_LAYERED](#) flag is also set, then the CUDA array is a collection of cubemaps and Depth must be a multiple of six.

5.29.1.17 #define CUDA_ARRAY3D_LAYERED 0x01

If set, the CUDA array is a collection of layers, where each layer is either a 1D or a 2D array and the Depth member of [CUDA_ARRAY3D_DESCRIPTOR](#) specifies the number of layers, not the depth of a 3D array.

5.29.1.18 #define CUDA_ARRAY3D_SURFACE_LDST 0x02

This flag must be set in order to bind a surface reference to the CUDA array

5.29.1.19 #define CUDA_ARRAY3D_TEXTURE_GATHER 0x08

This flag must be set in order to perform texture gather operations on a CUDA array.

5.29.1.20 #define CUDA_VERSION 5000

CUDA API version number

5.29.2 Typedef Documentation**5.29.2.1 typedef enum CUaddress_mode_enum CUaddress_mode**

Texture reference addressing modes

5.29.2.2 typedef struct CUarray_st* CUarray

CUDA array

5.29.2.3 typedef enum CUarray_cubemap_face_enum CUarray_cubemap_face

Array indices for cube faces

5.29.2.4 typedef enum CUarray_format_enum CUarray_format

Array formats

5.29.2.5 typedef enum CUcomputemode_enum CUcomputemode

Compute Modes

5.29.2.6 typedef struct CUctx_st* CUcontext

CUDA context

5.29.2.7 typedef enum CUctx_flags_enum CUctx_flags

Context creation flags

5.29.2.8 typedef struct CUDA_ARRAY3D_DESCRIPTOR_st CUDA_ARRAY3D_DESCRIPTOR

3D array descriptor

5.29.2.9 typedef struct CUDA_ARRAY_DESCRIPTOR_st CUDA_ARRAY_DESCRIPTOR

Array descriptor

5.29.2.10 typedef struct CUDA_MEMCPY2D_st CUDA_MEMCPY2D

2D memory copy parameters

5.29.2.11 typedef struct CUDA_MEMCPY3D_st CUDA_MEMCPY3D

3D memory copy parameters

5.29.2.12 typedef struct CUDA_MEMCPY3D_PEER_st CUDA_MEMCPY3D_PEER

3D memory cross-context copy parameters

5.29.2.13 typedef struct CUDA_POINTER_ATTRIBUTE_P2P_TOKENS_st CUDA_POINTER_ATTRIBUTE_P2P_TOKENS

GPU Direct v3 tokens

5.29.2.14 typedef struct CUDA_RESOURCE_DESC_st CUDA_RESOURCE_DESC

CUDA Resource descriptor

5.29.2.15 typedef struct CUDA_RESOURCE_VIEW_DESC_st CUDA_RESOURCE_VIEW_DESC

Resource view descriptor

5.29.2.16 typedef struct CUDA_TEXTURE_DESC_st CUDA_TEXTURE_DESC

Texture descriptor

5.29.2.17 typedef int CUdevice

CUDA device

5.29.2.18 typedef enum CUdevice_attribute_enum CUdevice_attribute

Device properties

5.29.2.19 typedef unsigned int CUdeviceptr

CUDA device pointer

5.29.2.20 typedef struct CUdevprop_st CUdevprop

Legacy device properties

5.29.2.21 typedef struct CUevent_st* CUevent

CUDA event

5.29.2.22 typedef enum CUevent_flags_enum CUevent_flags

Event creation flags

5.29.2.23 typedef enum CUfilter_mode_enum CUfilter_mode

Texture reference filtering modes

5.29.2.24 typedef enum CUfunc_cache_enum CUfunc_cache

Function cache configurations

5.29.2.25 typedef struct CUfunc_st* CUfunction

CUDA function

5.29.2.26 typedef enum CUfunction_attribute_enum CUfunction_attribute

Function properties

5.29.2.27 typedef enum CUgraphicsMapResourceFlags_enum CUgraphicsMapResourceFlags

Flags for mapping and unmapping interop resources

5.29.2.28 typedef enum CUgraphicsRegisterFlags_enum CUgraphicsRegisterFlags

Flags to register a graphics resource

5.29.2.29 typedef struct CUgraphicsResource_st* CUgraphicsResource

CUDA graphics interop resource

5.29.2.30 typedef struct CUipcEventHandle_st CUipcEventHandle

CUDA IPC event handle

5.29.2.31 typedef enum CUipcMem_flags_enum CUipcMem_flags

CUDA Ipc Mem Flags

5.29.2.32 typedef struct CUipcMemHandle_st CUipcMemHandle

CUDA IPC mem handle

5.29.2.33 typedef enum CUjit_fallback_enum CUjit_fallback

Cubin matching fallback strategies

5.29.2.34 typedef enum CUjit_option_enum CUjit_option

Online compiler options

5.29.2.35 typedef enum CUjit_target_enum CUjit_target

Online compilation targets

5.29.2.36 typedef enum CUlimit_enum CUlimit

Limits

5.29.2.37 typedef enum CUmemorytype_enum CUmemorytype

Memory types

5.29.2.38 typedef struct CUmipmappedArray_st* CUmipmappedArray

CUDA mipmapped array

5.29.2.39 typedef struct CUmod_st* CUmodule

CUDA module

5.29.2.40 typedef enum CUpointer_attribute_enum CUpointer_attribute

Pointer information

5.29.2.41 typedef enum CUresourcetype_enum CUresourcetype

Resource types

5.29.2.42 typedef enum CUresourceViewFormat_enum CUresourceViewFormat

Resource view format

5.29.2.43 typedef enum cudaError_enum CUresult

Error codes

5.29.2.44 typedef enum CUsharedconfig_enum CUsharedconfig

Shared memory configurations

5.29.2.45 typedef struct CUstream_st* CUstream

CUDA stream

5.29.2.46 typedef enum CUstream_flags_enum CUstream_flags

Stream creation flags

5.29.2.47 typedef void(CUDA_CB * CUstreamCallback)(CUstream hStream, CUresult status, void *userData)

CUDA stream callback

Parameters:

hStream The stream the callback was added to, as passed to [cuStreamAddCallback](#). May be NULL.

status [CUDA_SUCCESS](#) or any persistent error on the stream.

userData User parameter provided at registration.

5.29.2.48 typedef unsigned long long CUsurfObject

CUDA surface object

5.29.2.49 typedef struct CUsurfref_st* CUsurfref

CUDA surface reference

5.29.2.50 typedef unsigned long long CUtexObject

CUDA texture object

5.29.2.51 typedef struct CUtexref_st* CUtexref

CUDA texture reference

5.29.3 Enumeration Type Documentation**5.29.3.1 enum CUaddress_mode_enum**

Texture reference addressing modes

Enumerator:

CU_TR_ADDRESS_MODE_WRAP Wrapping address mode
CU_TR_ADDRESS_MODE_CLAMP Clamp to edge address mode
CU_TR_ADDRESS_MODE_MIRROR Mirror address mode
CU_TR_ADDRESS_MODE_BORDER Border address mode

5.29.3.2 enum CUarray_cubemap_face_enum

Array indices for cube faces

Enumerator:

CU_CUBEMAP_FACE_POSITIVE_X Positive X face of cubemap
CU_CUBEMAP_FACE_NEGATIVE_X Negative X face of cubemap
CU_CUBEMAP_FACE_POSITIVE_Y Positive Y face of cubemap
CU_CUBEMAP_FACE_NEGATIVE_Y Negative Y face of cubemap
CU_CUBEMAP_FACE_POSITIVE_Z Positive Z face of cubemap
CU_CUBEMAP_FACE_NEGATIVE_Z Negative Z face of cubemap

5.29.3.3 enum CUarray_format_enum

Array formats

Enumerator:

CU_AD_FORMAT_UNSIGNED_INT8 Unsigned 8-bit integers
CU_AD_FORMAT_UNSIGNED_INT16 Unsigned 16-bit integers
CU_AD_FORMAT_UNSIGNED_INT32 Unsigned 32-bit integers
CU_AD_FORMAT_SIGNED_INT8 Signed 8-bit integers
CU_AD_FORMAT_SIGNED_INT16 Signed 16-bit integers
CU_AD_FORMAT_SIGNED_INT32 Signed 32-bit integers
CU_AD_FORMAT_HALF 16-bit floating point
CU_AD_FORMAT_FLOAT 32-bit floating point

5.29.3.4 enum CUcomputemode_enum

Compute Modes

Enumerator:

CU_COMPUTEMODE_DEFAULT Default compute mode (Multiple contexts allowed per device)
CU_COMPUTEMODE_EXCLUSIVE Compute-exclusive-thread mode (Only one context used by a single thread can be present on this device at a time)
CU_COMPUTEMODE_PROHIBITED Compute-prohibited mode (No contexts can be created on this device at this time)
CU_COMPUTEMODE_EXCLUSIVE_PROCESS Compute-exclusive-process mode (Only one context used by a single process can be present on this device at a time)

5.29.3.5 enum CUctx_flags_enum

Context creation flags

Enumerator:

CU_CTX_SCHED_AUTO Automatic scheduling
CU_CTX_SCHED_SPIN Set spin as default scheduling
CU_CTX_SCHED_YIELD Set yield as default scheduling
CU_CTX_SCHED_BLOCKING_SYNC Set blocking synchronization as default scheduling
CU_CTX_BLOCKING_SYNC Set blocking synchronization as default scheduling

Deprecated

This flag was deprecated as of CUDA 4.0 and was replaced with [CU_CTX_SCHED_BLOCKING_SYNC](#).

CU_CTX_MAP_HOST Support mapped pinned allocations
CU_CTX_LMEM_RESIZE_TO_MAX Keep local memory allocation after launch

5.29.3.6 enum cudaError_enum

Error codes

Enumerator:

CUDA_SUCCESS The API call returned with no errors. In the case of query calls, this can also mean that the operation being queried is complete (see [cuEventQuery\(\)](#) and [cuStreamQuery\(\)](#)).

CUDA_ERROR_INVALID_VALUE This indicates that one or more of the parameters passed to the API call is not within an acceptable range of values.

CUDA_ERROR_OUT_OF_MEMORY The API call failed because it was unable to allocate enough memory to perform the requested operation.

CUDA_ERROR_NOT_INITIALIZED This indicates that the CUDA driver has not been initialized with [cuInit\(\)](#) or that initialization has failed.

CUDA_ERROR_DEINITIALIZED This indicates that the CUDA driver is in the process of shutting down.

CUDA_ERROR_PROFILER_DISABLED This indicates profiler is not initialized for this run. This can happen when the application is running with external profiling tools like visual profiler.

Deprecated

CUDA_ERROR_PROFILER_NOT_INITIALIZED This error return is deprecated as of CUDA 5.0. It is no longer an error to attempt to enable/disable the profiling via [cuProfilerStart](#) or [cuProfilerStop](#) without initialization.

Deprecated

CUDA_ERROR_PROFILER_ALREADY_STARTED This error return is deprecated as of CUDA 5.0. It is no longer an error to call [cuProfilerStart\(\)](#) when profiling is already enabled.

Deprecated

CUDA_ERROR_PROFILER_ALREADY_STOPPED This error return is deprecated as of CUDA 5.0. It is no longer an error to call [cuProfilerStop\(\)](#) when profiling is already disabled.

CUDA_ERROR_NO_DEVICE This indicates that no CUDA-capable devices were detected by the installed CUDA driver.

CUDA_ERROR_INVALID_DEVICE This indicates that the device ordinal supplied by the user does not correspond to a valid CUDA device.

CUDA_ERROR_INVALID_IMAGE This indicates that the device kernel image is invalid. This can also indicate an invalid CUDA module.

CUDA_ERROR_INVALID_CONTEXT This most frequently indicates that there is no context bound to the current thread. This can also be returned if the context passed to an API call is not a valid handle (such as a context that has had [cuCtxDestroy\(\)](#) invoked on it). This can also be returned if a user mixes different API versions (i.e. 3010 context with 3020 API calls). See [cuCtxGetApiVersion\(\)](#) for more details.

CUDA_ERROR_CONTEXT_ALREADY_CURRENT This indicated that the context being supplied as a parameter to the API call was already the active context.

Deprecated

This error return is deprecated as of CUDA 3.2. It is no longer an error to attempt to push the active context via [cuCtxPushCurrent\(\)](#).

CUDA_ERROR_MAP_FAILED This indicates that a map or register operation has failed.

CUDA_ERROR_UNMAP_FAILED This indicates that an unmap or unregister operation has failed.

CUDA_ERROR_ARRAY_IS_MAPPED This indicates that the specified array is currently mapped and thus cannot be destroyed.

CUDA_ERROR_ALREADY_MAPPED This indicates that the resource is already mapped.

- CUDA_ERROR_NO_BINARY_FOR_GPU*** This indicates that there is no kernel image available that is suitable for the device. This can occur when a user specifies code generation options for a particular CUDA source file that do not include the corresponding device configuration.
- CUDA_ERROR_ALREADY_ACQUIRED*** This indicates that a resource has already been acquired.
- CUDA_ERROR_NOT_MAPPED*** This indicates that a resource is not mapped.
- CUDA_ERROR_NOT_MAPPED_AS_ARRAY*** This indicates that a mapped resource is not available for access as an array.
- CUDA_ERROR_NOT_MAPPED_AS_POINTER*** This indicates that a mapped resource is not available for access as a pointer.
- CUDA_ERROR_ECC_UNCORRECTABLE*** This indicates that an uncorrectable ECC error was detected during execution.
- CUDA_ERROR_UNSUPPORTED_LIMIT*** This indicates that the [CUlimit](#) passed to the API call is not supported by the active device.
- CUDA_ERROR_CONTEXT_ALREADY_IN_USE*** This indicates that the [CUcontext](#) passed to the API call can only be bound to a single CPU thread at a time but is already bound to a CPU thread.
- CUDA_ERROR_PEER_ACCESS_UNSUPPORTED*** This indicates that peer access is not supported across the given devices.
- CUDA_ERROR_INVALID_SOURCE*** This indicates that the device kernel source is invalid.
- CUDA_ERROR_FILE_NOT_FOUND*** This indicates that the file specified was not found.
- CUDA_ERROR_SHARED_OBJECT_SYMBOL_NOT_FOUND*** This indicates that a link to a shared object failed to resolve.
- CUDA_ERROR_SHARED_OBJECT_INIT_FAILED*** This indicates that initialization of a shared object failed.
- CUDA_ERROR_OPERATING_SYSTEM*** This indicates that an OS call failed.
- CUDA_ERROR_INVALID_HANDLE*** This indicates that a resource handle passed to the API call was not valid. Resource handles are opaque types like [CUstream](#) and [CUevent](#).
- CUDA_ERROR_NOT_FOUND*** This indicates that a named symbol was not found. Examples of symbols are global/constant variable names, texture names, and surface names.
- CUDA_ERROR_NOT_READY*** This indicates that asynchronous operations issued previously have not completed yet. This result is not actually an error, but must be indicated differently than [CUDA_SUCCESS](#) (which indicates completion). Calls that may return this value include [cuEventQuery\(\)](#) and [cuStreamQuery\(\)](#).
- CUDA_ERROR_LAUNCH_FAILED*** An exception occurred on the device while executing a kernel. Common causes include dereferencing an invalid device pointer and accessing out of bounds shared memory. The context cannot be used, so it must be destroyed (and a new one should be created). All existing device memory allocations from this context are invalid and must be reconstructed if the program is to continue using CUDA.
- CUDA_ERROR_LAUNCH_OUT_OF_RESOURCES*** This indicates that a launch did not occur because it did not have appropriate resources. This error usually indicates that the user has attempted to pass too many arguments to the device kernel, or the kernel launch specifies too many threads for the kernel's register count. Passing arguments of the wrong size (i.e. a 64-bit pointer when a 32-bit int is expected) is equivalent to passing too many arguments and can also result in this error.
- CUDA_ERROR_LAUNCH_TIMEOUT*** This indicates that the device kernel took too long to execute. This can only occur if timeouts are enabled - see the device attribute [CU_DEVICE_ATTRIBUTE_KERNEL_EXEC_TIMEOUT](#) for more information. The context cannot be used (and must be destroyed similar to [CUDA_ERROR_LAUNCH_FAILED](#)). All existing device memory allocations from this context are invalid and must be reconstructed if the program is to continue using CUDA.

- CUDA_ERROR_LAUNCH_INCOMPATIBLE_TEXTURING*** This error indicates a kernel launch that uses an incompatible texturing mode.
- CUDA_ERROR_PEER_ACCESS_ALREADY_ENABLED*** This error indicates that a call to `cuCtxEnablePeerAccess()` is trying to re-enable peer access to a context which has already had peer access to it enabled.
- CUDA_ERROR_PEER_ACCESS_NOT_ENABLED*** This error indicates that `cuCtxDisablePeerAccess()` is trying to disable peer access which has not been enabled yet via `cuCtxEnablePeerAccess()`.
- CUDA_ERROR_PRIMARY_CONTEXT_ACTIVE*** This error indicates that the primary context for the specified device has already been initialized.
- CUDA_ERROR_CONTEXT_IS_DESTROYED*** This error indicates that the context current to the calling thread has been destroyed using `cuCtxDestroy`, or is a primary context which has not yet been initialized.
- CUDA_ERROR_ASSERT*** A device-side assert triggered during kernel execution. The context cannot be used anymore, and must be destroyed. All existing device memory allocations from this context are invalid and must be reconstructed if the program is to continue using CUDA.
- CUDA_ERROR_TOO_MANY_PEERS*** This error indicates that the hardware resources required to enable peer access have been exhausted for one or more of the devices passed to `cuCtxEnablePeerAccess()`.
- CUDA_ERROR_HOST_MEMORY_ALREADY_REGISTERED*** This error indicates that the memory range passed to `cuMemHostRegister()` has already been registered.
- CUDA_ERROR_HOST_MEMORY_NOT_REGISTERED*** This error indicates that the pointer passed to `cuMemHostUnregister()` does not correspond to any currently registered memory region.
- CUDA_ERROR_NOT_PERMITTED*** This error indicates that the attempted operation is not permitted.
- CUDA_ERROR_NOT_SUPPORTED*** This error indicates that the attempted operation is not supported on the current system or device.
- CUDA_ERROR_UNKNOWN*** This indicates that an unknown internal error has occurred.

5.29.3.7 enum CUdevice_attribute_enum

Device properties

Enumerator:

- CU_DEVICE_ATTRIBUTE_MAX_THREADS_PER_BLOCK*** Maximum number of threads per block
- CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_X*** Maximum block dimension X
- CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Y*** Maximum block dimension Y
- CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Z*** Maximum block dimension Z
- CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_X*** Maximum grid dimension X
- CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Y*** Maximum grid dimension Y
- CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Z*** Maximum grid dimension Z
- CU_DEVICE_ATTRIBUTE_MAX_SHARED_MEMORY_PER_BLOCK*** Maximum shared memory available per block in bytes
- CU_DEVICE_ATTRIBUTE_SHARED_MEMORY_PER_BLOCK*** Deprecated, use `CU_DEVICE_ATTRIBUTE_MAX_SHARED_MEMORY_PER_BLOCK`
- CU_DEVICE_ATTRIBUTE_TOTAL_CONSTANT_MEMORY*** Memory available on device for `__constant__` variables in a CUDA C kernel in bytes
- CU_DEVICE_ATTRIBUTE_WARP_SIZE*** Warp size in threads
- CU_DEVICE_ATTRIBUTE_MAX_PITCH*** Maximum pitch in bytes allowed by memory copies

CU_DEVICE_ATTRIBUTE_MAX_REGISTERS_PER_BLOCK Maximum number of 32-bit registers available per block

CU_DEVICE_ATTRIBUTE_REGISTERS_PER_BLOCK Deprecated, use ***CU_DEVICE_ATTRIBUTE_MAX_REGISTERS_PER_BLOCK***

CU_DEVICE_ATTRIBUTE_CLOCK_RATE Peak clock frequency in kilohertz

CU_DEVICE_ATTRIBUTE_TEXTURE_ALIGNMENT Alignment requirement for textures

CU_DEVICE_ATTRIBUTE_GPU_OVERLAP Device can possibly copy memory and execute a kernel concurrently. Deprecated. Use instead ***CU_DEVICE_ATTRIBUTE_ASYNC_ENGINE_COUNT***.

CU_DEVICE_ATTRIBUTE_MULTIPROCESSOR_COUNT Number of multiprocessors on device

CU_DEVICE_ATTRIBUTE_KERNEL_EXEC_TIMEOUT Specifies whether there is a run time limit on kernels

CU_DEVICE_ATTRIBUTE_INTEGRATED Device is integrated with host memory

CU_DEVICE_ATTRIBUTE_CAN_MAP_HOST_MEMORY Device can map host memory into CUDA address space

CU_DEVICE_ATTRIBUTE_COMPUTE_MODE Compute mode (See [CUcomputemode](#) for details)

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_WIDTH Maximum 1D texture width

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_WIDTH Maximum 2D texture width

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_HEIGHT Maximum 2D texture height

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_WIDTH Maximum 3D texture width

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_HEIGHT Maximum 3D texture height

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_DEPTH Maximum 3D texture depth

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_WIDTH Maximum 2D layered texture width

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_HEIGHT Maximum 2D layered texture height

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_LAYERS Maximum layers in a 2D layered texture

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_ARRAY_WIDTH Deprecated, use ***CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_WIDTH***

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_ARRAY_HEIGHT Deprecated, use ***CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_HEIGHT***

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_ARRAY_NUMSLICES Deprecated, use ***CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_LAYERS***

CU_DEVICE_ATTRIBUTE_SURFACE_ALIGNMENT Alignment requirement for surfaces

CU_DEVICE_ATTRIBUTE_CONCURRENT_KERNELS Device can possibly execute multiple kernels concurrently

CU_DEVICE_ATTRIBUTE_ECC_ENABLED Device has ECC support enabled

CU_DEVICE_ATTRIBUTE_PCI_BUS_ID PCI bus ID of the device

CU_DEVICE_ATTRIBUTE_PCI_DEVICE_ID PCI device ID of the device

CU_DEVICE_ATTRIBUTE_TCC_DRIVER Device is using TCC driver model

CU_DEVICE_ATTRIBUTE_MEMORY_CLOCK_RATE Peak memory clock frequency in kilohertz

CU_DEVICE_ATTRIBUTE_GLOBAL_MEMORY_BUS_WIDTH Global memory bus width in bits

CU_DEVICE_ATTRIBUTE_L2_CACHE_SIZE Size of L2 cache in bytes

CU_DEVICE_ATTRIBUTE_MAX_THREADS_PER_MULTIPROCESSOR Maximum resident threads per multiprocessor

- CU_DEVICE_ATTRIBUTE_ASYNC_ENGINE_COUNT*** Number of asynchronous engines
- CU_DEVICE_ATTRIBUTE_UNIFIED_ADDRESSING*** Device shares a unified address space with the host
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LAYERED_WIDTH*** Maximum 1D layered texture width
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LAYERED_LAYERS*** Maximum layers in a 1D layered texture
- CU_DEVICE_ATTRIBUTE_CAN_TEX2D_GATHER*** Deprecated, do not use.
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_GATHER_WIDTH*** Maximum 2D texture width if `CUDA_ARRAY3D_TEXTURE_GATHER` is set
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_GATHER_HEIGHT*** Maximum 2D texture height if `CUDA_ARRAY3D_TEXTURE_GATHER` is set
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_WIDTH_ALTERNATE*** Alternate maximum 3D texture width
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_HEIGHT_ALTERNATE*** Alternate maximum 3D texture height
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_DEPTH_ALTERNATE*** Alternate maximum 3D texture depth
- CU_DEVICE_ATTRIBUTE_PCI_DOMAIN_ID*** PCI domain ID of the device
- CU_DEVICE_ATTRIBUTE_TEXTURE_PITCH_ALIGNMENT*** Pitch alignment requirement for textures
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_WIDTH*** Maximum cubemap texture width/height
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_LAYERED_WIDTH*** Maximum cubemap layered texture width/height
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_LAYERED_LAYERS*** Maximum layers in a cubemap layered texture
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_WIDTH*** Maximum 1D surface width
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_WIDTH*** Maximum 2D surface width
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_HEIGHT*** Maximum 2D surface height
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_WIDTH*** Maximum 3D surface width
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_HEIGHT*** Maximum 3D surface height
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_DEPTH*** Maximum 3D surface depth
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_LAYERED_WIDTH*** Maximum 1D layered surface width
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_LAYERED_LAYERS*** Maximum layers in a 1D layered surface
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_WIDTH*** Maximum 2D layered surface width
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_HEIGHT*** Maximum 2D layered surface height
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_LAYERS*** Maximum layers in a 2D layered surface
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_WIDTH*** Maximum cubemap surface width
- CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_LAYERED_WIDTH*** Maximum cubemap layered surface width

CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_LAYERED_LAYERS Maximum layers in a cubemap layered surface

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LINEAR_WIDTH Maximum 1D linear texture width

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_WIDTH Maximum 2D linear texture width

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_HEIGHT Maximum 2D linear texture height

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_PITCH Maximum 2D linear texture pitch in bytes

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_MIPMAPPED_WIDTH Maximum mipmapped 2D texture width

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_MIPMAPPED_HEIGHT Maximum mipmapped 2D texture height

CU_DEVICE_ATTRIBUTE_COMPUTE_CAPABILITY_MAJOR Major compute capability version number

CU_DEVICE_ATTRIBUTE_COMPUTE_CAPABILITY_MINOR Minor compute capability version number

CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_MIPMAPPED_WIDTH Maximum mipmapped 1D texture width

5.29.3.8 enum CUevent_flags_enum

Event creation flags

Enumerator:

CU_EVENT_DEFAULT Default event flag

CU_EVENT_BLOCKING_SYNC Event uses blocking synchronization

CU_EVENT_DISABLE_TIMING Event will not record timing data

CU_EVENT_INTERPROCESS Event is suitable for interprocess use. ***CU_EVENT_DISABLE_TIMING*** must be set

5.29.3.9 enum CUfilter_mode_enum

Texture reference filtering modes

Enumerator:

CU_TR_FILTER_MODE_POINT Point filter mode

CU_TR_FILTER_MODE_LINEAR Linear filter mode

5.29.3.10 enum CUfunc_cache_enum

Function cache configurations

Enumerator:

CU_FUNC_CACHE_PREFER_NONE no preference for shared memory or L1 (default)

CU_FUNC_CACHE_PREFER_SHARED prefer larger shared memory and smaller L1 cache

CU_FUNC_CACHE_PREFER_L1 prefer larger L1 cache and smaller shared memory

CU_FUNC_CACHE_PREFER_EQUAL prefer equal sized L1 cache and shared memory

5.29.3.11 enum CUfunction_attribute_enum

Function properties

Enumerator:

CU_FUNC_ATTRIBUTE_MAX_THREADS_PER_BLOCK The maximum number of threads per block, beyond which a launch of the function would fail. This number depends on both the function and the device on which the function is currently loaded.

CU_FUNC_ATTRIBUTE_SHARED_SIZE_BYTES The size in bytes of statically-allocated shared memory required by this function. This does not include dynamically-allocated shared memory requested by the user at runtime.

CU_FUNC_ATTRIBUTE_CONST_SIZE_BYTES The size in bytes of user-allocated constant memory required by this function.

CU_FUNC_ATTRIBUTE_LOCAL_SIZE_BYTES The size in bytes of local memory used by each thread of this function.

CU_FUNC_ATTRIBUTE_NUM_REGS The number of registers used by each thread of this function.

CU_FUNC_ATTRIBUTE_PTX_VERSION The PTX virtual architecture version for which the function was compiled. This value is the major PTX version * 10 + the minor PTX version, so a PTX version 1.3 function would return the value 13. Note that this may return the undefined value of 0 for cubins compiled prior to CUDA 3.0.

CU_FUNC_ATTRIBUTE_BINARY_VERSION The binary architecture version for which the function was compiled. This value is the major binary version * 10 + the minor binary version, so a binary version 1.3 function would return the value 13. Note that this will return a value of 10 for legacy cubins that do not have a properly-encoded binary architecture version.

5.29.3.12 enum CUgraphicsMapResourceFlags_enum

Flags for mapping and unmapping interop resources

5.29.3.13 enum CUgraphicsRegisterFlags_enum

Flags to register a graphics resource

5.29.3.14 enum CUipcMem_flags_enum

CUDA Ipc Mem Flags

Enumerator:

CU_IPC_MEM_LAZY_ENABLE_PEER_ACCESS Automatically enable peer access between remote devices as needed

5.29.3.15 enum CUjit_fallback_enum

Cubin matching fallback strategies

Enumerator:

CU_PREFER_PTX Prefer to compile ptx

CU_PREFER_BINARY Prefer to fall back to compatible binary code

5.29.3.16 enum CUjit_option_enum

Online compiler options

Enumerator:

CU_JIT_MAX_REGISTERS Max number of registers that a thread may use.

Option type: unsigned int

CU_JIT_THREADS_PER_BLOCK IN: Specifies minimum number of threads per block to target compilation for

OUT: Returns the number of threads the compiler actually targeted. This restricts the resource utilization for the compiler (e.g. max registers) such that a block with the given number of threads should be able to launch based on register limitations. Note, this option does not currently take into account any other resource limitations, such as shared memory utilization.

Option type: unsigned int

CU_JIT_WALL_TIME Returns a float value in the option of the wall clock time, in milliseconds, spent creating the cubin

Option type: float

CU_JIT_INFO_LOG_BUFFER Pointer to a buffer in which to print any log messages from PTXAS that are informational in nature (the buffer size is specified via option [CU_JIT_INFO_LOG_BUFFER_SIZE_BYTES](#))

Option type: char*

CU_JIT_INFO_LOG_BUFFER_SIZE_BYTES IN: Log buffer size in bytes. Log messages will be capped at this size (including null terminator)

OUT: Amount of log buffer filled with messages

Option type: unsigned int

CU_JIT_ERROR_LOG_BUFFER Pointer to a buffer in which to print any log messages from PTXAS that reflect errors (the buffer size is specified via option [CU_JIT_ERROR_LOG_BUFFER_SIZE_BYTES](#))

Option type: char*

CU_JIT_ERROR_LOG_BUFFER_SIZE_BYTES IN: Log buffer size in bytes. Log messages will be capped at this size (including null terminator)

OUT: Amount of log buffer filled with messages

Option type: unsigned int

CU_JIT_OPTIMIZATION_LEVEL Level of optimizations to apply to generated code (0 - 4), with 4 being the default and highest level of optimizations.

Option type: unsigned int

CU_JIT_TARGET_FROM_CUCONTEXT No option value required. Determines the target based on the current attached context (default)

Option type: No option value needed

CU_JIT_TARGET Target is chosen based on supplied [CUjit_target_enum](#).

Option type: unsigned int for enumerated type [CUjit_target_enum](#)

CU_JIT_FALLBACK_STRATEGY Specifies choice of fallback strategy if matching cubin is not found. Choice is based on supplied [CUjit_fallback_enum](#).

Option type: unsigned int for enumerated type [CUjit_fallback_enum](#)

5.29.3.17 enum CUjit_target_enum

Online compilation targets

Enumerator:

CU_TARGET_COMPUTE_10 Compute device class 1.0

CU_TARGET_COMPUTE_11 Compute device class 1.1

CU_TARGET_COMPUTE_12 Compute device class 1.2

CU_TARGET_COMPUTE_13 Compute device class 1.3

CU_TARGET_COMPUTE_20 Compute device class 2.0

CU_TARGET_COMPUTE_21 Compute device class 2.1

CU_TARGET_COMPUTE_30 Compute device class 3.0

CU_TARGET_COMPUTE_35 Compute device class 3.5

5.29.3.18 enum CUlimit_enum

Limits

Enumerator:

CU_LIMIT_STACK_SIZE GPU thread stack size

CU_LIMIT_PRINTF_FIFO_SIZE GPU printf FIFO size

CU_LIMIT_MALLOC_HEAP_SIZE GPU malloc heap size

CU_LIMIT_DEV_RUNTIME_SYNC_DEPTH GPU device runtime launch synchronize depth

CU_LIMIT_DEV_RUNTIME_PENDING_LAUNCH_COUNT GPU device runtime pending launch count

5.29.3.19 enum CUmemorytype_enum

Memory types

Enumerator:

CU_MEMORYTYPE_HOST Host memory

CU_MEMORYTYPE_DEVICE Device memory

CU_MEMORYTYPE_ARRAY Array memory

CU_MEMORYTYPE_UNIFIED Unified device or host memory

5.29.3.20 enum CUpointer_attribute_enum

Pointer information

Enumerator:

- CU_POINTER_ATTRIBUTE_CONTEXT* The [CUcontext](#) on which a pointer was allocated or registered
- CU_POINTER_ATTRIBUTE_MEMORY_TYPE* The [CUmemorytype](#) describing the physical location of a pointer
- CU_POINTER_ATTRIBUTE_DEVICE_POINTER* The address at which a pointer's memory may be accessed on the device
- CU_POINTER_ATTRIBUTE_HOST_POINTER* The address at which a pointer's memory may be accessed on the host
- CU_POINTER_ATTRIBUTE_P2P_TOKENS* A pair of tokens for use with the `nv-p2p.h` Linux kernel interface

5.29.3.21 enum CUresourcetype_enum

Resource types

Enumerator:

- CU_RESOURCE_TYPE_ARRAY* Array resource
- CU_RESOURCE_TYPE_MIPMAPPED_ARRAY* Mipmapped array resource
- CU_RESOURCE_TYPE_LINEAR* Linear resource
- CU_RESOURCE_TYPE_PITCH2D* Pitch 2D resource

5.29.3.22 enum CUresourceViewFormat_enum

Resource view format

Enumerator:

- CU_RES_VIEW_FORMAT_NONE* No resource view format (use underlying resource format)
- CU_RES_VIEW_FORMAT_UINT_1X8* 1 channel unsigned 8-bit integers
- CU_RES_VIEW_FORMAT_UINT_2X8* 2 channel unsigned 8-bit integers
- CU_RES_VIEW_FORMAT_UINT_4X8* 4 channel unsigned 8-bit integers
- CU_RES_VIEW_FORMAT_SINT_1X8* 1 channel signed 8-bit integers
- CU_RES_VIEW_FORMAT_SINT_2X8* 2 channel signed 8-bit integers
- CU_RES_VIEW_FORMAT_SINT_4X8* 4 channel signed 8-bit integers
- CU_RES_VIEW_FORMAT_UINT_1X16* 1 channel unsigned 16-bit integers
- CU_RES_VIEW_FORMAT_UINT_2X16* 2 channel unsigned 16-bit integers
- CU_RES_VIEW_FORMAT_UINT_4X16* 4 channel unsigned 16-bit integers
- CU_RES_VIEW_FORMAT_SINT_1X16* 1 channel signed 16-bit integers
- CU_RES_VIEW_FORMAT_SINT_2X16* 2 channel signed 16-bit integers
- CU_RES_VIEW_FORMAT_SINT_4X16* 4 channel signed 16-bit integers
- CU_RES_VIEW_FORMAT_UINT_1X32* 1 channel unsigned 32-bit integers
- CU_RES_VIEW_FORMAT_UINT_2X32* 2 channel unsigned 32-bit integers

CU_RES_VIEW_FORMAT_UINT_4X32 4 channel unsigned 32-bit integers
CU_RES_VIEW_FORMAT_SINT_1X32 1 channel signed 32-bit integers
CU_RES_VIEW_FORMAT_SINT_2X32 2 channel signed 32-bit integers
CU_RES_VIEW_FORMAT_SINT_4X32 4 channel signed 32-bit integers
CU_RES_VIEW_FORMAT_FLOAT_1X16 1 channel 16-bit floating point
CU_RES_VIEW_FORMAT_FLOAT_2X16 2 channel 16-bit floating point
CU_RES_VIEW_FORMAT_FLOAT_4X16 4 channel 16-bit floating point
CU_RES_VIEW_FORMAT_FLOAT_1X32 1 channel 32-bit floating point
CU_RES_VIEW_FORMAT_FLOAT_2X32 2 channel 32-bit floating point
CU_RES_VIEW_FORMAT_FLOAT_4X32 4 channel 32-bit floating point
CU_RES_VIEW_FORMAT_UNSIGNED_BC1 Block compressed 1
CU_RES_VIEW_FORMAT_UNSIGNED_BC2 Block compressed 2
CU_RES_VIEW_FORMAT_UNSIGNED_BC3 Block compressed 3
CU_RES_VIEW_FORMAT_UNSIGNED_BC4 Block compressed 4 unsigned
CU_RES_VIEW_FORMAT_SIGNED_BC4 Block compressed 4 signed
CU_RES_VIEW_FORMAT_UNSIGNED_BC5 Block compressed 5 unsigned
CU_RES_VIEW_FORMAT_SIGNED_BC5 Block compressed 5 signed
CU_RES_VIEW_FORMAT_UNSIGNED_BC6H Block compressed 6 unsigned half-float
CU_RES_VIEW_FORMAT_SIGNED_BC6H Block compressed 6 signed half-float
CU_RES_VIEW_FORMAT_UNSIGNED_BC7 Block compressed 7

5.29.3.23 enum CUsharedconfig_enum

Shared memory configurations

Enumerator:

CU_SHARED_MEM_CONFIG_DEFAULT_BANK_SIZE set default shared memory bank size
CU_SHARED_MEM_CONFIG_FOUR_BYTE_BANK_SIZE set shared memory bank width to four bytes
CU_SHARED_MEM_CONFIG_EIGHT_BYTE_BANK_SIZE set shared memory bank width to eight bytes

5.29.3.24 enum CUstream_flags_enum

Stream creation flags

Enumerator:

CU_STREAM_DEFAULT Default stream flag
CU_STREAM_NON_BLOCKING Stream does not synchronize with stream 0 (the NULL stream)

5.30 Initialization

Functions

- [CUresult cuInit](#) (unsigned int *Flags*)
Initialize the CUDA driver API.

5.30.1 Detailed Description

This section describes the initialization functions of the low-level CUDA driver application programming interface.

5.30.2 Function Documentation

5.30.2.1 CUresult cuInit (unsigned int *Flags*)

Initializes the driver API and must be called before any other function from the driver API. Currently, the `Flags` parameter must be 0. If `cuInit()` has not been called, any function from the driver API will return `CUDA_ERROR_NOT_INITIALIZED`.

Parameters:

Flags - Initialization flag for CUDA.

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_INVALID_VALUE`, `CUDA_ERROR_INVALID_DEVICE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

5.31 Version Management

Functions

- [CUresult cuDriverGetVersion](#) (int *driverVersion)

Returns the CUDA driver version.

5.31.1 Detailed Description

This section describes the version management functions of the low-level CUDA driver application programming interface.

5.31.2 Function Documentation

5.31.2.1 CUresult cuDriverGetVersion (int * driverVersion)

Returns in *driverVersion the version number of the installed CUDA driver. This function automatically returns [CUDA_ERROR_INVALID_VALUE](#) if the driverVersion argument is NULL.

Parameters:

driverVersion - Returns the CUDA driver version

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

5.32 Device Management

Functions

- [CUresult cuDeviceGet](#) ([CUdevice](#) *device, int ordinal)
Returns a handle to a compute device.
- [CUresult cuDeviceGetAttribute](#) (int *pi, [CUdevice_attribute](#) attrib, [CUdevice](#) dev)
Returns information about the device.
- [CUresult cuDeviceGetCount](#) (int *count)
Returns the number of compute-capable devices.
- [CUresult cuDeviceGetName](#) (char *name, int len, [CUdevice](#) dev)
Returns an identifier string for the device.
- [CUresult cuDeviceTotalMem](#) (size_t *bytes, [CUdevice](#) dev)
Returns the total amount of memory on the device.

5.32.1 Detailed Description

This section describes the device management functions of the low-level CUDA driver application programming interface.

5.32.2 Function Documentation

5.32.2.1 [CUresult cuDeviceGet](#) ([CUdevice](#) *device, int ordinal)

Returns in *device a device handle given an ordinal in the range [0, [cuDeviceGetCount\(\)-1](#)].

Parameters:

- device* - Returned device handle
- ordinal* - Device number to get handle for

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_DEVICE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGetAttribute](#), [cuDeviceGetCount](#), [cuDeviceGetName](#), [cuDeviceTotalMem](#)

5.32.2.2 CUresult cuDeviceGetAttribute (int *pi, CUdevice_attribute attrib, CUdevice dev)

Returns in *pi the integer value of the attribute `attrib` on device `dev`. The supported attributes are:

- `CU_DEVICE_ATTRIBUTE_MAX_THREADS_PER_BLOCK`: Maximum number of threads per block;
- `CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_X`: Maximum x-dimension of a block;
- `CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Y`: Maximum y-dimension of a block;
- `CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Z`: Maximum z-dimension of a block;
- `CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_X`: Maximum x-dimension of a grid;
- `CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Y`: Maximum y-dimension of a grid;
- `CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Z`: Maximum z-dimension of a grid;
- `CU_DEVICE_ATTRIBUTE_MAX_SHARED_MEMORY_PER_BLOCK`: Maximum amount of shared memory available to a thread block in bytes; this amount is shared by all thread blocks simultaneously resident on a multiprocessor;
- `CU_DEVICE_ATTRIBUTE_TOTAL_CONSTANT_MEMORY`: Memory available on device for `__constant_` variables in a CUDA C kernel in bytes;
- `CU_DEVICE_ATTRIBUTE_WARP_SIZE`: Warp size in threads;
- `CU_DEVICE_ATTRIBUTE_MAX_PITCH`: Maximum pitch in bytes allowed by the memory copy functions that involve memory regions allocated through `cuMemAllocPitch()`;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_WIDTH`: Maximum 1D texture width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LINEAR_WIDTH`: Maximum width for a 1D texture bound to linear memory;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_MIPMAPPED_WIDTH`: Maximum mipmapped 1D texture width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_WIDTH`: Maximum 2D texture width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_HEIGHT`: Maximum 2D texture height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_WIDTH`: Maximum width for a 2D texture bound to linear memory;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_HEIGHT`: Maximum height for a 2D texture bound to linear memory;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_PITCH`: Maximum pitch in bytes for a 2D texture bound to linear memory;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_MIPMAPPED_WIDTH`: Maximum mipmapped 2D texture width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_MIPMAPPED_HEIGHT`: Maximum mipmapped 2D texture height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_WIDTH`: Maximum 3D texture width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_HEIGHT`: Maximum 3D texture height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_DEPTH`: Maximum 3D texture depth;

- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_WIDTH_ALTERNATE`: Alternate maximum 3D texture width, 0 if no alternate maximum 3D texture size is supported;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_HEIGHT_ALTERNATE`: Alternate maximum 3D texture height, 0 if no alternate maximum 3D texture size is supported;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE3D_DEPTH_ALTERNATE`: Alternate maximum 3D texture depth, 0 if no alternate maximum 3D texture size is supported;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_WIDTH`: Maximum cubemap texture width or height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LAYERED_WIDTH`: Maximum 1D layered texture width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_LAYERED_LAYERS`: Maximum layers in a 1D layered texture;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_WIDTH`: Maximum 2D layered texture width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_HEIGHT`: Maximum 2D layered texture height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LAYERED_LAYERS`: Maximum layers in a 2D layered texture;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_LAYERED_WIDTH`: Maximum cubemap layered texture width or height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_LAYERED_LAYERS`: Maximum layers in a cubemap layered texture;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_WIDTH`: Maximum 1D surface width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_WIDTH`: Maximum 2D surface width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_HEIGHT`: Maximum 2D surface height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_WIDTH`: Maximum 3D surface width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_HEIGHT`: Maximum 3D surface height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE3D_DEPTH`: Maximum 3D surface depth;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_LAYERED_WIDTH`: Maximum 1D layered surface width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE1D_LAYERED_LAYERS`: Maximum layers in a 1D layered surface;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_WIDTH`: Maximum 2D layered surface width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_HEIGHT`: Maximum 2D layered surface height;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACE2D_LAYERED_LAYERS`: Maximum layers in a 2D layered surface;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_WIDTH`: Maximum cubemap surface width;

- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_LAYERED_WIDTH`: Maximum cubemap layered surface width;
- `CU_DEVICE_ATTRIBUTE_MAXIMUM_SURFACECUBEMAP_LAYERED_LAYERS`: Maximum layers in a cubemap layered surface;
- `CU_DEVICE_ATTRIBUTE_MAX_REGISTERS_PER_BLOCK`: Maximum number of 32-bit registers available to a thread block; this number is shared by all thread blocks simultaneously resident on a multiprocessor;
- `CU_DEVICE_ATTRIBUTE_CLOCK_RATE`: Peak clock frequency in kilohertz;
- `CU_DEVICE_ATTRIBUTE_TEXTURE_ALIGNMENT`: Alignment requirement; texture base addresses aligned to textureAlign bytes do not need an offset applied to texture fetches;
- `CU_DEVICE_ATTRIBUTE_TEXTURE_PITCH_ALIGNMENT`: Pitch alignment requirement for 2D texture references bound to pitched memory;
- `CU_DEVICE_ATTRIBUTE_GPU_OVERLAP`: 1 if the device can concurrently copy memory between host and device while executing a kernel, or 0 if not;
- `CU_DEVICE_ATTRIBUTE_MULTIPROCESSOR_COUNT`: Number of multiprocessors on the device;
- `CU_DEVICE_ATTRIBUTE_KERNEL_EXEC_TIMEOUT`: 1 if there is a run time limit for kernels executed on the device, or 0 if not;
- `CU_DEVICE_ATTRIBUTE_INTEGRATED`: 1 if the device is integrated with the memory subsystem, or 0 if not;
- `CU_DEVICE_ATTRIBUTE_CAN_MAP_HOST_MEMORY`: 1 if the device can map host memory into the CUDA address space, or 0 if not;
- `CU_DEVICE_ATTRIBUTE_COMPUTE_MODE`: Compute mode that device is currently in. Available modes are as follows:
 - `CU_COMPUTEMODE_DEFAULT`: Default mode - Device is not restricted and can have multiple CUDA contexts present at a single time.
 - `CU_COMPUTEMODE_EXCLUSIVE`: Compute-exclusive mode - Device can have only one CUDA context present on it at a time.
 - `CU_COMPUTEMODE_PROHIBITED`: Compute-prohibited mode - Device is prohibited from creating new CUDA contexts.
 - `CU_COMPUTEMODE_EXCLUSIVE_PROCESS`: Compute-exclusive-process mode - Device can have only one context used by a single process at a time.
- `CU_DEVICE_ATTRIBUTE_CONCURRENT_KERNELS`: 1 if the device supports executing multiple kernels within the same context simultaneously, or 0 if not. It is not guaranteed that multiple kernels will be resident on the device concurrently so this feature should not be relied upon for correctness;
- `CU_DEVICE_ATTRIBUTE_ECC_ENABLED`: 1 if error correction is enabled on the device, 0 if error correction is disabled or not supported by the device;
- `CU_DEVICE_ATTRIBUTE_PCI_BUS_ID`: PCI bus identifier of the device;
- `CU_DEVICE_ATTRIBUTE_PCI_DEVICE_ID`: PCI device (also known as slot) identifier of the device;
- `CU_DEVICE_ATTRIBUTE_TCC_DRIVER`: 1 if the device is using a TCC driver. TCC is only available on Tesla hardware running Windows Vista or later;
- `CU_DEVICE_ATTRIBUTE_MEMORY_CLOCK_RATE`: Peak memory clock frequency in kilohertz;
- `CU_DEVICE_ATTRIBUTE_GLOBAL_MEMORY_BUS_WIDTH`: Global memory bus width in bits;

- [CU_DEVICE_ATTRIBUTE_L2_CACHE_SIZE](#): Size of L2 cache in bytes. 0 if the device doesn't have L2 cache;
- [CU_DEVICE_ATTRIBUTE_MAX_THREADS_PER_MULTIPROCESSOR](#): Maximum resident threads per multiprocessor;
- [CU_DEVICE_ATTRIBUTE_UNIFIED_ADDRESSING](#): 1 if the device shares a unified address space with the host, or 0 if not;
- [CU_DEVICE_ATTRIBUTE_COMPUTE_CAPABILITY_MAJOR](#): Major compute capability version number;
- [CU_DEVICE_ATTRIBUTE_COMPUTE_CAPABILITY_MINOR](#): Minor compute capability version number;

Parameters:

pi - Returned device attribute value
attrib - Device attribute to query
dev - Device handle

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_DEVICE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGetCount](#), [cuDeviceGetName](#), [cuDeviceGet](#), [cuDeviceTotalMem](#)

5.32.2.3 CUresult cuDeviceGetCount (int * count)

Returns in *count the number of devices with compute capability greater than or equal to 1.0 that are available for execution. If there is no such device, [cuDeviceGetCount\(\)](#) returns 0.

Parameters:

count - Returned number of compute-capable devices

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGetAttribute](#), [cuDeviceGetName](#), [cuDeviceGet](#), [cuDeviceTotalMem](#)

5.32.2.4 CUresult cuDeviceGetName (char * name, int len, CUdevice dev)

Returns an ASCII string identifying the device `dev` in the NULL-terminated string pointed to by `name`. `len` specifies the maximum length of the string that may be returned.

Parameters:

- name* - Returned identifier string for the device
- len* - Maximum length of string to store in *name*
- dev* - Device to get identifier string for

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_DEVICE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGetAttribute](#), [cuDeviceGetCount](#), [cuDeviceGet](#), [cuDeviceTotalMem](#)

5.32.2.5 CUresult cuDeviceTotalMem (size_t * bytes, CUdevice dev)

Returns in `*bytes` the total amount of memory available on the device `dev` in bytes.

Parameters:

- bytes* - Returned memory available on device in bytes
- dev* - Device handle

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_DEVICE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGetAttribute](#), [cuDeviceGetCount](#), [cuDeviceGetName](#), [cuDeviceGet](#),

5.33 Device Management [DEPRECATED]

Functions

- [CUresult cuDeviceComputeCapability](#) (int *major, int *minor, CUdevice dev)
Returns the compute capability of the device.
- [CUresult cuDeviceGetProperties](#) (CUdevprop *prop, CUdevice dev)
Returns properties for a selected device.

5.33.1 Detailed Description

This section describes the device management functions of the low-level CUDA driver application programming interface.

5.33.2 Function Documentation

5.33.2.1 CUresult cuDeviceComputeCapability (int *major, int *minor, CUdevice dev)

Deprecated

This function was deprecated as of CUDA 5.0 and its functionality superseded by [cuDeviceGetAttribute\(\)](#).

Returns in *major and *minor the major and minor revision numbers that define the compute capability of the device dev.

Parameters:

- major* - Major revision number
- minor* - Minor revision number
- dev* - Device handle

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_DEVICE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGetAttribute](#), [cuDeviceGetCount](#), [cuDeviceGetName](#), [cuDeviceGet](#), [cuDeviceTotalMem](#)

5.33.2.2 CUresult cuDeviceGetProperties (CUdevprop *prop, CUdevice dev)

Deprecated

This function was deprecated as of CUDA 5.0 and replaced by [cuDeviceGetAttribute\(\)](#).

Returns in **prop* the properties of device *dev*. The [CUdevprop](#) structure is defined as:

```
typedef struct CUdevprop_st {
    int maxThreadsPerBlock;
    int maxThreadsDim[3];
    int maxGridSize[3];
    int sharedMemPerBlock;
    int totalConstantMemory;
    int SIMDWidth;
    int memPitch;
    int regsPerBlock;
    int clockRate;
    int textureAlign
} CUdevprop;
```

where:

- `maxThreadsPerBlock` is the maximum number of threads per block;
- `maxThreadsDim[3]` is the maximum sizes of each dimension of a block;
- `maxGridSize[3]` is the maximum sizes of each dimension of a grid;
- `sharedMemPerBlock` is the total amount of shared memory available per block in bytes;
- `totalConstantMemory` is the total amount of constant memory available on the device in bytes;
- `SIMDWidth` is the warp size;
- `memPitch` is the maximum pitch allowed by the memory copy functions that involve memory regions allocated through [cuMemAllocPitch\(\)](#);
- `regsPerBlock` is the total number of registers available per block;
- `clockRate` is the clock frequency in kilohertz;
- `textureAlign` is the alignment requirement; texture base addresses that are aligned to `textureAlign` bytes do not need an offset applied to texture fetches.

Parameters:

prop - Returned properties of device

dev - Device to get properties for

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_DEVICE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGetAttribute](#), [cuDeviceGetCount](#), [cuDeviceGetName](#), [cuDeviceGet](#), [cuDeviceTotalMem](#)

5.34 Context Management

Functions

- **CUresult cuCtxCreate** (CUcontext *pctx, unsigned int flags, CUdevice dev)
Create a CUDA context.
- **CUresult cuCtxDestroy** (CUcontext ctx)
Destroy a CUDA context.
- **CUresult cuCtxGetApiVersion** (CUcontext ctx, unsigned int *version)
Gets the context's API version.
- **CUresult cuCtxGetCacheConfig** (CUfunc_cache *pconfig)
Returns the preferred cache configuration for the current context.
- **CUresult cuCtxGetCurrent** (CUcontext *pctx)
Returns the CUDA context bound to the calling CPU thread.
- **CUresult cuCtxGetDevice** (CUdevice *device)
Returns the device ID for the current context.
- **CUresult cuCtxGetLimit** (size_t *pvalue, CUlimit limit)
Returns resource limits.
- **CUresult cuCtxGetSharedMemConfig** (CUsharedconfig *pConfig)
Returns the current shared memory configuration for the current context.
- **CUresult cuCtxPopCurrent** (CUcontext *pctx)
Pops the current CUDA context from the current CPU thread.
- **CUresult cuCtxPushCurrent** (CUcontext ctx)
Pushes a context on the current CPU thread.
- **CUresult cuCtxSetCacheConfig** (CUfunc_cache config)
Sets the preferred cache configuration for the current context.
- **CUresult cuCtxSetCurrent** (CUcontext ctx)
Binds the specified CUDA context to the calling CPU thread.
- **CUresult cuCtxSetLimit** (CUlimit limit, size_t value)
Set resource limits.
- **CUresult cuCtxSetSharedMemConfig** (CUsharedconfig config)
Sets the shared memory configuration for the current context.
- **CUresult cuCtxSynchronize** (void)
Block for a context's tasks to complete.

5.34.1 Detailed Description

This section describes the context management functions of the low-level CUDA driver application programming interface.

5.34.2 Function Documentation

5.34.2.1 CUresult cuCtxCreate (CUcontext *pctx, unsigned int flags, CUdevice dev)

Creates a new CUDA context and associates it with the calling thread. The `flags` parameter is described below. The context is created with a usage count of 1 and the caller of `cuCtxCreate()` must call `cuCtxDestroy()` or when done using the context. If a context is already current to the thread, it is supplanted by the newly created context and may be restored by a subsequent call to `cuCtxPopCurrent()`.

The three LSBs of the `flags` parameter can be used to control how the OS thread, which owns the CUDA context at the time of an API call, interacts with the OS scheduler when waiting for results from the GPU. Only one of the scheduling flags can be set when creating a context.

- **CU_CTX_SCHED_AUTO**: The default value if the `flags` parameter is zero, uses a heuristic based on the number of active CUDA contexts in the process C and the number of logical processors in the system P . If $C > P$, then CUDA will yield to other OS threads when waiting for the GPU, otherwise CUDA will not yield while waiting for results and actively spin on the processor.
- **CU_CTX_SCHED_SPIN**: Instruct CUDA to actively spin when waiting for results from the GPU. This can decrease latency when waiting for the GPU, but may lower the performance of CPU threads if they are performing work in parallel with the CUDA thread.
- **CU_CTX_SCHED_YIELD**: Instruct CUDA to yield its thread when waiting for results from the GPU. This can increase latency when waiting for the GPU, but can increase the performance of CPU threads performing work in parallel with the GPU.
- **CU_CTX_SCHED_BLOCKING_SYNC**: Instruct CUDA to block the CPU thread on a synchronization primitive when waiting for the GPU to finish work.
- **CU_CTX_BLOCKING_SYNC**: Instruct CUDA to block the CPU thread on a synchronization primitive when waiting for the GPU to finish work.
Deprecated: This flag was deprecated as of CUDA 4.0 and was replaced with **CU_CTX_SCHED_BLOCKING_SYNC**.
- **CU_CTX_MAP_HOST**: Instruct CUDA to support mapped pinned allocations. This flag must be set in order to allocate pinned host memory that is accessible to the GPU.
- **CU_CTX_LMEM_RESIZE_TO_MAX**: Instruct CUDA to not reduce local memory after resizing local memory for a kernel. This can prevent thrashing by local memory allocations when launching many kernels with high local memory usage at the cost of potentially increased memory usage.

Context creation will fail with **CUDA_ERROR_UNKNOWN** if the compute mode of the device is **CU_COMPUTEMODE_PROHIBITED**. Similarly, context creation will also fail with **CUDA_ERROR_UNKNOWN** if the compute mode for the device is set to **CU_COMPUTEMODE_EXCLUSIVE** and there is already an active context on the device. The function `cuDeviceGetAttribute()` can be used with **CU_DEVICE_ATTRIBUTE_COMPUTE_MODE** to determine the compute mode of the device. The `nvidia-smi` tool can be used to set the compute mode for devices. Documentation for `nvidia-smi` can be obtained by passing a `-h` option to it.

Parameters:

ptx - Returned context handle of the new context
flags - Context creation flags
dev - Device to create context on

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_DEVICE](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxDestroy](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxGetLimit](#), [cuCtxPopCurrent](#), [cuCtxPushCurrent](#), [cuCtxSetCacheConfig](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#)

5.34.2.2 CUresult cuCtxDestroy (CUcontext ctx)

Destroys the CUDA context specified by `ctx`. The context `ctx` will be destroyed regardless of how many threads it is current to. It is the responsibility of the calling function to ensure that no API call issues using `ctx` while [cuCtxDestroy\(\)](#) is executing.

If `ctx` is current to the calling thread then `ctx` will also be popped from the current thread's context stack (as though [cuCtxPopCurrent\(\)](#) were called). If `ctx` is current to other threads, then `ctx` will remain current to those threads, and attempting to access `ctx` from those threads will result in the error [CUDA_ERROR_CONTEXT_IS_DESTROYED](#).

Parameters:

ctx - Context to destroy

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxGetLimit](#), [cuCtxPopCurrent](#), [cuCtxPushCurrent](#), [cuCtxSetCacheConfig](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#)

5.34.2.3 CUresult cuCtxGetApiVersion (CUcontext ctx, unsigned int * version)

Returns a version number in `version` corresponding to the capabilities of the context (e.g. 3010 or 3020), which library developers can use to direct callers to a specific API version. If `ctx` is NULL, returns the API version used to create the currently bound context.

Note that new API versions are only introduced when context capabilities are changed that break binary compatibility, so the API version and driver version may be different. For example, it is valid for the API version to be 3020 while the driver version is 4020.

Parameters:

ctx - Context to check
version - Pointer to version

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuCtxCreate, cuCtxDestroy, cuCtxGetDevice, cuCtxGetLimit, cuCtxPopCurrent, cuCtxPushCurrent, cuCtxSetCacheConfig, cuCtxSetLimit, cuCtxSynchronize

5.34.2.4 CUresult cuCtxGetCacheConfig (CUfunc_cache * pconfig)

On devices where the L1 cache and shared memory use the same hardware resources, this function returns through `pconfig` the preferred cache configuration for the current context. This is only a preference. The driver will use the requested configuration if possible, but it is free to choose a different configuration if required to execute functions.

This will return a `pconfig` of `CU_FUNC_CACHE_PREFER_NONE` on devices where the size of the L1 cache and shared memory are fixed.

The supported cache configurations are:

- `CU_FUNC_CACHE_PREFER_NONE`: no preference for shared memory or L1 (default)
- `CU_FUNC_CACHE_PREFER_SHARED`: prefer larger shared memory and smaller L1 cache
- `CU_FUNC_CACHE_PREFER_L1`: prefer larger L1 cache and smaller shared memory
- `CU_FUNC_CACHE_PREFER_EQUAL`: prefer equal sized L1 cache and shared memory

Parameters:

pconfig - Returned cache configuration

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuCtxCreate, cuCtxDestroy, cuCtxGetApiVersion, cuCtxGetDevice, cuCtxGetLimit, cuCtxPopCurrent, cuCtxPushCurrent, cuCtxSetCacheConfig, cuCtxSetLimit, cuCtxSynchronize, cuFuncSetCacheConfig

5.34.2.5 CUresult cuCtxGetCurrent (CUcontext * *pctx*)

Returns in **pctx* the CUDA context bound to the calling CPU thread. If no context is bound to the calling CPU thread then **pctx* is set to NULL and [CUDA_SUCCESS](#) is returned.

Parameters:

pctx - Returned context handle

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxSetCurrent](#), [cuCtxCreate](#), [cuCtxDestroy](#)

5.34.2.6 CUresult cuCtxGetDevice (CUdevice * *device*)

Returns in **device* the ordinal of the current context's device.

Parameters:

device - Returned device ID for the current context

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxDestroy](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetLimit](#), [cuCtxPopCurrent](#), [cuCtxPushCurrent](#), [cuCtxSetCacheConfig](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#)

5.34.2.7 CUresult cuCtxGetLimit (size_t * *pvalue*, CUlimit *limit*)

Returns in **pvalue* the current size of *limit*. The supported [CUlimit](#) values are:

- [CU_LIMIT_STACK_SIZE](#): stack size in bytes of each GPU thread.
- [CU_LIMIT_PRINTF_FIFO_SIZE](#): size in bytes of the FIFO used by the `printf()` device system call.
- [CU_LIMIT_MALLOC_HEAP_SIZE](#): size in bytes of the heap used by the `malloc()` and `free()` device system calls.
- [CU_LIMIT_DEV_RUNTIME_SYNC_DEPTH](#): maximum grid depth at which a thread can issue the device runtime call `cudaDeviceSynchronize()` to wait on child grid launches to complete.

- [CU_LIMIT_DEV_RUNTIME_PENDING_LAUNCH_COUNT](#): maximum number of outstanding device runtime launches that can be made from this context.

Parameters:

limit - Limit to query

pvalue - Returned size of limit

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_UNSUPPORTED_LIMIT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxDestroy](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxPopCurrent](#), [cuCtxPushCurrent](#), [cuCtxSetCacheConfig](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#)

5.34.2.8 CUresult cuCtxGetSharedMemConfig (CUsharedconfig * pConfig)

This function will return in `pConfig` the current size of shared memory banks in the current context. On devices with configurable shared memory banks, [cuCtxSetSharedMemConfig](#) can be used to change this setting, so that all subsequent kernel launches will by default use the new bank size. When [cuCtxGetSharedMemConfig](#) is called on devices without configurable shared memory, it will return the fixed bank size of the hardware.

The returned bank configurations can be either:

- [CU_SHARED_MEM_CONFIG_FOUR_BYTE_BANK_SIZE](#): shared memory bank width is four bytes.
- [CU_SHARED_MEM_CONFIG_EIGHT_BYTE_BANK_SIZE](#): shared memory bank width will eight bytes.

Parameters:

pConfig - returned shared memory configuration

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxDestroy](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxGetLimit](#), [cuCtxPopCurrent](#), [cuCtxPushCurrent](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#), [cuCtxGetSharedMemConfig](#), [cuFuncSetCacheConfig](#),

5.34.2.9 CUresult cuCtxPopCurrent (CUcontext * *pctx*)

Pops the current CUDA context from the CPU thread and passes back the old context handle in **pctx*. That context may then be made current to a different CPU thread by calling [cuCtxPushCurrent\(\)](#).

If a context was current to the CPU thread before [cuCtxCreate\(\)](#) or [cuCtxPushCurrent\(\)](#) was called, this function makes that context current to the CPU thread again.

Parameters:

pctx - Returned new context handle

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxDestroy](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxGetLimit](#), [cuCtxPushCurrent](#), [cuCtxSetCacheConfig](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#)

5.34.2.10 CUresult cuCtxPushCurrent (CUcontext *ctx*)

Pushes the given context *ctx* onto the CPU thread's stack of current contexts. The specified context becomes the CPU thread's current context, so all CUDA functions that operate on the current context are affected.

The previous current context may be made current again by calling [cuCtxDestroy\(\)](#) or [cuCtxPopCurrent\(\)](#).

Parameters:

ctx - Context to push

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxDestroy](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxGetLimit](#), [cuCtxPopCurrent](#), [cuCtxSetCacheConfig](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#)

5.34.2.11 CUresult cuCtxSetCacheConfig (CUfunc_cache *config*)

On devices where the L1 cache and shared memory use the same hardware resources, this sets through *config* the preferred cache configuration for the current context. This is only a preference. The driver will use the requested

configuration if possible, but it is free to choose a different configuration if required to execute the function. Any function preference set via `cuFuncSetCacheConfig()` will be preferred over this context-wide setting. Setting the context-wide cache configuration to `CU_FUNC_CACHE_PREFER_NONE` will cause subsequent kernel launches to prefer to not change the cache configuration unless required to launch the kernel.

This setting does nothing on devices where the size of the L1 cache and shared memory are fixed.

Launching a kernel with a different preference than the most recent preference setting may insert a device-side synchronization point.

The supported cache configurations are:

- `CU_FUNC_CACHE_PREFER_NONE`: no preference for shared memory or L1 (default)
- `CU_FUNC_CACHE_PREFER_SHARED`: prefer larger shared memory and smaller L1 cache
- `CU_FUNC_CACHE_PREFER_L1`: prefer larger L1 cache and smaller shared memory
- `CU_FUNC_CACHE_PREFER_EQUAL`: prefer equal sized L1 cache and shared memory

Parameters:

config - Requested cache configuration

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuCtxCreate`, `cuCtxDestroy`, `cuCtxGetApiVersion`, `cuCtxGetCacheConfig`, `cuCtxGetDevice`, `cuCtxGetLimit`, `cuCtxPopCurrent`, `cuCtxPushCurrent`, `cuCtxSetLimit`, `cuCtxSynchronize`, `cuFuncSetCacheConfig`

5.34.2.12 CUresult cuCtxSetCurrent (CUcontext ctx)

Binds the specified CUDA context to the calling CPU thread. If `ctx` is `NULL` then the CUDA context previously bound to the calling CPU thread is unbound and `CUDA_SUCCESS` is returned.

If there exists a CUDA context stack on the calling CPU thread, this will replace the top of that stack with `ctx`. If `ctx` is `NULL` then this will be equivalent to popping the top of the calling CPU thread's CUDA context stack (or a no-op if the calling CPU thread's CUDA context stack is empty).

Parameters:

ctx - Context to bind to the calling CPU thread

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxGetCurrent](#), [cuCtxCreate](#), [cuCtxDestroy](#)

5.34.2.13 CUresult cuCtxSetLimit (CUlimit *limit*, size_t *value*)

Setting *limit* to *value* is a request by the application to update the current limit maintained by the context. The driver is free to modify the requested value to meet h/w requirements (this could be clamping to minimum or maximum values, rounding up to nearest element size, etc). The application can use [cuCtxGetLimit\(\)](#) to find out exactly what the limit has been set to.

Setting each [CUlimit](#) has its own specific restrictions, so each is discussed here.

- [CU_LIMIT_STACK_SIZE](#) controls the stack size in bytes of each GPU thread. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [CUDA_ERROR_UNSUPPORTED_LIMIT](#) being returned.
- [CU_LIMIT_PRINTF_FIFO_SIZE](#) controls the size in bytes of the FIFO used by the `printf()` device system call. Setting [CU_LIMIT_PRINTF_FIFO_SIZE](#) must be performed before launching any kernel that uses the `printf()` device system call, otherwise [CUDA_ERROR_INVALID_VALUE](#) will be returned. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [CUDA_ERROR_UNSUPPORTED_LIMIT](#) being returned.
- [CU_LIMIT_MALLOC_HEAP_SIZE](#) controls the size in bytes of the heap used by the `malloc()` and `free()` device system calls. Setting [CU_LIMIT_MALLOC_HEAP_SIZE](#) must be performed before launching any kernel that uses the `malloc()` or `free()` device system calls, otherwise [CUDA_ERROR_INVALID_VALUE](#) will be returned. This limit is only applicable to devices of compute capability 2.0 and higher. Attempting to set this limit on devices of compute capability less than 2.0 will result in the error [CUDA_ERROR_UNSUPPORTED_LIMIT](#) being returned.
- [CU_LIMIT_DEV_RUNTIME_SYNC_DEPTH](#) controls the maximum nesting depth of a grid at which a thread can safely call `cudaDeviceSynchronize()`. Setting this limit must be performed before any launch of a kernel that uses the device runtime and calls `cudaDeviceSynchronize()` above the default sync depth, two levels of grids. Calls to `cudaDeviceSynchronize()` will fail with error code [cudaErrorSyncDepthExceeded](#) if the limitation is violated. This limit can be set smaller than the default or up the maximum launch depth of 24. When setting this limit, keep in mind that additional levels of sync depth require the driver to reserve large amounts of device memory which can no longer be used for user allocations. If these reservations of device memory fail, `cuCtxSetLimit` will return [CUDA_ERROR_OUT_OF_MEMORY](#), and the limit can be reset to a lower value. This limit is only applicable to devices of compute capability 3.5 and higher. Attempting to set this limit on devices of compute capability less than 3.5 will result in the error [CUDA_ERROR_UNSUPPORTED_LIMIT](#) being returned.
- [CU_LIMIT_DEV_RUNTIME_PENDING_LAUNCH_COUNT](#) controls the maximum number of outstanding device runtime launches that can be made from the current context. A grid is outstanding from the point of launch up until the grid is known to have been completed. Device runtime launches which violate this limitation fail and return [cudaErrorLaunchPendingCountExceeded](#) when `cudaGetLastError()` is called after launch. If more pending launches than the default (2048 launches) are needed for a module using the device runtime, this limit can be increased. Keep in mind that being able to sustain additional pending launches will require the driver to reserve larger amounts of device memory upfront which can no longer be used for allocations. If these reservations fail, `cuCtxSetLimit` will return [CUDA_ERROR_OUT_OF_MEMORY](#), and the limit can be reset to a lower value. This limit is only applicable to devices of compute capability 3.5 and higher. Attempting to set this limit on devices of compute capability less than 3.5 will result in the error [CUDA_ERROR_UNSUPPORTED_LIMIT](#) being returned.

Parameters:

limit - Limit to set
value - Size of limit

Returns:

CUDA_SUCCESS, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_UNSUPPORTED_LIMIT, CUDA_ERROR_OUT_OF_MEMORY

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxDestroy](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxGetLimit](#), [cuCtxPopCurrent](#), [cuCtxPushCurrent](#), [cuCtxSetCacheConfig](#), [cuCtxSynchronize](#)

5.34.2.14 CUresult cuCtxSetSharedMemConfig (CUsharedconfig *config*)

On devices with configurable shared memory banks, this function will set the context's shared memory bank size which is used for subsequent kernel launches.

Changing the shared memory configuration between launches may insert a device side synchronization point between those launches.

Changing the shared memory bank size will not increase shared memory usage or affect occupancy of kernels, but may have major effects on performance. Larger bank sizes will allow for greater potential bandwidth to shared memory, but will change what kinds of accesses to shared memory will result in bank conflicts.

This function will do nothing on devices with fixed shared memory bank size.

The supported bank configurations are:

- [CU_SHARED_MEM_CONFIG_DEFAULT_BANK_SIZE](#): set bank width to the default initial setting (currently, four bytes).
- [CU_SHARED_MEM_CONFIG_FOUR_BYTE_BANK_SIZE](#): set shared memory bank width to be natively four bytes.
- [CU_SHARED_MEM_CONFIG_EIGHT_BYTE_BANK_SIZE](#): set shared memory bank width to be natively eight bytes.

Parameters:

config - requested shared memory configuration

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxDestroy](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxGetLimit](#), [cuCtxPopCurrent](#), [cuCtxPushCurrent](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#), [cuCtxGetSharedMemConfig](#), [cuFuncSetCacheConfig](#),

5.34.2.15 CUresult cuCtxSynchronize (void)

Blocks until the device has completed all preceding requested tasks. `cuCtxSynchronize()` returns an error if one of the preceding tasks failed. If the context was created with the `CU_CTX_SCHED_BLOCKING_SYNC` flag, the CPU thread will block until the GPU context has finished its work.

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuCtxCreate`, `cuCtxDestroy`, `cuCtxGetApiVersion`, `cuCtxGetCacheConfig`, `cuCtxGetDevice`, `cuCtxGetLimit`, `cuCtxPopCurrent`, `cuCtxPushCurrent`, `cuCtxSetCacheConfig`, `cuCtxSetLimit`

5.35 Context Management [DEPRECATED]

Functions

- [CUresult cuCtxAttach](#) (CUcontext *pctx, unsigned int flags)

Increment a context's usage-count.

- [CUresult cuCtxDetach](#) (CUcontext ctx)

Decrement a context's usage-count.

5.35.1 Detailed Description

This section describes the deprecated context management functions of the low-level CUDA driver application programming interface.

5.35.2 Function Documentation

5.35.2.1 CUresult cuCtxAttach (CUcontext * pctx, unsigned int flags)

Deprecated

Note that this function is deprecated and should not be used.

Increments the usage count of the context and passes back a context handle in *pctx that must be passed to [cuCtxDetach\(\)](#) when the application is done with the context. [cuCtxAttach\(\)](#) fails if there is no context current to the thread.

Currently, the `flags` parameter must be 0.

Parameters:

pctx - Returned context handle of the current context

flags - Context attach flags (must be 0)

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuCtxDestroy](#), [cuCtxDetach](#), [cuCtxGetApiVersion](#), [cuCtxGetCacheConfig](#), [cuCtxGetDevice](#), [cuCtxGetLimit](#), [cuCtxPopCurrent](#), [cuCtxPushCurrent](#), [cuCtxSetCacheConfig](#), [cuCtxSetLimit](#), [cuCtxSynchronize](#)

5.35.2.2 CUresult cuCtxDetach (CUcontext *ctx*)

Deprecated

Note that this function is deprecated and should not be used.

Decrements the usage count of the context `ctx`, and destroys the context if the usage count goes to 0. The context must be a handle that was passed back by `cuCtxCreate()` or `cuCtxAttach()`, and must be current to the calling thread.

Parameters:

ctx - Context to destroy

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuCtxCreate`, `cuCtxDestroy`, `cuCtxGetApiVersion`, `cuCtxGetCacheConfig`, `cuCtxGetDevice`, `cuCtxGetLimit`, `cuCtxPopCurrent`, `cuCtxPushCurrent`, `cuCtxSetCacheConfig`, `cuCtxSetLimit`, `cuCtxSynchronize`

5.36 Module Management

Functions

- **CUresult cuModuleGetFunction** (CUfunction *hfunc, CUmodule hmod, const char *name)
Returns a function handle.
- **CUresult cuModuleGetGlobal** (CUdeviceptr *dptr, size_t *bytes, CUmodule hmod, const char *name)
Returns a global pointer from a module.
- **CUresult cuModuleGetSurfRef** (CUSurfref *pSurfRef, CUmodule hmod, const char *name)
Returns a handle to a surface reference.
- **CUresult cuModuleGetTexRef** (CUTexref *pTexRef, CUmodule hmod, const char *name)
Returns a handle to a texture reference.
- **CUresult cuModuleLoad** (CUmodule *module, const char *fname)
Loads a compute module.
- **CUresult cuModuleLoadData** (CUmodule *module, const void *image)
Load a module's data.
- **CUresult cuModuleLoadDataEx** (CUmodule *module, const void *image, unsigned int numOptions, CUjit_option *options, void **optionValues)
Load a module's data with options.
- **CUresult cuModuleLoadFatBinary** (CUmodule *module, const void *fatCubin)
Load a module's data.
- **CUresult cuModuleUnload** (CUmodule hmod)
Unloads a module.

5.36.1 Detailed Description

This section describes the module management functions of the low-level CUDA driver application programming interface.

5.36.2 Function Documentation

5.36.2.1 CUresult cuModuleGetFunction (CUfunction * hfunc, CUmodule hmod, const char * name)

Returns in *hfunc the handle of the function of name name located in module hmod. If no function of that name exists, cuModuleGetFunction() returns CUDA_ERROR_NOT_FOUND.

Parameters:

- hfunc* - Returned function handle
- hmod* - Module to retrieve function from
- name* - Name of function to retrieve

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_FOUND](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetGlobal](#), [cuModuleGetTexRef](#), [cuModuleLoad](#), [cuModuleLoadData](#), [cuModuleLoadDataEx](#), [cuModuleLoadFatBinary](#), [cuModuleUnload](#)

5.36.2.2 CUresult cuModuleGetGlobal (CUdeviceptr * *dptr*, size_t * *bytes*, CUmodule *hmod*, const char * *name*)

Returns in *dptr* and *bytes* the base pointer and size of the global of name *name* located in module *hmod*. If no variable of that name exists, [cuModuleGetGlobal\(\)](#) returns [CUDA_ERROR_NOT_FOUND](#). Both parameters *dptr* and *bytes* are optional. If one of them is NULL, it is ignored.

Parameters:

dptr - Returned global device pointer
bytes - Returned global size in bytes
hmod - Module to retrieve global from
name - Name of global to retrieve

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_FOUND](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetFunction](#), [cuModuleGetTexRef](#), [cuModuleLoad](#), [cuModuleLoadData](#), [cuModuleLoadDataEx](#), [cuModuleLoadFatBinary](#), [cuModuleUnload](#)

5.36.2.3 CUresult cuModuleGetSurfRef (CUSurfref * *pSurfRef*, CUmodule *hmod*, const char * *name*)

Returns in *pSurfRef* the handle of the surface reference of name *name* in the module *hmod*. If no surface reference of that name exists, [cuModuleGetSurfRef\(\)](#) returns [CUDA_ERROR_NOT_FOUND](#).

Parameters:

pSurfRef - Returned surface reference
hmod - Module to retrieve surface reference from
name - Name of surface reference to retrieve

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_NOT_FOUND

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetFunction](#), [cuModuleGetGlobal](#), [cuModuleGetTexRef](#), [cuModuleLoad](#), [cuModuleLoadData](#), [cuModuleLoadDataEx](#), [cuModuleLoadFatBinary](#), [cuModuleUnload](#)

5.36.2.4 CUresult cuModuleGetTexRef (CUtexref * *pTexRef*, CUmodule *hmod*, const char * *name*)

Returns in **pTexRef* the handle of the texture reference of name *name* in the module *hmod*. If no texture reference of that name exists, [cuModuleGetTexRef\(\)](#) returns [CUDA_ERROR_NOT_FOUND](#). This texture reference handle should not be destroyed, since it will be destroyed when the module is unloaded.

Parameters:

pTexRef - Returned texture reference
hmod - Module to retrieve texture reference from
name - Name of texture reference to retrieve

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_NOT_FOUND

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetFunction](#), [cuModuleGetGlobal](#), [cuModuleGetSurfRef](#), [cuModuleLoad](#), [cuModuleLoadData](#), [cuModuleLoadDataEx](#), [cuModuleLoadFatBinary](#), [cuModuleUnload](#)

5.36.2.5 CUresult cuModuleLoad (CUmodule * *module*, const char * *fname*)

Takes a filename *fname* and loads the corresponding module *module* into the current context. The CUDA driver API does not attempt to lazily allocate the resources needed by a module; if the memory for functions and data (constant and global) needed by the module cannot be allocated, [cuModuleLoad\(\)](#) fails. The file should be a *cubin* file as output by **nvcc**, or a *PTX* file either as output by **nvcc** or handwritten, or a *fatbin* file as output by **nvcc** from toolchain 4.0 or later.

Parameters:

module - Returned module
fname - Filename of module to load

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_NOT_FOUND, CUDA_ERROR_OUT_OF_MEMORY, CUDA_ERROR_FILE_NOT_FOUND, CUDA_ERROR_SHARED_OBJECT_SYMBOL_NOT_FOUND, CUDA_ERROR_SHARED_OBJECT_INIT_FAILED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetFunction](#), [cuModuleGetGlobal](#), [cuModuleGetTexRef](#), [cuModuleLoadData](#), [cuModuleLoadDataEx](#), [cuModuleLoadFatBinary](#), [cuModuleUnload](#)

5.36.2.6 CUresult cuModuleLoadData (CUmodule * module, const void * image)

Takes a pointer *image* and loads the corresponding module *module* into the current context. The pointer may be obtained by mapping a *cubin* or *PTX* or *fatbin* file, passing a *cubin* or *PTX* or *fatbin* file as a NULL-terminated text string, or incorporating a *cubin* or *fatbin* object into the executable resources and using operating system calls such as Windows `FindResource()` to obtain the pointer.

Parameters:

module - Returned module
image - Module data to load

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_OUT_OF_MEMORY, CUDA_ERROR_SHARED_OBJECT_SYMBOL_NOT_FOUND, CUDA_ERROR_SHARED_OBJECT_INIT_FAILED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetFunction](#), [cuModuleGetGlobal](#), [cuModuleGetTexRef](#), [cuModuleLoad](#), [cuModuleLoadDataEx](#), [cuModuleLoadFatBinary](#), [cuModuleUnload](#)

5.36.2.7 CUresult cuModuleLoadDataEx (CUmodule * module, const void * image, unsigned int numOptions, CUjit_option * options, void ** optionValues)

Takes a pointer *image* and loads the corresponding module *module* into the current context. The pointer may be obtained by mapping a *cubin* or *PTX* or *fatbin* file, passing a *cubin* or *PTX* or *fatbin* file as a NULL-terminated text string, or incorporating a *cubin* or *fatbin* object into the executable resources and using operating system calls such as Windows `FindResource()` to obtain the pointer. Options are passed as an array via *options* and any corresponding parameters are passed in *optionValues*. The number of total options is supplied via *numOptions*. Any outputs will be returned via *optionValues*. Supported options are (types for the option values are specified in parentheses after the option name):

- `CU_JIT_MAX_REGISTERS`: (unsigned int) input specifies the maximum number of registers per thread;
- `CU_JIT_THREADS_PER_BLOCK`: (unsigned int) input specifies number of threads per block to target compilation for; output returns the number of threads the compiler actually targeted;
- `CU_JIT_WALL_TIME`: (float) output returns the float value of wall clock time, in milliseconds, spent compiling the *PTX* code;
- `CU_JIT_INFO_LOG_BUFFER`: (char*) input is a pointer to a buffer in which to print any informational log messages from *PTX* assembly (the buffer size is specified via option `CU_JIT_INFO_LOG_BUFFER_SIZE_BYTES`);
- `CU_JIT_INFO_LOG_BUFFER_SIZE_BYTES`: (unsigned int) input is the size in bytes of the buffer; output is the number of bytes filled with messages;
- `CU_JIT_ERROR_LOG_BUFFER`: (char*) input is a pointer to a buffer in which to print any error log messages from *PTX* assembly (the buffer size is specified via option `CU_JIT_ERROR_LOG_BUFFER_SIZE_BYTES`);
- `CU_JIT_ERROR_LOG_BUFFER_SIZE_BYTES`: (unsigned int) input is the size in bytes of the buffer; output is the number of bytes filled with messages;
- `CU_JIT_OPTIMIZATION_LEVEL`: (unsigned int) input is the level of optimization to apply to generated code (0 - 4), with 4 being the default and highest level;
- `CU_JIT_TARGET_FROM_CUCONTEXT`: (No option value) causes compilation target to be determined based on current attached context (default);
- `CU_JIT_TARGET`: (unsigned int for enumerated type `CUjit_target_enum`) input is the compilation target based on supplied `CUjit_target_enum`; possible values are:
 - `CU_TARGET_COMPUTE_10`
 - `CU_TARGET_COMPUTE_11`
 - `CU_TARGET_COMPUTE_12`
 - `CU_TARGET_COMPUTE_13`
 - `CU_TARGET_COMPUTE_20`
- `CU_JIT_FALLBACK_STRATEGY`: (unsigned int for enumerated type `CUjit_fallback_enum`) chooses fallback strategy if matching cubin is not found; possible values are:
 - `CU_PREFER_PTX`
 - `CU_PREFER_BINARY`

Parameters:

module - Returned module

image - Module data to load

numOptions - Number of options

options - Options for JIT

optionValues - Option values for JIT

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`, `CUDA_ERROR_OUT_OF_MEMORY`, `CUDA_ERROR_NO_BINARY_FOR_GPU`, `CUDA_ERROR_SHARED_OBJECT_SYMBOL_NOT_FOUND`, `CUDA_ERROR_SHARED_OBJECT_INIT_FAILED`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetFunction](#), [cuModuleGetGlobal](#), [cuModuleGetTexRef](#), [cuModuleLoad](#), [cuModuleLoadData](#), [cuModuleLoadFatBinary](#), [cuModuleUnload](#)

5.36.2.8 CUresult cuModuleLoadFatBinary (CUmodule * *module*, const void * *fatCubin*)

Takes a pointer *fatCubin* and loads the corresponding module *module* into the current context. The pointer represents a *fat binary* object, which is a collection of different *cubin* and/or *PTX* files, all representing the same device code, but compiled and optimized for different architectures.

Prior to CUDA 4.0, there was no documented API for constructing and using fat binary objects by programmers. Starting with CUDA 4.0, fat binary objects can be constructed by providing the *-fatbin option* to **nvcc**. More information can be found in the **nvcc** document.

Parameters:

module - Returned module
fatCubin - Fat binary to load

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_FOUND](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_NO_BINARY_FOR_GPU](#), [CUDA_ERROR_SHARED_OBJECT_SYMBOL_NOT_FOUND](#), [CUDA_ERROR_SHARED_OBJECT_INIT_FAILED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetFunction](#), [cuModuleGetGlobal](#), [cuModuleGetTexRef](#), [cuModuleLoad](#), [cuModuleLoadData](#), [cuModuleLoadDataEx](#), [cuModuleUnload](#)

5.36.2.9 CUresult cuModuleUnload (CUmodule *hmod*)

Unloads a module *hmod* from the current context.

Parameters:

hmod - Module to unload

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuModuleGetFunction](#), [cuModuleGetGlobal](#), [cuModuleGetTexRef](#), [cuModuleLoad](#), [cuModuleLoadData](#), [cuModuleLoadDataEx](#), [cuModuleLoadFatBinary](#)

5.37 Memory Management

Functions

- **CUresult cuArray3DCreate** (CUarray *pHandle, const CUDA_ARRAY3D_DESCRIPTOR *pAllocateArray)
Creates a 3D CUDA array.
- **CUresult cuArray3DGetDescriptor** (CUDA_ARRAY3D_DESCRIPTOR *pArrayDescriptor, CUarray hArray)
Get a 3D CUDA array descriptor.
- **CUresult cuArrayCreate** (CUarray *pHandle, const CUDA_ARRAY_DESCRIPTOR *pAllocateArray)
Creates a 1D or 2D CUDA array.
- **CUresult cuArrayDestroy** (CUarray hArray)
Destroys a CUDA array.
- **CUresult cuArrayGetDescriptor** (CUDA_ARRAY_DESCRIPTOR *pArrayDescriptor, CUarray hArray)
Get a 1D or 2D CUDA array descriptor.
- **CUresult cuDeviceGetByPCIBusId** (CUdevice *dev, char *pciBusId)
Returns a handle to a compute device.
- **CUresult cuDeviceGetPCIBusId** (char *pciBusId, int len, CUdevice dev)
Returns a PCI Bus Id string for the device.
- **CUresult cuIpcCloseMemHandle** (CUdeviceptr dptr)
Close memory mapped with cuIpcOpenMemHandle.
- **CUresult cuIpcGetEventHandle** (CUipcEventHandle *pHandle, CUevent event)
Gets an interprocess handle for a previously allocated event.
- **CUresult cuIpcGetMemHandle** (CUipcMemHandle *pHandle, CUdeviceptr dptr)
- **CUresult cuIpcOpenEventHandle** (CUevent *phEvent, CUipcEventHandle handle)
Opens an interprocess event handle for use in the current process.
- **CUresult cuIpcOpenMemHandle** (CUdeviceptr *pdptr, CUipcMemHandle handle, unsigned int Flags)
- **CUresult cuMemAlloc** (CUdeviceptr *dptr, size_t bytesize)
Allocates device memory.
- **CUresult cuMemAllocHost** (void **pp, size_t bytesize)
Allocates page-locked host memory.
- **CUresult cuMemAllocPitch** (CUdeviceptr *dptr, size_t *pPitch, size_t WidthInBytes, size_t Height, unsigned int ElementSizeBytes)
Allocates pitched device memory.
- **CUresult cuMemcpy** (CUdeviceptr dst, CUdeviceptr src, size_t ByteCount)
Copies memory.

- **CUresult cuMemcpy2D** (const **CUDA_MEMCPY2D** *pCopy)
Copies memory for 2D arrays.
- **CUresult cuMemcpy2DAsync** (const **CUDA_MEMCPY2D** *pCopy, **CUstream** hStream)
Copies memory for 2D arrays.
- **CUresult cuMemcpy2DUnaligned** (const **CUDA_MEMCPY2D** *pCopy)
Copies memory for 2D arrays.
- **CUresult cuMemcpy3D** (const **CUDA_MEMCPY3D** *pCopy)
Copies memory for 3D arrays.
- **CUresult cuMemcpy3DAsync** (const **CUDA_MEMCPY3D** *pCopy, **CUstream** hStream)
Copies memory for 3D arrays.
- **CUresult cuMemcpy3DPeer** (const **CUDA_MEMCPY3D_PEER** *pCopy)
Copies memory between contexts.
- **CUresult cuMemcpy3DPeerAsync** (const **CUDA_MEMCPY3D_PEER** *pCopy, **CUstream** hStream)
Copies memory between contexts asynchronously.
- **CUresult cuMemcpyAsync** (**CUdeviceptr** dst, **CUdeviceptr** src, size_t ByteCount, **CUstream** hStream)
Copies memory asynchronously.
- **CUresult cuMemcpyAtoA** (**CUarray** dstArray, size_t dstOffset, **CUarray** srcArray, size_t srcOffset, size_t ByteCount)
Copies memory from Array to Array.
- **CUresult cuMemcpyAtoD** (**CUdeviceptr** dstDevice, **CUarray** srcArray, size_t srcOffset, size_t ByteCount)
Copies memory from Array to Device.
- **CUresult cuMemcpyAtoH** (void *dstHost, **CUarray** srcArray, size_t srcOffset, size_t ByteCount)
Copies memory from Array to Host.
- **CUresult cuMemcpyAtoHAsync** (void *dstHost, **CUarray** srcArray, size_t srcOffset, size_t ByteCount, **CUstream** hStream)
Copies memory from Array to Host.
- **CUresult cuMemcpyDtoA** (**CUarray** dstArray, size_t dstOffset, **CUdeviceptr** srcDevice, size_t ByteCount)
Copies memory from Device to Array.
- **CUresult cuMemcpyDtoD** (**CUdeviceptr** dstDevice, **CUdeviceptr** srcDevice, size_t ByteCount)
Copies memory from Device to Device.
- **CUresult cuMemcpyDtoDAsync** (**CUdeviceptr** dstDevice, **CUdeviceptr** srcDevice, size_t ByteCount, **CUstream** hStream)
Copies memory from Device to Device.
- **CUresult cuMemcpyDtoH** (void *dstHost, **CUdeviceptr** srcDevice, size_t ByteCount)
Copies memory from Device to Host.

- [CUresult cuMemcpyDtoHAsync](#) (void *dstHost, [CUdeviceptr](#) srcDevice, size_t ByteCount, [CUstream](#) hStream)
Copies memory from Device to Host.
- [CUresult cuMemcpyHtoA](#) ([CUarray](#) dstArray, size_t dstOffset, const void *srcHost, size_t ByteCount)
Copies memory from Host to Array.
- [CUresult cuMemcpyHtoAAsync](#) ([CUarray](#) dstArray, size_t dstOffset, const void *srcHost, size_t ByteCount, [CUstream](#) hStream)
Copies memory from Host to Array.
- [CUresult cuMemcpyHtoD](#) ([CUdeviceptr](#) dstDevice, const void *srcHost, size_t ByteCount)
Copies memory from Host to Device.
- [CUresult cuMemcpyHtoDAsync](#) ([CUdeviceptr](#) dstDevice, const void *srcHost, size_t ByteCount, [CUstream](#) hStream)
Copies memory from Host to Device.
- [CUresult cuMemcpyPeer](#) ([CUdeviceptr](#) dstDevice, [CUcontext](#) dstContext, [CUdeviceptr](#) srcDevice, [CUcontext](#) srcContext, size_t ByteCount)
Copies device memory between two contexts.
- [CUresult cuMemcpyPeerAsync](#) ([CUdeviceptr](#) dstDevice, [CUcontext](#) dstContext, [CUdeviceptr](#) srcDevice, [CUcontext](#) srcContext, size_t ByteCount, [CUstream](#) hStream)
Copies device memory between two contexts asynchronously.
- [CUresult cuMemFree](#) ([CUdeviceptr](#) dptr)
Frees device memory.
- [CUresult cuMemFreeHost](#) (void *p)
Frees page-locked host memory.
- [CUresult cuMemGetAddressRange](#) ([CUdeviceptr](#) *pbase, size_t *psize, [CUdeviceptr](#) dptr)
Get information on memory allocations.
- [CUresult cuMemGetInfo](#) (size_t *free, size_t *total)
Gets free and total memory.
- [CUresult cuMemHostAlloc](#) (void **pp, size_t bytesize, unsigned int Flags)
Allocates page-locked host memory.
- [CUresult cuMemHostGetDevicePointer](#) ([CUdeviceptr](#) *pdptr, void *p, unsigned int Flags)
Passes back device pointer of mapped pinned memory.
- [CUresult cuMemHostGetFlags](#) (unsigned int *pFlags, void *p)
Passes back flags that were used for a pinned allocation.
- [CUresult cuMemHostRegister](#) (void *p, size_t bytesize, unsigned int Flags)
Registers an existing host memory range for use by CUDA.

- **CUresult cuMemHostUnregister** (void *p)
Unregisters a memory range that was registered with cuMemHostRegister.
- **CUresult cuMemsetD16** (CUdeviceptr dstDevice, unsigned short us, size_t N)
Initializes device memory.
- **CUresult cuMemsetD16Async** (CUdeviceptr dstDevice, unsigned short us, size_t N, CUstream hStream)
Sets device memory.
- **CUresult cuMemsetD2D16** (CUdeviceptr dstDevice, size_t dstPitch, unsigned short us, size_t Width, size_t Height)
Initializes device memory.
- **CUresult cuMemsetD2D16Async** (CUdeviceptr dstDevice, size_t dstPitch, unsigned short us, size_t Width, size_t Height, CUstream hStream)
Sets device memory.
- **CUresult cuMemsetD2D32** (CUdeviceptr dstDevice, size_t dstPitch, unsigned int ui, size_t Width, size_t Height)
Initializes device memory.
- **CUresult cuMemsetD2D32Async** (CUdeviceptr dstDevice, size_t dstPitch, unsigned int ui, size_t Width, size_t Height, CUstream hStream)
Sets device memory.
- **CUresult cuMemsetD2D8** (CUdeviceptr dstDevice, size_t dstPitch, unsigned char uc, size_t Width, size_t Height)
Initializes device memory.
- **CUresult cuMemsetD2D8Async** (CUdeviceptr dstDevice, size_t dstPitch, unsigned char uc, size_t Width, size_t Height, CUstream hStream)
Sets device memory.
- **CUresult cuMemsetD32** (CUdeviceptr dstDevice, unsigned int ui, size_t N)
Initializes device memory.
- **CUresult cuMemsetD32Async** (CUdeviceptr dstDevice, unsigned int ui, size_t N, CUstream hStream)
Sets device memory.
- **CUresult cuMemsetD8** (CUdeviceptr dstDevice, unsigned char uc, size_t N)
Initializes device memory.
- **CUresult cuMemsetD8Async** (CUdeviceptr dstDevice, unsigned char uc, size_t N, CUstream hStream)
Sets device memory.
- **CUresult cuMipmappedArrayCreate** (CUmipmappedArray *pHandle, const CUDA_ARRAY3D_DESCRIPTOR *pMipmappedArrayDesc, unsigned int numMipmapLevels)
Creates a CUDA mipmapped array.
- **CUresult cuMipmappedArrayDestroy** (CUmipmappedArray hMipmappedArray)
Destroys a CUDA mipmapped array.

- [CUresult cuMipmappedArrayGetLevel](#) ([CUarray](#) *pLevelArray, [CUmipmappedArray](#) hMipmappedArray, unsigned int level)

Gets a mipmap level of a CUDA mipmapped array.

5.37.1 Detailed Description

This section describes the memory management functions of the low-level CUDA driver application programming interface.

5.37.2 Function Documentation

5.37.2.1 [CUresult cuArray3DCreate](#) ([CUarray](#) *pHandle, const [CUDA_ARRAY3D_DESCRIPTOR](#) *pAllocateArray)

Creates a CUDA array according to the [CUDA_ARRAY3D_DESCRIPTOR](#) structure pAllocateArray and returns a handle to the new CUDA array in *pHandle. The [CUDA_ARRAY3D_DESCRIPTOR](#) is defined as:

```
typedef struct {
    unsigned int Width;
    unsigned int Height;
    unsigned int Depth;
    CUarray_format Format;
    unsigned int NumChannels;
    unsigned int Flags;
} CUDA_ARRAY3D_DESCRIPTOR;
```

where:

- Width, Height, and Depth are the width, height, and depth of the CUDA array (in elements); the following types of CUDA arrays can be allocated:
 - A 1D array is allocated if Height and Depth extents are both zero.
 - A 2D array is allocated if only Depth extent is zero.
 - A 3D array is allocated if all three extents are non-zero.
 - A 1D layered CUDA array is allocated if only Height is zero and the [CUDA_ARRAY3D_LAYERED](#) flag is set. Each layer is a 1D array. The number of layers is determined by the depth extent.
 - A 2D layered CUDA array is allocated if all three extents are non-zero and the [CUDA_ARRAY3D_LAYERED](#) flag is set. Each layer is a 2D array. The number of layers is determined by the depth extent.
 - A cubemap CUDA array is allocated if all three extents are non-zero and the [CUDA_ARRAY3D_CUBEMAP](#) flag is set. Width must be equal to Height, and Depth must be six. A cubemap is a special type of 2D layered CUDA array, where the six layers represent the six faces of a cube. The order of the six layers in memory is the same as that listed in [CUarray_cubemap_face](#).
 - A cubemap layered CUDA array is allocated if all three extents are non-zero, and both, [CUDA_ARRAY3D_CUBEMAP](#) and [CUDA_ARRAY3D_LAYERED](#) flags are set. Width must be equal to Height, and Depth must be a multiple of six. A cubemap layered CUDA array is a special type of 2D layered CUDA array that consists of a collection of cubemaps. The first six layers represent the first cubemap, the next six layers form the second cubemap, and so on.
- Format specifies the format of the elements; [CUarray_format](#) is defined as:

```
typedef enum CUarray_format_enum {
    CU_AD_FORMAT_UNSIGNED_INT8 = 0x01,
    CU_AD_FORMAT_UNSIGNED_INT16 = 0x02,
    CU_AD_FORMAT_UNSIGNED_INT32 = 0x03,
    CU_AD_FORMAT_SIGNED_INT8 = 0x08,
    CU_AD_FORMAT_SIGNED_INT16 = 0x09,
    CU_AD_FORMAT_SIGNED_INT32 = 0x0a,
    CU_AD_FORMAT_HALF = 0x10,
    CU_AD_FORMAT_FLOAT = 0x20
} CUarray_format;
```

- NumChannels specifies the number of packed components per CUDA array element; it may be 1, 2, or 4;

- Flags may be set to

- [CUDA_ARRAY3D_LAYERED](#) to enable creation of layered CUDA arrays. If this flag is set, Depth specifies the number of layers, not the depth of a 3D array.
- [CUDA_ARRAY3D_SURFACE_LDST](#) to enable surface references to be bound to the CUDA array. If this flag is not set, `cuSurfRefSetArray` will fail when attempting to bind the CUDA array to a surface reference.
- [CUDA_ARRAY3D_CUBEMAP](#) to enable creation of cubemaps. If this flag is set, Width must be equal to Height, and Depth must be six. If the [CUDA_ARRAY3D_LAYERED](#) flag is also set, then Depth must be a multiple of six.
- [CUDA_ARRAY3D_TEXTURE_GATHER](#) to indicate that the CUDA array will be used for texture gather. Texture gather can only be performed on 2D CUDA arrays.

Width, Height and Depth must meet certain size requirements as listed in the following table. All values are specified in elements. Note that for brevity's sake, the full name of the device attribute is not specified. For ex., TEXTURE1D_WIDTH refers to the device attribute [CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_WIDTH](#).

Note that 2D CUDA arrays have different size requirements if the [CUDA_ARRAY3D_TEXTURE_GATHER](#) flag is set. Width and Height must not be greater than [CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_GATHER_WIDTH](#) and [CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_GATHER_HEIGHT](#) respectively, in that case.

CUDA array type	Valid extents that must always be met {(width range in elements), (height range), (depth range)}	Valid extents with CUDA_- ARRAY3D_SURFACE_LDST set {(width range in elements), (height range), (depth range)}
1D	{ (1,TEXTURE1D_WIDTH), 0, 0 }	{ (1,SURFACE1D_WIDTH), 0, 0 }
2D	{ (1,TEXTURE2D_WIDTH), (1,TEXTURE2D_HEIGHT), 0 }	{ (1,SURFACE2D_WIDTH), (1,SURFACE2D_HEIGHT), 0 }
3D	{ (1,TEXTURE3D_WIDTH), (1,TEXTURE3D_HEIGHT), (1,TEXTURE3D_DEPTH) } OR { (1,TEXTURE3D_WIDTH_ALTERNATE), (1,TEXTURE3D_HEIGHT_ALTERNATE), (1,TEXTURE3D_DEPTH_ALTERNATE) }	{ (1,SURFACE3D_WIDTH), (1,SURFACE3D_HEIGHT), (1,SURFACE3D_DEPTH) }
1D Layered	{ (1,TEXTURE1D_LAYERED_WIDTH), 0, (1,TEXTURE1D_LAYERED_LAYERS) }	{ (1,SURFACE1D_LAYERED_WIDTH), 0, (1,SURFACE1D_LAYERED_LAYERS) }
2D Layered	{ (1,TEXTURE2D_LAYERED_WIDTH), (1,TEXTURE2D_LAYERED_HEIGHT), (1,TEXTURE2D_LAYERED_LAYERS) }	{ (1,SURFACE2D_LAYERED_WIDTH), (1,SURFACE2D_LAYERED_HEIGHT), (1,SURFACE2D_LAYERED_LAYERS) }
Cubemap	{ (1,TEXTURECUBEMAP_WIDTH), (1,TEXTURECUBEMAP_WIDTH), 6 }	{ (1,SURFACECUBEMAP_WIDTH), (1,SURFACECUBEMAP_WIDTH), 6 }
Cubemap Layered	{ (1,TEXTURECUBEMAP_LAYERED_- WIDTH), (1,TEXTURECUBEMAP_LAYERED_- WIDTH), (1,TEXTURECUBEMAP_LAYERED_- LAYERS) }	{ (1,SURFACECUBEMAP_LAYERED_- WIDTH), (1,SURFACECUBEMAP_LAYERED_- WIDTH), (1,SURFACECUBEMAP_LAYERED_- LAYERS) }

Here are examples of CUDA array descriptions:

Description for a CUDA array of 2048 floats:

```

CUDA_ARRAY3D_DESCRIPTOR desc;
desc.Format = CU_AD_FORMAT_FLOAT;
desc.NumChannels = 1;
desc.Width = 2048;
desc.Height = 0;
desc.Depth = 0;

```

Description for a 64 x 64 CUDA array of floats:

```

CUDA_ARRAY3D_DESCRIPTOR desc;
desc.Format = CU_AD_FORMAT_FLOAT;
desc.NumChannels = 1;
desc.Width = 64;
desc.Height = 64;
desc.Depth = 0;

```

Description for a width x height x depth CUDA array of 64-bit, 4x16-bit float16's:

```

CUDA_ARRAY3D_DESCRIPTOR desc;
desc.FormatFlags = CU_AD_FORMAT_HALF;
desc.NumChannels = 4;
desc.Width = width;
desc.Height = height;
desc.Depth = depth;

```

Parameters:

pHandle - Returned array
pAllocateArray - 3D array descriptor

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_OUT_OF_MEMORY, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.2 CUresult cuArray3DGetDescriptor (CUDA_ARRAY3D_DESCRIPTOR * pArrayDescriptor, CUarray hArray)

Returns in *pArrayDescriptor a descriptor containing information on the format and dimensions of the CUDA array hArray. It is useful for subroutines that have been passed a CUDA array, but need to know the CUDA array parameters for validation or other purposes.

This function may be called on 1D and 2D arrays, in which case the Height and/or Depth members of the descriptor struct will be set to 0.

Parameters:

pArrayDescriptor - Returned 3D array descriptor
hArray - 3D array to get descriptor of

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_INVALID_HANDLE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.3 CUresult cuArrayCreate (CUarray *pHandle, const CUDA_ARRAY_DESCRIPTOR *pAllocateArray)

Creates a CUDA array according to the [CUDA_ARRAY_DESCRIPTOR](#) structure `pAllocateArray` and returns a handle to the new CUDA array in `*pHandle`. The [CUDA_ARRAY_DESCRIPTOR](#) is defined as:

```
typedef struct {
    unsigned int Width;
    unsigned int Height;
    CUarray_format Format;
    unsigned int NumChannels;
} CUDA_ARRAY_DESCRIPTOR;
```

where:

- `Width`, and `Height` are the width, and height of the CUDA array (in elements); the CUDA array is one-dimensional if height is 0, two-dimensional otherwise;
- `Format` specifies the format of the elements; [CUarray_format](#) is defined as:

```
typedef enum CUarray_format_enum {
    CU_AD_FORMAT_UNSIGNED_INT8 = 0x01,
    CU_AD_FORMAT_UNSIGNED_INT16 = 0x02,
    CU_AD_FORMAT_UNSIGNED_INT32 = 0x03,
    CU_AD_FORMAT_SIGNED_INT8 = 0x08,
    CU_AD_FORMAT_SIGNED_INT16 = 0x09,
    CU_AD_FORMAT_SIGNED_INT32 = 0x0a,
    CU_AD_FORMAT_HALF = 0x10,
    CU_AD_FORMAT_FLOAT = 0x20
} CUarray_format;
```

- `NumChannels` specifies the number of packed components per CUDA array element; it may be 1, 2, or 4;

Here are examples of CUDA array descriptions:

Description for a CUDA array of 2048 floats:

```
CUDA_ARRAY_DESCRIPTOR desc;
desc.Format = CU_AD_FORMAT_FLOAT;
desc.NumChannels = 1;
desc.Width = 2048;
desc.Height = 1;
```

Description for a 64 x 64 CUDA array of floats:

```
CUDA_ARRAY_DESCRIPTOR desc;
desc.Format = CU_AD_FORMAT_FLOAT;
desc.NumChannels = 1;
desc.Width = 64;
desc.Height = 64;
```

Description for a width x height CUDA array of 64-bit, 4x16-bit float16's:

```
CUDA_ARRAY_DESCRIPTOR desc;
desc.FormatFlags = CU_AD_FORMAT_HALF;
desc.NumChannels = 4;
desc.Width = width;
desc.Height = height;
```

Description for a `width x height` CUDA array of 16-bit elements, each of which is two 8-bit unsigned chars:

```

CUDA_ARRAY_DESCRIPTOR arrayDesc;
desc.FormatFlags = CU_AD_FORMAT_UNSIGNED_INT8;
desc.NumChannels = 2;
desc.Width = width;
desc.Height = height;

```

Parameters:

pHandle - Returned array
pAllocateArray - Array descriptor

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_OUT_OF_MEMORY, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.4 CUresult cuArrayDestroy (CUarray hArray)

Destroys the CUDA array `hArray`.

Parameters:

hArray - Array to destroy

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_ARRAY_IS_MAPPED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.5 CUresult cuArrayGetDescriptor (CUDA_ARRAY_DESCRIPTOR * *pArrayDescriptor*, CUarray *hArray*)

Returns in **pArrayDescriptor* a descriptor containing information on the format and dimensions of the CUDA array *hArray*. It is useful for subroutines that have been passed a CUDA array, but need to know the CUDA array parameters for validation or other purposes.

Parameters:

pArrayDescriptor - Returned array descriptor
hArray - Array to get descriptor of

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.6 CUresult cuDeviceGetByPCIBusId (CUdevice * *dev*, char * *pciBusId*)

Returns in **device* a device handle given a PCI bus ID string.

Parameters:

dev - Returned device handle
pciBusId - String in one of the following forms: [domain]:[bus]:[device].[function] [domain]:[bus]:[device][bus]:[device].[function] where *domain*, *bus*, *device*, and *function* are all hexadecimal values

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_DEVICE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGet](#), [cuDeviceGetAttribute](#), [cuDeviceGetPCIBusId](#)

5.37.2.7 CUresult cuDeviceGetPCIBusId (char * pciBusId, int len, CUdevice dev)

Returns an ASCII string identifying the device `dev` in the NULL-terminated string pointed to by `pciBusId`. `len` specifies the maximum length of the string that may be returned.

Parameters:

- `pciBusId` - Returned identifier string for the device in the following format [domain]:[bus]:[device].[function] where `domain`, `bus`, `device`, and `function` are all hexadecimal values. `pciBusId` should be large enough to store 13 characters including the NULL-terminator.
- `len` - Maximum length of string to store in name
- `dev` - Device to get identifier string for

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_DEVICE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceGet](#), [cuDeviceGetAttribute](#), [cuDeviceGetByPCIBusId](#)

5.37.2.8 CUresult cuIpcCloseMemHandle (CUdeviceptr dptr)

Unmaps memory returned by [cuIpcOpenMemHandle](#). The original allocation in the exporting process as well as imported mappings in other processes will be unaffected.

Any resources used to enable peer access will be freed if this is the last mapping using them.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

- `dptr` - Device pointer returned by [cuIpcOpenMemHandle](#)

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_MAP_FAILED](#), [CUDA_ERROR_INVALID_HANDLE](#),

See also:

[cuMemAlloc](#), [cuMemFree](#), [cuIpcGetEventHandle](#), [cuIpcOpenEventHandle](#), [cuIpcGetMemHandle](#), [cuIpcOpenMemHandle](#),

5.37.2.9 CUresult cuIpcGetEventHandle (CUipcEventHandle * pHandle, CUevent event)

Takes as input a previously allocated event. This event must have been created with the [CU_EVENT_INTERPROCESS](#) and [CU_EVENT_DISABLE_TIMING](#) flags set. This opaque handle may be copied into other processes and opened with [cuIpcOpenEventHandle](#) to allow efficient hardware synchronization between GPU work in different processes.

After the event has been opened in the importing process, [cuEventRecord](#), [cuEventSynchronize](#), [cuStreamWaitEvent](#) and [cuEventQuery](#) may be used in either process. Performing operations on the imported event after the exported event has been freed with [cuEventDestroy](#) will result in undefined behavior.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

pHandle - Pointer to a user allocated [CUipcEventHandle](#) in which to return the opaque event handle
event - Event allocated with [CU_EVENT_INTERPROCESS](#) and [CU_EVENT_DISABLE_TIMING](#) flags.

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_MAP_FAILED](#)

See also:

[cuEventCreate](#), [cuEventDestroy](#), [cuEventSynchronize](#), [cuEventQuery](#), [cuStreamWaitEvent](#), [cuIpcOpenEventHandle](#), [cuIpcGetMemHandle](#), [cuIpcOpenMemHandle](#), [cuIpcCloseMemHandle](#)

5.37.2.10 CUresult cuIpcGetMemHandle (CUipcMemHandle * pHandle, CUdeviceptr dptr)

/brief Gets an interprocess memory handle for an existing device memory allocation

Takes a pointer to the base of an existing device memory allocation created with [cuMemAlloc](#) and exports it for use in another process. This is a lightweight operation and may be called multiple times on an allocation without adverse effects.

If a region of memory is freed with [cuMemFree](#) and a subsequent call to [cuMemAlloc](#) returns memory with the same device address, [cuIpcGetMemHandle](#) will return a unique handle for the new memory.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

pHandle - Pointer to user allocated [CUipcMemHandle](#) to return the handle in.
dptr - Base pointer to previously allocated device memory

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_MAP_FAILED](#),

See also:

[cuMemAlloc](#), [cuMemFree](#), [cuIpcGetEventHandle](#), [cuIpcOpenEventHandle](#), [cuIpcOpenMemHandle](#), [cuIpcCloseMemHandle](#)

5.37.2.11 CUresult cuIpcOpenEventHandle (CUevent * phEvent, CUipcEventHandle handle)

Opens an interprocess event handle exported from another process with [cuIpcGetEventHandle](#). This function returns a [CUevent](#) that behaves like a locally created event with the [CU_EVENT_DISABLE_TIMING](#) flag specified. This event must be freed with [cuEventDestroy](#).

Performing operations on the imported event after the exported event has been freed with [cuEventDestroy](#) will result in undefined behavior.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

phEvent - Returns the imported event
handle - Interprocess handle to open

Returns:

CUDA_SUCCESS, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_MAP_FAILED, CUDA_ERROR_PEER_ACCESS_UNSUPPORTED, CUDA_ERROR_INVALID_HANDLE

See also:

cuEventCreate, cuEventDestroy, cuEventSynchronize, cuEventQuery, cuStreamWaitEvent, cuIpcGetEventHandle, cuIpcGetMemHandle, cuIpcOpenMemHandle, cuIpcCloseMemHandle

5.37.2.12 CUresult cuIpcOpenMemHandle (CUdeviceptr * *pdptr*, CUipcMemHandle *handle*, unsigned int *Flags*)

/brief Opens an interprocess memory handle exported from another process and returns a device pointer usable in the local process.

Maps memory exported from another process with [cuIpcGetMemHandle](#) into the current device address space. For contexts on different devices [cuIpcOpenMemHandle](#) can attempt to enable peer access between the devices as if the user called [cuCtxEnablePeerAccess](#). This behavior is controlled by the [CU_IPC_MEM_LAZY_ENABLE_PEER_ACCESS](#) flag. [cuDeviceCanAccessPeer](#) can determine if a mapping is possible.

Contexts that may open CUipcMemHandles are restricted in the following way. CUipcMemHandles from each CUdevice in a given process may only be opened by one CUcontext per CUdevice per other process.

Memory returned from [cuIpcOpenMemHandle](#) must be freed with [cuIpcCloseMemHandle](#).

Calling [cuMemFree](#) on an exported memory region before calling [cuIpcCloseMemHandle](#) in the importing context will result in undefined behavior.

IPC functionality is restricted to devices with support for unified addressing on Linux operating systems.

Parameters:

pdptr - Returned device pointer
handle - CUipcMemHandle to open
Flags - Flags for this operation. Must be specified as [CU_IPC_MEM_LAZY_ENABLE_PEER_ACCESS](#)

Returns:

CUDA_SUCCESS, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_MAP_FAILED, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_TOO_MANY_PEERS

See also:

cuMemAlloc, cuMemFree, cuIpcGetEventHandle, cuIpcOpenEventHandle, cuIpcGetMemHandle, cuIpcCloseMemHandle, cuCtxEnablePeerAccess, cuDeviceCanAccessPeer,

5.37.2.13 CUresult cuMemAlloc (CUdeviceptr * *dptr*, size_t *bytesize*)

Allocates *bytesize* bytes of linear memory on the device and returns in **dptr* a pointer to the allocated memory. The allocated memory is suitably aligned for any kind of variable. The memory is not cleared. If *bytesize* is 0, [cuMemAlloc\(\)](#) returns [CUDA_ERROR_INVALID_VALUE](#).

Parameters:

dptr - Returned device pointer
bytesize - Requested allocation size in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.14 CUresult cuMemAllocHost (void ** pp, size_t bytesize)

Allocates *bytesize* bytes of host memory that is page-locked and accessible to the device. The driver tracks the virtual memory ranges allocated with this function and automatically accelerates calls to functions such as [cuMemcpy\(\)](#). Since the memory can be accessed directly by the device, it can be read or written with much higher bandwidth than pageable memory obtained with functions such as `malloc()`. Allocating excessive amounts of memory with [cuMemAllocHost\(\)](#) may degrade system performance, since it reduces the amount of memory available to the system for paging. As a result, this function is best used sparingly to allocate staging areas for data exchange between host and device.

Note all host memory allocated using [cuMemHostAlloc\(\)](#) will automatically be immediately accessible to all contexts on all devices which support unified addressing (as may be queried using [CU_DEVICE_ATTRIBUTE_UNIFIED_ADDRESSING](#)). The device pointer that may be used to access this host memory from those contexts is always equal to the returned host pointer **pp*. See [Unified Addressing](#) for additional details.

Parameters:

pp - Returned host pointer to page-locked memory
bytesize - Requested allocation size in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#),

[cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.15 CUresult cuMemAllocPitch (CUdeviceptr * *dptr*, size_t * *pPitch*, size_t *WidthInBytes*, size_t *Height*, unsigned int *ElementSizeBytes*)

Allocates at least `WidthInBytes * Height` bytes of linear memory on the device and returns in `*dptr` a pointer to the allocated memory. The function may pad the allocation to ensure that corresponding pointers in any given row will continue to meet the alignment requirements for coalescing as the address is updated from row to row. `ElementSizeBytes` specifies the size of the largest reads and writes that will be performed on the memory range. `ElementSizeBytes` may be 4, 8 or 16 (since coalesced memory transactions are not possible on other data sizes). If `ElementSizeBytes` is smaller than the actual read/write size of a kernel, the kernel will run correctly, but possibly at reduced speed. The pitch returned in `*pPitch` by [cuMemAllocPitch\(\)](#) is the width in bytes of the allocation. The intended usage of pitch is as a separate parameter of the allocation, used to compute addresses within the 2D array. Given the row and column of an array element of type `T`, the address is computed as:

```
T* pElement = (T*)((char*)BaseAddress + Row * Pitch) + Column;
```

The pitch returned by [cuMemAllocPitch\(\)](#) is guaranteed to work with [cuMemcpy2D\(\)](#) under all circumstances. For allocations of 2D arrays, it is recommended that programmers consider performing pitch allocations using [cuMemAllocPitch\(\)](#). Due to alignment restrictions in the hardware, this is especially true if the application will be performing 2D memory copies between different regions of device memory (whether linear memory or CUDA arrays).

The byte alignment of the pitch returned by [cuMemAllocPitch\(\)](#) is guaranteed to match or exceed the alignment requirement for texture binding with [cuTexRefSetAddress2D\(\)](#).

Parameters:

dptr - Returned device pointer
pPitch - Returned pitch of allocation in bytes
WidthInBytes - Requested allocation width in bytes
Height - Requested allocation height in rows
ElementSizeBytes - Size of largest reads/writes for range

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.16 CUresult cuMemcpy (CUdeviceptr *dst*, CUdeviceptr *src*, size_t *ByteCount*)

Copies data between two pointers. *dst* and *src* are base pointers of the destination and source, respectively. *ByteCount* specifies the number of bytes to copy. Note that this function infers the type of the transfer (host to host, host to device, device to device, or device to host) from the pointer values. This function is only allowed in contexts which support unified addressing. Note that this function is synchronous.

Parameters:

dst - Destination unified virtual address space pointer

src - Source unified virtual address space pointer

ByteCount - Size of memory copy in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.17 CUresult cuMemcpy2D (const CUDA_MEMCPY2D * *pCopy*)

Perform a 2D memory copy according to the parameters specified in *pCopy*. The [CUDA_MEMCPY2D](#) structure is defined as:

```
typedef struct CUDA_MEMCPY2D_st {
    unsigned int srcXInBytes, srcY;
    CUmemorytype srcMemoryType;
    const void *srcHost;
    CUdeviceptr srcDevice;
    CUarray srcArray;
    unsigned int srcPitch;

    unsigned int dstXInBytes, dstY;
    CUmemorytype dstMemoryType;
    void *dstHost;
    CUdeviceptr dstDevice;
    CUarray dstArray;
    unsigned int dstPitch;

    unsigned int WidthInBytes;
    unsigned int Height;
} CUDA_MEMCPY2D;
```

where:

- `srcMemoryType` and `dstMemoryType` specify the type of memory of the source and destination, respectively; `CUmemorytype_enum` is defined as:

```
typedef enum CUmemorytype_enum {
    CU_MEMORYTYPE_HOST = 0x01,
    CU_MEMORYTYPE_DEVICE = 0x02,
    CU_MEMORYTYPE_ARRAY = 0x03,
    CU_MEMORYTYPE_UNIFIED = 0x04
} CUmemorytype;
```

If `srcMemoryType` is `CU_MEMORYTYPE_UNIFIED`, `srcDevice` and `srcPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `srcArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `srcMemoryType` is `CU_MEMORYTYPE_HOST`, `srcHost` and `srcPitch` specify the (host) base address of the source data and the bytes per row to apply. `srcArray` is ignored.

If `srcMemoryType` is `CU_MEMORYTYPE_DEVICE`, `srcDevice` and `srcPitch` specify the (device) base address of the source data and the bytes per row to apply. `srcArray` is ignored.

If `srcMemoryType` is `CU_MEMORYTYPE_ARRAY`, `srcArray` specifies the handle of the source data. `srcHost`, `srcDevice` and `srcPitch` are ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_HOST`, `dstHost` and `dstPitch` specify the (host) base address of the destination data and the bytes per row to apply. `dstArray` is ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_UNIFIED`, `dstDevice` and `dstPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `dstArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `dstMemoryType` is `CU_MEMORYTYPE_DEVICE`, `dstDevice` and `dstPitch` specify the (device) base address of the destination data and the bytes per row to apply. `dstArray` is ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_ARRAY`, `dstArray` specifies the handle of the destination data. `dstHost`, `dstDevice` and `dstPitch` are ignored.

- `srcXInBytes` and `srcY` specify the base address of the source data for the copy.

For host pointers, the starting address is

```
void* Start = (void*)((char*)srcHost+srcY*srcPitch + srcXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr Start = srcDevice+srcY*srcPitch+srcXInBytes;
```

For CUDA arrays, `srcXInBytes` must be evenly divisible by the array element size.

- `dstXInBytes` and `dstY` specify the base address of the destination data for the copy.

For host pointers, the base address is

```
void* dstStart = (void*)((char*)dstHost+dstY*dstPitch + dstXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr dstStart = dstDevice+dstY*dstPitch+dstXInBytes;
```

For CUDA arrays, `dstXInBytes` must be evenly divisible by the array element size.

- `WidthInBytes` and `Height` specify the width (in bytes) and height of the 2D copy being performed.
- If specified, `srcPitch` must be greater than or equal to `WidthInBytes + srcXInBytes`, and `dstPitch` must be greater than or equal to `WidthInBytes + dstXInBytes`.

`cuMemcpy2D()` returns an error if any pitch is greater than the maximum allowed (`CU_DEVICE_ATTRIBUTE_MAX_PITCH`). `cuMemAllocPitch()` passes back pitches that always work with `cuMemcpy2D()`. On intra-device memory copies (device to device, CUDA array to device, CUDA array to CUDA array), `cuMemcpy2D()` may fail for pitches not computed by `cuMemAllocPitch()`. `cuMemcpy2DUnaligned()` does not have this restriction, but may run significantly slower in the cases where `cuMemcpy2D()` would have returned an error code.

Parameters:

pCopy - Parameters for the memory copy

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuArray3DCreate`, `cuArray3DGetDescriptor`, `cuArrayCreate`, `cuArrayDestroy`, `cuArrayGetDescriptor`, `cuMemAlloc`, `cuMemAllocHost`, `cuMemAllocPitch`, `cuMemcpy2DAsync`, `cuMemcpy2DUnaligned`, `cuMemcpy3D`, `cuMemcpy3DAsync`, `cuMemcpyAtoA`, `cuMemcpyAtoD`, `cuMemcpyAtoH`, `cuMemcpyAtoHAsync`, `cuMemcpyDtoA`, `cuMemcpyDtoD`, `cuMemcpyDtoDAsync`, `cuMemcpyDtoH`, `cuMemcpyDtoHAsync`, `cuMemcpyHtoA`, `cuMemcpyHtoAAsync`, `cuMemcpyHtoD`, `cuMemcpyHtoDAsync`, `cuMemFree`, `cuMemFreeHost`, `cuMemGetAddressRange`, `cuMemGetInfo`, `cuMemHostAlloc`, `cuMemHostGetDevicePointer`, `cuMemsetD2D8`, `cuMemsetD2D16`, `cuMemsetD2D32`, `cuMemsetD8`, `cuMemsetD16`, `cuMemsetD32`

5.37.2.18 CUresult cuMemcpy2DAsync (const CUDA_MEMCPY2D *pCopy, CUstream hStream)

Perform a 2D memory copy according to the parameters specified in `pCopy`. The `CUDA_MEMCPY2D` structure is defined as:

```
typedef struct CUDA_MEMCPY2D_st {
    unsigned int srcXInBytes, srcY;
    CUMemorytype srcMemoryType;
    const void *srcHost;
    CUdeviceptr srcDevice;
    CUarray srcArray;
    unsigned int srcPitch;
    unsigned int dstXInBytes, dstY;
    CUMemorytype dstMemoryType;
    void *dstHost;
    CUdeviceptr dstDevice;
    CUarray dstArray;
    unsigned int dstPitch;
    unsigned int WidthInBytes;
    unsigned int Height;
} CUDA_MEMCPY2D;
```

where:

- `srcMemoryType` and `dstMemoryType` specify the type of memory of the source and destination, respectively; `CUMemorytype_enum` is defined as:

```
typedef enum CUMemorytype_enum {
    CU_MEMORYTYPE_HOST = 0x01,
    CU_MEMORYTYPE_DEVICE = 0x02,
    CU_MEMORYTYPE_ARRAY = 0x03,
    CU_MEMORYTYPE_UNIFIED = 0x04
} CUMemorytype;
```

If `srcMemoryType` is `CU_MEMORYTYPE_HOST`, `srcHost` and `srcPitch` specify the (host) base address of the source data and the bytes per row to apply. `srcArray` is ignored.

If `srcMemoryType` is `CU_MEMORYTYPE_UNIFIED`, `srcDevice` and `srcPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `srcArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `srcMemoryType` is `CU_MEMORYTYPE_DEVICE`, `srcDevice` and `srcPitch` specify the (device) base address of the source data and the bytes per row to apply. `srcArray` is ignored.

If `srcMemoryType` is `CU_MEMORYTYPE_ARRAY`, `srcArray` specifies the handle of the source data. `srcHost`, `srcDevice` and `srcPitch` are ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_UNIFIED`, `dstDevice` and `dstPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `dstArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `dstMemoryType` is `CU_MEMORYTYPE_HOST`, `dstHost` and `dstPitch` specify the (host) base address of the destination data and the bytes per row to apply. `dstArray` is ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_DEVICE`, `dstDevice` and `dstPitch` specify the (device) base address of the destination data and the bytes per row to apply. `dstArray` is ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_ARRAY`, `dstArray` specifies the handle of the destination data. `dstHost`, `dstDevice` and `dstPitch` are ignored.

- `srcXInBytes` and `srcY` specify the base address of the source data for the copy.

For host pointers, the starting address is

```
void* Start = (void*)((char*)srcHost+srcY*srcPitch + srcXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr Start = srcDevice+srcY*srcPitch+srcXInBytes;
```

For CUDA arrays, `srcXInBytes` must be evenly divisible by the array element size.

- `dstXInBytes` and `dstY` specify the base address of the destination data for the copy.

For host pointers, the base address is

```
void* dstStart = (void*)((char*)dstHost+dstY*dstPitch + dstXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr dstStart = dstDevice+dstY*dstPitch+dstXInBytes;
```

For CUDA arrays, `dstXInBytes` must be evenly divisible by the array element size.

- `WidthInBytes` and `Height` specify the width (in bytes) and height of the 2D copy being performed.
- If specified, `srcPitch` must be greater than or equal to `WidthInBytes + srcXInBytes`, and `dstPitch` must be greater than or equal to `WidthInBytes + dstXInBytes`.
- If specified, `srcPitch` must be greater than or equal to `WidthInBytes + srcXInBytes`, and `dstPitch` must be greater than or equal to `WidthInBytes + dstXInBytes`.
- If specified, `srcHeight` must be greater than or equal to `Height + srcY`, and `dstHeight` must be greater than or equal to `Height + dstY`.

`cuMemcpy2D()` returns an error if any pitch is greater than the maximum allowed (`CU_DEVICE_ATTRIBUTE_MAX_PITCH`). `cuMemAllocPitch()` passes back pitches that always work with `cuMemcpy2D()`. On intra-device memory copies (device to device, CUDA array to device, CUDA array to CUDA array), `cuMemcpy2D()` may fail for pitches not computed by `cuMemAllocPitch()`. `cuMemcpy2DUnaligned()` does not have this restriction, but may run significantly slower in the cases where `cuMemcpy2D()` would have returned an error code.

`cuMemcpy2DAsync()` is asynchronous and can optionally be associated to a stream by passing a non-zero `hStream` argument. It only works on page-locked host memory and returns an error if a pointer to pageable memory is passed as input.

Parameters:

pCopy - Parameters for the memory copy

hStream - Stream identifier

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuArray3DCreate`, `cuArray3DGetDescriptor`, `cuArrayCreate`, `cuArrayDestroy`, `cuArrayGetDescriptor`, `cuMemAlloc`, `cuMemAllocHost`, `cuMemAllocPitch`, `cuMemcpy2D`, `cuMemcpy2DUnaligned`, `cuMemcpy3D`, `cuMemcpy3DAsync`, `cuMemcpyAtoA`, `cuMemcpyAtoD`, `cuMemcpyAtoH`, `cuMemcpyAtoHAsync`, `cuMemcpyDtoA`, `cuMemcpyDtoD`, `cuMemcpyDtoDAsync`, `cuMemcpyDtoH`, `cuMemcpyDtoHAsync`, `cuMemcpyHtoA`, `cuMemcpyHtoAAsync`, `cuMemcpyHtoD`, `cuMemcpyHtoDAsync`, `cuMemFree`, `cuMemFreeHost`, `cuMemGetAddressRange`, `cuMemGetInfo`, `cuMemHostAlloc`, `cuMemHostGetDevicePointer`, `cuMemsetD2D8`, `cuMemsetD2D8Async`, `cuMemsetD2D16`, `cuMemsetD2D16Async`, `cuMemsetD2D32`, `cuMemsetD2D32Async`, `cuMemsetD8`, `cuMemsetD8Async`, `cuMemsetD16`, `cuMemsetD16Async`, `cuMemsetD32`, `cuMemsetD32Async`

5.37.2.19 CUresult cuMemcpy2DUnaligned (const CUDA_MEMCPY2D * pCopy)

Perform a 2D memory copy according to the parameters specified in `pCopy`. The `CUDA_MEMCPY2D` structure is defined as:

```
typedef struct CUDA_MEMCPY2D_st {
    unsigned int srcXInBytes, srcY;
    CUmemorytype srcMemoryType;
    const void *srcHost;
    CUdeviceptr srcDevice;
    CUarray srcArray;
    unsigned int srcPitch;
    unsigned int dstXInBytes, dstY;
    CUmemorytype dstMemoryType;
    void *dstHost;
    CUdeviceptr dstDevice;
    CUarray dstArray;
    unsigned int dstPitch;
    unsigned int WidthInBytes;
    unsigned int Height;
} CUDA_MEMCPY2D;
```

where:

- `srcMemoryType` and `dstMemoryType` specify the type of memory of the source and destination, respectively; `CUmemorytype_enum` is defined as:

```
typedef enum CUMemoryType_enum {
    CU_MEMORYTYPE_HOST = 0x01,
    CU_MEMORYTYPE_DEVICE = 0x02,
    CU_MEMORYTYPE_ARRAY = 0x03,
    CU_MEMORYTYPE_UNIFIED = 0x04
} CUMemoryType;
```

If `srcMemoryType` is `CU_MEMORYTYPE_UNIFIED`, `srcDevice` and `srcPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `srcArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `srcMemoryType` is `CU_MEMORYTYPE_HOST`, `srcHost` and `srcPitch` specify the (host) base address of the source data and the bytes per row to apply. `srcArray` is ignored.

If `srcMemoryType` is `CU_MEMORYTYPE_DEVICE`, `srcDevice` and `srcPitch` specify the (device) base address of the source data and the bytes per row to apply. `srcArray` is ignored.

If `srcMemoryType` is `CU_MEMORYTYPE_ARRAY`, `srcArray` specifies the handle of the source data. `srcHost`, `srcDevice` and `srcPitch` are ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_UNIFIED`, `dstDevice` and `dstPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `dstArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `dstMemoryType` is `CU_MEMORYTYPE_HOST`, `dstHost` and `dstPitch` specify the (host) base address of the destination data and the bytes per row to apply. `dstArray` is ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_DEVICE`, `dstDevice` and `dstPitch` specify the (device) base address of the destination data and the bytes per row to apply. `dstArray` is ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_ARRAY`, `dstArray` specifies the handle of the destination data. `dstHost`, `dstDevice` and `dstPitch` are ignored.

- `srcXInBytes` and `srcY` specify the base address of the source data for the copy.

For host pointers, the starting address is

```
void* Start = (void*)((char*)srcHost+srcY*srcPitch + srcXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr Start = srcDevice+srcY*srcPitch+srcXInBytes;
```

For CUDA arrays, `srcXInBytes` must be evenly divisible by the array element size.

- `dstXInBytes` and `dstY` specify the base address of the destination data for the copy.

For host pointers, the base address is

```
void* dstStart = (void*)((char*)dstHost+dstY*dstPitch + dstXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr dstStart = dstDevice+dstY*dstPitch+dstXInBytes;
```

For CUDA arrays, `dstXInBytes` must be evenly divisible by the array element size.

- `WidthInBytes` and `Height` specify the width (in bytes) and height of the 2D copy being performed.
- If specified, `srcPitch` must be greater than or equal to `WidthInBytes + srcXInBytes`, and `dstPitch` must be greater than or equal to `WidthInBytes + dstXInBytes`.

`cuMemcpy2D()` returns an error if any pitch is greater than the maximum allowed (`CU_DEVICE_ATTRIBUTE_MAX_PITCH`). `cuMemAllocPitch()` passes back pitches that always work with `cuMemcpy2D()`. On intra-device memory copies (device to device, CUDA array to device, CUDA array to CUDA array), `cuMemcpy2D()` may fail for pitches not computed by `cuMemAllocPitch()`. `cuMemcpy2DUnaligned()` does not have this restriction, but may run significantly slower in the cases where `cuMemcpy2D()` would have returned an error code.

Parameters:

pCopy - Parameters for the memory copy

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuArray3DCreate`, `cuArray3DGetDescriptor`, `cuArrayCreate`, `cuArrayDestroy`, `cuArrayGetDescriptor`, `cuMemAlloc`, `cuMemAllocHost`, `cuMemAllocPitch`, `cuMemcpy2D`, `cuMemcpy2DAsync`, `cuMemcpy3D`, `cuMemcpy3DAsync`, `cuMemcpyAtoA`, `cuMemcpyAtoD`, `cuMemcpyAtoH`, `cuMemcpyAtoHAsync`, `cuMemcpyDtoA`, `cuMemcpyDtoD`, `cuMemcpyDtoDAsync`, `cuMemcpyDtoH`, `cuMemcpyDtoHAsync`, `cuMemcpyHtoA`, `cuMemcpyHtoAAsync`, `cuMemcpyHtoD`, `cuMemcpyHtoDAsync`, `cuMemFree`, `cuMemFreeHost`, `cuMemGetAddressRange`, `cuMemGetInfo`, `cuMemHostAlloc`, `cuMemHostGetDevicePointer`, `cuMemsetD2D8`, `cuMemsetD2D16`, `cuMemsetD2D32`, `cuMemsetD8`, `cuMemsetD16`, `cuMemsetD32`

5.37.2.20 CUresult cuMemcpy3D (const CUDA_MEMCPY3D * pCopy)

Perform a 3D memory copy according to the parameters specified in `pCopy`. The `CUDA_MEMCPY3D` structure is defined as:

```

typedef struct CUDA_MEMCPY3D_st {

    unsigned int srcXInBytes, srcY, srcZ;
    unsigned int srcLOD;
    CUMemorytype srcMemoryType;
        const void *srcHost;
        CUdeviceptr srcDevice;
        CUarray srcArray;
        unsigned int srcPitch; // ignored when src is array
        unsigned int srcHeight; // ignored when src is array; may be 0 if Depth==1

    unsigned int dstXInBytes, dstY, dstZ;
    unsigned int dstLOD;
    CUMemorytype dstMemoryType;
        void *dstHost;
        CUdeviceptr dstDevice;
        CUarray dstArray;
        unsigned int dstPitch; // ignored when dst is array
        unsigned int dstHeight; // ignored when dst is array; may be 0 if Depth==1

    unsigned int WidthInBytes;
    unsigned int Height;
    unsigned int Depth;
} CUDA_MEMCPY3D;

```

where:

- `srcMemoryType` and `dstMemoryType` specify the type of memory of the source and destination, respectively; [CUMemorytype_enum](#) is defined as:

```

typedef enum CUMemorytype_enum {
    CU_MEMORYTYPE_HOST = 0x01,
    CU_MEMORYTYPE_DEVICE = 0x02,
    CU_MEMORYTYPE_ARRAY = 0x03,
    CU_MEMORYTYPE_UNIFIED = 0x04
} CUMemorytype;

```

If `srcMemoryType` is [CU_MEMORYTYPE_UNIFIED](#), `srcDevice` and `srcPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `srcArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `srcMemoryType` is [CU_MEMORYTYPE_HOST](#), `srcHost`, `srcPitch` and `srcHeight` specify the (host) base address of the source data, the bytes per row, and the height of each 2D slice of the 3D array. `srcArray` is ignored.

If `srcMemoryType` is [CU_MEMORYTYPE_DEVICE](#), `srcDevice`, `srcPitch` and `srcHeight` specify the (device) base address of the source data, the bytes per row, and the height of each 2D slice of the 3D array. `srcArray` is ignored.

If `srcMemoryType` is [CU_MEMORYTYPE_ARRAY](#), `srcArray` specifies the handle of the source data. `srcHost`, `srcDevice`, `srcPitch` and `srcHeight` are ignored.

If `dstMemoryType` is [CU_MEMORYTYPE_UNIFIED](#), `dstDevice` and `dstPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `dstArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `dstMemoryType` is `CU_MEMORYTYPE_HOST`, `dstHost` and `dstPitch` specify the (host) base address of the destination data, the bytes per row, and the height of each 2D slice of the 3D array. `dstArray` is ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_DEVICE`, `dstDevice` and `dstPitch` specify the (device) base address of the destination data, the bytes per row, and the height of each 2D slice of the 3D array. `dstArray` is ignored.

If `dstMemoryType` is `CU_MEMORYTYPE_ARRAY`, `dstArray` specifies the handle of the destination data. `dstHost`, `dstDevice`, `dstPitch` and `dstHeight` are ignored.

- `srcXInBytes`, `srcY` and `srcZ` specify the base address of the source data for the copy.

For host pointers, the starting address is

```
void* Start = (void*)((char*)srcHost+(srcZ*srcHeight+srcY)*srcPitch + srcXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr Start = srcDevice+(srcZ*srcHeight+srcY)*srcPitch+srcXInBytes;
```

For CUDA arrays, `srcXInBytes` must be evenly divisible by the array element size.

- `dstXInBytes`, `dstY` and `dstZ` specify the base address of the destination data for the copy.

For host pointers, the base address is

```
void* dstStart = (void*)((char*)dstHost+(dstZ*dstHeight+dstY)*dstPitch + dstXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr dstStart = dstDevice+(dstZ*dstHeight+dstY)*dstPitch+dstXInBytes;
```

For CUDA arrays, `dstXInBytes` must be evenly divisible by the array element size.

- `WidthInBytes`, `Height` and `Depth` specify the width (in bytes), height and depth of the 3D copy being performed.
- If specified, `srcPitch` must be greater than or equal to `WidthInBytes + srcXInBytes`, and `dstPitch` must be greater than or equal to `WidthInBytes + dstXInBytes`.
- If specified, `srcHeight` must be greater than or equal to `Height + srcY`, and `dstHeight` must be greater than or equal to `Height + dstY`.

`cuMemcpy3D()` returns an error if any pitch is greater than the maximum allowed (`CU_DEVICE_ATTRIBUTE_MAX_PITCH`).

The `srcLOD` and `dstLOD` members of the `CUDA_MEMCPY3D` structure must be set to 0.

Parameters:

pCopy - Parameters for the memory copy

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.21 CUresult cuMemcpy3DAsync (const CUDA_MEMCPY3D * pCopy, CUstream hStream)

Perform a 3D memory copy according to the parameters specified in *pCopy*. The [CUDA_MEMCPY3D](#) structure is defined as:

```
typedef struct CUDA_MEMCPY3D_st {
    unsigned int srcXInBytes, srcY, srcZ;
    unsigned int srcLOD;
    CUMemorytype srcMemoryType;
    const void *srcHost;
    CUdeviceptr srcDevice;
    CUarray srcArray;
    unsigned int srcPitch; // ignored when src is array
    unsigned int srcHeight; // ignored when src is array; may be 0 if Depth==1

    unsigned int dstXInBytes, dstY, dstZ;
    unsigned int dstLOD;
    CUMemorytype dstMemoryType;
    void *dstHost;
    CUdeviceptr dstDevice;
    CUarray dstArray;
    unsigned int dstPitch; // ignored when dst is array
    unsigned int dstHeight; // ignored when dst is array; may be 0 if Depth==1

    unsigned int WidthInBytes;
    unsigned int Height;
    unsigned int Depth;
} CUDA_MEMCPY3D;
```

where:

- *srcMemoryType* and *dstMemoryType* specify the type of memory of the source and destination, respectively; [CUMemorytype_enum](#) is defined as:

```
typedef enum CUMemorytype_enum {
```

```

    CU_MEMORYTYPE_HOST = 0x01,
    CU_MEMORYTYPE_DEVICE = 0x02,
    CU_MEMORYTYPE_ARRAY = 0x03,
    CU_MEMORYTYPE_UNIFIED = 0x04
} CUmemorytype;

```

If `srcMemoryType` is [CU_MEMORYTYPE_UNIFIED](#), `srcDevice` and `srcPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `srcArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `srcMemoryType` is [CU_MEMORYTYPE_HOST](#), `srcHost`, `srcPitch` and `srcHeight` specify the (host) base address of the source data, the bytes per row, and the height of each 2D slice of the 3D array. `srcArray` is ignored.

If `srcMemoryType` is [CU_MEMORYTYPE_DEVICE](#), `srcDevice`, `srcPitch` and `srcHeight` specify the (device) base address of the source data, the bytes per row, and the height of each 2D slice of the 3D array. `srcArray` is ignored.

If `srcMemoryType` is [CU_MEMORYTYPE_ARRAY](#), `srcArray` specifies the handle of the source data. `srcHost`, `srcDevice`, `srcPitch` and `srcHeight` are ignored.

If `dstMemoryType` is [CU_MEMORYTYPE_UNIFIED](#), `dstDevice` and `dstPitch` specify the (unified virtual address space) base address of the source data and the bytes per row to apply. `dstArray` is ignored. This value may be used only if unified addressing is supported in the calling context.

If `dstMemoryType` is [CU_MEMORYTYPE_HOST](#), `dstHost` and `dstPitch` specify the (host) base address of the destination data, the bytes per row, and the height of each 2D slice of the 3D array. `dstArray` is ignored.

If `dstMemoryType` is [CU_MEMORYTYPE_DEVICE](#), `dstDevice` and `dstPitch` specify the (device) base address of the destination data, the bytes per row, and the height of each 2D slice of the 3D array. `dstArray` is ignored.

If `dstMemoryType` is [CU_MEMORYTYPE_ARRAY](#), `dstArray` specifies the handle of the destination data. `dstHost`, `dstDevice`, `dstPitch` and `dstHeight` are ignored.

- `srcXInBytes`, `srcY` and `srcZ` specify the base address of the source data for the copy.

For host pointers, the starting address is

```
void* Start = (void*)((char*)srcHost+(srcZ*srcHeight+srcY)*srcPitch + srcXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr Start = srcDevice+(srcZ*srcHeight+srcY)*srcPitch+srcXInBytes;
```

For CUDA arrays, `srcXInBytes` must be evenly divisible by the array element size.

- `dstXInBytes`, `dstY` and `dstZ` specify the base address of the destination data for the copy.

For host pointers, the base address is

```
void* dstStart = (void*)((char*)dstHost+(dstZ*dstHeight+dstY)*dstPitch + dstXInBytes);
```

For device pointers, the starting address is

```
CUdeviceptr dstStart = dstDevice+(dstZ*dstHeight+dstY)*dstPitch+dstXInBytes;
```

For CUDA arrays, `dstXInBytes` must be evenly divisible by the array element size.

- `WidthInBytes`, `Height` and `Depth` specify the width (in bytes), height and depth of the 3D copy being performed.
- If specified, `srcPitch` must be greater than or equal to `WidthInBytes + srcXInBytes`, and `dstPitch` must be greater than or equal to `WidthInBytes + dstXInBytes`.
- If specified, `srcHeight` must be greater than or equal to `Height + srcY`, and `dstHeight` must be greater than or equal to `Height + dstY`.

`cuMemcpy3D()` returns an error if any pitch is greater than the maximum allowed (`CU_DEVICE_ATTRIBUTE_MAX_PITCH`).

`cuMemcpy3DAsync()` is asynchronous and can optionally be associated to a stream by passing a non-zero `hStream` argument. It only works on page-locked host memory and returns an error if a pointer to pageable memory is passed as input.

The `srcLOD` and `dstLOD` members of the `CUDA_MEMCPY3D` structure must be set to 0.

Parameters:

- pCopy* - Parameters for the memory copy
- hStream* - Stream identifier

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuArray3DCreate`, `cuArray3DGetDescriptor`, `cuArrayCreate`, `cuArrayDestroy`, `cuArrayGetDescriptor`, `cuMemAlloc`, `cuMemAllocHost`, `cuMemAllocPitch`, `cuMemcpy2D`, `cuMemcpy2DAsync`, `cuMemcpy2DUnaligned`, `cuMemcpy3D`, `cuMemcpyAtoA`, `cuMemcpyAtoD`, `cuMemcpyAtoH`, `cuMemcpyAtoHAsync`, `cuMemcpyDtoA`, `cuMemcpyDtoD`, `cuMemcpyDtoDAsync`, `cuMemcpyDtoH`, `cuMemcpyDtoHAsync`, `cuMemcpyHtoA`, `cuMemcpyHtoAAsync`, `cuMemcpyHtoD`, `cuMemcpyHtoDAsync`, `cuMemFree`, `cuMemFreeHost`, `cuMemGetAddressRange`, `cuMemGetInfo`, `cuMemHostAlloc`, `cuMemHostGetDevicePointer`, `cuMemsetD2D8`, `cuMemsetD2D8Async`, `cuMemsetD2D16`, `cuMemsetD2D16Async`, `cuMemsetD2D32`, `cuMemsetD2D32Async`, `cuMemsetD8`, `cuMemsetD8Async`, `cuMemsetD16`, `cuMemsetD16Async`, `cuMemsetD32`, `cuMemsetD32Async`

5.37.2.22 CUresult cuMemcpy3DPeer (const CUDA_MEMCPY3D_PEER * pCopy)

Perform a 3D memory copy according to the parameters specified in `pCopy`. See the definition of the [CUDA_MEMCPY3D_PEER](#) structure for documentation of its parameters.

Note that this function is synchronous with respect to the host only if the source or destination memory is of type [CU_MEMORYTYPE_HOST](#). Note also that this copy is serialized with respect all pending and future asynchronous work in to the current context, the copy's source context, and the copy's destination context (use [cuMemcpy3DPeerAsync](#) to avoid this synchronization).

Parameters:

pCopy - Parameters for the memory copy

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMemcpyDtoD](#), [cuMemcpyPeer](#), [cuMemcpyDtoDAsync](#), [cuMemcpyPeerAsync](#), [cuMemcpy3DPeerAsync](#)

5.37.2.23 CUresult cuMemcpy3DPeerAsync (const CUDA_MEMCPY3D_PEER * pCopy, CUstream hStream)

Perform a 3D memory copy according to the parameters specified in `pCopy`. See the definition of the [CUDA_MEMCPY3D_PEER](#) structure for documentation of its parameters.

Parameters:

pCopy - Parameters for the memory copy

hStream - Stream identifier

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMemcpyDtoD](#), [cuMemcpyPeer](#), [cuMemcpyDtoDAsync](#), [cuMemcpyPeerAsync](#), [cuMemcpy3DPeerAsync](#)

5.37.2.24 CUresult cuMemcpyAsync (CUdeviceptr dst, CUdeviceptr src, size_t ByteCount, CUstream hStream)

Copies data between two pointers. `dst` and `src` are base pointers of the destination and source, respectively. `ByteCount` specifies the number of bytes to copy. Note that this function infers the type of the transfer (host to

host, host to device, device to device, or device to host) from the pointer values. This function is only allowed in contexts which support unified addressing. Note that this function is asynchronous and can optionally be associated to a stream by passing a non-zero `hStream` argument

Parameters:

dst - Destination unified virtual address space pointer

src - Source unified virtual address space pointer

ByteCount - Size of memory copy in bytes

hStream - Stream identifier

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.25 CUresult cuMemcpyAtoA (CUarray dstArray, size_t dstOffset, CUarray srcArray, size_t srcOffset, size_t ByteCount)

Copies from one 1D CUDA array to another. `dstArray` and `srcArray` specify the handles of the destination and source CUDA arrays for the copy, respectively. `dstOffset` and `srcOffset` specify the destination and source offsets in bytes into the CUDA arrays. `ByteCount` is the number of bytes to be copied. The size of the elements in the CUDA arrays need not be the same format, but the elements must be the same size; and count must be evenly divisible by that size.

Parameters:

dstArray - Destination array

dstOffset - Offset in bytes of destination array

srcArray - Source array

srcOffset - Offset in bytes of source array

ByteCount - Size of memory copy in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.26 CUresult cuMemcpyAtoD (CUdeviceptr *dstDevice*, CUarray *srcArray*, size_t *srcOffset*, size_t *ByteCount*)

Copies from one 1D CUDA array to device memory. *dstDevice* specifies the base pointer of the destination and must be naturally aligned with the CUDA array elements. *srcArray* and *srcOffset* specify the CUDA array handle and the offset in bytes into the array where the copy is to begin. *ByteCount* specifies the number of bytes to copy and must be evenly divisible by the array element size.

Parameters:

dstDevice - Destination device pointer
srcArray - Source array
srcOffset - Offset in bytes of source array
ByteCount - Size of memory copy in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.27 CUresult cuMemcpyAtoH (void * *dstHost*, CUarray *srcArray*, size_t *srcOffset*, size_t *ByteCount*)

Copies from one 1D CUDA array to host memory. *dstHost* specifies the base pointer of the destination. *srcArray* and *srcOffset* specify the CUDA array handle and starting offset in bytes of the source data. *ByteCount* specifies the number of bytes to copy.

Parameters:

dstHost - Destination device pointer
srcArray - Source array
srcOffset - Offset in bytes of source array
ByteCount - Size of memory copy in bytes

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.28 CUresult cuMemcpyAtoHAsync (void * *dstHost*, CUarray *srcArray*, size_t *srcOffset*, size_t *ByteCount*, CUstream *hStream*)

Copies from one 1D CUDA array to host memory. *dstHost* specifies the base pointer of the destination. *srcArray* and *srcOffset* specify the CUDA array handle and starting offset in bytes of the source data. *ByteCount* specifies the number of bytes to copy.

[cuMemcpyAtoHAsync\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero *stream* argument. It only works on page-locked host memory and returns an error if a pointer to pageable memory is passed as input.

Parameters:

dstHost - Destination pointer
srcArray - Source array
srcOffset - Offset in bytes of source array
ByteCount - Size of memory copy in bytes
hStream - Stream identifier

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.29 CUresult cuMemcpyDtoA (CUarray *dstArray*, size_t *dstOffset*, CUdeviceptr *srcDevice*, size_t *ByteCount*)

Copies from device memory to a 1D CUDA array. *dstArray* and *dstOffset* specify the CUDA array handle and starting index of the destination data. *srcDevice* specifies the base pointer of the source. *ByteCount* specifies the number of bytes to copy.

Parameters:

dstArray - Destination array
dstOffset - Offset in bytes of destination array
srcDevice - Source device pointer
ByteCount - Size of memory copy in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.30 CUresult cuMemcpyDtoD (CUdeviceptr *dstDevice*, CUdeviceptr *srcDevice*, size_t *ByteCount*)

Copies from device memory to device memory. *dstDevice* and *srcDevice* are the base pointers of the destination and source, respectively. *ByteCount* specifies the number of bytes to copy. Note that this function is asynchronous.

Parameters:

dstDevice - Destination device pointer
srcDevice - Source device pointer

ByteCount - Size of memory copy in bytes

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.31 CUresult cuMemcpyDtoDAsync (CUdeviceptr *dstDevice*, CUdeviceptr *srcDevice*, size_t *ByteCount*, CUstream *hStream*)

Copies from device memory to device memory. *dstDevice* and *srcDevice* are the base pointers of the destination and source, respectively. *ByteCount* specifies the number of bytes to copy. Note that this function is asynchronous and can optionally be associated to a stream by passing a non-zero *hStream* argument

Parameters:

dstDevice - Destination device pointer
srcDevice - Source device pointer
ByteCount - Size of memory copy in bytes
hStream - Stream identifier

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D8Async, cuMemsetD2D16, cuMemsetD2D16Async, cuMemsetD2D32, cuMemsetD2D32Async, cuMemsetD8, cuMemsetD8Async, cuMemsetD16, cuMemsetD16Async, cuMemsetD32, cuMemsetD32Async

5.37.2.32 CUresult cuMemcpyDtoH (void * *dstHost*, CUdeviceptr *srcDevice*, size_t *ByteCount*)

Copies from device to host memory. *dstHost* and *srcDevice* specify the base pointers of the destination and source, respectively. *ByteCount* specifies the number of bytes to copy. Note that this function is synchronous.

Parameters:

dstHost - Destination host pointer
srcDevice - Source device pointer
ByteCount - Size of memory copy in bytes

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.33 CUresult cuMemcpyDtoHAsync (void * *dstHost*, CUdeviceptr *srcDevice*, size_t *ByteCount*, CUstream *hStream*)

Copies from device to host memory. *dstHost* and *srcDevice* specify the base pointers of the destination and source, respectively. *ByteCount* specifies the number of bytes to copy.

[cuMemcpyDtoHAsync\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero *hStream* argument. It only works on page-locked memory and returns an error if a pointer to pageable memory is passed as input.

Parameters:

dstHost - Destination host pointer
srcDevice - Source device pointer
ByteCount - Size of memory copy in bytes
hStream - Stream identifier

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.34 **CUresult cuMemcpyHtoA** (CUarray *dstArray*, size_t *dstOffset*, const void * *srcHost*, size_t *ByteCount*)

Copies from host memory to a 1D CUDA array. *dstArray* and *dstOffset* specify the CUDA array handle and starting offset in bytes of the destination data. *pSrc* specifies the base address of the source. *ByteCount* specifies the number of bytes to copy.

Parameters:

dstArray - Destination array
dstOffset - Offset in bytes of destination array
srcHost - Source host pointer
ByteCount - Size of memory copy in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.35 **CUresult cuMemcpyHtoAAsync** (CUarray *dstArray*, size_t *dstOffset*, const void * *srcHost*, size_t *ByteCount*, CUstream *hStream*)

Copies from host memory to a 1D CUDA array. *dstArray* and *dstOffset* specify the CUDA array handle and starting offset in bytes of the destination data. *srcHost* specifies the base address of the source. *ByteCount* specifies the number of bytes to copy.

[cuMemcpyHtoAAsync\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero *hStream* argument. It only works on page-locked memory and returns an error if a pointer to pageable memory is passed as input.

Parameters:

dstArray - Destination array
dstOffset - Offset in bytes of destination array
srcHost - Source host pointer
ByteCount - Size of memory copy in bytes
hStream - Stream identifier

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D8Async, cuMemsetD2D16, cuMemsetD2D16Async, cuMemsetD2D32, cuMemsetD2D32Async, cuMemsetD8, cuMemsetD8Async, cuMemsetD16, cuMemsetD16Async, cuMemsetD32, cuMemsetD32Async

5.37.2.36 CUresult cuMemcpyHtoD (CUdeviceptr dstDevice, const void * srcHost, size_t ByteCount)

Copies from host memory to device memory. *dstDevice* and *srcHost* are the base addresses of the destination and source, respectively. *ByteCount* specifies the number of bytes to copy. Note that this function is synchronous.

Parameters:

dstDevice - Destination device pointer
srcHost - Source host pointer
ByteCount - Size of memory copy in bytes

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.37 CUresult cuMemcpyHtoDAsync (CUdeviceptr *dstDevice*, const void * *srcHost*, size_t *ByteCount*, CUstream *hStream*)

Copies from host memory to device memory. *dstDevice* and *srcHost* are the base addresses of the destination and source, respectively. *ByteCount* specifies the number of bytes to copy.

[cuMemcpyHtoDAsync\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero *hStream* argument. It only works on page-locked memory and returns an error if a pointer to pageable memory is passed as input.

Parameters:

- dstDevice* - Destination device pointer
- srcHost* - Source host pointer
- ByteCount* - Size of memory copy in bytes
- hStream* - Stream identifier

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.38 CUresult cuMemcpyPeer (CUdeviceptr *dstDevice*, CUcontext *dstContext*, CUdeviceptr *srcDevice*, CUcontext *srcContext*, size_t *ByteCount*)

Copies from device memory in one context to device memory in another context. *dstDevice* is the base device pointer of the destination memory and *dstContext* is the destination context. *srcDevice* is the base device pointer of the source memory and *srcContext* is the source pointer. *ByteCount* specifies the number of bytes to copy.

Note that this function is asynchronous with respect to the host, but serialized with respect all pending and future asynchronous work in to the current context, *srcContext*, and *dstContext* (use [cuMemcpyPeerAsync](#) to avoid this synchronization).

Parameters:

- dstDevice* - Destination device pointer
- dstContext* - Destination context
- srcDevice* - Source device pointer

srcContext - Source context

ByteCount - Size of memory copy in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMemcpyDtoD](#), [cuMemcpy3DPeer](#), [cuMemcpyDtoDAsync](#), [cuMemcpyPeerAsync](#), [cuMemcpy3DPeerAsync](#)

5.37.2.39 CUresult cuMemcpyPeerAsync (CUdeviceptr *dstDevice*, CUcontext *dstContext*, CUdeviceptr *srcDevice*, CUcontext *srcContext*, size_t *ByteCount*, CUstream *hStream*)

Copies from device memory in one context to device memory in another context. *dstDevice* is the base device pointer of the destination memory and *dstContext* is the destination context. *srcDevice* is the base device pointer of the source memory and *srcContext* is the source pointer. *ByteCount* specifies the number of bytes to copy. Note that this function is asynchronous with respect to the host and all work in other streams in other devices.

Parameters:

dstDevice - Destination device pointer

dstContext - Destination context

srcDevice - Source device pointer

srcContext - Source context

ByteCount - Size of memory copy in bytes

hStream - Stream identifier

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMemcpyDtoD](#), [cuMemcpyPeer](#), [cuMemcpy3DPeer](#), [cuMemcpyDtoDAsync](#), [cuMemcpy3DPeerAsync](#)

5.37.2.40 CUresult cuMemFree (CUdeviceptr *dptr*)

Frees the memory space pointed to by *dptr*, which must have been returned by a previous call to [cuMemAlloc\(\)](#) or [cuMemAllocPitch\(\)](#).

Parameters:

dptr - Pointer to memory to free

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.41 CUresult cuMemFreeHost (void * p)

Frees the memory space pointed to by *p*, which must have been returned by a previous call to [cuMemAllocHost\(\)](#).

Parameters:

p - Pointer to memory to free

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.42 CUresult cuMemGetAddressRange (CUdeviceptr * pbase, size_t * psize, CUdeviceptr dptr)

Returns the base address in *pbase* and size in *psize* of the allocation by [cuMemAlloc\(\)](#) or [cuMemAllocPitch\(\)](#) that contains the input pointer *dptr*. Both parameters *pbase* and *psize* are optional. If one of them is NULL, it is ignored.

Parameters:

pbase - Returned base address

psize - Returned size of device memory allocation

dptr - Device pointer to query

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.43 CUresult cuMemGetInfo (size_t *free, size_t *total)

Returns in **free* and **total* respectively, the free and total amount of memory available for allocation by the CUDA context, in bytes.

Parameters:

free - Returned free memory in bytes

total - Returned total memory in bytes

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D16, cuMemsetD2D32, cuMemsetD8, cuMemsetD16, cuMemsetD32

5.37.2.44 CUresult cuMemHostAlloc (void ** pp, size_t bytesize, unsigned int Flags)

Allocates `bytesize` bytes of host memory that is page-locked and accessible to the device. The driver tracks the virtual memory ranges allocated with this function and automatically accelerates calls to functions such as `cuMemcpyHtoD()`. Since the memory can be accessed directly by the device, it can be read or written with much higher bandwidth than pageable memory obtained with functions such as `malloc()`. Allocating excessive amounts of pinned memory may degrade system performance, since it reduces the amount of memory available to the system for paging. As a result, this function is best used sparingly to allocate staging areas for data exchange between host and device.

The `Flags` parameter enables different options to be specified that affect the allocation, as follows.

- `CU_MEMHOSTALLOC_PORTABLE`: The memory returned by this call will be considered as pinned memory by all CUDA contexts, not just the one that performed the allocation.
- `CU_MEMHOSTALLOC_DEVICEMAP`: Maps the allocation into the CUDA address space. The device pointer to the memory may be obtained by calling `cuMemHostGetDevicePointer()`. This feature is available only on GPUs with compute capability greater than or equal to 1.1.
- `CU_MEMHOSTALLOC_WRITECOMBINED`: Allocates the memory as write-combined (WC). WC memory can be transferred across the PCI Express bus more quickly on some system configurations, but cannot be read efficiently by most CPUs. WC memory is a good option for buffers that will be written by the CPU and read by the GPU via mapped pinned memory or host->device transfers.

All of these flags are orthogonal to one another: a developer may allocate memory that is portable, mapped and/or write-combined with no restrictions.

The CUDA context must have been created with the `CU_CTX_MAP_HOST` flag in order for the `CU_MEMHOSTALLOC_DEVICEMAP` flag to have any effect.

The `CU_MEMHOSTALLOC_DEVICEMAP` flag may be specified on CUDA contexts for devices that do not support mapped pinned memory. The failure is deferred to `cuMemHostGetDevicePointer()` because the memory may be mapped into other CUDA contexts via the `CU_MEMHOSTALLOC_PORTABLE` flag.

The memory allocated by this function must be freed with `cuMemFreeHost()`.

Note all host memory allocated using `cuMemHostAlloc()` will automatically be immediately accessible to all contexts on all devices which support unified addressing (as may be queried using `CU_DEVICE_ATTRIBUTE_UNIFIED_ADDRESSING`). Unless the flag `CU_MEMHOSTALLOC_WRITECOMBINED` is specified, the device pointer that may be used to access this host memory from those contexts is always equal to the returned host pointer `*pp`. If the flag `CU_MEMHOSTALLOC_WRITECOMBINED` is specified, then the function `cuMemHostGetDevicePointer()` must be used to query the device pointer, even if the context supports unified addressing. See [Unified Addressing](#) for additional details.

Parameters:

pp - Returned host pointer to page-locked memory

bytesize - Requested allocation size in bytes

Flags - Flags for allocation request

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`, `CUDA_ERROR_OUT_OF_MEMORY`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.45 CUresult cuMemHostGetDevicePointer (CUdeviceptr * *pdptr*, void * *p*, unsigned int *Flags*)

Passes back the device pointer `pdptr` corresponding to the mapped, pinned host buffer `p` allocated by [cuMemHostAlloc](#).

`cuMemHostGetDevicePointer()` will fail if the `CU_MEMHOSTALLOC_DEVICEMAP` flag was not specified at the time the memory was allocated, or if the function is called on a GPU that does not support mapped pinned memory.

`Flags` provides for future releases. For now, it must be set to 0.

Parameters:

pdptr - Returned device pointer

p - Host pointer

Flags - Options (must be 0)

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD16](#), [cuMemsetD32](#)

5.37.2.46 CUresult cuMemHostGetFlags (unsigned int * *pFlags*, void * *p*)

Passes back the flags `pFlags` that were specified when allocating the pinned host buffer `p` allocated by [cuMemHostAlloc](#).

`cuMemHostGetFlags()` will fail if the pointer does not reside in an allocation performed by [cuMemAllocHost\(\)](#) or [cuMemHostAlloc\(\)](#).

Parameters:

pFlags - Returned flags word

p - Host pointer

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMemAllocHost](#), [cuMemHostAlloc](#)

5.37.2.47 CUresult cuMemHostRegister (void **p*, size_t *bytesize*, unsigned int *Flags*)

Page-locks the memory range specified by *p* and *bytesize* and maps it for the device(s) as specified by *Flags*. This memory range also is added to the same tracking mechanism as [cuMemHostAlloc](#) to automatically accelerate calls to functions such as [cuMemcpyHtoD\(\)](#). Since the memory can be accessed directly by the device, it can be read or written with much higher bandwidth than pageable memory that has not been registered. Page-locking excessive amounts of memory may degrade system performance, since it reduces the amount of memory available to the system for paging. As a result, this function is best used sparingly to register staging areas for data exchange between host and device.

This function has limited support on Mac OS X. OS 10.7 or higher is required.

The *Flags* parameter enables different options to be specified that affect the allocation, as follows.

- [CU_MEMHOSTREGISTER_PORTABLE](#): The memory returned by this call will be considered as pinned memory by all CUDA contexts, not just the one that performed the allocation.
- [CU_MEMHOSTREGISTER_DEVICEMAP](#): Maps the allocation into the CUDA address space. The device pointer to the memory may be obtained by calling [cuMemHostGetDevicePointer\(\)](#). This feature is available only on GPUs with compute capability greater than or equal to 1.1.

All of these flags are orthogonal to one another: a developer may page-lock memory that is portable or mapped with no restrictions.

The CUDA context must have been created with the [CU_CTX_MAP_HOST](#) flag in order for the [CU_MEMHOSTREGISTER_DEVICEMAP](#) flag to have any effect.

The [CU_MEMHOSTREGISTER_DEVICEMAP](#) flag may be specified on CUDA contexts for devices that do not support mapped pinned memory. The failure is deferred to [cuMemHostGetDevicePointer\(\)](#) because the memory may be mapped into other CUDA contexts via the [CU_MEMHOSTREGISTER_PORTABLE](#) flag.

The memory page-locked by this function must be unregistered with [cuMemHostUnregister\(\)](#).

Parameters:

p - Host pointer to memory to page-lock

bytesize - Size in bytes of the address range to page-lock

Flags - Flags for allocation request

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_OUT_OF_MEMORY, CUDA_ERROR_HOST_MEMORY_ALREADY_REGISTERED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMemHostUnregister](#), [cuMemHostGetFlags](#), [cuMemHostGetDevicePointer](#)

5.37.2.48 CUresult cuMemHostUnregister (void * *p*)

Unmaps the memory range whose base address is specified by *p*, and makes it pageable again.

The base address must be the same one specified to [cuMemHostRegister\(\)](#).

Parameters:

p - Host pointer to memory to unregister

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_OUT_OF_MEMORY, CUDA_ERROR_HOST_MEMORY_NOT_REGISTERED,

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMemHostRegister](#)

5.37.2.49 CUresult cuMemsetD16 (CUdeviceptr *dstDevice*, unsigned short *us*, size_t *N*)

Sets the memory range of *N* 16-bit values to the specified value *us*. The *dstDevice* pointer must be two byte aligned.

Note that this function is asynchronous with respect to the host unless *dstDevice* refers to pinned host memory.

Parameters:

dstDevice - Destination device pointer

us - Value to set

N - Number of elements

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.50 CUresult cuMemsetD16Async (CUdeviceptr *dstDevice*, unsigned short *us*, size_t *N*, CUstream *hStream*)

Sets the memory range of *N* 16-bit values to the specified value *us*. The *dstDevice* pointer must be two byte aligned.

[cuMemsetD16Async\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero *stream* argument.

Parameters:

dstDevice - Destination device pointer

us - Value to set

N - Number of elements

hStream - Stream identifier

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.51 CUresult cuMemsetD2D16 (CUdeviceptr *dstDevice*, size_t *dstPitch*, unsigned short *us*, size_t *Width*, size_t *Height*)

Sets the 2D memory range of *Width* 16-bit values to the specified value *us*. *Height* specifies the number of rows to set, and *dstPitch* specifies the number of bytes between each row. The *dstDevice* pointer and *dstPitch* offset must be two byte aligned. This function performs fastest when the pitch is one that has been passed back by [cuMemAllocPitch\(\)](#).

Note that this function is asynchronous with respect to the host unless *dstDevice* refers to pinned host memory.

Parameters:

- dstDevice* - Destination device pointer
- dstPitch* - Pitch of destination device pointer
- us* - Value to set
- Width* - Width of row
- Height* - Number of rows

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.52 CUresult cuMemsetD2D16Async (CUdeviceptr *dstDevice*, size_t *dstPitch*, unsigned short *us*, size_t *Width*, size_t *Height*, CUstream *hStream*)

Sets the 2D memory range of *Width* 16-bit values to the specified value *us*. *Height* specifies the number of rows to set, and *dstPitch* specifies the number of bytes between each row. The *dstDevice* pointer and *dstPitch* offset must be two byte aligned. This function performs fastest when the pitch is one that has been passed back by [cuMemAllocPitch\(\)](#).

[cuMemsetD2D16Async\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero *stream* argument.

Parameters:

- dstDevice* - Destination device pointer
- dstPitch* - Pitch of destination device pointer
- us* - Value to set

Width - Width of row

Height - Number of rows

hStream - Stream identifier

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.53 CUresult cuMemsetD2D32 (CUdeviceptr *dstDevice*, size_t *dstPitch*, unsigned int *ui*, size_t *Width*, size_t *Height*)

Sets the 2D memory range of *Width* 32-bit values to the specified value *ui*. *Height* specifies the number of rows to set, and *dstPitch* specifies the number of bytes between each row. The *dstDevice* pointer and *dstPitch* offset must be four byte aligned. This function performs fastest when the pitch is one that has been passed back by [cuMemAllocPitch\(\)](#).

Note that this function is asynchronous with respect to the host unless *dstDevice* refers to pinned host memory.

Parameters:

dstDevice - Destination device pointer

dstPitch - Pitch of destination device pointer

ui - Value to set

Width - Width of row

Height - Number of rows

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#),

[cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.54 CUresult cuMemsetD2D32Async (CUdeviceptr *dstDevice*, size_t *dstPitch*, unsigned int *ui*, size_t *Width*, size_t *Height*, CUstream *hStream*)

Sets the 2D memory range of *Width* 32-bit values to the specified value *ui*. *Height* specifies the number of rows to set, and *dstPitch* specifies the number of bytes between each row. The *dstDevice* pointer and *dstPitch* offset must be four byte aligned. This function performs fastest when the pitch is one that has been passed back by [cuMemAllocPitch\(\)](#).

[cuMemsetD2D32Async\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero *stream* argument.

Parameters:

dstDevice - Destination device pointer
dstPitch - Pitch of destination device pointer
ui - Value to set
Width - Width of row
Height - Number of rows
hStream - Stream identifier

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.55 CUresult cuMemsetD2D8 (CUdeviceptr *dstDevice*, size_t *dstPitch*, unsigned char *uc*, size_t *Width*, size_t *Height*)

Sets the 2D memory range of *Width* 8-bit values to the specified value *uc*. *Height* specifies the number of rows to set, and *dstPitch* specifies the number of bytes between each row. This function performs fastest when the pitch is one that has been passed back by [cuMemAllocPitch\(\)](#).

Note that this function is asynchronous with respect to the host unless *dstDevice* refers to pinned host memory.

Parameters:

dstDevice - Destination device pointer
dstPitch - Pitch of destination device pointer
uc - Value to set
Width - Width of row
Height - Number of rows

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8Async, cuMemsetD2D16, cuMemsetD2D16Async, cuMemsetD2D32, cuMemsetD2D32Async, cuMemsetD8, cuMemsetD8Async, cuMemsetD16, cuMemsetD16Async, cuMemsetD32, cuMemsetD32Async

5.37.2.56 CUresult cuMemsetD2D8Async (CUdeviceptr *dstDevice*, size_t *dstPitch*, unsigned char *uc*, size_t *Width*, size_t *Height*, CUstream *hStream*)

Sets the 2D memory range of *Width* 8-bit values to the specified value *uc*. *Height* specifies the number of rows to set, and *dstPitch* specifies the number of bytes between each row. This function performs fastest when the pitch is one that has been passed back by [cuMemAllocPitch\(\)](#).

[cuMemsetD2D8Async\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero `stream` argument.

Parameters:

dstDevice - Destination device pointer
dstPitch - Pitch of destination device pointer
uc - Value to set
Width - Width of row
Height - Number of rows
hStream - Stream identifier

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32](#), [cuMemsetD32Async](#)

5.37.2.57 CUresult cuMemsetD32 (CUdeviceptr *dstDevice*, unsigned int *ui*, size_t *N*)

Sets the memory range of *N* 32-bit values to the specified value *ui*. The *dstDevice* pointer must be four byte aligned.

Note that this function is asynchronous with respect to the host unless *dstDevice* refers to pinned host memory.

Parameters:

dstDevice - Destination device pointer
ui - Value to set
N - Number of elements

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuArray3DCreate](#), [cuArray3DGetDescriptor](#), [cuArrayCreate](#), [cuArrayDestroy](#), [cuArrayGetDescriptor](#), [cuMemAlloc](#), [cuMemAllocHost](#), [cuMemAllocPitch](#), [cuMemcpy2D](#), [cuMemcpy2DAsync](#), [cuMemcpy2DUnaligned](#), [cuMemcpy3D](#), [cuMemcpy3DAsync](#), [cuMemcpyAtoA](#), [cuMemcpyAtoD](#), [cuMemcpyAtoH](#), [cuMemcpyAtoHAsync](#), [cuMemcpyDtoA](#), [cuMemcpyDtoD](#), [cuMemcpyDtoDAsync](#), [cuMemcpyDtoH](#), [cuMemcpyDtoHAsync](#), [cuMemcpyHtoA](#), [cuMemcpyHtoAAsync](#), [cuMemcpyHtoD](#), [cuMemcpyHtoDAsync](#), [cuMemFree](#), [cuMemFreeHost](#), [cuMemGetAddressRange](#), [cuMemGetInfo](#), [cuMemHostAlloc](#), [cuMemHostGetDevicePointer](#), [cuMemsetD2D8](#), [cuMemsetD2D8Async](#), [cuMemsetD2D16](#), [cuMemsetD2D16Async](#), [cuMemsetD2D32](#), [cuMemsetD2D32Async](#), [cuMemsetD8](#), [cuMemsetD8Async](#), [cuMemsetD16](#), [cuMemsetD16Async](#), [cuMemsetD32Async](#)

5.37.2.58 CUresult cuMemsetD32Async (CUdeviceptr *dstDevice*, unsigned int *ui*, size_t *N*, CUstream *hStream*)

Sets the memory range of *N* 32-bit values to the specified value *ui*. The *dstDevice* pointer must be four byte aligned.

[cuMemsetD32Async\(\)](#) is asynchronous and can optionally be associated to a stream by passing a non-zero *stream* argument.

Parameters:

dstDevice - Destination device pointer

ui - Value to set
N - Number of elements
hStream - Stream identifier

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D8Async, cuMemsetD2D16, cuMemsetD2D16Async, cuMemsetD2D32, cuMemsetD2D32Async, cuMemsetD8, cuMemsetD8Async, cuMemsetD16, cuMemsetD16Async, cuMemsetD32

5.37.2.59 CUresult cuMemsetD8 (CUdeviceptr *dstDevice*, unsigned char *uc*, size_t *N*)

Sets the memory range of *N* 8-bit values to the specified value *uc*.

Note that this function is asynchronous with respect to the host unless *dstDevice* refers to pinned host memory.

Parameters:

dstDevice - Destination device pointer
uc - Value to set
N - Number of elements

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuArray3DCreate, cuArray3DGetDescriptor, cuArrayCreate, cuArrayDestroy, cuArrayGetDescriptor, cuMemAlloc, cuMemAllocHost, cuMemAllocPitch, cuMemcpy2D, cuMemcpy2DAsync, cuMemcpy2DUnaligned, cuMemcpy3D, cuMemcpy3DAsync, cuMemcpyAtoA, cuMemcpyAtoD, cuMemcpyAtoH, cuMemcpyAtoHAsync, cuMemcpyDtoA, cuMemcpyDtoD, cuMemcpyDtoDAsync, cuMemcpyDtoH, cuMemcpyDtoHAsync, cuMemcpyHtoA, cuMemcpyHtoAAsync, cuMemcpyHtoD, cuMemcpyHtoDAsync, cuMemFree, cuMemFreeHost, cuMemGetAddressRange, cuMemGetInfo, cuMemHostAlloc, cuMemHostGetDevicePointer, cuMemsetD2D8, cuMemsetD2D8Async, cuMemsetD2D16, cuMemsetD2D16Async, cuMemsetD2D32, cuMemsetD2D32Async, cuMemsetD8Async, cuMemsetD16, cuMemsetD16Async, cuMemsetD32, cuMemsetD32Async

5.37.2.60 **CUresult cuMemsetD8Async** (CUdeviceptr *dstDevice*, unsigned char *uc*, size_t *N*, CUstream *hStream*)

Sets the memory range of *N* 8-bit values to the specified value *uc*.

`cuMemsetD8Async()` is asynchronous and can optionally be associated to a stream by passing a non-zero `stream` argument.

Parameters:

dstDevice - Destination device pointer

uc - Value to set

N - Number of elements

hStream - Stream identifier

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuArray3DCreate`, `cuArray3DGetDescriptor`, `cuArrayCreate`, `cuArrayDestroy`, `cuArrayGetDescriptor`, `cuMemAlloc`, `cuMemAllocHost`, `cuMemAllocPitch`, `cuMemcpy2D`, `cuMemcpy2DAsync`, `cuMemcpy2DUnaligned`, `cuMemcpy3D`, `cuMemcpy3DAsync`, `cuMemcpyAtoA`, `cuMemcpyAtoD`, `cuMemcpyAtoH`, `cuMemcpyAtoHAsync`, `cuMemcpyDtoA`, `cuMemcpyDtoD`, `cuMemcpyDtoDAsync`, `cuMemcpyDtoH`, `cuMemcpyDtoHAsync`, `cuMemcpyHtoA`, `cuMemcpyHtoAAsync`, `cuMemcpyHtoD`, `cuMemcpyHtoDAsync`, `cuMemFree`, `cuMemFreeHost`, `cuMemGetAddressRange`, `cuMemGetInfo`, `cuMemHostAlloc`, `cuMemHostGetDevicePointer`, `cuMemsetD2D8`, `cuMemsetD2D8Async`, `cuMemsetD2D16`, `cuMemsetD2D16Async`, `cuMemsetD2D32`, `cuMemsetD2D32Async`, `cuMemsetD8`, `cuMemsetD16`, `cuMemsetD16Async`, `cuMemsetD32`, `cuMemsetD32Async`

5.37.2.61 **CUresult cuMipmappedArrayCreate** (CUmipmappedArray * *pHandle*, const CUDA_ARRAY3D_DESCRIPTOR * *pMipmappedArrayDesc*, unsigned int *numMipmapLevels*)

Creates a CUDA mipmapped array according to the `CUDA_ARRAY3D_DESCRIPTOR` structure `pMipmappedArrayDesc` and returns a handle to the new CUDA mipmapped array in `*pHandle`. `numMipmapLevels` specifies the number of mipmap levels to be allocated. This value is clamped to the range $[1, 1 + \text{floor}(\log_2(\max(\text{width}, \text{height}, \text{depth})))]$.

The `CUDA_ARRAY3D_DESCRIPTOR` is defined as:

```
typedef struct {
    unsigned int Width;
    unsigned int Height;
    unsigned int Depth;
    CUarray_format Format;
    unsigned int NumChannels;
    unsigned int Flags;
} CUDA_ARRAY3D_DESCRIPTOR;
```

where:

- `Width`, `Height`, and `Depth` are the width, height, and depth of the CUDA array (in elements); the following types of CUDA arrays can be allocated:
 - A 1D mipmapped array is allocated if `Height` and `Depth` extents are both zero.
 - A 2D mipmapped array is allocated if only `Depth` extent is zero.
 - A 3D mipmapped array is allocated if all three extents are non-zero.
 - A 1D layered CUDA mipmapped array is allocated if only `Height` is zero and the `CUDA_ARRAY3D_LAYERED` flag is set. Each layer is a 1D array. The number of layers is determined by the depth extent.
 - A 2D layered CUDA mipmapped array is allocated if all three extents are non-zero and the `CUDA_ARRAY3D_LAYERED` flag is set. Each layer is a 2D array. The number of layers is determined by the depth extent.
 - A cubemap CUDA mipmapped array is allocated if all three extents are non-zero and the `CUDA_ARRAY3D_CUBEMAP` flag is set. `Width` must be equal to `Height`, and `Depth` must be six. A cubemap is a special type of 2D layered CUDA array, where the six layers represent the six faces of a cube. The order of the six layers in memory is the same as that listed in `CUarray_cubemap_face`.
 - A cubemap layered CUDA mipmapped array is allocated if all three extents are non-zero, and both, `CUDA_ARRAY3D_CUBEMAP` and `CUDA_ARRAY3D_LAYERED` flags are set. `Width` must be equal to `Height`, and `Depth` must be a multiple of six. A cubemap layered CUDA array is a special type of 2D layered CUDA array that consists of a collection of cubemaps. The first six layers represent the first cubemap, the next six layers form the second cubemap, and so on.
- `Format` specifies the format of the elements; `CUarray_format` is defined as:

```
typedef enum CUarray_format_enum {
    CU_AD_FORMAT_UNSIGNED_INT8 = 0x01,
    CU_AD_FORMAT_UNSIGNED_INT16 = 0x02,
    CU_AD_FORMAT_UNSIGNED_INT32 = 0x03,
    CU_AD_FORMAT_SIGNED_INT8 = 0x08,
    CU_AD_FORMAT_SIGNED_INT16 = 0x09,
    CU_AD_FORMAT_SIGNED_INT32 = 0x0a,
    CU_AD_FORMAT_HALF = 0x10,
    CU_AD_FORMAT_FLOAT = 0x20
} CUarray_format;
```

- `NumChannels` specifies the number of packed components per CUDA array element; it may be 1, 2, or 4;
- Flags may be set to
 - `CUDA_ARRAY3D_LAYERED` to enable creation of layered CUDA mipmapped arrays. If this flag is set, `Depth` specifies the number of layers, not the depth of a 3D array.
 - `CUDA_ARRAY3D_SURFACE_LDST` to enable surface references to be bound to individual mipmap levels of the CUDA mipmapped array. If this flag is not set, `cuSurfRefSetArray` will fail when attempting to bind a mipmap level of the CUDA mipmapped array to a surface reference.
 - `CUDA_ARRAY3D_CUBEMAP` to enable creation of mipmapped cubemaps. If this flag is set, `Width` must be equal to `Height`, and `Depth` must be six. If the `CUDA_ARRAY3D_LAYERED` flag is also set, then `Depth` must be a multiple of six.
 - `CUDA_ARRAY3D_TEXTURE_GATHER` to indicate that the CUDA mipmapped array will be used for texture gather. Texture gather can only be performed on 2D CUDA mipmapped arrays.

`Width`, `Height` and `Depth` must meet certain size requirements as listed in the following table. All values are specified in elements. Note that for brevity's sake, the full name of the device attribute is not specified. For ex., `TEXTURE1D_MIPMAPPED_WIDTH` refers to the device attribute `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE1D_MIPMAPPED_WIDTH`.

CUDA array type	Valid extents that must always be met {(width range in elements), (height range), (depth range)}
1D	{ (1,TEXTURE1D_MIPMAPPED_WIDTH), 0, 0 }
2D	{ (1,TEXTURE2D_MIPMAPPED_WIDTH), (1,TEXTURE2D_MIPMAPPED_HEIGHT), 0 }
3D	{ (1,TEXTURE3D_WIDTH), (1,TEXTURE3D_HEIGHT), (1,TEXTURE3D_DEPTH) } OR { (1,TEXTURE3D_WIDTH_ALTERNATE), (1,TEXTURE3D_HEIGHT_ALTERNATE), (1,TEXTURE3D_DEPTH_ALTERNATE) }
1D Layered	{ (1,TEXTURE1D_LAYERED_WIDTH), 0, (1,TEXTURE1D_LAYERED_LAYERS) }
2D Layered	{ (1,TEXTURE2D_LAYERED_WIDTH), (1,TEXTURE2D_LAYERED_HEIGHT), (1,TEXTURE2D_LAYERED_LAYERS) }
Cubemap	{ (1,TEXTURECUBEMAP_WIDTH), (1,TEXTURECUBEMAP_WIDTH), 6 }
Cubemap Layered	{ (1,TEXTURECUBEMAP_LAYERED_WIDTH), (1,TEXTURECUBEMAP_LAYERED_WIDTH), (1,TEXTURECUBEMAP_LAYERED_LAYERS) }

Parameters:

- pHandle* - Returned mipmapped array
pMipmappedArrayDesc - mipmapped array descriptor
numMipmapLevels - Number of mipmap levels

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_OUT_OF_MEMORY, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMipmappedArrayDestroy](#), [cuMipmappedArrayGetLevel](#), [cuArrayCreate](#),

5.37.2.62 CUresult cuMipmappedArrayDestroy (CUmipmappedArray hMipmappedArray)

Destroys the CUDA mipmapped array hMipmappedArray.

Parameters:

- hMipmappedArray* - Mipmapped array to destroy

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_ARRAY_IS_MAPPED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMipmappedArrayCreate](#), [cuMipmappedArrayGetLevel](#), [cuArrayCreate](#),

5.37.2.63 CUresult cuMipmappedArrayGetLevel (CUarray * *pLevelArray*, CUmipmappedArray *hMipmappedArray*, unsigned int *level*)

Returns in **pLevelArray* a CUDA array that represents a single mipmap level of the CUDA mipmapped array *hMipmappedArray*.

If *level* is greater than the maximum number of levels in this mipmapped array, [CUDA_ERROR_INVALID_VALUE](#) is returned.

Parameters:

pLevelArray - Returned mipmap level CUDA array

hMipmappedArray - CUDA mipmapped array

level - Mipmap level

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMipmappedArrayCreate](#), [cuMipmappedArrayDestroy](#), [cuArrayCreate](#),

5.38 Unified Addressing

Functions

- [CUresult cuPointerGetAttribute](#) (void *data, [CUpointer_attribute](#) attribute, [CUdeviceptr](#) ptr)

Returns information about a pointer.

5.38.1 Detailed Description

This section describes the unified addressing functions of the low-level CUDA driver application programming interface.

5.38.2 Overview

CUDA devices can share a unified address space with the host. For these devices there is no distinction between a device pointer and a host pointer – the same pointer value may be used to access memory from the host program and from a kernel running on the device (with exceptions enumerated below).

5.38.3 Supported Platforms

Whether or not a device supports unified addressing may be queried by calling [cuDeviceGetAttribute\(\)](#) with the device attribute [CU_DEVICE_ATTRIBUTE_UNIFIED_ADDRESSING](#).

Unified addressing is automatically enabled in 64-bit processes on devices with compute capability greater than or equal to 2.0.

5.38.4 Looking Up Information from Pointer Values

It is possible to look up information about the memory which backs a pointer value. For instance, one may want to know if a pointer points to host or device memory. As another example, in the case of device memory, one may want to know on which CUDA device the memory resides. These properties may be queried using the function [cuPointerGetAttribute\(\)](#)

Since pointers are unique, it is not necessary to specify information about the pointers specified to the various copy functions in the CUDA API. The function [cuMemcpy\(\)](#) may be used to perform a copy between two pointers, ignoring whether they point to host or device memory (making [cuMemcpyHtoD\(\)](#), [cuMemcpyDtoD\(\)](#), and [cuMemcpyDtoH\(\)](#) unnecessary for devices supporting unified addressing). For multidimensional copies, the memory type [CU_MEMORYTYPE_UNIFIED](#) may be used to specify that the CUDA driver should infer the location of the pointer from its value.

5.38.5 Automatic Mapping of Host Allocated Host Memory

All host memory allocated in all contexts using [cuMemAllocHost\(\)](#) and [cuMemHostAlloc\(\)](#) is always directly accessible from all contexts on all devices that support unified addressing. This is the case regardless of whether or not the flags [CU_MEMHOSTALLOC_PORTABLE](#) and [CU_MEMHOSTALLOC_DEVICEMAP](#) are specified.

The pointer value through which allocated host memory may be accessed in kernels on all devices that support unified addressing is the same as the pointer value through which that memory is accessed on the host, so it is not necessary to call [cuMemHostGetDevicePointer\(\)](#) to get the device pointer for these allocations.

Note that this is not the case for memory allocated using the flag `CU_MEMHOSTALLOC_WRITECOMBINED`, as discussed below.

5.38.6 Automatic Registration of Peer Memory

Upon enabling direct access from a context that supports unified addressing to another peer context that supports unified addressing using `cuCtxEnablePeerAccess()` all memory allocated in the peer context using `cuMemAlloc()` and `cuMemAllocPitch()` will immediately be accessible by the current context. The device pointer value through which any peer memory may be accessed in the current context is the same pointer value through which that memory may be accessed in the peer context.

5.38.7 Exceptions, Disjoint Addressing

Not all memory may be accessed on devices through the same pointer value through which they are accessed on the host. These exceptions are host memory registered using `cuMemHostRegister()` and host memory allocated using the flag `CU_MEMHOSTALLOC_WRITECOMBINED`. For these exceptions, there exists a distinct host and device address for the memory. The device address is guaranteed to not overlap any valid host pointer range and is guaranteed to have the same value across all contexts that support unified addressing.

This device address may be queried using `cuMemHostGetDevicePointer()` when a context using unified addressing is current. Either the host or the unified device pointer value may be used to refer to this memory through `cuMemcpy()` and similar functions using the `CU_MEMORYTYPE_UNIFIED` memory type.

5.38.8 Function Documentation

5.38.8.1 `CUresult cuPointerGetAttribute (void * data, CUpointer_attribute attribute, CUdeviceptr ptr)`

The supported attributes are:

- `CU_POINTER_ATTRIBUTE_CONTEXT`:

Returns in `*data` the `CUcontext` in which `ptr` was allocated or registered. The type of `data` must be `CUcontext *`.

If `ptr` was not allocated by, mapped by, or registered with a `CUcontext` which uses unified virtual addressing then `CUDA_ERROR_INVALID_VALUE` is returned.

- `CU_POINTER_ATTRIBUTE_MEMORY_TYPE`:

Returns in `*data` the physical memory type of the memory that `ptr` addresses as a `CUmemorytype` enumerated value. The type of `data` must be unsigned int.

If `ptr` addresses device memory then `*data` is set to `CU_MEMORYTYPE_DEVICE`. The particular `CUdevice` on which the memory resides is the `CUdevice` of the `CUcontext` returned by the `CU_POINTER_ATTRIBUTE_CONTEXT` attribute of `ptr`.

If `ptr` addresses host memory then `*data` is set to `CU_MEMORYTYPE_HOST`.

If `ptr` was not allocated by, mapped by, or registered with a `CUcontext` which uses unified virtual addressing then `CUDA_ERROR_INVALID_VALUE` is returned.

If the current `CUcontext` does not support unified virtual addressing then `CUDA_ERROR_INVALID_CONTEXT` is returned.

- `CU_POINTER_ATTRIBUTE_DEVICE_POINTER`:

Returns in `*data` the device pointer value through which `ptr` may be accessed by kernels running in the current `CUcontext`. The type of `data` must be `CUdeviceptr *`.

If there exists no device pointer value through which kernels running in the current `CUcontext` may access `ptr` then `CUDA_ERROR_INVALID_VALUE` is returned.

If there is no current `CUcontext` then `CUDA_ERROR_INVALID_CONTEXT` is returned.

Except in the exceptional disjoint addressing cases discussed below, the value returned in `*data` will equal the input value `ptr`.

- `CU_POINTER_ATTRIBUTE_HOST_POINTER`:

Returns in `*data` the host pointer value through which `ptr` may be accessed by the host program. The type of `data` must be `void **`. If there exists no host pointer value through which the host program may directly access `ptr` then `CUDA_ERROR_INVALID_VALUE` is returned.

Except in the exceptional disjoint addressing cases discussed below, the value returned in `*data` will equal the input value `ptr`.

- `CU_POINTER_ATTRIBUTE_P2P_TOKENS`:

Returns in `*data` two tokens for use with the `nv-p2p.h` Linux kernel interface. `data` must be a struct of type `CUDA_POINTER_ATTRIBUTE_P2P_TOKENS`.

`ptr` must be a pointer to memory obtained from `cuMemAlloc()`. Note that `p2pToken` and `vaSpaceToken` are only valid for the lifetime of the source allocation. A subsequent allocation at the same address may return completely different tokens.

Note that for most allocations in the unified virtual address space the host and device pointer for accessing the allocation will be the same. The exceptions to this are

- user memory registered using `cuMemHostRegister`
- host memory allocated using `cuMemHostAlloc` with the `CU_MEMHOSTALLOC_WRITECOMBINED` flag
For these types of allocation there will exist separate, disjoint host and device addresses for accessing the allocation. In particular
- The host address will correspond to an invalid unmapped device address (which will result in an exception if accessed from the device)
- The device address will correspond to an invalid unmapped host address (which will result in an exception if accessed from the host). For these types of allocations, querying `CU_POINTER_ATTRIBUTE_HOST_POINTER` and `CU_POINTER_ATTRIBUTE_DEVICE_POINTER` may be used to retrieve the host and device addresses from either address.

Parameters:

data - Returned pointer attribute value

attribute - Pointer attribute to query

ptr - Pointer

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`, `CUDA_ERROR_INVALID_DEVICE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuMemAlloc](#), [cuMemFree](#), [cuMemAllocHost](#), [cuMemFreeHost](#), [cuMemHostAlloc](#), [cuMemHostRegister](#), [cuMemHostUnregister](#)

5.39 Stream Management

Functions

- [CUresult cuStreamAddCallback](#) ([CUstream](#) hStream, [CUstreamCallback](#) callback, void *userData, unsigned int flags)
Add a callback to a compute stream.
- [CUresult cuStreamCreate](#) ([CUstream](#) *phStream, unsigned int Flags)
Create a stream.
- [CUresult cuStreamDestroy](#) ([CUstream](#) hStream)
Destroys a stream.
- [CUresult cuStreamQuery](#) ([CUstream](#) hStream)
Determine status of a compute stream.
- [CUresult cuStreamSynchronize](#) ([CUstream](#) hStream)
Wait until a stream's tasks are completed.
- [CUresult cuStreamWaitEvent](#) ([CUstream](#) hStream, [CUevent](#) hEvent, unsigned int Flags)
Make a compute stream wait on an event.

5.39.1 Detailed Description

This section describes the stream management functions of the low-level CUDA driver application programming interface.

5.39.2 Function Documentation

5.39.2.1 [CUresult cuStreamAddCallback](#) ([CUstream](#) *hStream*, [CUstreamCallback](#) *callback*, void * *userData*, unsigned int *flags*)

Adds a callback to be called on the host after all currently enqueued items in the stream have completed. For each [cuStreamAddCallback](#) call, the callback will be executed exactly once. The callback will block later work in the stream until it is finished.

The callback may be passed [CUDA_SUCCESS](#) or an error code. In the event of a device error, all subsequently executed callbacks will receive an appropriate [CUresult](#).

Callbacks must not make any CUDA API calls. Attempting to use a CUDA API will result in [CUDA_ERROR_NOT_PERMITTED](#). Callbacks must not perform any synchronization that may depend on outstanding device work or other callbacks that are not mandated to run earlier. Callbacks without a mandated order (in independent streams) execute in undefined order and may be serialized.

This API requires compute capability 1.1 or greater. See [cuDeviceGetAttribute](#) or [cuDeviceGetProperties](#) to query compute capability. Attempting to use this API with earlier compute versions will return [CUDA_ERROR_NOT_SUPPORTED](#).

Parameters:

hStream - Stream to add callback to

callback - The function to call once preceding stream operations are complete

userData - User specified data to be passed to the callback function

flags - Reserved for future use, must be 0

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_SUPPORTED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuStreamCreate](#), [cuStreamQuery](#), [cuStreamSynchronize](#), [cuStreamWaitEvent](#), [cuStreamDestroy](#)

5.39.2.2 CUresult cuStreamCreate (CUstream * *phStream*, unsigned int *Flags*)

Creates a stream and returns a handle in `phStream`. The `Flags` argument determines behaviors of the stream. Valid values for `Flags` are:

- [CU_STREAM_DEFAULT](#): Default stream creation flag.
- [CU_STREAM_NON_BLOCKING](#): Specifies that work running in the created stream may run concurrently with work in stream 0 (the NULL stream), and that the created stream should perform no implicit synchronization with stream 0.

Parameters:

phStream - Returned newly created stream

Flags - Parameters for stream creation

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuStreamDestroy](#), [cuStreamWaitEvent](#), [cuStreamQuery](#), [cuStreamSynchronize](#), [cuStreamAddCallback](#)

5.39.2.3 CUresult cuStreamDestroy (CUstream *hStream*)

Destroys the stream specified by `hStream`.

In case the device is still doing work in the stream `hStream` when [cuStreamDestroy\(\)](#) is called, the function will return immediately and the resources associated with `hStream` will be released automatically once the device has completed all work in `hStream`.

Parameters:

hStream - Stream to destroy

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuStreamCreate](#), [cuStreamWaitEvent](#), [cuStreamQuery](#), [cuStreamSynchronize](#), [cuStreamAddCallback](#)

5.39.2.4 CUresult cuStreamQuery (CUstream hStream)

Returns [CUDA_SUCCESS](#) if all operations in the stream specified by *hStream* have completed, or [CUDA_ERROR_NOT_READY](#) if not.

Parameters:

hStream - Stream to query status of

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_NOT_READY

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuStreamCreate](#), [cuStreamWaitEvent](#), [cuStreamDestroy](#), [cuStreamSynchronize](#), [cuStreamAddCallback](#)

5.39.2.5 CUresult cuStreamSynchronize (CUstream hStream)

Waits until the device has completed all operations in the stream specified by *hStream*. If the context was created with the [CU_CTX_SCHED_BLOCKING_SYNC](#) flag, the CPU thread will block until the stream is finished with all of its tasks.

Parameters:

hStream - Stream to wait for

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuStreamCreate](#), [cuStreamDestroy](#), [cuStreamWaitEvent](#), [cuStreamQuery](#), [cuStreamAddCallback](#)

5.39.2.6 CUresult cuStreamWaitEvent (CUstream *hStream*, CUevent *hEvent*, unsigned int *Flags*)

Makes all future work submitted to *hStream* wait until *hEvent* reports completion before beginning execution. This synchronization will be performed efficiently on the device. The event *hEvent* may be from a different context than *hStream*, in which case this function will perform cross-device synchronization.

The stream *hStream* will wait only for the completion of the most recent host call to [cuEventRecord\(\)](#) on *hEvent*. Once this call has returned, any functions (including [cuEventRecord\(\)](#) and [cuEventDestroy\(\)](#)) may be called on *hEvent* again, and subsequent calls will not have any effect on *hStream*.

If *hStream* is 0 (the NULL stream) any future work submitted in any stream will wait for *hEvent* to complete before beginning execution. This effectively creates a barrier for all future work submitted to the context.

If [cuEventRecord\(\)](#) has not been called on *hEvent*, this call acts as if the record has already completed, and so is a functional no-op.

Parameters:

hStream - Stream to wait

hEvent - Event to wait on (may not be NULL)

Flags - Parameters for the operation (must be 0)

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuStreamCreate](#), [cuEventRecord](#), [cuStreamQuery](#), [cuStreamSynchronize](#), [cuStreamAddCallback](#), [cuStreamDestroy](#)

5.40 Event Management

Functions

- **CUresult cuEventCreate** (CUevent *phEvent, unsigned int Flags)
Creates an event.
- **CUresult cuEventDestroy** (CUevent hEvent)
Destroys an event.
- **CUresult cuEventElapsedTime** (float *pMilliseconds, CUevent hStart, CUevent hEnd)
Computes the elapsed time between two events.
- **CUresult cuEventQuery** (CUevent hEvent)
Queries an event's status.
- **CUresult cuEventRecord** (CUevent hEvent, CUstream hStream)
Records an event.
- **CUresult cuEventSynchronize** (CUevent hEvent)
Waits for an event to complete.

5.40.1 Detailed Description

This section describes the event management functions of the low-level CUDA driver application programming interface.

5.40.2 Function Documentation

5.40.2.1 CUresult cuEventCreate (CUevent *phEvent, unsigned int Flags)

Creates an event *phEvent with the flags specified via `Flags`. Valid flags include:

- **CU_EVENT_DEFAULT**: Default event creation flag.
- **CU_EVENT_BLOCKING_SYNC**: Specifies that the created event should use blocking synchronization. A CPU thread that uses `cuEventSynchronize()` to wait on an event created with this flag will block until the event has actually been recorded.
- **CU_EVENT_DISABLE_TIMING**: Specifies that the created event does not need to record timing data. Events created with this flag specified and the **CU_EVENT_BLOCKING_SYNC** flag not specified will provide the best performance when used with `cuStreamWaitEvent()` and `cuEventQuery()`.
- **CU_EVENT_INTERPROCESS**: Specifies that the created event may be used as an interprocess event by `cuIpcGetEventHandle()`. **CU_EVENT_INTERPROCESS** must be specified along with **CU_EVENT_DISABLE_TIMING**.

Parameters:

- phEvent* - Returns newly created event
- Flags* - Event creation flags

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuEventRecord](#), [cuEventQuery](#), [cuEventSynchronize](#), [cuEventDestroy](#), [cuEventElapsedTime](#)

5.40.2.2 CUresult cuEventDestroy (CUevent hEvent)

Destroys the event specified by `hEvent`.

In case `hEvent` has been recorded but has not yet been completed when `cuEventDestroy()` is called, the function will return immediately and the resources associated with `hEvent` will be released automatically once the device has completed `hEvent`.

Parameters:

hEvent - Event to destroy

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuEventCreate](#), [cuEventRecord](#), [cuEventQuery](#), [cuEventSynchronize](#), [cuEventElapsedTime](#)

5.40.2.3 CUresult cuEventElapsedTime (float *pMilliseconds, CUevent hStart, CUevent hEnd)

Computes the elapsed time between two events (in milliseconds with a resolution of around 0.5 microseconds).

If either event was last recorded in a non-NULL stream, the resulting time may be greater than expected (even if both used the same stream handle). This happens because the `cuEventRecord()` operation takes place asynchronously and there is no guarantee that the measured latency is actually just between the two events. Any number of other different stream operations could execute in between the two measured events, thus altering the timing in a significant way.

If `cuEventRecord()` has not been called on either event then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If `cuEventRecord()` has been called on both events but one or both of them has not yet been completed (that is, `cuEventQuery()` would return [CUDA_ERROR_NOT_READY](#) on at least one of the events), [CUDA_ERROR_NOT_READY](#) is returned. If either event was created with the [CU_EVENT_DISABLE_TIMING](#) flag, then this function will return [CUDA_ERROR_INVALID_HANDLE](#).

Parameters:

pMilliseconds - Time between `hStart` and `hEnd` in ms

hStart - Starting event

hEnd - Ending event

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_READY](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuEventCreate](#), [cuEventRecord](#), [cuEventQuery](#), [cuEventSynchronize](#), [cuEventDestroy](#)

5.40.2.4 CUresult cuEventQuery (CUevent hEvent)

Query the status of all device work preceding the most recent call to [cuEventRecord\(\)](#) (in the appropriate compute streams, as specified by the arguments to [cuEventRecord\(\)](#)).

If this work has successfully been completed by the device, or if [cuEventRecord\(\)](#) has not been called on *hEvent*, then [CUDA_SUCCESS](#) is returned. If this work has not yet been completed by the device then [CUDA_ERROR_NOT_READY](#) is returned.

Parameters:

hEvent - Event to query

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_READY](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuEventCreate](#), [cuEventRecord](#), [cuEventSynchronize](#), [cuEventDestroy](#), [cuEventElapsedTime](#)

5.40.2.5 CUresult cuEventRecord (CUevent hEvent, CUstream hStream)

Records an event. If *hStream* is non-zero, the event is recorded after all preceding operations in *hStream* have been completed; otherwise, it is recorded after all preceding operations in the CUDA context have been completed. Since operation is asynchronous, [cuEventQuery](#) and/or [cuEventSynchronize\(\)](#) must be used to determine when the event has actually been recorded.

If [cuEventRecord\(\)](#) has previously been called on *hEvent*, then this call will overwrite any existing state in *hEvent*. Any subsequent calls which examine the status of *hEvent* will only examine the completion of this most recent call to [cuEventRecord\(\)](#).

It is necessary that *hEvent* and *hStream* be created on the same context.

Parameters:

hEvent - Event to record
hStream - Stream to record event for

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuEventCreate](#), [cuEventQuery](#), [cuEventSynchronize](#), [cuStreamWaitEvent](#), [cuEventDestroy](#), [cuEventElapsedTime](#)

5.40.2.6 CUresult cuEventSynchronize (CUevent *hEvent*)

Wait until the completion of all device work preceding the most recent call to [cuEventRecord\(\)](#) (in the appropriate compute streams, as specified by the arguments to [cuEventRecord\(\)](#)).

If [cuEventRecord\(\)](#) has not been called on *hEvent*, [CUDA_SUCCESS](#) is returned immediately.

Waiting for an event that was created with the [CU_EVENT_BLOCKING_SYNC](#) flag will cause the calling CPU thread to block until the event has been completed by the device. If the [CU_EVENT_BLOCKING_SYNC](#) flag has not been set, then the CPU thread will busy-wait until the event has been completed by the device.

Parameters:

hEvent - Event to wait for

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuEventCreate](#), [cuEventRecord](#), [cuEventQuery](#), [cuEventDestroy](#), [cuEventElapsedTime](#)

5.41 Execution Control

Functions

- **CUresult cuFuncGetAttribute** (int *pi, CUfunction_attribute attrib, CUfunction hfunc)
Returns information about a function.
- **CUresult cuFuncSetCacheConfig** (CUfunction hfunc, CUfunc_cache config)
Sets the preferred cache configuration for a device function.
- **CUresult cuFuncSetSharedMemConfig** (CUfunction hfunc, CUsharedconfig config)
Sets the shared memory configuration for a device function.
- **CUresult cuLaunchKernel** (CUfunction f, unsigned int gridDimX, unsigned int gridDimY, unsigned int gridDimZ, unsigned int blockDimX, unsigned int blockDimY, unsigned int blockDimZ, unsigned int sharedMemBytes, CUstream hStream, void **kernelParams, void **extra)
Launches a CUDA function.

5.41.1 Detailed Description

This section describes the execution control functions of the low-level CUDA driver application programming interface.

5.41.2 Function Documentation

5.41.2.1 CUresult cuFuncGetAttribute (int * pi, CUfunction_attribute attrib, CUfunction hfunc)

Returns in *pi the integer value of the attribute attrib on the kernel given by hfunc. The supported attributes are:

- **CU_FUNC_ATTRIBUTE_MAX_THREADS_PER_BLOCK**: The maximum number of threads per block, beyond which a launch of the function would fail. This number depends on both the function and the device on which the function is currently loaded.
- **CU_FUNC_ATTRIBUTE_SHARED_SIZE_BYTES**: The size in bytes of statically-allocated shared memory per block required by this function. This does not include dynamically-allocated shared memory requested by the user at runtime.
- **CU_FUNC_ATTRIBUTE_CONST_SIZE_BYTES**: The size in bytes of user-allocated constant memory required by this function.
- **CU_FUNC_ATTRIBUTE_LOCAL_SIZE_BYTES**: The size in bytes of local memory used by each thread of this function.
- **CU_FUNC_ATTRIBUTE_NUM_REGS**: The number of registers used by each thread of this function.
- **CU_FUNC_ATTRIBUTE_PTX_VERSION**: The PTX virtual architecture version for which the function was compiled. This value is the major PTX version * 10 + the minor PTX version, so a PTX version 1.3 function would return the value 13. Note that this may return the undefined value of 0 for cubins compiled prior to CUDA 3.0.

- [CU_FUNC_ATTRIBUTE_BINARY_VERSION](#): The binary architecture version for which the function was compiled. This value is the major binary version * 10 + the minor binary version, so a binary version 1.3 function would return the value 13. Note that this will return a value of 10 for legacy cubins that do not have a properly-encoded binary architecture version.

Parameters:

- pi* - Returned attribute value
- attrib* - Attribute requested
- hfunc* - Function to query attribute of

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxGetCacheConfig](#), [cuCtxSetCacheConfig](#), [cuFuncSetCacheConfig](#), [cuLaunchKernel](#)

5.41.2.2 CUresult cuFuncSetCacheConfig (CUfunction *hfunc*, CUfunc_cache *config*)

On devices where the L1 cache and shared memory use the same hardware resources, this sets through *config* the preferred cache configuration for the device function *hfunc*. This is only a preference. The driver will use the requested configuration if possible, but it is free to choose a different configuration if required to execute *hfunc*. Any context-wide preference set via [cuCtxSetCacheConfig\(\)](#) will be overridden by this per-function setting unless the per-function setting is [CU_FUNC_CACHE_PREFER_NONE](#). In that case, the current context-wide setting will be used.

This setting does nothing on devices where the size of the L1 cache and shared memory are fixed.

Launching a kernel with a different preference than the most recent preference setting may insert a device-side synchronization point.

The supported cache configurations are:

- [CU_FUNC_CACHE_PREFER_NONE](#): no preference for shared memory or L1 (default)
- [CU_FUNC_CACHE_PREFER_SHARED](#): prefer larger shared memory and smaller L1 cache
- [CU_FUNC_CACHE_PREFER_L1](#): prefer larger L1 cache and smaller shared memory
- [CU_FUNC_CACHE_PREFER_EQUAL](#): prefer equal sized L1 cache and shared memory

Parameters:

- hfunc* - Kernel to configure cache for
- config* - Requested cache configuration

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxGetCacheConfig](#), [cuCtxSetCacheConfig](#), [cuFuncGetAttribute](#), [cuLaunchKernel](#)

5.41.2.3 CUresult cuFuncSetSharedMemConfig (CUfunction *hfunc*, CUsharedconfig *config*)

On devices with configurable shared memory banks, this function will force all subsequent launches of the specified device function to have the given shared memory bank size configuration. On any given launch of the function, the shared memory configuration of the device will be temporarily changed if needed to suit the function's preferred configuration. Changes in shared memory configuration between subsequent launches of functions, may introduce a device side synchronization point.

Any per-function setting of shared memory bank size set via [cuFuncSetSharedMemConfig](#) will override the context wide setting set with [cuCtxSetSharedMemConfig](#).

Changing the shared memory bank size will not increase shared memory usage or affect occupancy of kernels, but may have major effects on performance. Larger bank sizes will allow for greater potential bandwidth to shared memory, but will change what kinds of accesses to shared memory will result in bank conflicts.

This function will do nothing on devices with fixed shared memory bank size.

The supported bank configurations are:

- [CU_SHARED_MEM_CONFIG_DEFAULT_BANK_SIZE](#): use the context's shared memory configuration when launching this function.
- [CU_SHARED_MEM_CONFIG_FOUR_BYTE_BANK_SIZE](#): set shared memory bank width to be natively four bytes when launching this function.
- [CU_SHARED_MEM_CONFIG_EIGHT_BYTE_BANK_SIZE](#): set shared memory bank width to be natively eight bytes when launching this function.

Parameters:

hfunc - kernel to be given a shared memory config

config - requested shared memory configuration

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxGetCacheConfig](#), [cuCtxSetCacheConfig](#), [cuCtxGetSharedMemConfig](#), [cuCtxSetSharedMemConfig](#), [cuFuncGetAttribute](#), [cuLaunchKernel](#)

5.41.2.4 `CUresult cuLaunchKernel` (CUfunction *f*, unsigned int *gridDimX*, unsigned int *gridDimY*, unsigned int *gridDimZ*, unsigned int *blockDimX*, unsigned int *blockDimY*, unsigned int *blockDimZ*, unsigned int *sharedMemBytes*, CUstream *hStream*, void ** *kernelParams*, void ** *extra*)

Invokes the kernel *f* on a *gridDimX* x *gridDimY* x *gridDimZ* grid of blocks. Each block contains *blockDimX* x *blockDimY* x *blockDimZ* threads.

sharedMemBytes sets the amount of dynamic shared memory that will be available to each thread block.

`cuLaunchKernel()` can optionally be associated to a stream by passing a non-zero *hStream* argument.

Kernel parameters to *f* can be specified in one of two ways:

1) Kernel parameters can be specified via *kernelParams*. If *f* has *N* parameters, then *kernelParams* needs to be an array of *N* pointers. Each of *kernelParams*[0] through *kernelParams*[*N*-1] must point to a region of memory from which the actual kernel parameter will be copied. The number of kernel parameters and their offsets and sizes do not need to be specified as that information is retrieved directly from the kernel's image.

2) Kernel parameters can also be packaged by the application into a single buffer that is passed in via the *extra* parameter. This places the burden on the application of knowing each kernel parameter's size and alignment/padding within the buffer. Here is an example of using the *extra* parameter in this manner:

```
size_t argBufferSize;
char argBuffer[256];

// populate argBuffer and argBufferSize

void *config[] = {
    CU_LAUNCH_PARAM_BUFFER_POINTER, argBuffer,
    CU_LAUNCH_PARAM_BUFFER_SIZE,   &argBufferSize,
    CU_LAUNCH_PARAM_END
};
status = cuLaunchKernel(f, gx, gy, gz, bx, by, bz, sh, s, NULL, config);
```

The *extra* parameter exists to allow `cuLaunchKernel` to take additional less commonly used arguments. *extra* specifies a list of names of extra settings and their corresponding values. Each extra setting name is immediately followed by the corresponding value. The list must be terminated with either NULL or `CU_LAUNCH_PARAM_END`.

- `CU_LAUNCH_PARAM_END`, which indicates the end of the *extra* array;
- `CU_LAUNCH_PARAM_BUFFER_POINTER`, which specifies that the next value in *extra* will be a pointer to a buffer containing all the kernel parameters for launching kernel *f*;
- `CU_LAUNCH_PARAM_BUFFER_SIZE`, which specifies that the next value in *extra* will be a pointer to a `size_t` containing the size of the buffer specified with `CU_LAUNCH_PARAM_BUFFER_POINTER`;

The error `CUDA_ERROR_INVALID_VALUE` will be returned if kernel parameters are specified with both *kernelParams* and *extra* (i.e. both *kernelParams* and *extra* are non-NULL).

Calling `cuLaunchKernel()` sets persistent function state that is the same as function state set through the following deprecated APIs:

`cuFuncSetBlockShape()` `cuFuncSetSharedSize()` `cuParamSetSize()` `cuParamSeti()` `cuParamSetf()` `cuParamSetv()`

When the kernel *f* is launched via `cuLaunchKernel()`, the previous block shape, shared size and parameter info associated with *f* is overwritten.

Note that to use `cuLaunchKernel()`, the kernel *f* must either have been compiled with toolchain version 3.2 or later so that it will contain kernel parameter information, or have no kernel parameters. If either of these conditions is not met, then `cuLaunchKernel()` will return `CUDA_ERROR_INVALID_IMAGE`.

Parameters:

f - Kernel to launch

gridDimX - Width of grid in blocks

gridDimY - Height of grid in blocks

gridDimZ - Depth of grid in blocks

blockDimX - X dimension of each thread block

blockDimY - Y dimension of each thread block

blockDimZ - Z dimension of each thread block

sharedMemBytes - Dynamic shared-memory size per thread block in bytes

hStream - Stream identifier

kernelParams - Array of pointers to kernel parameters

extra - Extra options

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_INVALID_IMAGE, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_LAUNCH_FAILED, CUDA_ERROR_LAUNCH_OUT_OF_RESOURCES, CUDA_ERROR_LAUNCH_TIMEOUT, CUDA_ERROR_LAUNCH_INCOMPATIBLE_TEXTURING, CUDA_ERROR_SHARED_OBJECT_INIT_FAILED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxGetCacheConfig](#), [cuCtxSetCacheConfig](#), [cuFuncSetCacheConfig](#), [cuFuncGetAttribute](#),

5.42 Execution Control [DEPRECATED]

Functions

- [CUresult cuFuncSetBlockShape](#) ([CUfunction](#) hfunc, int x, int y, int z)
Sets the block-dimensions for the function.
- [CUresult cuFuncSetSharedSize](#) ([CUfunction](#) hfunc, unsigned int bytes)
Sets the dynamic shared-memory size for the function.
- [CUresult cuLaunch](#) ([CUfunction](#) f)
Launches a CUDA function.
- [CUresult cuLaunchGrid](#) ([CUfunction](#) f, int grid_width, int grid_height)
Launches a CUDA function.
- [CUresult cuLaunchGridAsync](#) ([CUfunction](#) f, int grid_width, int grid_height, [CUstream](#) hStream)
Launches a CUDA function.
- [CUresult cuParamSetf](#) ([CUfunction](#) hfunc, int offset, float value)
Adds a floating-point parameter to the function's argument list.
- [CUresult cuParamSeti](#) ([CUfunction](#) hfunc, int offset, unsigned int value)
Adds an integer parameter to the function's argument list.
- [CUresult cuParamSetSize](#) ([CUfunction](#) hfunc, unsigned int numbytes)
Sets the parameter size for the function.
- [CUresult cuParamSetTexRef](#) ([CUfunction](#) hfunc, int texunit, [CUTexref](#) hTexRef)
Adds a texture-reference to the function's argument list.
- [CUresult cuParamSetv](#) ([CUfunction](#) hfunc, int offset, void *ptr, unsigned int numbytes)
Adds arbitrary data to the function's argument list.

5.42.1 Detailed Description

This section describes the deprecated execution control functions of the low-level CUDA driver application programming interface.

5.42.2 Function Documentation

5.42.2.1 [CUresult cuFuncSetBlockShape](#) ([CUfunction](#) hfunc, int x, int y, int z)

Deprecated

Specifies the x, y, and z dimensions of the thread blocks that are created when the kernel given by hfunc is launched.

Parameters:

hfunc - Kernel to specify dimensions of
x - X dimension
y - Y dimension
z - Z dimension

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetSharedSize](#), [cuFuncSetCacheConfig](#), [cuFuncGetAttribute](#), [cuParamSetSize](#), [cuParamSeti](#), [cuParamSetf](#), [cuParamSetv](#), [cuLaunch](#), [cuLaunchGrid](#), [cuLaunchGridAsync](#), [cuLaunchKernel](#)

5.42.2.2 CUresult cuFuncSetSharedSize (CUfunction *hfunc*, unsigned int *bytes*)**Deprecated**

Sets through *bytes* the amount of dynamic shared memory that will be available to each thread block when the kernel given by *hfunc* is launched.

Parameters:

hfunc - Kernel to specify dynamic shared-memory size for
bytes - Dynamic shared-memory size per thread in bytes

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_INVALID_VALUE

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetBlockShape](#), [cuFuncSetCacheConfig](#), [cuFuncGetAttribute](#), [cuParamSetSize](#), [cuParamSeti](#), [cuParamSetf](#), [cuParamSetv](#), [cuLaunch](#), [cuLaunchGrid](#), [cuLaunchGridAsync](#), [cuLaunchKernel](#)

5.42.2.3 CUresult cuLaunch (CUfunction *f*)**Deprecated**

Invokes the kernel f on a $1 \times 1 \times 1$ grid of blocks. The block contains the number of threads specified by a previous call to [cuFuncSetBlockShape\(\)](#).

Parameters:

f - Kernel to launch

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_LAUNCH_FAILED](#), [CUDA_ERROR_LAUNCH_OUT_OF_RESOURCES](#), [CUDA_ERROR_LAUNCH_TIMEOUT](#), [CUDA_ERROR_LAUNCH_INCOMPATIBLE_TEXTURING](#), [CUDA_ERROR_SHARED_OBJECT_INIT_FAILED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetBlockShape](#), [cuFuncSetSharedSize](#), [cuFuncGetAttribute](#), [cuParamSetSize](#), [cuParamSetf](#), [cuParamSeti](#), [cuParamSetv](#), [cuLaunchGrid](#), [cuLaunchGridAsync](#), [cuLaunchKernel](#)

5.42.2.4 CUresult cuLaunchGrid (CUfunction f , int $grid_width$, int $grid_height$)

Deprecated

Invokes the kernel f on a $grid_width \times grid_height$ grid of blocks. Each block contains the number of threads specified by a previous call to [cuFuncSetBlockShape\(\)](#).

Parameters:

f - Kernel to launch

$grid_width$ - Width of grid in blocks

$grid_height$ - Height of grid in blocks

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_LAUNCH_FAILED](#), [CUDA_ERROR_LAUNCH_OUT_OF_RESOURCES](#), [CUDA_ERROR_LAUNCH_TIMEOUT](#), [CUDA_ERROR_LAUNCH_INCOMPATIBLE_TEXTURING](#), [CUDA_ERROR_SHARED_OBJECT_INIT_FAILED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetBlockShape](#), [cuFuncSetSharedSize](#), [cuFuncGetAttribute](#), [cuParamSetSize](#), [cuParamSetf](#), [cuParamSeti](#), [cuParamSetv](#), [cuLaunch](#), [cuLaunchGridAsync](#), [cuLaunchKernel](#)

5.42.2.5 CUresult cuLaunchGridAsync (CUfunction *f*, int *grid_width*, int *grid_height*, CUstream *hStream*)**Deprecated**

Invokes the kernel *f* on a *grid_width* x *grid_height* grid of blocks. Each block contains the number of threads specified by a previous call to [cuFuncSetBlockShape\(\)](#).

[cuLaunchGridAsync\(\)](#) can optionally be associated to a stream by passing a non-zero *hStream* argument.

Parameters:

f - Kernel to launch

grid_width - Width of grid in blocks

grid_height - Height of grid in blocks

hStream - Stream identifier

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_LAUNCH_FAILED](#), [CUDA_ERROR_LAUNCH_OUT_OF_RESOURCES](#), [CUDA_ERROR_LAUNCH_TIMEOUT](#), [CUDA_ERROR_LAUNCH_INCOMPATIBLE_TEXTURING](#), [CUDA_ERROR_SHARED_OBJECT_INIT_FAILED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetBlockShape](#), [cuFuncSetSharedSize](#), [cuFuncGetAttribute](#), [cuParamSetSize](#), [cuParamSetf](#), [cuParamSeti](#), [cuParamSetv](#), [cuLaunch](#), [cuLaunchGrid](#), [cuLaunchKernel](#)

5.42.2.6 CUresult cuParamSetf (CUfunction *hfunc*, int *offset*, float *value*)**Deprecated**

Sets a floating-point parameter that will be specified the next time the kernel corresponding to *hfunc* will be invoked. *offset* is a byte offset.

Parameters:

hfunc - Kernel to add parameter to

offset - Offset to add parameter to argument list

value - Value of parameter

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetBlockShape](#), [cuFuncSetSharedSize](#), [cuFuncGetAttribute](#), [cuParamSetSize](#), [cuParamSeti](#), [cuParamSetv](#), [cuLaunch](#), [cuLaunchGrid](#), [cuLaunchGridAsync](#), [cuLaunchKernel](#)

5.42.2.7 CUresult cuParamSeti (CUfunction *hfunc*, int *offset*, unsigned int *value*)**Deprecated**

Sets an integer parameter that will be specified the next time the kernel corresponding to `hfunc` will be invoked. `offset` is a byte offset.

Parameters:

hfunc - Kernel to add parameter to
offset - Offset to add parameter to argument list
value - Value of parameter

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetBlockShape](#), [cuFuncSetSharedSize](#), [cuFuncGetAttribute](#), [cuParamSetSize](#), [cuParamSetf](#), [cuParamSetv](#), [cuLaunch](#), [cuLaunchGrid](#), [cuLaunchGridAsync](#), [cuLaunchKernel](#)

5.42.2.8 CUresult cuParamSetSize (CUfunction *hfunc*, unsigned int *numbytes*)**Deprecated**

Sets through `numbytes` the total size in bytes needed by the function parameters of the kernel corresponding to `hfunc`.

Parameters:

hfunc - Kernel to set parameter size for
numbytes - Size of parameter list in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetBlockShape](#), [cuFuncSetSharedSize](#), [cuFuncGetAttribute](#), [cuParamSetf](#), [cuParamSeti](#), [cuParamSetv](#), [cuLaunch](#), [cuLaunchGrid](#), [cuLaunchGridAsync](#), [cuLaunchKernel](#)

5.42.2.9 CUresult cuParamSetTexRef (CUfunction *hfunc*, int *texunit*, CUtexref *hTexRef*)**Deprecated**

Makes the CUDA array or linear memory bound to the texture reference `hTexRef` available to a device program as a texture. In this version of CUDA, the texture-reference must be obtained via [cuModuleGetTexRef\(\)](#) and the `texunit` parameter must be set to [CU_PARAM_TR_DEFAULT](#).

Parameters:

- hfunc* - Kernel to add texture-reference to
- texunit* - Texture unit (must be [CU_PARAM_TR_DEFAULT](#))
- hTexRef* - Texture-reference to add to argument list

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

5.42.2.10 CUresult cuParamSetv (CUfunction *hfunc*, int *offset*, void * *ptr*, unsigned int *numbytes*)**Deprecated**

Copies an arbitrary amount of data (specified in `numbytes`) from `ptr` into the parameter space of the kernel corresponding to `hfunc`. `offset` is a byte offset.

Parameters:

- hfunc* - Kernel to add data to
- offset* - Offset to add data to argument list
- ptr* - Pointer to arbitrary data
- numbytes* - Size of data to copy in bytes

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuFuncSetBlockShape](#), [cuFuncSetSharedSize](#), [cuFuncGetAttribute](#), [cuParamSetSize](#), [cuParamSetf](#), [cuParamSeti](#), [cuLaunch](#), [cuLaunchGrid](#), [cuLaunchGridAsync](#), [cuLaunchKernel](#)

5.43 Texture Reference Management

Functions

- **CUresult cuTexRefGetAddress** (CUdeviceptr *pdptr, CUtexref hTexRef)
Gets the address associated with a texture reference.
- **CUresult cuTexRefGetAddressMode** (CUaddress_mode *pam, CUtexref hTexRef, int dim)
Gets the addressing mode used by a texture reference.
- **CUresult cuTexRefGetArray** (CUarray *phArray, CUtexref hTexRef)
Gets the array bound to a texture reference.
- **CUresult cuTexRefGetFilterMode** (CUfilter_mode *pfm, CUtexref hTexRef)
Gets the filter-mode used by a texture reference.
- **CUresult cuTexRefGetFlags** (unsigned int *pFlags, CUtexref hTexRef)
Gets the flags used by a texture reference.
- **CUresult cuTexRefGetFormat** (CUarray_format *pFormat, int *pNumChannels, CUtexref hTexRef)
Gets the format used by a texture reference.
- **CUresult cuTexRefGetMaxAnisotropy** (int *pmaxAniso, CUtexref hTexRef)
Gets the maximum anisotropy for a texture reference.
- **CUresult cuTexRefGetMipmapFilterMode** (CUfilter_mode *pfm, CUtexref hTexRef)
Gets the mipmap filtering mode for a texture reference.
- **CUresult cuTexRefGetMipmapLevelBias** (float *pbias, CUtexref hTexRef)
Gets the mipmap level bias for a texture reference.
- **CUresult cuTexRefGetMipmapLevelClamp** (float *pminMipmapLevelClamp, float *pmaxMipmapLevelClamp, CUtexref hTexRef)
Gets the min/max mipmap level clamps for a texture reference.
- **CUresult cuTexRefGetMipmappedArray** (CUmipmappedArray *phMipmappedArray, CUtexref hTexRef)
Gets the mipmapped array bound to a texture reference.
- **CUresult cuTexRefSetAddress** (size_t *pByteOffset, CUtexref hTexRef, CUdeviceptr dptr, size_t bytes)
Binds an address as a texture reference.
- **CUresult cuTexRefSetAddress2D** (CUtexref hTexRef, const CUDA_ARRAY_DESCRIPTOR *desc, CUdeviceptr dptr, size_t Pitch)
Binds an address as a 2D texture reference.
- **CUresult cuTexRefSetAddressMode** (CUtexref hTexRef, int dim, CUaddress_mode am)
Sets the addressing mode for a texture reference.
- **CUresult cuTexRefSetArray** (CUtexref hTexRef, CUarray hArray, unsigned int Flags)
Binds an array as a texture reference.

- [CUresult cuTexRefSetFilterMode](#) (CUtexref hTexRef, CUfilter_mode fm)
Sets the filtering mode for a texture reference.
- [CUresult cuTexRefSetFlags](#) (CUtexref hTexRef, unsigned int Flags)
Sets the flags for a texture reference.
- [CUresult cuTexRefSetFormat](#) (CUtexref hTexRef, CUarray_format fmt, int NumPackedComponents)
Sets the format for a texture reference.
- [CUresult cuTexRefSetMaxAnisotropy](#) (CUtexref hTexRef, unsigned int maxAniso)
Sets the maximum anisotropy for a texture reference.
- [CUresult cuTexRefSetMipmapFilterMode](#) (CUtexref hTexRef, CUfilter_mode fm)
Sets the mipmap filtering mode for a texture reference.
- [CUresult cuTexRefSetMipmapLevelBias](#) (CUtexref hTexRef, float bias)
Sets the mipmap level bias for a texture reference.
- [CUresult cuTexRefSetMipmapLevelClamp](#) (CUtexref hTexRef, float minMipmapLevelClamp, float maxMipmapLevelClamp)
Sets the mipmap min/max mipmap level clamps for a texture reference.
- [CUresult cuTexRefSetMipmappedArray](#) (CUtexref hTexRef, CUmipmappedArray hMipmappedArray, unsigned int Flags)
Binds a mipmapped array to a texture reference.

5.43.1 Detailed Description

This section describes the texture reference management functions of the low-level CUDA driver application programming interface.

5.43.2 Function Documentation

5.43.2.1 CUresult cuTexRefGetAddress (CUdeviceptr *pdptr, CUtexref hTexRef)

Returns in *pdptr the base address bound to the texture reference hTexRef, or returns [CUDA_ERROR_INVALID_VALUE](#) if the texture reference is not bound to any device memory range.

Parameters:

- pdptr* - Returned device address
- hTexRef* - Texture reference

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFilterMode](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddressMode](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.2 CUresult cuTexRefGetAddressMode (CUaddress_mode * pam, CUTexref hTexRef, int dim)

Returns in *pam the addressing mode corresponding to the dimension dim of the texture reference hTexRef. Currently, the only valid value for dim are 0 and 1.

Parameters:

pam - Returned addressing mode

hTexRef - Texture reference

dim - Dimension

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFilterMode](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.3 CUresult cuTexRefGetArray (CUarray * phArray, CUTexref hTexRef)

Returns in *phArray the CUDA array bound to the texture reference hTexRef, or returns [CUDA_ERROR_INVALID_VALUE](#) if the texture reference is not bound to any CUDA array.

Parameters:

phArray - Returned array

hTexRef - Texture reference

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFilterMode](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetAddressMode](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.4 CUresult cuTexRefGetFilterMode (CUfilter_mode * *pfm*, CUtexref *hTexRef*)

Returns in **pfm* the filtering mode of the texture reference *hTexRef*.

Parameters:

pfm - Returned filtering mode

hTexRef - Texture reference

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFilterMode, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.5 CUresult cuTexRefGetFlags (unsigned int * *pFlags*, CUtexref *hTexRef*)

Returns in **pFlags* the flags of the texture reference *hTexRef*.

Parameters:

pFlags - Returned flags

hTexRef - Texture reference

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFilterMode, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFormat

5.43.2.6 CUresult cuTexRefGetFormat (CUarray_format * *pFormat*, int * *pNumChannels*, CUtexref *hTexRef*)

Returns in **pFormat* and **pNumChannels* the format and number of components of the CUDA array bound to the texture reference *hTexRef*. If *pFormat* or *pNumChannels* is NULL, it will be ignored.

Parameters:

pFormat - Returned format

pNumChannels - Returned number of components

hTexRef - Texture reference

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFilterMode, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFlags

5.43.2.7 CUresult cuTexRefGetMaxAnisotropy (int * pmaxAniso, CUtexref hTexRef)

Returns the maximum anisotropy in `pmaxAniso` that's used when reading memory through the texture reference `hTexRef`.

Parameters:

pmaxAniso - Returned maximum anisotropy

hTexRef - Texture reference

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.8 CUresult cuTexRefGetMipmapFilterMode (CUfilter_mode * pfm, CUtexref hTexRef)

Returns the mipmap filtering mode in `pfm` that's used when reading memory through the texture reference `hTexRef`.

Parameters:

pfm - Returned mipmap filtering mode

hTexRef - Texture reference

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.9 CUresult cuTexRefGetMipmapLevelBias (float * *pbias*, CUtexref *hTexRef*)

Returns the mipmap level bias in *pBias* that's added to the specified mipmap level when reading memory through the texture reference *hTexRef*.

Parameters:

pbias - Returned mipmap level bias

hTexRef - Texture reference

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetAddressMode](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.10 CUresult cuTexRefGetMipmapLevelClamp (float * *pminMipmapLevelClamp*, float * *pmaxMipmapLevelClamp*, CUtexref *hTexRef*)

Returns the min/max mipmap level clamps in *pminMipmapLevelClamp* and *pmaxMipmapLevelClamp* that's used when reading memory through the texture reference *hTexRef*.

Parameters:

pminMipmapLevelClamp - Returned mipmap min level clamp

pmaxMipmapLevelClamp - Returned mipmap max level clamp

hTexRef - Texture reference

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetAddressMode](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.11 CUresult cuTexRefGetMipmappedArray (CUmipmappedArray * *phMipmappedArray*, CUtexref *hTexRef*)

Returns in **phMipmappedArray* the CUDA mipmapped array bound to the texture reference *hTexRef*, or returns [CUDA_ERROR_INVALID_VALUE](#) if the texture reference is not bound to any CUDA mipmapped array.

Parameters:

phMipmappedArray - Returned mipmapped array

hTexRef - Texture reference

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFilterMode, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetFilterMode, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.12 CUresult cuTexRefSetAddress (size_t * ByteOffset, CUtexref hTexRef, CUdeviceptr dptr, size_t bytes)

Binds a linear address range to the texture reference *hTexRef*. Any previous address or CUDA array state associated with the texture reference is superseded by this function. Any memory previously bound to *hTexRef* is unbound.

Since the hardware enforces an alignment requirement on texture base addresses, [cuTexRefSetAddress\(\)](#) passes back a byte offset in **ByteOffset* that must be applied to texture fetches in order to read from the desired memory. This offset must be divided by the texel size and passed to kernels that read from the texture so they can be applied to the [tex1Dfetch\(\)](#) function.

If the device memory pointer was returned from [cuMemAlloc\(\)](#), the offset is guaranteed to be 0 and NULL may be passed as the *ByteOffset* parameter.

The total number of elements (or texels) in the linear address range cannot exceed [CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTUREID_LINEAR_WIDTH](#). The number of elements is computed as (*bytes* / *bytesPerElement*), where *bytesPerElement* is determined from the data format and number of components set using [cuTexRefSetFormat\(\)](#).

Parameters:

ByteOffset - Returned byte offset

hTexRef - Texture reference to bind

dptr - Device pointer to bind

bytes - Size of memory to bind in bytes

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFilterMode, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.13 CUresult cuTexRefSetAddress2D (CUtexref hTexRef, const CUDA_ARRAY_DESCRIPTOR * desc, CUdeviceptr dptr, size_t Pitch)

Binds a linear address range to the texture reference *hTexRef*. Any previous address or CUDA array state associated with the texture reference is superseded by this function. Any memory previously bound to *hTexRef* is unbound.

Using a `tex2D()` function inside a kernel requires a call to either `cuTexRefSetArray()` to bind the corresponding texture reference to an array, or `cuTexRefSetAddress2D()` to bind the texture reference to linear memory.

Function calls to `cuTexRefSetFormat()` cannot follow calls to `cuTexRefSetAddress2D()` for the same texture reference.

It is required that `dptr` be aligned to the appropriate hardware-specific texture alignment. You can query this value using the device attribute `CU_DEVICE_ATTRIBUTE_TEXTURE_ALIGNMENT`. If an unaligned `dptr` is supplied, `CUDA_ERROR_INVALID_VALUE` is returned.

`Pitch` has to be aligned to the hardware-specific texture pitch alignment. This value can be queried using the device attribute `CU_DEVICE_ATTRIBUTE_TEXTURE_PITCH_ALIGNMENT`. If an unaligned `Pitch` is supplied, `CUDA_ERROR_INVALID_VALUE` is returned.

Width and Height, which are specified in elements (or texels), cannot exceed `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_WIDTH` and `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_HEIGHT` respectively. `Pitch`, which is specified in bytes, cannot exceed `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_PITCH`.

Parameters:

- hTexRef* - Texture reference to bind
- desc* - Descriptor of CUDA array
- dptr* - Device pointer to bind
- Pitch* - Line pitch in bytes

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

See also:

`cuTexRefSetAddress`, `cuTexRefSetAddressMode`, `cuTexRefSetArray`, `cuTexRefSetFilterMode`, `cuTexRefSetFlags`, `cuTexRefSetFormat`, `cuTexRefGetAddress`, `cuTexRefGetAddressMode`, `cuTexRefGetArray`, `cuTexRefGetFilterMode`, `cuTexRefGetFlags`, `cuTexRefGetFormat`

5.43.2.14 CUresult cuTexRefSetAddressMode (CUtexref hTexRef, int dim, CUaddress_mode am)

Specifies the addressing mode `am` for the given dimension `dim` of the texture reference `hTexRef`. If `dim` is zero, the addressing mode is applied to the first parameter of the functions used to fetch from the texture; if `dim` is 1, the second, and so on. `CUaddress_mode` is defined as:

```
typedef enum CUaddress_mode_enum {
    CU_TR_ADDRESS_MODE_WRAP = 0,
    CU_TR_ADDRESS_MODE_CLAMP = 1,
    CU_TR_ADDRESS_MODE_MIRROR = 2,
    CU_TR_ADDRESS_MODE_BORDER = 3
} CUaddress_mode;
```

Note that this call has no effect if `hTexRef` is bound to linear memory. Also, if the flag, `CU_TRSF_NORMALIZED_COORDINATES`, is not set, the only supported address mode is `CU_TR_ADDRESS_MODE_CLAMP`.

Parameters:

- hTexRef* - Texture reference
- dim* - Dimension

am - Addressing mode to set

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetArray, cuTexRefSetFilterMode, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.15 CUresult cuTexRefSetArray (CUtexref *hTexRef*, CUarray *hArray*, unsigned int *Flags*)

Binds the CUDA array *hArray* to the texture reference *hTexRef*. Any previous address or CUDA array state associated with the texture reference is superseded by this function. *Flags* must be set to [CU_TRSA_OVERRIDE_FORMAT](#). Any CUDA array previously bound to *hTexRef* is unbound.

Parameters:

hTexRef - Texture reference to bind
hArray - Array to bind
Flags - Options (must be [CU_TRSA_OVERRIDE_FORMAT](#))

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetFilterMode, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.16 CUresult cuTexRefSetFilterMode (CUtexref *hTexRef*, CUfilter_mode *fm*)

Specifies the filtering mode *fm* to be used when reading memory through the texture reference *hTexRef*. [CUfilter_mode_enum](#) is defined as:

```
typedef enum CUfilter_mode_enum {
    CU_TR_FILTER_MODE_POINT = 0,
    CU_TR_FILTER_MODE_LINEAR = 1
} CUfilter_mode;
```

Note that this call has no effect if *hTexRef* is bound to linear memory.

Parameters:

hTexRef - Texture reference
fm - Filtering mode to set

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetAddressMode](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.17 CUresult cuTexRefSetFlags (CUtexref *hTexRef*, unsigned int *Flags*)

Specifies optional flags via `Flags` to specify the behavior of data returned through the texture reference `hTexRef`. The valid flags are:

- [CU_TRSF_READ_AS_INTEGER](#), which suppresses the default behavior of having the texture promote integer data to floating point data in the range [0, 1]. Note that texture with 32-bit integer format would not be promoted, regardless of whether or not this flag is specified;
- [CU_TRSF_NORMALIZED_COORDINATES](#), which suppresses the default behavior of having the texture coordinates range from [0, Dim) where Dim is the width or height of the CUDA array. Instead, the texture coordinates [0, 1.0) reference the entire breadth of the array dimension;

Parameters:

hTexRef - Texture reference

Flags - Optional flags to set

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFilterMode](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetAddressMode](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.18 CUresult cuTexRefSetFormat (CUtexref *hTexRef*, CUarray_format *fmt*, int *NumPackedComponents*)

Specifies the format of the data to be read by the texture reference `hTexRef`. `fmt` and `NumPackedComponents` are exactly analogous to the `Format` and `NumChannels` members of the [CUDA_ARRAY_DESCRIPTOR](#) structure: They specify the format of each component and the number of components per array element.

Parameters:

hTexRef - Texture reference

fmt - Format to set

NumPackedComponents - Number of components per array element

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFilterMode, cuTexRefSetFlags, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.19 CUresult cuTexRefSetMaxAnisotropy (CUtexref *hTexRef*, unsigned int *maxAniso*)

Specifies the maximum anisotropy *maxAniso* to be used when reading memory through the texture reference *hTexRef*.

Note that this call has no effect if *hTexRef* is bound to linear memory.

Parameters:

hTexRef - Texture reference
maxAniso - Maximum anisotropy

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

cuTexRefSetAddress, cuTexRefSetAddress2D, cuTexRefSetAddressMode, cuTexRefSetArray, cuTexRefSetFlags, cuTexRefSetFormat, cuTexRefGetAddress, cuTexRefGetAddressMode, cuTexRefGetArray, cuTexRefGetFilterMode, cuTexRefGetFlags, cuTexRefGetFormat

5.43.2.20 CUresult cuTexRefSetMipmapFilterMode (CUtexref *hTexRef*, CUfilter_mode *fm*)

Specifies the mipmap filtering mode *fm* to be used when reading memory through the texture reference *hTexRef*. [CUfilter_mode_enum](#) is defined as:

```
typedef enum CUfilter_mode_enum {
    CU_TR_FILTER_MODE_POINT = 0,
    CU_TR_FILTER_MODE_LINEAR = 1
} CUfilter_mode;
```

Note that this call has no effect if *hTexRef* is not bound to a mipmapped array.

Parameters:

hTexRef - Texture reference
fm - Filtering mode to set

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetAddressMode](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.21 CUresult cuTexRefSetMipmapLevelBias (CUtexref *hTexRef*, float *bias*)

Specifies the mipmap level bias *bias* to be added to the specified mipmap level when reading memory through the texture reference *hTexRef*.

Note that this call has no effect if *hTexRef* is not bound to a mipmapped array.

Parameters:

hTexRef - Texture reference

bias - Mipmap level bias

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetAddressMode](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.22 CUresult cuTexRefSetMipmapLevelClamp (CUtexref *hTexRef*, float *minMipmapLevelClamp*, float *maxMipmapLevelClamp*)

Specifies the min/max mipmap level clamps, *minMipmapLevelClamp* and *maxMipmapLevelClamp* respectively, to be used when reading memory through the texture reference *hTexRef*.

Note that this call has no effect if *hTexRef* is not bound to a mipmapped array.

Parameters:

hTexRef - Texture reference

minMipmapLevelClamp - Mipmap min level clamp

maxMipmapLevelClamp - Mipmap max level clamp

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefSetAddress](#), [cuTexRefSetAddress2D](#), [cuTexRefSetAddressMode](#), [cuTexRefSetArray](#), [cuTexRefSetFlags](#), [cuTexRefSetFormat](#), [cuTexRefGetAddress](#), [cuTexRefGetAddressMode](#), [cuTexRefGetArray](#), [cuTexRefGetFilterMode](#), [cuTexRefGetFlags](#), [cuTexRefGetFormat](#)

5.43.2.23 CUresult cuTexRefSetMipmappedArray (CUtexref *hTexRef*, CUmipmappedArray *hMipmappedArray*, unsigned int *Flags*)

Binds the CUDA mipmapped array `hMipmappedArray` to the texture reference `hTexRef`. Any previous address or CUDA array state associated with the texture reference is superseded by this function. `Flags` must be set to `CU_TRSA_OVERRIDE_FORMAT`. Any CUDA array previously bound to `hTexRef` is unbound.

Parameters:

hTexRef - Texture reference to bind

hMipmappedArray - Mipmapped array to bind

Flags - Options (must be `CU_TRSA_OVERRIDE_FORMAT`)

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

See also:

`cuTexRefSetAddress`, `cuTexRefSetAddress2D`, `cuTexRefSetAddressMode`, `cuTexRefSetFilterMode`, `cuTexRefSetFlags`, `cuTexRefSetFormat`, `cuTexRefGetAddress`, `cuTexRefGetAddressMode`, `cuTexRefGetArray`, `cuTexRefGetFilterMode`, `cuTexRefGetFlags`, `cuTexRefGetFormat`

5.44 Texture Reference Management [DEPRECATED]

Functions

- [CUresult cuTexRefCreate](#) (CUtexref *pTexRef)
Creates a texture reference.
- [CUresult cuTexRefDestroy](#) (CUtexref hTexRef)
Destroys a texture reference.

5.44.1 Detailed Description

This section describes the deprecated texture reference management functions of the low-level CUDA driver application programming interface.

5.44.2 Function Documentation

5.44.2.1 CUresult cuTexRefCreate (CUtexref *pTexRef)

Deprecated

Creates a texture reference and returns its handle in *pTexRef. Once created, the application must call [cuTexRefSetArray\(\)](#) or [cuTexRefSetAddress\(\)](#) to associate the reference with allocated memory. Other texture reference functions are used to specify the format and interpretation (addressing, filtering, etc.) to be used when the memory is read through this texture reference.

Parameters:

pTexRef - Returned texture reference

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefDestroy](#)

5.44.2.2 CUresult cuTexRefDestroy (CUtexref hTexRef)

Deprecated

Destroys the texture reference specified by hTexRef.

Parameters:

hTexRef - Texture reference to destroy

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexRefCreate](#)

5.45 Surface Reference Management

Functions

- [CUresult cuSurfRefGetArray](#) ([CUarray *phArray](#), [CUSurfref hSurfRef](#))
Passes back the CUDA array bound to a surface reference.
- [CUresult cuSurfRefSetArray](#) ([CUSurfref hSurfRef](#), [CUarray hArray](#), [unsigned int Flags](#))
Sets the CUDA array for a surface reference.

5.45.1 Detailed Description

This section describes the surface reference management functions of the low-level CUDA driver application programming interface.

5.45.2 Function Documentation

5.45.2.1 [CUresult cuSurfRefGetArray](#) ([CUarray *phArray](#), [CUSurfref hSurfRef](#))

Returns in [*phArray](#) the CUDA array bound to the surface reference [hSurfRef](#), or returns [CUDA_ERROR_INVALID_VALUE](#) if the surface reference is not bound to any CUDA array.

Parameters:

- [phArray](#) - Surface reference handle
- [hSurfRef](#) - Surface reference handle

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuModuleGetSurfRef](#), [cuSurfRefSetArray](#)

5.45.2.2 [CUresult cuSurfRefSetArray](#) ([CUSurfref hSurfRef](#), [CUarray hArray](#), [unsigned int Flags](#))

Sets the CUDA array [hArray](#) to be read and written by the surface reference [hSurfRef](#). Any previous CUDA array state associated with the surface reference is superseded by this function. [Flags](#) must be set to 0. The [CUDA_ARRAY3D_SURFACE_LDST](#) flag must have been set for the CUDA array. Any CUDA array previously bound to [hSurfRef](#) is unbound.

Parameters:

- [hSurfRef](#) - Surface reference handle
- [hArray](#) - CUDA array handle
- [Flags](#) - set to 0

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuModuleGetSurfRef](#), [cuSurfRefGetArray](#)

5.46 Texture Object Management

Functions

- `CUresult cuTexObjectCreate` (`CUtexObject *pTexObject`, `const CUDA_RESOURCE_DESC *pResDesc`, `const CUDA_TEXTURE_DESC *pTexDesc`, `const CUDA_RESOURCE_VIEW_DESC *pResViewDesc`)
Creates a texture object.
- `CUresult cuTexObjectDestroy` (`CUtexObject texObject`)
Destroys a texture object.
- `CUresult cuTexObjectGetResourceDesc` (`CUDA_RESOURCE_DESC *pResDesc`, `CUtexObject texObject`)
Returns a texture object's resource descriptor.
- `CUresult cuTexObjectGetResourceViewDesc` (`CUDA_RESOURCE_VIEW_DESC *pResViewDesc`, `CUtexObject texObject`)
Returns a texture object's resource view descriptor.
- `CUresult cuTexObjectGetTextureDesc` (`CUDA_TEXTURE_DESC *pTexDesc`, `CUtexObject texObject`)
Returns a texture object's texture descriptor.

5.46.1 Detailed Description

This section describes the texture object management functions of the low-level CUDA driver application programming interface. The texture object API is only supported on devices of compute capability 3.0 or higher.

5.46.2 Function Documentation

5.46.2.1 `CUresult cuTexObjectCreate` (`CUtexObject *pTexObject`, `const CUDA_RESOURCE_DESC *pResDesc`, `const CUDA_TEXTURE_DESC *pTexDesc`, `const CUDA_RESOURCE_VIEW_DESC *pResViewDesc`)

Creates a texture object and returns it in `pTexObject`. `pResDesc` describes the data to texture from. `pTexDesc` describes how the data should be sampled. `pResViewDesc` is an optional argument that specifies an alternate format for the data described by `pResDesc`, and also describes the subresource region to restrict access to when texturing. `pResViewDesc` can only be specified if the type of resource is a CUDA array or a CUDA mipmapped array.

Texture objects are only supported on devices of compute capability 3.0 or higher.

The `CUDA_RESOURCE_DESC` structure is defined as:

```
typedef struct CUDA_RESOURCE_DESC_st
{
    CUresourcetype resType;

    union {
        struct {
            CUarray hArray;
        } array;
        struct {
            CUmipmappedArray hMipmappedArray;
        } mipmap;
        struct {
```

```

        CUdeviceptr devPtr;
        CUarray_format format;
        unsigned int numChannels;
        size_t sizeInBytes;
    } linear;
    struct {
        CUdeviceptr devPtr;
        CUarray_format format;
        unsigned int numChannels;
        size_t width;
        size_t height;
        size_t pitchInBytes;
    } pitch2D;
} res;

    unsigned int flags;
} CUDA_RESOURCE_DESC;

```

where:

- `CUDA_RESOURCE_DESC::resType` specifies the type of resource to texture from. `CUresourceType` is defined as:

```

typedef enum CUresourcetype_enum {
    CU_RESOURCE_TYPE_ARRAY          = 0x00,
    CU_RESOURCE_TYPE_MIPMAPPED_ARRAY = 0x01,
    CU_RESOURCE_TYPE_LINEAR         = 0x02,
    CU_RESOURCE_TYPE_PITCH2D       = 0x03
} CUresourcetype;

```

If `CUDA_RESOURCE_DESC::resType` is set to `CU_RESOURCE_TYPE_ARRAY`, `CUDA_RESOURCE_DESC::res::array::hArray` must be set to a valid CUDA array handle.

If `CUDA_RESOURCE_DESC::resType` is set to `CU_RESOURCE_TYPE_MIPMAPPED_ARRAY`, `CUDA_RESOURCE_DESC::res::mipmap::hMipmappedArray` must be set to a valid CUDA mipmapped array handle.

If `CUDA_RESOURCE_DESC::resType` is set to `CU_RESOURCE_TYPE_LINEAR`, `CUDA_RESOURCE_DESC::res::linear::devPtr` must be set to a valid device pointer, that is aligned to `CU_DEVICE_ATTRIBUTE_TEXTURE_ALIGNMENT`. `CUDA_RESOURCE_DESC::res::linear::format` and `CUDA_RESOURCE_DESC::res::linear::numChannels` describe the format of each component and the number of components per array element. `CUDA_RESOURCE_DESC::res::linear::sizeInBytes` specifies the size of the array in bytes. The total number of elements in the linear address range cannot exceed `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTUREID_LINEAR_WIDTH`. The number of elements is computed as $(\text{sizeInBytes} / (\text{sizeof}(\text{format}) * \text{numChannels}))$.

If `CUDA_RESOURCE_DESC::resType` is set to `CU_RESOURCE_TYPE_PITCH2D`, `CUDA_RESOURCE_DESC::res::pitch2D::devPtr` must be set to a valid device pointer, that is aligned to `CU_DEVICE_ATTRIBUTE_TEXTURE_ALIGNMENT`. `CUDA_RESOURCE_DESC::res::pitch2D::format` and `CUDA_RESOURCE_DESC::res::pitch2D::numChannels` describe the format of each component and the number of components per array element. `CUDA_RESOURCE_DESC::res::pitch2D::width` and `CUDA_RESOURCE_DESC::res::pitch2D::height` specify the width and height of the array in elements, and cannot exceed `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_WIDTH` and `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_HEIGHT` respectively. `CUDA_RESOURCE_DESC::res::pitch2D::pitchInBytes` specifies the pitch between two rows in bytes and has to be aligned to `CU_DEVICE_ATTRIBUTE_TEXTURE_PITCH_ALIGNMENT`. Pitch cannot exceed `CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURE2D_LINEAR_PITCH`.

- flags must be set to zero.

The `CUDA_TEXTURE_DESC` struct is defined as

```
typedef struct CUDA_TEXTURE_DESC_st {
    CUaddress_mode addressMode[3];
    CUfilter_mode filterMode;
    unsigned int flags;
    unsigned int maxAnisotropy;
    CUfilter_mode mipmapFilterMode;
    float mipmapLevelBias;
    float minMipmapLevelClamp;
    float maxMipmapLevelClamp;
} CUDA_TEXTURE_DESC;
```

where

- `CUDA_TEXTURE_DESC::addressMode` specifies the addressing mode for each dimension of the texture data. `CUaddress_mode` is defined as:

```
typedef enum CUaddress_mode_enum {
    CU_TR_ADDRESS_MODE_WRAP = 0,
    CU_TR_ADDRESS_MODE_CLAMP = 1,
    CU_TR_ADDRESS_MODE_MIRROR = 2,
    CU_TR_ADDRESS_MODE_BORDER = 3
} CUaddress_mode;
```

This is ignored if `CUDA_RESOURCE_DESC::resType` is `CU_RESOURCE_TYPE_LINEAR`. Also, if the flag, `CU_TRSF_NORMALIZED_COORDINATES` is not set, the only supported address mode is `CU_TR_ADDRESS_MODE_CLAMP`.

- `CUDA_TEXTURE_DESC::filterMode` specifies the filtering mode to be used when fetching from the texture. `CUfilter_mode` is defined as:

```
typedef enum CUfilter_mode_enum {
    CU_TR_FILTER_MODE_POINT = 0,
    CU_TR_FILTER_MODE_LINEAR = 1
} CUfilter_mode;
```

This is ignored if `CUDA_RESOURCE_DESC::resType` is `CU_RESOURCE_TYPE_LINEAR`.

- `CUDA_TEXTURE_DESC::flags` can be any combination of the following:
 - `CU_TRSF_READ_AS_INTEGER`, which suppresses the default behavior of having the texture promote integer data to floating point data in the range [0, 1]. Note that texture with 32-bit integer format would not be promoted, regardless of whether or not this flag is specified.
 - `CU_TRSF_NORMALIZED_COORDINATES`, which suppresses the default behavior of having the texture coordinates range from [0, Dim) where Dim is the width or height of the CUDA array. Instead, the texture coordinates [0, 1.0) reference the entire breadth of the array dimension; Note that for CUDA mipmapped arrays, this flag has to be set.
- `CUDA_TEXTURE_DESC::maxAnisotropy` specifies the maximum anisotropy ratio to be used when doing anisotropic filtering. This value will be clamped to the range [1,16].
- `CUDA_TEXTURE_DESC::mipmapFilterMode` specifies the filter mode when the calculated mipmap level lies between two defined mipmap levels.

- [CUDA_TEXTURE_DESC::mipmapLevelBias](#) specifies the offset to be applied to the calculated mipmap level.
- [CUDA_TEXTURE_DESC::minMipmapLevelClamp](#) specifies the lower end of the mipmap level range to clamp access to.
- [CUDA_TEXTURE_DESC::maxMipmapLevelClamp](#) specifies the upper end of the mipmap level range to clamp access to.

The [CUDA_RESOURCE_VIEW_DESC](#) struct is defined as

```
typedef struct CUDA_RESOURCE_VIEW_DESC_st
{
    CUresourceViewFormat format;
    size_t width;
    size_t height;
    size_t depth;
    unsigned int firstMipmapLevel;
    unsigned int lastMipmapLevel;
    unsigned int firstLayer;
    unsigned int lastLayer;
} CUDA_RESOURCE_VIEW_DESC;
```

where:

- [CUDA_RESOURCE_VIEW_DESC::format](#) specifies how the data contained in the CUDA array or CUDA mipmapped array should be interpreted. Note that this can incur a change in size of the texture data. If the resource view format is a block compressed format, then the underlying CUDA array or CUDA mipmapped array has to have a base of format [CU_AD_FORMAT_UNSIGNED_INT32](#) with 2 or 4 channels, depending on the block compressed format. For ex., BC1 and BC4 require the underlying CUDA array to have a format of [CU_AD_FORMAT_UNSIGNED_INT32](#) with 2 channels. The other BC formats require the underlying resource to have the same base format but with 4 channels.
- [CUDA_RESOURCE_VIEW_DESC::width](#) specifies the new width of the texture data. If the resource view format is a block compressed format, this value has to be 4 times the original width of the resource. For non block compressed formats, this value has to be equal to that of the original resource.
- [CUDA_RESOURCE_VIEW_DESC::height](#) specifies the new height of the texture data. If the resource view format is a block compressed format, this value has to be 4 times the original height of the resource. For non block compressed formats, this value has to be equal to that of the original resource.
- [CUDA_RESOURCE_VIEW_DESC::depth](#) specifies the new depth of the texture data. This value has to be equal to that of the original resource.
- [CUDA_RESOURCE_VIEW_DESC::firstMipmapLevel](#) specifies the most detailed mipmap level. This will be the new mipmap level zero. For non-mipmapped resources, this value has to be zero. [CUDA_TEXTURE_DESC::minMipmapLevelClamp](#) and [CUDA_TEXTURE_DESC::maxMipmapLevelClamp](#) will be relative to this value. For ex., if the firstMipmapLevel is set to 2, and a minMipmapLevelClamp of 1.2 is specified, then the actual minimum mipmap level clamp will be 3.2.
- [CUDA_RESOURCE_VIEW_DESC::lastMipmapLevel](#) specifies the least detailed mipmap level. For non-mipmapped resources, this value has to be zero.
- [CUDA_RESOURCE_VIEW_DESC::firstLayer](#) specifies the first layer index for layered textures. This will be the new layer zero. For non-layered resources, this value has to be zero.

- [CUDA_RESOURCE_VIEW_DESC::lastLayer](#) specifies the last layer index for layered textures. For non-layered resources, this value has to be zero.

Parameters:

- pTexObject* - Texture object to create
- pResDesc* - Resource descriptor
- pTexDesc* - Texture descriptor
- pResViewDesc* - Resource view descriptor

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexObjectDestroy](#)

5.46.2.2 CUresult cuTexObjectDestroy (CUtexObject *texObject*)

Destroys the texture object specified by `texObject`.

Parameters:

- texObject* - Texture object to destroy

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexObjectCreate](#)

5.46.2.3 CUresult cuTexObjectGetResourceDesc (CUDA_RESOURCE_DESC * *pResDesc*, CUtexObject *texObject*)

Returns the resource descriptor for the texture object specified by `texObject`.

Parameters:

- pResDesc* - Resource descriptor
- texObject* - Texture object

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuTexObjectCreate](#)

5.46.2.4 CUresult cuTexObjectGetResourceViewDesc (CUDA_RESOURCE_VIEW_DESC * *pResViewDesc*, CUtexObject *texObject*)

Returns the resource view descriptor for the texture object specified by `texObject`. If no resource view was set for `texObject`, the `CUDA_ERROR_INVALID_VALUE` is returned.

Parameters:

pResViewDesc - Resource view descriptor

texObject - Texture object

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

See also:

[cuTexObjectCreate](#)

5.46.2.5 CUresult cuTexObjectGetTextureDesc (CUDA_TEXTURE_DESC * *pTexDesc*, CUtexObject *texObject*)

Returns the texture descriptor for the texture object specified by `texObject`.

Parameters:

pTexDesc - Texture descriptor

texObject - Texture object

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`

See also:

[cuTexObjectCreate](#)

5.47 Surface Object Management

Functions

- [CUresult cuSurfObjectCreate](#) ([CUsurfObject](#) *pSurfObject, const [CUDA_RESOURCE_DESC](#) *pResDesc)
Creates a surface object.
- [CUresult cuSurfObjectDestroy](#) ([CUsurfObject](#) surfObject)
Destroys a surface object.
- [CUresult cuSurfObjectGetResourceDesc](#) ([CUDA_RESOURCE_DESC](#) *pResDesc, [CUsurfObject](#) surfObject)
Returns a surface object's resource descriptor.

5.47.1 Detailed Description

This section describes the surface object management functions of the low-level CUDA driver application programming interface. The surface object API is only supported on devices of compute capability 3.0 or higher.

5.47.2 Function Documentation

5.47.2.1 [CUresult cuSurfObjectCreate](#) ([CUsurfObject](#) *pSurfObject, const [CUDA_RESOURCE_DESC](#) *pResDesc)

Creates a surface object and returns it in `pSurfObject`. `pResDesc` describes the data to perform surface load/stores on. [CUDA_RESOURCE_DESC::resType](#) must be [CU_RESOURCE_TYPE_ARRAY](#) and [CUDA_RESOURCE_DESC::res::array::hArray](#) must be set to a valid CUDA array handle. [CUDA_RESOURCE_DESC::flags](#) must be set to zero.

Surface objects are only supported on devices of compute capability 3.0 or higher.

Parameters:

- pSurfObject* - Surface object to create
- pResDesc* - Resource descriptor

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

See also:

[cuSurfObjectDestroy](#)

5.47.2.2 [CUresult cuSurfObjectDestroy](#) ([CUsurfObject](#) surfObject)

Destroys the surface object specified by `surfObject`.

Parameters:

- surfObject* - Surface object to destroy

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

[cuSurfObjectCreate](#)

5.47.2.3 CUresult cuSurfObjectGetResourceDesc (CUDA_RESOURCE_DESC * *pResDesc*, CUsurfObject *surfObject*)

Returns the resource descriptor for the surface object specified by `surfObject`.

Parameters:

pResDesc - Resource descriptor

surfObject - Surface object

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE

See also:

[cuSurfObjectCreate](#)

5.48 Peer Context Memory Access

Functions

- [CUresult cuCtxDisablePeerAccess](#) ([CUcontext](#) peerContext)
Disables direct access to memory allocations in a peer context and unregisters any registered allocations.
- [CUresult cuCtxEnablePeerAccess](#) ([CUcontext](#) peerContext, unsigned int Flags)
Enables direct access to memory allocations in a peer context.
- [CUresult cuDeviceCanAccessPeer](#) (int *canAccessPeer, [CUdevice](#) dev, [CUdevice](#) peerDev)
Queries if a device may directly access a peer device's memory.

5.48.1 Detailed Description

This section describes the direct peer context memory access functions of the low-level CUDA driver application programming interface.

5.48.2 Function Documentation

5.48.2.1 CUresult cuCtxDisablePeerAccess (CUcontext peerContext)

Returns [CUDA_ERROR_PEER_ACCESS_NOT_ENABLED](#) if direct peer access has not yet been enabled from `peerContext` to the current context.

Returns [CUDA_ERROR_INVALID_CONTEXT](#) if there is no current context, or if `peerContext` is not a valid context.

Parameters:

peerContext - Peer context to disable direct access to

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_PEER_ACCESS_NOT_ENABLED](#), [CUDA_ERROR_INVALID_CONTEXT](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuDeviceCanAccessPeer](#), [cuCtxEnablePeerAccess](#)

5.48.2.2 CUresult cuCtxEnablePeerAccess (CUcontext peerContext, unsigned int Flags)

If both the current context and `peerContext` are on devices which support unified addressing (as may be queried using [CU_DEVICE_ATTRIBUTE_UNIFIED_ADDRESSING](#)) and same major compute capability, then on success all allocations from `peerContext` will immediately be accessible by the current context. See [Unified Addressing](#) for additional details.

Note that access granted by this call is unidirectional and that in order to access memory from the current context in `peerContext`, a separate symmetric call to `cuCtxEnablePeerAccess()` is required.

Returns `CUDA_ERROR_PEER_ACCESS_UNSUPPORTED` if `cuDeviceCanAccessPeer()` indicates that the `CUdevice` of the current context cannot directly access memory from the `CUdevice` of `peerContext`.

Returns `CUDA_ERROR_PEER_ACCESS_ALREADY_ENABLED` if direct access of `peerContext` from the current context has already been enabled.

Returns `CUDA_ERROR_TOO_MANY_PEERS` if direct peer access is not possible because hardware resources required for peer access have been exhausted.

Returns `CUDA_ERROR_INVALID_CONTEXT` if there is no current context, `peerContext` is not a valid context, or if the current context is `peerContext`.

Returns `CUDA_ERROR_INVALID_VALUE` if `Flags` is not 0.

Parameters:

peerContext - Peer context to enable direct access to from the current context

Flags - Reserved for future use and must be set to 0

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_PEER_ACCESS_ALREADY_ENABLED`, `CUDA_ERROR_TOO_MANY_PEERS`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_PEER_ACCESS_UNSUPPORTED`, `CUDA_ERROR_INVALID_VALUE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuDeviceCanAccessPeer`, `cuCtxDisablePeerAccess`

5.48.2.3 `CUresult cuDeviceCanAccessPeer (int * canAccessPeer, CUdevice dev, CUdevice peerDev)`

Returns in `*canAccessPeer` a value of 1 if contexts on `dev` are capable of directly accessing memory from contexts on `peerDev` and 0 otherwise. If direct access of `peerDev` from `dev` is possible, then access may be enabled on two specific contexts by calling `cuCtxEnablePeerAccess()`.

Parameters:

canAccessPeer - Returned access capability

dev - Device from which allocations on `peerDev` are to be directly accessed.

peerDev - Device on which the allocations to be directly accessed by `dev` reside.

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_DEVICE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuCtxEnablePeerAccess`, `cuCtxDisablePeerAccess`

5.49 Graphics Interoperability

Functions

- **CUresult cuGraphicsMapResources** (unsigned int count, CUgraphicsResource *resources, CUstream hStream)
Map graphics resources for access by CUDA.
- **CUresult cuGraphicsResourceGetMappedMipmappedArray** (CUmipmappedArray *pMipmappedArray, CUgraphicsResource resource)
Get a mipmapped array through which to access a mapped graphics resource.
- **CUresult cuGraphicsResourceGetMappedPointer** (CUdeviceptr *pDevPtr, size_t *pSize, CUgraphicsResource resource)
Get a device pointer through which to access a mapped graphics resource.
- **CUresult cuGraphicsResourceSetMapFlags** (CUgraphicsResource resource, unsigned int flags)
Set usage flags for mapping a graphics resource.
- **CUresult cuGraphicsSubResourceGetMappedArray** (CUarray *pArray, CUgraphicsResource resource, unsigned int arrayIndex, unsigned int mipLevel)
Get an array through which to access a subresource of a mapped graphics resource.
- **CUresult cuGraphicsUnmapResources** (unsigned int count, CUgraphicsResource *resources, CUstream hStream)
Unmap graphics resources.
- **CUresult cuGraphicsUnregisterResource** (CUgraphicsResource resource)
Unregisters a graphics resource for access by CUDA.

5.49.1 Detailed Description

This section describes the graphics interoperability functions of the low-level CUDA driver application programming interface.

5.49.2 Function Documentation

5.49.2.1 CUresult cuGraphicsMapResources (unsigned int count, CUgraphicsResource * resources, CUstream hStream)

Maps the `count` graphics resources in `resources` for access by CUDA.

The resources in `resources` may be accessed by CUDA until they are unmapped. The graphics API from which `resources` were registered should not access any resources while they are mapped by CUDA. If an application does so, the results are undefined.

This function provides the synchronization guarantee that any graphics calls issued before `cuGraphicsMapResources()` will complete before any subsequent CUDA work issued in `stream` begins.

If `resources` includes any duplicate entries then `CUDA_ERROR_INVALID_HANDLE` is returned. If any of `resources` are presently mapped for access by CUDA then `CUDA_ERROR_ALREADY_MAPPED` is returned.

Parameters:

count - Number of resources to map
resources - Resources to map for CUDA usage
hStream - Stream with which to synchronize

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_ALREADY_MAPPED, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceGetMappedPointer](#) [cuGraphicsSubResourceGetMappedArray](#) [cuGraphicsUnmapResources](#)

5.49.2.2 CUresult cuGraphicsResourceGetMappedMipmappedArray (CUmipmappedArray * pMipmappedArray, CUgraphicsResource resource)

Returns in *pMipmappedArray* a mipmapped array through which the mapped graphics resource *resource*. The value set in *pMipmappedArray* may change every time that *resource* is mapped.

If *resource* is not a texture then it cannot be accessed via a mipmapped array and [CUDA_ERROR_NOT_MAPPED_AS_ARRAY](#) is returned. If *resource* is not mapped then [CUDA_ERROR_NOT_MAPPED](#) is returned.

Parameters:

pMipmappedArray - Returned mipmapped array through which *resource* may be accessed
resource - Mapped resource to access

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_NOT_MAPPED, CUDA_ERROR_NOT_MAPPED_AS_ARRAY

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceGetMappedPointer](#)

5.49.2.3 CUresult cuGraphicsResourceGetMappedPointer (CUdeviceptr * pDevPtr, size_t * pSize, CUgraphicsResource resource)

Returns in *pDevPtr* a pointer through which the mapped graphics resource *resource* may be accessed. Returns in *pSize* the size of the memory in bytes which may be accessed from that pointer. The value set in *pPointer* may change every time that *resource* is mapped.

If *resource* is not a buffer then it cannot be accessed via a pointer and [CUDA_ERROR_NOT_MAPPED_AS_POINTER](#) is returned. If *resource* is not mapped then [CUDA_ERROR_NOT_MAPPED](#) is returned. *

Parameters:

pDevPtr - Returned pointer through which `resource` may be accessed
pSize - Returned size of the buffer accessible starting at `*pPointer`
resource - Mapped resource to access

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#) [CUDA_ERROR_NOT_MAPPED_AS_POINTER](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsMapResources](#), [cuGraphicsSubResourceGetMappedArray](#)

5.49.2.4 CUresult cuGraphicsResourceSetMapFlags (CUgraphicsResource *resource*, unsigned int *flags*)

Set `flags` for mapping the graphics resource `resource`.

Changes to `flags` will take effect the next time `resource` is mapped. The `flags` argument may be any of the following:

- [CU_GRAPHICS_MAP_RESOURCE_FLAGS_NONE](#): Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA kernels. This is the default value.
- [CU_GRAPHICS_MAP_RESOURCE_FLAGS_READONLY](#): Specifies that CUDA kernels which access this resource will not write to this resource.
- [CU_GRAPHICS_MAP_RESOURCE_FLAGS_WRITEDISCARD](#): Specifies that CUDA kernels which access this resource will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

If `resource` is presently mapped for access by CUDA then [CUDA_ERROR_ALREADY_MAPPED](#) is returned. If `flags` is not one of the above values then [CUDA_ERROR_INVALID_VALUE](#) is returned.

Parameters:

resource - Registered resource to set flags for
flags - Parameters for resource mapping

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_ALREADY_MAPPED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsMapResources](#)

5.49.2.5 CUresult cuGraphicsSubResourceGetMappedArray (CUarray * *pArray*, CUgraphicsResource *resource*, unsigned int *arrayIndex*, unsigned int *mipLevel*)

Returns in **pArray* an array through which the subresource of the mapped graphics resource *resource* which corresponds to array index *arrayIndex* and mipmap level *mipLevel* may be accessed. The value set in **pArray* may change every time that *resource* is mapped.

If *resource* is not a texture then it cannot be accessed via an array and [CUDA_ERROR_NOT_MAPPED_AS_ARRAY](#) is returned. If *arrayIndex* is not a valid array index for *resource* then [CUDA_ERROR_INVALID_VALUE](#) is returned. If *mipLevel* is not a valid mipmap level for *resource* then [CUDA_ERROR_INVALID_VALUE](#) is returned. If *resource* is not mapped then [CUDA_ERROR_NOT_MAPPED](#) is returned.

Parameters:

pArray - Returned array through which a subresource of *resource* may be accessed

resource - Mapped resource to access

arrayIndex - Array index for array textures or cubemap face index as defined by [CUarray_cubemap_face](#) for cubemap textures for the subresource to access

mipLevel - Mipmap level for the subresource to access

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#), [CUDA_ERROR_NOT_MAPPED_AS_ARRAY](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceGetMappedPointer](#)

5.49.2.6 CUresult cuGraphicsUnmapResources (unsigned int *count*, CUgraphicsResource * *resources*, CUstream *hStream*)

Unmaps the *count* graphics resources in *resources*.

Once unmapped, the resources in *resources* may not be accessed by CUDA until they are mapped again.

This function provides the synchronization guarantee that any CUDA work issued in *stream* before [cuGraphicsUnmapResources\(\)](#) will complete before any subsequently issued graphics work begins.

If *resources* includes any duplicate entries then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If any of *resources* are not presently mapped for access by CUDA then [CUDA_ERROR_NOT_MAPPED](#) is returned.

Parameters:

count - Number of resources to unmap

resources - Resources to unmap

hStream - Stream with which to synchronize

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsMapResources](#)

5.49.2.7 CUresult cuGraphicsUnregisterResource (CUgraphicsResource *resource*)

Unregisters the graphics resource `resource` so it is not accessible by CUDA unless registered again.

If `resource` is invalid then [CUDA_ERROR_INVALID_HANDLE](#) is returned.

Parameters:

resource - Resource to unregister

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsD3D9RegisterResource](#), [cuGraphicsD3D10RegisterResource](#), [cuGraphicsD3D11RegisterResource](#), [cuGraphicsGLRegisterBuffer](#), [cuGraphicsGLRegisterImage](#)

5.50 Profiler Control

Functions

- [CUresult cuProfilerInitialize](#) (const char *configFile, const char *outputFile, CUoutput_mode outputMode)
Initialize the profiling.
- [CUresult cuProfilerStart](#) (void)
Enable profiling.
- [CUresult cuProfilerStop](#) (void)
Disable profiling.

5.50.1 Detailed Description

This section describes the profiler control functions of the low-level CUDA driver application programming interface.

5.50.2 Function Documentation

5.50.2.1 [CUresult cuProfilerInitialize](#) (const char * *configFile*, const char * *outputFile*, CUoutput_mode *outputMode*)

Using this API user can initialize the CUDA profiler by specifying the configuration file, output file and output file format. This API is generally used to profile different set of counters by looping the kernel launch. The `configFile` parameter can be used to select profiling options including profiler counters. Refer to the "Compute Command Line Profiler User Guide" for supported profiler options and counters.

Limitation: The CUDA profiler cannot be initialized with this API if another profiling tool is already active, as indicated by the [CUDA_ERROR_PROFILER_DISABLED](#) return code.

Typical usage of the profiling APIs is as follows:

for each set of counters/options

```
{  
cuProfilerInitialize(); //Initialize profiling, set the counters or options in the config file  
...  
cuProfilerStart();  
// code to be profiled  
cuProfilerStop();  
...  
cuProfilerStart();  
// code to be profiled  
cuProfilerStop();  
...  
}
```

Parameters:

configFile - Name of the config file that lists the counters/options for profiling.

outputFile - Name of the outputFile where the profiling results will be stored.

outputMode - outputMode, can be CU_OUT_KEY_VALUE_PAIR or CU_OUT_CSV.

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_PROFILER_DISABLED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuProfilerStart](#), [cuProfilerStop](#)

5.50.2.2 CUresult cuProfilerStart (void)

Enables profile collection by the active profiling tool. If profiling is already enabled, then [cuProfilerStart\(\)](#) has no effect.

[cuProfilerStart](#) and [cuProfilerStop](#) APIs are used to programmatically control the profiling granularity by allowing profiling to be done only on selective pieces of code.

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_CONTEXT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuProfilerInitialize](#), [cuProfilerStop](#)

5.50.2.3 CUresult cuProfilerStop (void)

Disables profile collection by the active profiling tool. If profiling is already disabled, then [cuProfilerStop\(\)](#) has no effect.

[cuProfilerStart](#) and [cuProfilerStop](#) APIs are used to programmatically control the profiling granularity by allowing profiling to be done only on selective pieces of code.

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_CONTEXT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuProfilerInitialize](#), [cuProfilerStart](#)

5.51 OpenGL Interoperability

Modules

- [OpenGL Interoperability \[DEPRECATED\]](#)

Typedefs

- typedef enum [CUGLDeviceList_enum](#) CUGLDeviceList

Enumerations

- enum [CUGLDeviceList_enum](#) {
 [CU_GL_DEVICE_LIST_ALL](#) = 0x01,
 [CU_GL_DEVICE_LIST_CURRENT_FRAME](#) = 0x02,
 [CU_GL_DEVICE_LIST_NEXT_FRAME](#) = 0x03 }

Functions

- [CUresult cuGLGetDevices](#) (unsigned int *pCudaDeviceCount, [CUdevice](#) *pCudaDevices, unsigned int cudaDeviceCount, [CUGLDeviceList](#) deviceList)
Gets the CUDA devices associated with the current OpenGL context.
- [CUresult cuGraphicsGLRegisterBuffer](#) ([CUgraphicsResource](#) *pCudaResource, GLuint buffer, unsigned int Flags)
Registers an OpenGL buffer object.
- [CUresult cuGraphicsGLRegisterImage](#) ([CUgraphicsResource](#) *pCudaResource, GLuint image, GLenum target, unsigned int Flags)
Register an OpenGL texture or renderbuffer object.
- [CUresult cuWGLGetDevice](#) ([CUdevice](#) *pDevice, HGPUNV hGpu)
Gets the CUDA device associated with hGpu.

5.51.1 Detailed Description

This section describes the OpenGL interoperability functions of the low-level CUDA driver application programming interface. Note that mapping of OpenGL resources is performed with the graphics API agnostic, resource mapping interface described in [Graphics Interopability](#).

5.51.2 Typedef Documentation

5.51.2.1 typedef enum CUGLDeviceList_enum CUGLDeviceList

CUDA devices corresponding to an OpenGL device

5.51.3 Enumeration Type Documentation

5.51.3.1 enum CUGLDeviceList_enum

CUDA devices corresponding to an OpenGL device

Enumerator:

CU_GL_DEVICE_LIST_ALL The CUDA devices for all GPUs used by the current OpenGL context

CU_GL_DEVICE_LIST_CURRENT_FRAME The CUDA devices for the GPUs used by the current OpenGL context in its currently rendering frame

CU_GL_DEVICE_LIST_NEXT_FRAME The CUDA devices for the GPUs to be used by the current OpenGL context in the next frame

5.51.4 Function Documentation

5.51.4.1 CUresult cuGLGetDevices (unsigned int *pCudaDeviceCount, CUdevice *pCudaDevices, unsigned int cudaDeviceCount, CUGLDeviceList deviceList)

Returns in *pCudaDeviceCount the number of CUDA-compatible devices corresponding to the current OpenGL context. Also returns in *pCudaDevices at most cudaDeviceCount of the CUDA-compatible devices corresponding to the current OpenGL context. If any of the GPUs being used by the current OpenGL context are not CUDA capable then the call will return CUDA_ERROR_NO_DEVICE.

The deviceList argument may be any of the following:

- [CU_GL_DEVICE_LIST_ALL](#): Query all devices used by the current OpenGL context.
- [CU_GL_DEVICE_LIST_CURRENT_FRAME](#): Query the devices used by the current OpenGL context to render the current frame (in SLI).
- [CU_GL_DEVICE_LIST_NEXT_FRAME](#): Query the devices used by the current OpenGL context to render the next frame (in SLI). Note that this is a prediction, it can't be guaranteed that this is correct in all cases.

Parameters:

pCudaDeviceCount - Returned number of CUDA devices.

pCudaDevices - Returned CUDA devices.

cudaDeviceCount - The size of the output device array pCudaDevices.

deviceList - The set of devices to return.

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_NO_DEVICE](#), [CUDA_ERROR_INVALID_VALUE](#) [CUDA_ERROR_INVALID_CONTEXT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuWGLGetDevice](#)

5.51.4.2 CUresult cuGraphicsGLRegisterBuffer (CUgraphicsResource * pCudaResource, GLuint buffer, unsigned int Flags)

Registers the buffer object specified by `buffer` for access by CUDA. A handle to the registered object is returned as `pCudaResource`. The register flags `Flags` specify the intended usage, as follows:

- `CU_GRAPHICS_REGISTER_FLAGS_NONE`: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA. This is the default value.
- `CU_GRAPHICS_REGISTER_FLAGS_READ_ONLY`: Specifies that CUDA will not write to this resource.
- `CU_GRAPHICS_REGISTER_FLAGS_WRITE_DISCARD`: Specifies that CUDA will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

Parameters:

pCudaResource - Pointer to the returned object handle

buffer - name of buffer object to be registered

Flags - Register flags

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_ALREADY_MAPPED](#), [CUDA_ERROR_INVALID_CONTEXT](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnregisterResource](#), [cuGraphicsMapResources](#), [cuGraphicsResourceGetMappedPointer](#)

5.51.4.3 CUresult cuGraphicsGLRegisterImage (CUgraphicsResource * pCudaResource, GLuint image, GLenum target, unsigned int Flags)

Registers the texture or renderbuffer object specified by `image` for access by CUDA. A handle to the registered object is returned as `pCudaResource`.

`target` must match the type of the object, and must be one of `GL_TEXTURE_2D`, `GL_TEXTURE_RECTANGLE`, `GL_TEXTURE_CUBE_MAP`, `GL_TEXTURE_3D`, `GL_TEXTURE_2D_ARRAY`, or `GL_RENDERBUFFER`.

The register flags `Flags` specify the intended usage, as follows:

- `CU_GRAPHICS_REGISTER_FLAGS_NONE`: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA. This is the default value.
- `CU_GRAPHICS_REGISTER_FLAGS_READ_ONLY`: Specifies that CUDA will not write to this resource.
- `CU_GRAPHICS_REGISTER_FLAGS_WRITE_DISCARD`: Specifies that CUDA will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.
- `CU_GRAPHICS_REGISTER_FLAGS_SURFACE_LDST`: Specifies that CUDA will bind this resource to a surface reference.

- `CU_GRAPHICS_REGISTER_FLAGS_TEXTURE_GATHER`: Specifies that CUDA will perform texture gather operations on this resource.

The following image formats are supported. For brevity's sake, the list is abbreviated. For ex., `{GL_R, GL_RG} X {8, 16}` would expand to the following 4 formats `{GL_R8, GL_R16, GL_RG8, GL_RG16}` :

- `GL_RED, GL_RG, GL_RGBA, GL_LUMINANCE, GL_ALPHA, GL_LUMINANCE_ALPHA, GL_INTENSITY`
- `{GL_R, GL_RG, GL_RGBA} X {8, 16, 16F, 32F, 8UI, 16UI, 32UI, 8I, 16I, 32I}`
- `{GL_LUMINANCE, GL_ALPHA, GL_LUMINANCE_ALPHA, GL_INTENSITY} X {8, 16, 16F_ARB, 32F_ARB, 8UI_EXT, 16UI_EXT, 32UI_EXT, 8I_EXT, 16I_EXT, 32I_EXT}`

The following image classes are currently disallowed:

- Textures with borders
- Multisampled renderbuffers

Parameters:

- *pCudaResource* - Pointer to the returned object handle
- *image* - name of texture or renderbuffer object to be registered
- *target* - Identifies the type of object specified by *image*
- *Flags* - Register flags

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_ALREADY_MAPPED](#), [CUDA_ERROR_INVALID_CONTEXT](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnregisterResource](#), [cuGraphicsMapResources](#), [cuGraphicsSubResourceGetMappedArray](#)

5.51.4.4 CUresult cuWGLGetDevice (CUdevice *pDevice, HGPUNV hGpu)

Returns in *pDevice the CUDA device associated with a hGpu, if applicable.

Parameters:

- *pDevice* - Device associated with hGpu
- *hGpu* - Handle to a GPU, as queried via `WGL_NV_gpu_affinity()`

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGLMapBufferObject](#), [cuGLRegisterBufferObject](#), [cuGLUnmapBufferObject](#), [cuGLUnregisterBufferObject](#), [cuGLUnmapBufferObjectAsync](#), [cuGLSetBufferObjectMapFlags](#)

5.52 OpenGL Interoperability [DEPRECATED]

Typedefs

- typedef enum [CUGLmap_flags_enum](#) [CUGLmap_flags](#)

Enumerations

- enum [CUGLmap_flags_enum](#)

Functions

- [CUresult cuGLCtxCreate](#) ([CUcontext](#) *pCtx, unsigned int Flags, [CUdevice](#) device)
Create a CUDA context for interoperability with OpenGL.
- [CUresult cuGLInit](#) (void)
Initializes OpenGL interoperability.
- [CUresult cuGLMapBufferObject](#) ([CUdeviceptr](#) *dptr, size_t *size, GLuint buffer)
Maps an OpenGL buffer object.
- [CUresult cuGLMapBufferObjectAsync](#) ([CUdeviceptr](#) *dptr, size_t *size, GLuint buffer, [CUstream](#) hStream)
Maps an OpenGL buffer object.
- [CUresult cuGLRegisterBufferObject](#) (GLuint buffer)
Registers an OpenGL buffer object.
- [CUresult cuGLSetBufferObjectMapFlags](#) (GLuint buffer, unsigned int Flags)
Set the map flags for an OpenGL buffer object.
- [CUresult cuGLUnmapBufferObject](#) (GLuint buffer)
Unmaps an OpenGL buffer object.
- [CUresult cuGLUnmapBufferObjectAsync](#) (GLuint buffer, [CUstream](#) hStream)
Unmaps an OpenGL buffer object.
- [CUresult cuGLUnregisterBufferObject](#) (GLuint buffer)
Unregister an OpenGL buffer object.

5.52.1 Detailed Description

This section describes deprecated OpenGL interoperability functionality.

5.52.2 Typedef Documentation

5.52.2.1 typedef enum [CUGLmap_flags_enum](#) [CUGLmap_flags](#)

Flags to map or unmap a resource

5.52.3 Enumeration Type Documentation

5.52.3.1 enum CUGLmap_flags_enum

Flags to map or unmap a resource

5.52.4 Function Documentation

5.52.4.1 CUresult cuGLCtxCreate (CUcontext *pCtx, unsigned int Flags, CUdevice device)

Deprecated

This function is deprecated as of Cuda 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA context with an OpenGL context in order to achieve maximum interoperability performance.

Parameters:

pCtx - Returned CUDA context

Flags - Options for CUDA context creation

device - Device on which to create the context

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_OUT_OF_MEMORY

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuCtxCreate, cuGLInit, cuGLMapBufferObject, cuGLRegisterBufferObject, cuGLUnmapBufferObject, cuGLUnregisterBufferObject, cuGLMapBufferObjectAsync, cuGLUnmapBufferObjectAsync, cuGLSetBufferObjectMapFlags, cuWGLGetDevice

5.52.4.2 CUresult cuGLInit (void)

Deprecated

This function is deprecated as of Cuda 3.0.

Initializes OpenGL interoperability. This function is deprecated and calling it is no longer required. It may fail if the needed OpenGL driver facilities are not available.

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGLMapBufferObject](#), [cuGLRegisterBufferObject](#), [cuGLUnmapBufferObject](#), [cuGLUnregisterBufferObject](#), [cuGLMapBufferObjectAsync](#), [cuGLUnmapBufferObjectAsync](#), [cuGLSetBufferObjectMapFlags](#), [cuWGLGetDevice](#)

5.52.4.3 CUresult cuGLMapBufferObject (CUdeviceptr * *dptr*, size_t * *size*, GLuint *buffer*)**Deprecated**

This function is deprecated as of Cuda 3.0.

Maps the buffer object specified by `buffer` into the address space of the current CUDA context and returns in `*dptr` and `*size` the base pointer and size of the resulting mapping.

There must be a valid OpenGL context bound to the current thread when this function is called. This must be the same context, or a member of the same shareGroup, as the context that was bound when the buffer was registered.

All streams in the current CUDA context are synchronized with the current GL context.

Parameters:

dptr - Returned mapped base pointer
size - Returned size of mapping
buffer - The name of the buffer object to map

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_MAP_FAILED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsMapResources](#)

5.52.4.4 CUresult cuGLMapBufferObjectAsync (CUdeviceptr * *dptr*, size_t * *size*, GLuint *buffer*, CUstream *hStream*)**Deprecated**

This function is deprecated as of Cuda 3.0.

Maps the buffer object specified by `buffer` into the address space of the current CUDA context and returns in `*dptr` and `*size` the base pointer and size of the resulting mapping.

There must be a valid OpenGL context bound to the current thread when this function is called. This must be the same context, or a member of the same shareGroup, as the context that was bound when the buffer was registered.

Stream `hStream` in the current CUDA context is synchronized with the current GL context.

Parameters:

dptr - Returned mapped base pointer

size - Returned size of mapping

buffer - The name of the buffer object to map

hStream - Stream to synchronize

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_MAP_FAILED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsMapResources](#)

5.52.4.5 CUresult cuGLRegisterBufferObject (GLuint *buffer*)

Deprecated

This function is deprecated as of Cuda 3.0.

Registers the buffer object specified by *buffer* for access by CUDA. This function must be called before CUDA can map the buffer object. There must be a valid OpenGL context bound to the current thread when this function is called, and the buffer name is resolved by that context.

Parameters:

buffer - The name of the buffer object to register.

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_ALREADY_MAPPED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsGLRegisterBuffer](#)

5.52.4.6 CUresult cuGLSetBufferObjectMapFlags (GLuint *buffer*, unsigned int *Flags*)

Deprecated

This function is deprecated as of Cuda 3.0.

Sets the map flags for the buffer object specified by *buffer*.

Changes to *Flags* will take effect the next time *buffer* is mapped. The *Flags* argument may be any of the following:

- `CU_GL_MAP_RESOURCE_FLAGS_NONE`: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA kernels. This is the default value.
- `CU_GL_MAP_RESOURCE_FLAGS_READ_ONLY`: Specifies that CUDA kernels which access this resource will not write to this resource.
- `CU_GL_MAP_RESOURCE_FLAGS_WRITE_DISCARD`: Specifies that CUDA kernels which access this resource will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

If `buffer` has not been registered for use with CUDA, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If `buffer` is presently mapped for access by CUDA, then [CUDA_ERROR_ALREADY_MAPPED](#) is returned.

There must be a valid OpenGL context bound to the current thread when this function is called. This must be the same context, or a member of the same `shareGroup`, as the context that was bound when the buffer was registered.

Parameters:

buffer - Buffer object to unmap

Flags - Map flags

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_ALREADY_MAPPED](#), [CUDA_ERROR_INVALID_CONTEXT](#),

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceSetMapFlags](#)

5.52.4.7 CUresult cuGLUnmapBufferObject (GLuint *buffer*)

Deprecated

This function is deprecated as of Cuda 3.0.

Unmaps the buffer object specified by `buffer` for access by CUDA.

There must be a valid OpenGL context bound to the current thread when this function is called. This must be the same context, or a member of the same `shareGroup`, as the context that was bound when the buffer was registered.

All streams in the current CUDA context are synchronized with the current GL context.

Parameters:

buffer - Buffer object to unmap

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnmapResources](#)

5.52.4.8 CUresult cuGLUnmapBufferObjectAsync (GLuint *buffer*, CUSTream *hStream*)**Deprecated**

This function is deprecated as of Cuda 3.0.

Unmaps the buffer object specified by *buffer* for access by CUDA.

There must be a valid OpenGL context bound to the current thread when this function is called. This must be the same context, or a member of the same *shareGroup*, as the context that was bound when the buffer was registered.

Stream *hStream* in the current CUDA context is synchronized with the current GL context.

Parameters:

buffer - Name of the buffer object to unmap

hStream - Stream to synchronize

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnmapResources](#)

5.52.4.9 CUresult cuGLUnregisterBufferObject (GLuint *buffer*)**Deprecated**

This function is deprecated as of Cuda 3.0.

Unregisters the buffer object specified by *buffer*. This releases any resources associated with the registered buffer. After this call, the buffer may no longer be mapped for access by CUDA.

There must be a valid OpenGL context bound to the current thread when this function is called. This must be the same context, or a member of the same *shareGroup*, as the context that was bound when the buffer was registered.

Parameters:

buffer - Name of the buffer object to unregister

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnregisterResource](#)

5.53 Direct3D 9 Interoperability

Modules

- [Direct3D 9 Interoperability \[DEPRECATED\]](#)

Typedefs

- typedef enum [CUd3d9DeviceList_enum](#) [CUd3d9DeviceList](#)

Enumerations

- enum [CUd3d9DeviceList_enum](#) {
[CU_D3D9_DEVICE_LIST_ALL](#) = 0x01,
[CU_D3D9_DEVICE_LIST_CURRENT_FRAME](#) = 0x02,
[CU_D3D9_DEVICE_LIST_NEXT_FRAME](#) = 0x03 }

Functions

- [CUresult cuD3D9CtxCreate](#) ([CUcontext](#) *pCtx, [CUdevice](#) *pCudaDevice, unsigned int Flags, [IDirect3DDevice9](#) *pD3DDevice)
Create a CUDA context for interoperability with Direct3D 9.
- [CUresult cuD3D9CtxCreateOnDevice](#) ([CUcontext](#) *pCtx, unsigned int flags, [IDirect3DDevice9](#) *pD3DDevice, [CUdevice](#) cudaDevice)
Create a CUDA context for interoperability with Direct3D 9.
- [CUresult cuD3D9GetDevice](#) ([CUdevice](#) *pCudaDevice, const char *pszAdapterName)
Gets the CUDA device corresponding to a display adapter.
- [CUresult cuD3D9GetDevices](#) (unsigned int *pCudaDeviceCount, [CUdevice](#) *pCudaDevices, unsigned int cudaDeviceCount, [IDirect3DDevice9](#) *pD3D9Device, [CUd3d9DeviceList](#) deviceList)
Gets the CUDA devices corresponding to a Direct3D 9 device.
- [CUresult cuD3D9GetDirect3DDevice](#) ([IDirect3DDevice9](#) **ppD3DDevice)
Get the Direct3D 9 device against which the current CUDA context was created.
- [CUresult cuGraphicsD3D9RegisterResource](#) ([CUgraphicsResource](#) *pCudaResource, [IDirect3DResource9](#) *pD3DResource, unsigned int Flags)
Register a Direct3D 9 resource for access by CUDA.

5.53.1 Detailed Description

This section describes the Direct3D 9 interoperability functions of the low-level CUDA driver application programming interface. Note that mapping of Direct3D 9 resources is performed with the graphics API agnostic, resource mapping interface described in [Graphics Interopability](#).

5.53.2 Typedef Documentation

5.53.2.1 typedef enum CUd3d9DeviceList_enum CUd3d9DeviceList

CUDA devices corresponding to a D3D9 device

5.53.3 Enumeration Type Documentation

5.53.3.1 enum CUd3d9DeviceList_enum

CUDA devices corresponding to a D3D9 device

Enumerator:

CU_D3D9_DEVICE_LIST_ALL The CUDA devices for all GPUs used by a D3D9 device

CU_D3D9_DEVICE_LIST_CURRENT_FRAME The CUDA devices for the GPUs used by a D3D9 device in its currently rendering frame

CU_D3D9_DEVICE_LIST_NEXT_FRAME The CUDA devices for the GPUs to be used by a D3D9 device in the next frame

5.53.4 Function Documentation

5.53.4.1 CUresult cuD3D9CtxCreate (CUcontext * pCtx, CUdevice * pCudaDevice, unsigned int Flags, IDirect3DDevice9 * pD3DDevice)

Creates a new CUDA context, enables interoperability for that context with the Direct3D device `pD3DDevice`, and associates the created CUDA context with the calling thread. The created [CUcontext](#) will be returned in `*pCtx`. Direct3D resources from this device may be registered and mapped through the lifetime of this CUDA context. If `pCudaDevice` is non-NULL then the [CUdevice](#) on which this CUDA context was created will be returned in `*pCudaDevice`.

On success, this call will increase the internal reference count on `pD3DDevice`. This reference count will be decremented upon destruction of this context through [cuCtxDestroy\(\)](#). This context will cease to function if `pD3DDevice` is destroyed or encounters an error.

Note that this function is never required for correct functionality. Use of this function will result in accelerated interoperability only when the operating system is Windows Vista or Windows 7, and the device `pD3DDevice` is not an `IDirect3DDevice9Ex`. In all other circumstances, this function is not necessary.

Parameters:

pCtx - Returned newly created CUDA context

pCudaDevice - Returned pointer to the device on which the context was created

Flags - Context creation flags (see [cuCtxCreate\(\)](#) for details)

pD3DDevice - Direct3D device to create interoperability context with

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D9GetDevice](#), [cuGraphicsD3D9RegisterResource](#)

5.53.4.2 CUresult cuD3D9CtxCreateOnDevice (CUcontext * pCtx, unsigned int flags, IDirect3DDevice9 * pD3DDevice, CUdevice cudaDevice)

Creates a new CUDA context, enables interoperability for that context with the Direct3D device `pD3DDevice`, and associates the created CUDA context with the calling thread. The created [CUcontext](#) will be returned in `*pCtx`. Direct3D resources from this device may be registered and mapped through the lifetime of this CUDA context.

On success, this call will increase the internal reference count on `pD3DDevice`. This reference count will be decremented upon destruction of this context through [cuCtxDestroy\(\)](#). This context will cease to function if `pD3DDevice` is destroyed or encounters an error.

Note that this function is never required for correct functionality. Use of this function will result in accelerated interoperability only when the operating system is Windows Vista or Windows 7, and the device `pD3DDevice` is not an `IDirect3DDevice9Ex`. In all other circumstances, this function is not necessary.

Parameters:

pCtx - Returned newly created CUDA context

flags - Context creation flags (see [cuCtxCreate\(\)](#) for details)

pD3DDevice - Direct3D device to create interoperability context with

cudaDevice - The CUDA device on which to create the context. This device must be among the devices returned when querying `CU_D3D9_DEVICES_ALL` from [cuD3D9GetDevices](#).

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D9GetDevices](#), [cuGraphicsD3D9RegisterResource](#)

5.53.4.3 CUresult cuD3D9GetDevice (CUdevice * pCudaDevice, const char * pszAdapterName)

Returns in `*pCudaDevice` the CUDA-compatible device corresponding to the adapter name `pszAdapterName` obtained from `EnumDisplayDevices()` or `IDirect3D9::GetAdapterIdentifier()`.

If no device on the adapter with name `pszAdapterName` is CUDA-compatible, then the call will fail.

Parameters:

pCudaDevice - Returned CUDA device corresponding to `pszAdapterName`

pszAdapterName - Adapter name to query for device

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_FOUND](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D9CtxCreate](#)

5.53.4.4 CUresult cuD3D9GetDevices (unsigned int * *pCudaDeviceCount*, CUdevice * *pCudaDevices*, unsigned int *cudaDeviceCount*, IDirect3DDevice9 * *pD3D9Device*, CUd3d9DeviceList *deviceList*)

Returns in *pCudaDeviceCount* the number of CUDA-compatible device corresponding to the Direct3D 9 device *pD3D9Device*. Also returns in *pCudaDevices* at most *cudaDeviceCount* of the the CUDA-compatible devices corresponding to the Direct3D 9 device *pD3D9Device*.

If any of the GPUs being used to render *pDevice* are not CUDA capable then the call will return [CUDA_ERROR_NO_DEVICE](#).

Parameters:

pCudaDeviceCount - Returned number of CUDA devices corresponding to *pD3D9Device*

pCudaDevices - Returned CUDA devices corresponding to *pD3D9Device*

cudaDeviceCount - The size of the output device array *pCudaDevices*

pD3D9Device - Direct3D 9 device to query for CUDA devices

deviceList - The set of devices to return. This set may be [CU_D3D9_DEVICE_LIST_ALL](#) for all devices, [CU_D3D9_DEVICE_LIST_CURRENT_FRAME](#) for the devices used to render the current frame (in SLI), or [CU_D3D9_DEVICE_LIST_NEXT_FRAME](#) for the devices used to render the next frame (in SLI).

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_NO_DEVICE](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_FOUND](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D9CtxCreate](#)

5.53.4.5 CUresult cuD3D9GetDirect3DDevice (IDirect3DDevice9 ** *ppD3DDevice*)

Returns in *ppD3DDevice* the Direct3D device against which this CUDA context was created in [cuD3D9CtxCreate\(\)](#).

Parameters:

ppD3DDevice - Returned Direct3D device corresponding to CUDA context

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D9GetDevice](#)

5.53.4.6 CUresult cuGraphicsD3D9RegisterResource (CUgraphicsResource * pCudaResource, IDirect3DResource9 * pD3DResource, unsigned int Flags)

Registers the Direct3D 9 resource `pD3DResource` for access by CUDA and returns a CUDA handle to `pD3DResource` in `pCudaResource`. The handle returned in `pCudaResource` may be used to map and unmap this resource until it is unregistered. On success this call will increase the internal reference count on `pD3DResource`. This reference count will be decremented when this resource is unregistered through [cuGraphicsUnregisterResource\(\)](#).

This call is potentially high-overhead and should not be called every frame in interactive applications.

The type of `pD3DResource` must be one of the following.

- `IDirect3DVertexBuffer9`: may be accessed through a device pointer
- `IDirect3DIndexBuffer9`: may be accessed through a device pointer
- `IDirect3DSurface9`: may be accessed through an array. Only stand-alone objects of type `IDirect3DSurface9` may be explicitly shared. In particular, individual mipmap levels and faces of cube maps may not be registered directly. To access individual surfaces associated with a texture, one must register the base texture object.
- `IDirect3DBaseTexture9`: individual surfaces on this texture may be accessed through an array.

The `Flags` argument may be used to specify additional parameters at register time. The valid values for this parameter are

- `CU_GRAPHICS_REGISTER_FLAGS_NONE`: Specifies no hints about how this resource will be used.
- `CU_GRAPHICS_REGISTER_FLAGS_SURFACE_LDST`: Specifies that CUDA will bind this resource to a surface reference.
- `CU_GRAPHICS_REGISTER_FLAGS_TEXTURE_GATHER`: Specifies that CUDA will perform texture gather operations on this resource.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

If Direct3D interoperability is not initialized for this context using [cuD3D9CtxCreate](#) then `CUDA_ERROR_INVALID_CONTEXT` is returned. If `pD3DResource` is of incorrect type or is already registered then `CUDA_ERROR_INVALID_HANDLE` is returned. If `pD3DResource` cannot be registered then `CUDA_ERROR_UNKNOWN` is returned. If `Flags` is not one of the above specified value then `CUDA_ERROR_INVALID_VALUE` is returned.

Parameters:

pCudaResource - Returned graphics resource handle

pD3DResource - Direct3D resource to register

Flags - Parameters for resource registration

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D9CtxCreate](#), [cuGraphicsUnregisterResource](#), [cuGraphicsMapResources](#), [cuGraphicsSubResourceGetMappedArray](#), [cuGraphicsResourceGetMappedPointer](#)

5.54 Direct3D 9 Interoperability [DEPRECATED]

Typedefs

- typedef enum [CUd3d9map_flags_enum](#) [CUd3d9map_flags](#)
- typedef enum [CUd3d9register_flags_enum](#) [CUd3d9register_flags](#)

Enumerations

- enum [CUd3d9map_flags_enum](#)
- enum [CUd3d9register_flags_enum](#)

Functions

- [CUresult cuD3D9MapResources](#) (unsigned int count, IDirect3DResource9 **ppResource)
Map Direct3D resources for access by CUDA.
- [CUresult cuD3D9RegisterResource](#) (IDirect3DResource9 *pResource, unsigned int Flags)
Register a Direct3D resource for access by CUDA.
- [CUresult cuD3D9ResourceGetMappedArray](#) (CUarray *pArray, IDirect3DResource9 *pResource, unsigned int Face, unsigned int Level)
Get an array through which to access a subresource of a Direct3D resource which has been mapped for access by CUDA.
- [CUresult cuD3D9ResourceGetMappedPitch](#) (size_t *pPitch, size_t *pPitchSlice, IDirect3DResource9 *pResource, unsigned int Face, unsigned int Level)
Get the pitch of a subresource of a Direct3D resource which has been mapped for access by CUDA.
- [CUresult cuD3D9ResourceGetMappedPointer](#) (CUdeviceptr *pDevPtr, IDirect3DResource9 *pResource, unsigned int Face, unsigned int Level)
Get the pointer through which to access a subresource of a Direct3D resource which has been mapped for access by CUDA.
- [CUresult cuD3D9ResourceGetMappedSize](#) (size_t *pSize, IDirect3DResource9 *pResource, unsigned int Face, unsigned int Level)
Get the size of a subresource of a Direct3D resource which has been mapped for access by CUDA.
- [CUresult cuD3D9ResourceGetSurfaceDimensions](#) (size_t *pWidth, size_t *pHeight, size_t *pDepth, IDirect3DResource9 *pResource, unsigned int Face, unsigned int Level)
Get the dimensions of a registered surface.
- [CUresult cuD3D9ResourceSetMapFlags](#) (IDirect3DResource9 *pResource, unsigned int Flags)
Set usage flags for mapping a Direct3D resource.
- [CUresult cuD3D9UnmapResources](#) (unsigned int count, IDirect3DResource9 **ppResource)
Unmaps Direct3D resources.
- [CUresult cuD3D9UnregisterResource](#) (IDirect3DResource9 *pResource)
Unregister a Direct3D resource.

5.54.1 Detailed Description

This section describes deprecated Direct3D 9 interoperability functionality.

5.54.2 Typedef Documentation

5.54.2.1 typedef enum CUd3d9map_flags_enum CUd3d9map_flags

Flags to map or unmap a resource

5.54.2.2 typedef enum CUd3d9register_flags_enum CUd3d9register_flags

Flags to register a resource

5.54.3 Enumeration Type Documentation

5.54.3.1 enum CUd3d9map_flags_enum

Flags to map or unmap a resource

5.54.3.2 enum CUd3d9register_flags_enum

Flags to register a resource

5.54.4 Function Documentation

5.54.4.1 CUresult cuD3D9MapResources (unsigned int *count*, IDirect3DResource9 ** *ppResource*)

Deprecated

This function is deprecated as of Cuda 3.0.

Maps the *count* Direct3D resources in *ppResource* for access by CUDA.

The resources in *ppResource* may be accessed in CUDA kernels until they are unmapped. Direct3D should not access any resources while they are mapped by CUDA. If an application does so the results are undefined.

This function provides the synchronization guarantee that any Direct3D calls issued before `cuD3D9MapResources()` will complete before any CUDA kernels issued after `cuD3D9MapResources()` begin.

If any of *ppResource* have not been registered for use with CUDA or if *ppResource* contains any duplicate entries, then `CUDA_ERROR_INVALID_HANDLE` is returned. If any of *ppResource* are presently mapped for access by CUDA, then `CUDA_ERROR_ALREADY_MAPPED` is returned.

Parameters:

count - Number of resources in *ppResource*

ppResource - Resources to map for CUDA usage

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_ALREADY_MAPPED, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsMapResources](#)

5.54.4.2 CUresult cuD3D9RegisterResource (IDirect3DResource9 *pResource, unsigned int Flags)**Deprecated**

This function is deprecated as of Cuda 3.0.

Registers the Direct3D resource `pResource` for access by CUDA.

If this call is successful, then the application will be able to map and unmap this resource until it is unregistered through [cuD3D9UnregisterResource\(\)](#). Also on success, this call will increase the internal reference count on `pResource`. This reference count will be decremented when this resource is unregistered through [cuD3D9UnregisterResource\(\)](#).

This call is potentially high-overhead and should not be called every frame in interactive applications.

The type of `pResource` must be one of the following.

- `IDirect3DVertexBuffer9`: Cannot be used with `Flags` set to `CU_D3D9_REGISTER_FLAGS_ARRAY`.
- `IDirect3DIndexBuffer9`: Cannot be used with `Flags` set to `CU_D3D9_REGISTER_FLAGS_ARRAY`.
- `IDirect3DSurface9`: Only stand-alone objects of type `IDirect3DSurface9` may be explicitly shared. In particular, individual mipmap levels and faces of cube maps may not be registered directly. To access individual surfaces associated with a texture, one must register the base texture object. For restrictions on the `Flags` parameter, see type `IDirect3DBaseTexture9`.
- `IDirect3DBaseTexture9`: When a texture is registered, all surfaces associated with the all mipmap levels of all faces of the texture will be accessible to CUDA.

The `Flags` argument specifies the mechanism through which CUDA will access the Direct3D resource. The following values are allowed.

- `CU_D3D9_REGISTER_FLAGS_NONE`: Specifies that CUDA will access this resource through a [CUDevicePtr](#). The pointer, size, and (for textures), pitch for each subresource of this allocation may be queried through [cuD3D9ResourceGetMappedPointer\(\)](#), [cuD3D9ResourceGetMappedSize\(\)](#), and [cuD3D9ResourceGetMappedPitch\(\)](#) respectively. This option is valid for all resource types.
- `CU_D3D9_REGISTER_FLAGS_ARRAY`: Specifies that CUDA will access this resource through a [CUarray](#) queried on a sub-resource basis through [cuD3D9ResourceGetMappedArray\(\)](#). This option is only valid for resources of type `IDirect3DSurface9` and subtypes of `IDirect3DBaseTexture9`.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Any resources allocated in D3DPOOL_SYSTEMMEM or D3DPOOL_MANAGED may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

If Direct3D interoperability is not initialized on this context, then [CUDA_ERROR_INVALID_CONTEXT](#) is returned. If `pResource` is of incorrect type (e.g. is a non-stand-alone IDirect3DSurface9) or is already registered, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If `pResource` cannot be registered then [CUDA_ERROR_UNKNOWN](#) is returned.

Parameters:

pResource - Resource to register for CUDA access

Flags - Flags for resource registration

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsD3D9RegisterResource](#)

5.54.4.3 CUresult cuD3D9ResourceGetMappedArray (CUarray *pArray, IDirect3DResource9 *pResource, unsigned int Face, unsigned int Level)

Deprecated

This function is deprecated as of Cuda 3.0.

Returns in `*pArray` an array through which the subresource of the mapped Direct3D resource `pResource` which corresponds to `Face` and `Level` may be accessed. The value set in `pArray` may change every time that `pResource` is mapped.

If `pResource` is not registered then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If `pResource` was not registered with usage flags `CU_D3D9_REGISTER_FLAGS_ARRAY` then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If `pResource` is not mapped then [CUDA_ERROR_NOT_MAPPED](#) is returned.

For usage requirements of `Face` and `Level` parameters, see [cuD3D9ResourceGetMappedPointer\(\)](#).

Parameters:

pArray - Returned array corresponding to subresource

pResource - Mapped resource to access

Face - Face of resource to access

Level - Level of resource to access

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_NOT_MAPPED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsSubResourceGetMappedArray](#)

5.54.4.4 CUresult cuD3D9ResourceGetMappedPitch (size_t * pPitch, size_t * pPitchSlice, IDirect3DResource9 * pResource, unsigned int Face, unsigned int Level)

Deprecated

This function is deprecated as of Cuda 3.0.

Returns in *pPitch and *pPitchSlice the pitch and Z-slice pitch of the subresource of the mapped Direct3D resource pResource, which corresponds to Face and Level. The values set in pPitch and pPitchSlice may change every time that pResource is mapped.

The pitch and Z-slice pitch values may be used to compute the location of a sample on a surface as follows.

For a 2D surface, the byte offset of the sample at position **x**, **y** from the base pointer of the surface is:

y * pitch + (bytes per pixel) * x

For a 3D surface, the byte offset of the sample at position **x**, **y**, **z** from the base pointer of the surface is:

z* slicePitch + y * pitch + (bytes per pixel) * x

Both parameters pPitch and pPitchSlice are optional and may be set to NULL.

If pResource is not of type IDirect3DBaseTexture9 or one of its sub-types or if pResource has not been registered for use with CUDA, then [cudaErrorInvalidResourceHandle](#) is returned. If pResource was not registered with usage flags CU_D3D9_REGISTER_FLAGS_NONE, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If pResource is not mapped for access by CUDA then [CUDA_ERROR_NOT_MAPPED](#) is returned.

For usage requirements of Face and Level parameters, see [cuD3D9ResourceGetMappedPointer\(\)](#).

Parameters:

pPitch - Returned pitch of subresource

pPitchSlice - Returned Z-slice pitch of subresource

pResource - Mapped resource to access

Face - Face of resource to access

Level - Level of resource to access

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_NOT_MAPPED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsSubResourceGetMappedArray](#)

5.54.4.5 CUresult cuD3D9ResourceGetMappedPointer (CUdeviceptr * *pDevPtr*, IDirect3DResource9 * *pResource*, unsigned int *Face*, unsigned int *Level*)

Deprecated

This function is deprecated as of Cuda 3.0.

Returns in **pDevPtr* the base pointer of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *Face* and *Level*. The value set in *pDevPtr* may change every time that *pResource* is mapped.

If *pResource* is not registered, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* was not registered with usage flags `CU_D3D9_REGISTER_FLAGS_NONE`, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* is not mapped, then [CUDA_ERROR_NOT_MAPPED](#) is returned.

If *pResource* is of type `IDirect3DCubeTexture9`, then *Face* must one of the values enumerated by type `D3DCUBEMAP_FACES`. For all other types *Face* must be 0. If *Face* is invalid, then [CUDA_ERROR_INVALID_VALUE](#) is returned.

If *pResource* is of type `IDirect3DBaseTexture9`, then *Level* must correspond to a valid mipmap level. At present only mipmap level 0 is supported. For all other types *Level* must be 0. If *Level* is invalid, then [CUDA_ERROR_INVALID_VALUE](#) is returned.

Parameters:

pDevPtr - Returned pointer corresponding to subresource

pResource - Mapped resource to access

Face - Face of resource to access

Level - Level of resource to access

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceGetMappedPointer](#)

5.54.4.6 CUresult cuD3D9ResourceGetMappedSize (size_t * *pSize*, IDirect3DResource9 * *pResource*, unsigned int *Face*, unsigned int *Level*)

Deprecated

This function is deprecated as of Cuda 3.0.

Returns in *pSize* the size of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *Face* and *Level*. The value set in *pSize* may change every time that *pResource* is mapped.

If *pResource* has not been registered for use with CUDA, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* was not registered with usage flags `CU_D3D9_REGISTER_FLAGS_NONE`, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* is not mapped for access by CUDA, then [CUDA_ERROR_NOT_MAPPED](#) is returned.

For usage requirements of *Face* and *Level* parameters, see [cuD3D9ResourceGetMappedPointer](#).

Parameters:

- pSize* - Returned size of subresource
- pResource* - Mapped resource to access
- Face* - Face of resource to access
- Level* - Level of resource to access

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceGetMappedPointer](#)

5.54.4.7 CUresult cuD3D9ResourceGetSurfaceDimensions (size_t * *pWidth*, size_t * *pHeight*, size_t * *pDepth*, IDirect3DResource9 * *pResource*, unsigned int *Face*, unsigned int *Level*)

Deprecated

This function is deprecated as of Cuda 3.0.

Returns in *pWidth*, *pHeight*, and *pDepth* the dimensions of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *Face* and *Level*.

Because anti-aliased surfaces may have multiple samples per pixel, it is possible that the dimensions of a resource will be an integer factor larger than the dimensions reported by the Direct3D runtime.

The parameters *pWidth*, *pHeight*, and *pDepth* are optional. For 2D surfaces, the value returned in *pDepth* will be 0.

If *pResource* is not of type `IDirect3DBaseTexture9` or `IDirect3DSurface9` or if *pResource* has not been registered for use with CUDA, then [CUDA_ERROR_INVALID_HANDLE](#) is returned.

For usage requirements of *Face* and *Level* parameters, see [cuD3D9ResourceGetMappedPointer\(\)](#).

Parameters:

- pWidth* - Returned width of surface
- pHeight* - Returned height of surface
- pDepth* - Returned depth of surface

pResource - Registered resource to access

Face - Face of resource to access

Level - Level of resource to access

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsSubResourceGetMappedArray](#)

5.54.4.8 CUresult cuD3D9ResourceSetMapFlags (IDirect3DResource9 *pResource, unsigned int Flags)

Deprecated

This function is deprecated as of Cuda 3.0.

Set `Flags` for mapping the Direct3D resource `pResource`.

Changes to `Flags` will take effect the next time `pResource` is mapped. The `Flags` argument may be any of the following:

- `CU_D3D9_MAPRESOURCE_FLAGS_NONE`: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA kernels. This is the default value.
- `CU_D3D9_MAPRESOURCE_FLAGS_READONLY`: Specifies that CUDA kernels which access this resource will not write to this resource.
- `CU_D3D9_MAPRESOURCE_FLAGS_WRITEDISCARD`: Specifies that CUDA kernels which access this resource will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

If `pResource` has not been registered for use with CUDA, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If `pResource` is presently mapped for access by CUDA, then [CUDA_ERROR_ALREADY_MAPPED](#) is returned.

Parameters:

pResource - Registered resource to set flags for

Flags - Parameters for resource mapping

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_ALREADY_MAPPED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceSetMapFlags](#)

5.54.4.9 CUresult cuD3D9UnmapResources (unsigned int *count*, IDirect3DResource9 ** *ppResource*)**Deprecated**

This function is deprecated as of Cuda 3.0.

Unmaps the *count* Direct3D resources in *ppResource*.

This function provides the synchronization guarantee that any CUDA kernels issued before [cuD3D9UnmapResources\(\)](#) will complete before any Direct3D calls issued after [cuD3D9UnmapResources\(\)](#) begin.

If any of *ppResource* have not been registered for use with CUDA or if *ppResource* contains any duplicate entries, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If any of *ppResource* are not presently mapped for access by CUDA, then [CUDA_ERROR_NOT_MAPPED](#) is returned.

Parameters:

count - Number of resources to unmap for CUDA

ppResource - Resources to unmap for CUDA

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnmapResources](#)

5.54.4.10 CUresult cuD3D9UnregisterResource (IDirect3DResource9 * *pResource*)**Deprecated**

This function is deprecated as of Cuda 3.0.

Unregisters the Direct3D resource *pResource* so it is not accessible by CUDA unless registered again.

If *pResource* is not registered, then [CUDA_ERROR_INVALID_HANDLE](#) is returned.

Parameters:

pResource - Resource to unregister

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnregisterResource](#)

5.55 Direct3D 10 Interoperability

Modules

- [Direct3D 10 Interoperability \[DEPRECATED\]](#)

Typedefs

- typedef enum [CUd3d10DeviceList_enum](#) [CUd3d10DeviceList](#)

Enumerations

- enum [CUd3d10DeviceList_enum](#) {
[CU_D3D10_DEVICE_LIST_ALL](#) = 0x01,
[CU_D3D10_DEVICE_LIST_CURRENT_FRAME](#) = 0x02,
[CU_D3D10_DEVICE_LIST_NEXT_FRAME](#) = 0x03 }

Functions

- [CUresult cuD3D10GetDevice](#) ([CUdevice](#) *pCudaDevice, [IDXGIAdapter](#) *pAdapter)
Gets the CUDA device corresponding to a display adapter.
- [CUresult cuD3D10GetDevices](#) (unsigned int *pCudaDeviceCount, [CUdevice](#) *pCudaDevices, unsigned int cudaDeviceCount, [ID3D10Device](#) *pD3D10Device, [CUd3d10DeviceList](#) deviceList)
Gets the CUDA devices corresponding to a Direct3D 10 device.
- [CUresult cuGraphicsD3D10RegisterResource](#) ([CUgraphicsResource](#) *pCudaResource, [ID3D10Resource](#) *pD3DResource, unsigned int Flags)
Register a Direct3D 10 resource for access by CUDA.

5.55.1 Detailed Description

This section describes the Direct3D 10 interoperability functions of the low-level CUDA driver application programming interface. Note that mapping of Direct3D 10 resources is performed with the graphics API agnostic, resource mapping interface described in [Graphics Interoperability](#).

5.55.2 Typedef Documentation

5.55.2.1 typedef enum CUd3d10DeviceList_enum CUd3d10DeviceList

CUDA devices corresponding to a D3D10 device

5.55.3 Enumeration Type Documentation

5.55.3.1 enum CUd3d10DeviceList_enum

CUDA devices corresponding to a D3D10 device

Enumerator:

CUDA_D3D10_DEVICE_LIST_ALL The CUDA devices for all GPUs used by a D3D10 device

CUDA_D3D10_DEVICE_LIST_CURRENT_FRAME The CUDA devices for the GPUs used by a D3D10 device in its currently rendering frame

CUDA_D3D10_DEVICE_LIST_NEXT_FRAME The CUDA devices for the GPUs to be used by a D3D10 device in the next frame

5.55.4 Function Documentation**5.55.4.1 CUresult cuD3D10GetDevice (CUdevice * pCudaDevice, IDXGIAdapter * pAdapter)**

Returns in *pCudaDevice the CUDA-compatible device corresponding to the adapter pAdapter obtained from IDXGIFactory::EnumAdapters.

If no device on pAdapter is CUDA-compatible then the call will fail.

Parameters:

pCudaDevice - Returned CUDA device corresponding to pAdapter

pAdapter - Adapter to query for CUDA device

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_FOUND](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D10GetDevices](#)

5.55.4.2 CUresult cuD3D10GetDevices (unsigned int * pCudaDeviceCount, CUdevice * pCudaDevices, unsigned int cudaDeviceCount, IDXGIAdapter * pD3D10Device, CUd3d10DeviceList deviceList)

Returns in *pCudaDeviceCount the number of CUDA-compatible device corresponding to the Direct3D 10 device pD3D10Device. Also returns in *pCudaDevices at most cudaDeviceCount of the the CUDA-compatible devices corresponding to the Direct3D 10 device pD3D10Device.

If any of the GPUs being used to render pDevice are not CUDA capable then the call will return [CUDA_ERROR_NO_DEVICE](#).

Parameters:

pCudaDeviceCount - Returned number of CUDA devices corresponding to pD3D10Device

pCudaDevices - Returned CUDA devices corresponding to pD3D10Device

cudaDeviceCount - The size of the output device array pCudaDevices

pD3D10Device - Direct3D 10 device to query for CUDA devices

deviceList - The set of devices to return. This set may be [CUDA_D3D10_DEVICE_LIST_ALL](#) for all devices, [CUDA_D3D10_DEVICE_LIST_CURRENT_FRAME](#) for the devices used to render the current frame (in SLI), or [CUDA_D3D10_DEVICE_LIST_NEXT_FRAME](#) for the devices used to render the next frame (in SLI).

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_NO_DEVICE, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_NOT_FOUND, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D10GetDevice](#)

5.55.4.3 CUresult cuGraphicsD3D10RegisterResource (CUgraphicsResource * pCudaResource, ID3D10Resource * pD3DResource, unsigned int Flags)

Registers the Direct3D 10 resource `pD3DResource` for access by CUDA and returns a CUDA handle to `pD3DResource` in `pCudaResource`. The handle returned in `pCudaResource` may be used to map and unmap this resource until it is unregistered. On success this call will increase the internal reference count on `pD3DResource`. This reference count will be decremented when this resource is unregistered through [cuGraphicsUnregisterResource\(\)](#).

This call is potentially high-overhead and should not be called every frame in interactive applications.

The type of `pD3DResource` must be one of the following.

- ID3D10Buffer: may be accessed through a device pointer.
- ID3D10Texture1D: individual subresources of the texture may be accessed via arrays
- ID3D10Texture2D: individual subresources of the texture may be accessed via arrays
- ID3D10Texture3D: individual subresources of the texture may be accessed via arrays

The `Flags` argument may be used to specify additional parameters at register time. The valid values for this parameter are

- CU_GRAPHICS_REGISTER_FLAGS_NONE: Specifies no hints about how this resource will be used.
- CU_GRAPHICS_REGISTER_FLAGS_SURFACE_LDST: Specifies that CUDA will bind this resource to a surface reference.
- CU_GRAPHICS_REGISTER_FLAGS_TEXTURE_GATHER: Specifies that CUDA will perform texture gather operations on this resource.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

If `pD3DResource` is of incorrect type or is already registered then `CUDA_ERROR_INVALID_HANDLE` is returned. If `pD3DResource` cannot be registered then `CUDA_ERROR_UNKNOWN` is returned. If `Flags` is not one of the above specified value then `CUDA_ERROR_INVALID_VALUE` is returned.

Parameters:

pCudaResource - Returned graphics resource handle

pD3DResource - Direct3D resource to register

Flags - Parameters for resource registration

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`, `CUDA_ERROR_INVALID_HANDLE`, `CUDA_ERROR_OUT_OF_MEMORY`, `CUDA_ERROR_UNKNOWN`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

`cuGraphicsUnregisterResource`, `cuGraphicsMapResources`, `cuGraphicsSubResourceGetMappedArray`, `cuGraphicsResourceGetMappedPointer`

5.56 Direct3D 10 Interoperability [DEPRECATED]

Typedefs

- typedef enum [CUDA3D10map_flags_enum](#) [CUDA3D10map_flags](#)
- typedef enum [CUDA3D10register_flags_enum](#) [CUDA3D10register_flags](#)

Enumerations

- enum [CUDA3D10map_flags_enum](#)
- enum [CUDA3D10register_flags_enum](#)

Functions

- [CUresult cuD3D10CtxCreate](#) ([CUcontext](#) *pCtx, [CUdevice](#) *pCudaDevice, unsigned int Flags, [ID3D10Device](#) *pD3DDevice)

Create a CUDA context for interoperability with Direct3D 10.
- [CUresult cuD3D10CtxCreateOnDevice](#) ([CUcontext](#) *pCtx, unsigned int flags, [ID3D10Device](#) *pD3DDevice, [CUdevice](#) cudaDevice)

Create a CUDA context for interoperability with Direct3D 10.
- [CUresult cuD3D10GetDirect3DDevice](#) ([ID3D10Device](#) **ppD3DDevice)

Get the Direct3D 10 device against which the current CUDA context was created.
- [CUresult cuD3D10MapResources](#) (unsigned int count, [ID3D10Resource](#) **ppResources)

Map Direct3D resources for access by CUDA.
- [CUresult cuD3D10RegisterResource](#) ([ID3D10Resource](#) *pResource, unsigned int Flags)

Register a Direct3D resource for access by CUDA.
- [CUresult cuD3D10ResourceGetMappedArray](#) ([CUarray](#) *pArray, [ID3D10Resource](#) *pResource, unsigned int SubResource)

Get an array through which to access a subresource of a Direct3D resource which has been mapped for access by CUDA.
- [CUresult cuD3D10ResourceGetMappedPitch](#) (size_t *pPitch, size_t *pPitchSlice, [ID3D10Resource](#) *pResource, unsigned int SubResource)

Get the pitch of a subresource of a Direct3D resource which has been mapped for access by CUDA.
- [CUresult cuD3D10ResourceGetMappedPointer](#) ([CUdeviceptr](#) *pDevPtr, [ID3D10Resource](#) *pResource, unsigned int SubResource)

Get a pointer through which to access a subresource of a Direct3D resource which has been mapped for access by CUDA.
- [CUresult cuD3D10ResourceGetMappedSize](#) (size_t *pSize, [ID3D10Resource](#) *pResource, unsigned int SubResource)

Get the size of a subresource of a Direct3D resource which has been mapped for access by CUDA.
- [CUresult cuD3D10ResourceGetSurfaceDimensions](#) (size_t *pWidth, size_t *pHeight, size_t *pDepth, [ID3D10Resource](#) *pResource, unsigned int SubResource)

Get the dimensions of a registered surface.

- **CUresult cuD3D10ResourceSetMapFlags** (ID3D10Resource *pResource, unsigned int Flags)
Set usage flags for mapping a Direct3D resource.
- **CUresult cuD3D10UnmapResources** (unsigned int count, ID3D10Resource **ppResources)
Unmap Direct3D resources.
- **CUresult cuD3D10UnregisterResource** (ID3D10Resource *pResource)
Unregister a Direct3D resource.

5.56.1 Detailed Description

This section describes deprecated Direct3D 10 interoperability functionality.

5.56.2 Typedef Documentation

5.56.2.1 typedef enum CUD3D10map_flags_enum CUD3D10map_flags

Flags to map or unmap a resource

5.56.2.2 typedef enum CUD3D10register_flags_enum CUD3D10register_flags

Flags to register a resource

5.56.3 Enumeration Type Documentation

5.56.3.1 enum CUD3D10map_flags_enum

Flags to map or unmap a resource

5.56.3.2 enum CUD3D10register_flags_enum

Flags to register a resource

5.56.4 Function Documentation

5.56.4.1 CUresult cuD3D10CtxCreate (CUcontext *pCtx, CUdevice *pCudaDevice, unsigned int Flags, ID3D10Device *pD3DDevice)

Deprecated

This function is deprecated as of Cuda 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA context with a D3D10 device in order to achieve maximum interoperability performance.

Parameters:

pCtx - Returned newly created CUDA context
pCudaDevice - Returned pointer to the device on which the context was created
Flags - Context creation flags (see [cuCtxCreate\(\)](#) for details)
pD3DDevice - Direct3D device to create interoperability context with

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D10GetDevice](#), [cuGraphicsD3D10RegisterResource](#)

5.56.4.2 CUresult cuD3D10CtxCreateOnDevice (CUcontext * *pCtx*, unsigned int *flags*, ID3D10Device * *pD3DDevice*, CUdevice *cudaDevice*)

Deprecated

This function is deprecated as of Cuda 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA context with a D3D10 device in order to achieve maximum interoperability performance.

Parameters:

pCtx - Returned newly created CUDA context
flags - Context creation flags (see [cuCtxCreate\(\)](#) for details)
pD3DDevice - Direct3D device to create interoperability context with
cudaDevice - The CUDA device on which to create the context. This device must be among the devices returned when querying CU_D3D10_DEVICES_ALL from [cuD3D10GetDevices](#).

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D10GetDevices](#), [cuGraphicsD3D10RegisterResource](#)

5.56.4.3 CUresult cuD3D10GetDirect3DDevice (ID3D10Device ** ppD3DDevice)

Deprecated

This function is deprecated as of Cuda 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA context with a D3D10 device in order to achieve maximum interoperability performance.

Parameters:

ppD3DDevice - Returned Direct3D device corresponding to CUDA context

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D10GetDevice](#)

5.56.4.4 CUresult cuD3D10MapResources (unsigned int count, ID3D10Resource ** ppResources)

Deprecated

This function is deprecated as of Cuda 3.0.

Maps the *count* Direct3D resources in *ppResources* for access by CUDA.

The resources in *ppResources* may be accessed in CUDA kernels until they are unmapped. Direct3D should not access any resources while they are mapped by CUDA. If an application does so, the results are undefined.

This function provides the synchronization guarantee that any Direct3D calls issued before [cuD3D10MapResources\(\)](#) will complete before any CUDA kernels issued after [cuD3D10MapResources\(\)](#) begin.

If any of *ppResources* have not been registered for use with CUDA or if *ppResources* contains any duplicate entries, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If any of *ppResources* are presently mapped for access by CUDA, then [CUDA_ERROR_ALREADY_MAPPED](#) is returned.

Parameters:

count - Number of resources to map for CUDA

ppResources - Resources to map for CUDA

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_ALREADY_MAPPED, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsMapResources](#)

5.56.4.5 CUresult cuD3D10RegisterResource (ID3D10Resource *pResource, unsigned int Flags)

Deprecated

This function is deprecated as of Cuda 3.0.

Registers the Direct3D resource `pResource` for access by CUDA.

If this call is successful, then the application will be able to map and unmap this resource until it is unregistered through [cuD3D10UnregisterResource\(\)](#). Also on success, this call will increase the internal reference count on `pResource`. This reference count will be decremented when this resource is unregistered through [cuD3D10UnregisterResource\(\)](#).

This call is potentially high-overhead and should not be called every frame in interactive applications.

The type of `pResource` must be one of the following.

- ID3D10Buffer: Cannot be used with `Flags` set to `CU_D3D10_REGISTER_FLAGS_ARRAY`.
- ID3D10Texture1D: No restrictions.
- ID3D10Texture2D: No restrictions.
- ID3D10Texture3D: No restrictions.

The `Flags` argument specifies the mechanism through which CUDA will access the Direct3D resource. The following values are allowed.

- `CU_D3D10_REGISTER_FLAGS_NONE`: Specifies that CUDA will access this resource through a [CUdeviceptr](#). The pointer, size, and (for textures), pitch for each subresource of this allocation may be queried through [cuD3D10ResourceGetMappedPointer\(\)](#), [cuD3D10ResourceGetMappedSize\(\)](#), and [cuD3D10ResourceGetMappedPitch\(\)](#) respectively. This option is valid for all resource types.
- `CU_D3D10_REGISTER_FLAGS_ARRAY`: Specifies that CUDA will access this resource through a [CUarray](#) queried on a sub-resource basis through [cuD3D10ResourceGetMappedArray\(\)](#). This option is only valid for resources of type `ID3D10Texture1D`, `ID3D10Texture2D`, and `ID3D10Texture3D`.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.
- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

If Direct3D interoperability is not initialized on this context then [CUDA_ERROR_INVALID_CONTEXT](#) is returned. If `pResource` is of incorrect type or is already registered, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If `pResource` cannot be registered, then [CUDA_ERROR_UNKNOWN](#) is returned.

Parameters:

pResource - Resource to register
Flags - Parameters for resource registration

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_OUT_OF_MEMORY, CUDA_ERROR_UNKNOWN

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsD3D10RegisterResource](#)

5.56.4.6 CUresult cuD3D10ResourceGetMappedArray (CUarray *pArray, ID3D10Resource *pResource, unsigned int SubResource)**Deprecated**

This function is deprecated as of Cuda 3.0.

Returns in *pArray an array through which the subresource of the mapped Direct3D resource pResource, which corresponds to SubResource may be accessed. The value set in pArray may change every time that pResource is mapped.

If pResource is not registered, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If pResource was not registered with usage flags CU_D3D10_REGISTER_FLAGS_ARRAY, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If pResource is not mapped, then [CUDA_ERROR_NOT_MAPPED](#) is returned.

For usage requirements of the SubResource parameter, see [cuD3D10ResourceGetMappedPointer\(\)](#).

Parameters:

pArray - Returned array corresponding to subresource
pResource - Mapped resource to access
SubResource - Subresource of pResource to access

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_NOT_MAPPED

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsSubResourceGetMappedArray](#)

5.56.4.7 CUresult cuD3D10ResourceGetMappedPitch (size_t * *pPitch*, size_t * *pPitchSlice*, ID3D10Resource * *pResource*, unsigned int *SubResource*)

Deprecated

This function is deprecated as of Cuda 3.0.

Returns in *pPitch* and *pPitchSlice* the pitch and Z-slice pitch of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *SubResource*. The values set in *pPitch* and *pPitchSlice* may change every time that *pResource* is mapped.

The pitch and Z-slice pitch values may be used to compute the location of a sample on a surface as follows.

For a 2D surface, the byte offset of the sample at position *x*, *y* from the base pointer of the surface is:

$$y * \text{pitch} + (\text{bytes per pixel}) * x$$

For a 3D surface, the byte offset of the sample at position *x*, *y*, *z* from the base pointer of the surface is:

$$z * \text{slicePitch} + y * \text{pitch} + (\text{bytes per pixel}) * x$$

Both parameters *pPitch* and *pPitchSlice* are optional and may be set to NULL.

If *pResource* is not of type IDirect3DBaseTexture10 or one of its sub-types or if *pResource* has not been registered for use with CUDA, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* was not registered with usage flags CU_D3D10_REGISTER_FLAGS_NONE, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* is not mapped for access by CUDA, then [CUDA_ERROR_NOT_MAPPED](#) is returned.

For usage requirements of the *SubResource* parameter, see [cuD3D10ResourceGetMappedPointer\(\)](#).

Parameters:

pPitch - Returned pitch of subresource

pPitchSlice - Returned Z-slice pitch of subresource

pResource - Mapped resource to access

SubResource - Subresource of *pResource* to access

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsSubResourceGetMappedArray](#)

5.56.4.8 CUresult cuD3D10ResourceGetMappedPointer (CUdeviceptr * *pDevPtr*, ID3D10Resource * *pResource*, unsigned int *SubResource*)

Deprecated

This function is deprecated as of Cuda 3.0.

Returns in *pDevPtr* the base pointer of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *SubResource*. The value set in *pDevPtr* may change every time that *pResource* is mapped.

If *pResource* is not registered, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* was not registered with usage flags `CU_D3D10_REGISTER_FLAGS_NONE`, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* is not mapped, then [CUDA_ERROR_NOT_MAPPED](#) is returned.

If *pResource* is of type `ID3D10Buffer`, then *SubResource* must be 0. If *pResource* is of any other type, then the value of *SubResource* must come from the subresource calculation in `D3D10CalcSubResource()`.

Parameters:

pDevPtr - Returned pointer corresponding to subresource

pResource - Mapped resource to access

SubResource - Subresource of *pResource* to access

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceGetMappedPointer](#)

5.56.4.9 CUresult cuD3D10ResourceGetMappedSize (size_t *pSize, ID3D10Resource *pResource, unsigned int SubResource)

Deprecated

This function is deprecated as of Cuda 3.0.

Returns in *pSize* the size of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *SubResource*. The value set in *pSize* may change every time that *pResource* is mapped.

If *pResource* has not been registered for use with CUDA, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* was not registered with usage flags `CU_D3D10_REGISTER_FLAGS_NONE`, then [CUDA_ERROR_INVALID_HANDLE](#) is returned. If *pResource* is not mapped for access by CUDA, then [CUDA_ERROR_NOT_MAPPED](#) is returned.

For usage requirements of the *SubResource* parameter, see [cuD3D10ResourceGetMappedPointer\(\)](#).

Parameters:

pSize - Returned size of subresource

pResource - Mapped resource to access

SubResource - Subresource of *pResource* to access

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceGetMappedPointer](#)

5.56.4.10 `CUresult cuD3D10ResourceGetSurfaceDimensions (size_t * pWidth, size_t * pHeight, size_t * pDepth, ID3D10Resource * pResource, unsigned int SubResource)`**Deprecated**

This function is deprecated as of Cuda 3.0.

Returns in *pWidth*, *pHeight*, and *pDepth* the dimensions of the subresource of the mapped Direct3D resource *pResource*, which corresponds to *SubResource*.

Because anti-aliased surfaces may have multiple samples per pixel, it is possible that the dimensions of a resource will be an integer factor larger than the dimensions reported by the Direct3D runtime.

The parameters *pWidth*, *pHeight*, and *pDepth* are optional. For 2D surfaces, the value returned in *pDepth* will be 0.

If *pResource* is not of type `IDirect3DBaseTexture10` or `IDirect3DSurface10` or if *pResource* has not been registered for use with CUDA, then `CUDA_ERROR_INVALID_HANDLE` is returned.

For usage requirements of the *SubResource* parameter, see [cuD3D10ResourceGetMappedPointer\(\)](#).

Parameters:

- pWidth* - Returned width of surface
- pHeight* - Returned height of surface
- pDepth* - Returned depth of surface
- pResource* - Registered resource to access
- SubResource* - Subresource of *pResource* to access

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`, `CUDA_ERROR_INVALID_HANDLE`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsSubResourceGetMappedArray](#)

5.56.4.11 `CUresult cuD3D10ResourceSetMapFlags (ID3D10Resource * pResource, unsigned int Flags)`**Deprecated**

This function is deprecated as of Cuda 3.0.

Set flags for mapping the Direct3D resource `pResource`.

Changes to flags will take effect the next time `pResource` is mapped. The `Flags` argument may be any of the following.

- `CUDA_D3D10_MAPRESOURCE_FLAGS_NONE`: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA kernels. This is the default value.
- `CUDA_D3D10_MAPRESOURCE_FLAGS_READONLY`: Specifies that CUDA kernels which access this resource will not write to this resource.
- `CUDA_D3D10_MAPRESOURCE_FLAGS_WRITEDISCARD`: Specifies that CUDA kernels which access this resource will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

If `pResource` has not been registered for use with CUDA, then `CUDA_ERROR_INVALID_HANDLE` is returned. If `pResource` is presently mapped for access by CUDA then `CUDA_ERROR_ALREADY_MAPPED` is returned.

Parameters:

pResource - Registered resource to set flags for

Flags - Parameters for resource mapping

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`, `CUDA_ERROR_INVALID_HANDLE`, `CUDA_ERROR_ALREADY_MAPPED`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsResourceSetMapFlags](#)

5.56.4.12 `CUresult cuD3D10UnmapResources (unsigned int count, ID3D10Resource ** ppResources)`

Deprecated

This function is deprecated as of Cuda 3.0.

Unmaps the `count` Direct3D resources in `ppResources`.

This function provides the synchronization guarantee that any CUDA kernels issued before `cuD3D10UnmapResources()` will complete before any Direct3D calls issued after `cuD3D10UnmapResources()` begin.

If any of `ppResources` have not been registered for use with CUDA or if `ppResources` contains any duplicate entries, then `CUDA_ERROR_INVALID_HANDLE` is returned. If any of `ppResources` are not presently mapped for access by CUDA, then `CUDA_ERROR_NOT_MAPPED` is returned.

Parameters:

count - Number of resources to unmap for CUDA

ppResources - Resources to unmap for CUDA

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_NOT_MAPPED](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnmapResources](#)

5.56.4.13 CUresult cuD3D10UnregisterResource (ID3D10Resource * *pResource*)

Deprecated

This function is deprecated as of Cuda 3.0.

Unregisters the Direct3D resource *pResource* so it is not accessible by CUDA unless registered again.

If *pResource* is not registered, then [CUDA_ERROR_INVALID_HANDLE](#) is returned.

Parameters:

pResource - Resources to unregister

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_HANDLE](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnregisterResource](#)

5.57 Direct3D 11 Interoperability

Modules

- [Direct3D 11 Interoperability \[DEPRECATED\]](#)

Typedefs

- typedef enum [CUd3d11DeviceList_enum](#) [CUd3d11DeviceList](#)

Enumerations

- enum [CUd3d11DeviceList_enum](#) {
[CU_D3D11_DEVICE_LIST_ALL](#) = 0x01,
[CU_D3D11_DEVICE_LIST_CURRENT_FRAME](#) = 0x02,
[CU_D3D11_DEVICE_LIST_NEXT_FRAME](#) = 0x03 }

Functions

- [CUresult cuD3D11GetDevice](#) ([CUdevice](#) *pCudaDevice, [IDXGIAdapter](#) *pAdapter)
Gets the CUDA device corresponding to a display adapter.
- [CUresult cuD3D11GetDevices](#) (unsigned int *pCudaDeviceCount, [CUdevice](#) *pCudaDevices, unsigned int cudaDeviceCount, [ID3D11Device](#) *pD3D11Device, [CUd3d11DeviceList](#) deviceList)
Gets the CUDA devices corresponding to a Direct3D 11 device.
- [CUresult cuGraphicsD3D11RegisterResource](#) ([CUgraphicsResource](#) *pCudaResource, [ID3D11Resource](#) *pD3DResource, unsigned int Flags)
Register a Direct3D 11 resource for access by CUDA.

5.57.1 Detailed Description

This section describes the Direct3D 11 interoperability functions of the low-level CUDA driver application programming interface. Note that mapping of Direct3D 11 resources is performed with the graphics API agnostic, resource mapping interface described in [Graphics Interoperability](#).

5.57.2 Typedef Documentation

5.57.2.1 typedef enum [CUd3d11DeviceList_enum](#) [CUd3d11DeviceList](#)

CUDA devices corresponding to a D3D11 device

5.57.3 Enumeration Type Documentation

5.57.3.1 enum [CUd3d11DeviceList_enum](#)

CUDA devices corresponding to a D3D11 device

Enumerator:

CUDA_D3D11_DEVICE_LIST_ALL The CUDA devices for all GPUs used by a D3D11 device

CUDA_D3D11_DEVICE_LIST_CURRENT_FRAME The CUDA devices for the GPUs used by a D3D11 device in its currently rendering frame

CUDA_D3D11_DEVICE_LIST_NEXT_FRAME The CUDA devices for the GPUs to be used by a D3D11 device in the next frame

5.57.4 Function Documentation**5.57.4.1 CUresult cuD3D11GetDevice (CUdevice * pCudaDevice, IDXGIAdapter * pAdapter)**

Returns in *pCudaDevice* the CUDA-compatible device corresponding to the adapter *pAdapter* obtained from `IDXGIFactory::EnumAdapters`.

If no device on *pAdapter* is CUDA-compatible the call will return [CUDA_ERROR_NO_DEVICE](#).

Parameters:

pCudaDevice - Returned CUDA device corresponding to *pAdapter*

pAdapter - Adapter to query for CUDA device

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_NO_DEVICE](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_FOUND](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D11GetDevices](#)

5.57.4.2 CUresult cuD3D11GetDevices (unsigned int * pCudaDeviceCount, CUdevice * pCudaDevices, unsigned int cudaDeviceCount, ID3D11Device * pD3D11Device, CUd3d11DeviceList deviceList)

Returns in *pCudaDeviceCount* the number of CUDA-compatible device corresponding to the Direct3D 11 device *pD3D11Device*. Also returns in *pCudaDevices* at most *cudaDeviceCount* of the the CUDA-compatible devices corresponding to the Direct3D 11 device *pD3D11Device*.

If any of the GPUs being used to render *pDevice* are not CUDA capable then the call will return [CUDA_ERROR_NO_DEVICE](#).

Parameters:

pCudaDeviceCount - Returned number of CUDA devices corresponding to *pD3D11Device*

pCudaDevices - Returned CUDA devices corresponding to *pD3D11Device*

cudaDeviceCount - The size of the output device array *pCudaDevices*

pD3D11Device - Direct3D 11 device to query for CUDA devices

deviceList - The set of devices to return. This set may be [CU_D3D11_DEVICE_LIST_ALL](#) for all devices, [CU_D3D11_DEVICE_LIST_CURRENT_FRAME](#) for the devices used to render the current frame (in SLI), or [CU_D3D11_DEVICE_LIST_NEXT_FRAME](#) for the devices used to render the next frame (in SLI).

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_NO_DEVICE](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_NOT_FOUND](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D11GetDevice](#)

5.57.4.3 CUresult cuGraphicsD3D11RegisterResource (CUgraphicsResource * pCudaResource, ID3D11Resource * pD3DResource, unsigned int Flags)

Registers the Direct3D 11 resource `pD3DResource` for access by CUDA and returns a CUDA handle to `pD3DResource` in `pCudaResource`. The handle returned in `pCudaResource` may be used to map and unmap this resource until it is unregistered. On success this call will increase the internal reference count on `pD3DResource`. This reference count will be decremented when this resource is unregistered through [cuGraphicsUnregisterResource\(\)](#).

This call is potentially high-overhead and should not be called every frame in interactive applications.

The type of `pD3DResource` must be one of the following.

- `ID3D11Buffer`: may be accessed through a device pointer.
- `ID3D11Texture1D`: individual subresources of the texture may be accessed via arrays
- `ID3D11Texture2D`: individual subresources of the texture may be accessed via arrays
- `ID3D11Texture3D`: individual subresources of the texture may be accessed via arrays

The `Flags` argument may be used to specify additional parameters at register time. The valid values for this parameter are

- `CU_GRAPHICS_REGISTER_FLAGS_NONE`: Specifies no hints about how this resource will be used.
- `CU_GRAPHICS_REGISTER_FLAGS_SURFACE_LDST`: Specifies that CUDA will bind this resource to a surface reference.
- `CU_GRAPHICS_REGISTER_FLAGS_TEXTURE_GATHER`: Specifies that CUDA will perform texture gather operations on this resource.

Not all Direct3D resources of the above types may be used for interoperability with CUDA. The following are some limitations.

- The primary rendertarget may not be registered with CUDA.
- Resources allocated as shared may not be registered with CUDA.

- Textures which are not of a format which is 1, 2, or 4 channels of 8, 16, or 32-bit integer or floating-point data cannot be shared.
- Surfaces of depth or stencil formats cannot be shared.

If `pD3DResource` is of incorrect type or is already registered then `CUDA_ERROR_INVALID_HANDLE` is returned. If `pD3DResource` cannot be registered then `CUDA_ERROR_UNKNOWN` is returned. If `Flags` is not one of the above specified value then `CUDA_ERROR_INVALID_VALUE` is returned.

Parameters:

pCudaResource - Returned graphics resource handle

pD3DResource - Direct3D resource to register

Flags - Parameters for resource registration

Returns:

`CUDA_SUCCESS`, `CUDA_ERROR_DEINITIALIZED`, `CUDA_ERROR_NOT_INITIALIZED`, `CUDA_ERROR_INVALID_CONTEXT`, `CUDA_ERROR_INVALID_VALUE`, `CUDA_ERROR_INVALID_HANDLE`, `CUDA_ERROR_OUT_OF_MEMORY`, `CUDA_ERROR_UNKNOWN`

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuGraphicsUnregisterResource](#), [cuGraphicsMapResources](#), [cuGraphicsSubResourceGetMappedArray](#), [cuGraphicsResourceGetMappedPointer](#)

5.58 Direct3D 11 Interoperability [DEPRECATED]

Functions

- [CUresult cuD3D11CtxCreate](#) (CUcontext *pCtx, CUdevice *pCudaDevice, unsigned int Flags, ID3D11Device *pD3DDevice)
Create a CUDA context for interoperability with Direct3D 11.
- [CUresult cuD3D11CtxCreateOnDevice](#) (CUcontext *pCtx, unsigned int flags, ID3D11Device *pD3DDevice, CUdevice cudaDevice)
Create a CUDA context for interoperability with Direct3D 11.
- [CUresult cuD3D11GetDirect3DDevice](#) (ID3D11Device **ppD3DDevice)
Get the Direct3D 11 device against which the current CUDA context was created.

5.58.1 Detailed Description

This section describes deprecated Direct3D 11 interoperability functionality.

5.58.2 Function Documentation

5.58.2.1 CUresult cuD3D11CtxCreate (CUcontext * pCtx, CUdevice * pCudaDevice, unsigned int Flags, ID3D11Device * pD3DDevice)

Deprecated

This function is deprecated as of Cuda 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA context with a D3D11 device in order to achieve maximum interoperability performance.

Parameters:

pCtx - Returned newly created CUDA context

pCudaDevice - Returned pointer to the device on which the context was created

Flags - Context creation flags (see [cuCtxCreate\(\)](#) for details)

pD3DDevice - Direct3D device to create interoperability context with

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D11GetDevice](#), [cuGraphicsD3D11RegisterResource](#)

5.58.2.2 CUresult cuD3D11CtxCreateOnDevice (CUcontext * *pCtx*, unsigned int *flags*, ID3D11Device * *pD3DDevice*, CUdevice *cudaDevice*)

Deprecated

This function is deprecated as of Cuda 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA context with a D3D11 device in order to achieve maximum interoperability performance.

Parameters:

pCtx - Returned newly created CUDA context

flags - Context creation flags (see [cuCtxCreate\(\)](#) for details)

pD3DDevice - Direct3D device to create interoperability context with

cudaDevice - The CUDA device on which to create the context. This device must be among the devices returned when querying CU_D3D11_DEVICES_ALL from [cuD3D11GetDevices](#).

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_VALUE](#), [CUDA_ERROR_OUT_OF_MEMORY](#), [CUDA_ERROR_UNKNOWN](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D11GetDevices](#), [cuGraphicsD3D11RegisterResource](#)

5.58.2.3 CUresult cuD3D11GetDirect3DDevice (ID3D11Device ** *ppD3DDevice*)

Deprecated

This function is deprecated as of Cuda 5.0.

This function is deprecated and should no longer be used. It is no longer necessary to associate a CUDA context with a D3D11 device in order to achieve maximum interoperability performance.

Parameters:

ppD3DDevice - Returned Direct3D device corresponding to CUDA context

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuD3D11GetDevice](#)

5.59 VDPAU Interoperability

Functions

- **CUresult cuGraphicsVDPAURegisterOutputSurface** (**CUgraphicsResource** *pCudaResource, VdpOutputSurface vdpSurface, unsigned int flags)
Registers a VDPAU VdpOutputSurface object.
- **CUresult cuGraphicsVDPAURegisterVideoSurface** (**CUgraphicsResource** *pCudaResource, VdpVideoSurface vdpSurface, unsigned int flags)
Registers a VDPAU VdpVideoSurface object.
- **CUresult cuVDPAUCtxCreate** (**CUcontext** *pCtx, unsigned int flags, **CUdevice** device, VdpDevice vdpDevice, VdpGetProcAddress *vdpGetProcAddress)
Create a CUDA context for interoperability with VDPAU.
- **CUresult cuVDPAUGetDevice** (**CUdevice** *pDevice, VdpDevice vdpDevice, VdpGetProcAddress *vdpGetProcAddress)
Gets the CUDA device associated with a VDPAU device.

5.59.1 Detailed Description

This section describes the VDPAU interoperability functions of the low-level CUDA driver application programming interface.

5.59.2 Function Documentation

5.59.2.1 CUresult cuGraphicsVDPAURegisterOutputSurface (CUgraphicsResource *pCudaResource, VdpOutputSurface vdpSurface, unsigned int flags)

Registers the VdpOutputSurface specified by vdpSurface for access by CUDA. A handle to the registered object is returned as pCudaResource. The surface's intended usage is specified using flags, as follows:

- **CU_GRAPHICS_MAP_RESOURCE_FLAGS_NONE**: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA. This is the default value.
- **CU_GRAPHICS_MAP_RESOURCE_FLAGS_READ_ONLY**: Specifies that CUDA will not write to this resource.
- **CU_GRAPHICS_MAP_RESOURCE_FLAGS_WRITE_DISCARD**: Specifies that CUDA will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

The VdpOutputSurface is presented as an array of subresources that may be accessed using pointers returned by [cuGraphicsSubResourceGetMappedArray](#). The exact number of valid arrayIndex values depends on the VDPAU surface format. The mapping is shown in the table below. mipLevel must be 0.

VdpRGBAFormat	arrayIndex	Size	Format	Content
VDP_RGBA_FORMAT_B8G8R8A8	0	w x h	ARGB8	Entire surface
VDP_RGBA_FORMAT_R10G10B10A2	0	w x h	A2BGR10	Entire surface

Parameters:

pCudaResource - Pointer to the returned object handle

vdpSurface - The VdpOutputSurface to be registered

flags - Map flags

Returns:

CUDA_SUCCESS, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_ALREADY_MAPPED, CUDA_ERROR_INVALID_CONTEXT,

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuVDPACtxCreate](#), [cuGraphicsVDPAURegisterVideoSurface](#), [cuGraphicsUnregisterResource](#), [cuGraphicsResourceSetMapFlags](#), [cuGraphicsMapResources](#), [cuGraphicsUnmapResources](#), [cuGraphicsSubResourceGetMappedArray](#), [cuVDPAGetDevice](#)

5.59.2.2 CUresult cuGraphicsVDPAURegisterVideoSurface (CUgraphicsResource * pCudaResource, VdpVideoSurface vdpSurface, unsigned int flags)

Registers the VdpVideoSurface specified by vdpSurface for access by CUDA. A handle to the registered object is returned as pCudaResource. The surface's intended usage is specified using flags, as follows:

- CU_GRAPHICS_MAP_RESOURCE_FLAGS_NONE: Specifies no hints about how this resource will be used. It is therefore assumed that this resource will be read from and written to by CUDA. This is the default value.
- CU_GRAPHICS_MAP_RESOURCE_FLAGS_READ_ONLY: Specifies that CUDA will not write to this resource.
- CU_GRAPHICS_MAP_RESOURCE_FLAGS_WRITE_DISCARD: Specifies that CUDA will not read from this resource and will write over the entire contents of the resource, so none of the data previously stored in the resource will be preserved.

The VdpVideoSurface is presented as an array of subresources that may be accessed using pointers returned by [cuGraphicsSubResourceGetMappedArray](#). The exact number of valid arrayIndex values depends on the VDP AU surface format. The mapping is shown in the table below. mipLevel must be 0.

VdpChromaType	arrayIndex	Size	Format	Content
VDP_CHROMA_TYPE_420	0	w x h/2	R8	Top-field luma
	1	w x h/2	R8	Bottom-field luma
	2	w/2 x h/4	R8G8	Top-field chroma
	3	w/2 x h/4	R8G8	Bottom-field chroma
VDP_CHROMA_TYPE_422	0	w x h/2	R8	Top-field luma
	1	w x h/2	R8	Bottom-field luma
	2	w/2 x h/2	R8G8	Top-field chroma
	3	w/2 x h/2	R8G8	Bottom-field chroma

Parameters:

pCudaResource - Pointer to the returned object handle

vdpSurface - The VdpVideoSurface to be registered

flags - Map flags

Returns:

CUDA_SUCCESS, CUDA_ERROR_INVALID_HANDLE, CUDA_ERROR_ALREADY_MAPPED, CUDA_ERROR_INVALID_CONTEXT,

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuCtxCreate, cuVDPAUCtxCreate, cuGraphicsVDPAURegisterOutputSurface, cuGraphicsUnregisterResource, cuGraphicsResourceSetMapFlags, cuGraphicsMapResources, cuGraphicsUnmapResources, cuGraphicsSubResourceGetMappedArray, cuVDPAUGetDevice

5.59.2.3 CUresult cuVDPAUCtxCreate (CUcontext * *pCtx*, unsigned int *flags*, CUdevice *device*, VdpDevice *vdpDevice*, VdpGetProcAddress * *vdpGetProcAddress*)

Creates a new CUDA context, initializes VDPAU interoperability, and associates the CUDA context with the calling thread. It must be called before performing any other VDPAU interoperability operations. It may fail if the needed VDPAU driver facilities are not available. For usage of the *flags* parameter, see [cuCtxCreate\(\)](#).

Parameters:

pCtx - Returned CUDA context

flags - Options for CUDA context creation

device - Device on which to create the context

vdpDevice - The VdpDevice to interop with

vdpGetProcAddress - VDPAU's VdpGetProcAddress function pointer

Returns:

CUDA_SUCCESS, CUDA_ERROR_DEINITIALIZED, CUDA_ERROR_NOT_INITIALIZED, CUDA_ERROR_INVALID_CONTEXT, CUDA_ERROR_INVALID_VALUE, CUDA_ERROR_OUT_OF_MEMORY

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

cuCtxCreate, cuGraphicsVDPAURegisterVideoSurface, cuGraphicsVDPAURegisterOutputSurface, cuGraphicsUnregisterResource, cuGraphicsResourceSetMapFlags, cuGraphicsMapResources, cuGraphicsUnmapResources, cuGraphicsSubResourceGetMappedArray, cuVDPAUGetDevice

5.59.2.4 CUresult cuVDPAUGetDevice (CUdevice * *pDevice*, VdpDevice *vdpDevice*, VdpGetProcAddress * *vdpGetProcAddress*)

Returns in **pDevice* the CUDA device associated with a *vdpDevice*, if applicable.

Parameters:

pDevice - Device associated with vdpDevice

vdpDevice - A VdpDevice handle

vdpGetProcAddress - VDPAU's VdpGetProcAddress function pointer

Returns:

[CUDA_SUCCESS](#), [CUDA_ERROR_DEINITIALIZED](#), [CUDA_ERROR_NOT_INITIALIZED](#), [CUDA_ERROR_INVALID_CONTEXT](#), [CUDA_ERROR_INVALID_VALUE](#)

Note:

Note that this function may also return error codes from previous, asynchronous launches.

See also:

[cuCtxCreate](#), [cuVDPAUCtxCreate](#), [cuGraphicsVDPAURegisterVideoSurface](#), [cuGraphicsVDPAURegisterOutputSurface](#), [cuGraphicsUnregisterResource](#), [cuGraphicsResourceSetMapFlags](#), [cuGraphicsMapResources](#), [cuGraphicsUnmapResources](#), [cuGraphicsSubResourceGetMappedArray](#)

5.60 Mathematical Functions

CUDA mathematical functions are always available in device code. Some functions are also available in host code as indicated.

Note that floating-point functions are overloaded for different argument types. For example, the `log()` function has the following prototypes:

```
double log(double x);  
float log(float x);  
float logf(float x);
```

5.61 Single Precision Mathematical Functions

Functions

- `__device__ float acosf` (float x)
Calculate the arc cosine of the input argument.
- `__device__ float acoshf` (float x)
Calculate the nonnegative arc hyperbolic cosine of the input argument.
- `__device__ float asinf` (float x)
Calculate the arc sine of the input argument.
- `__device__ float asinhf` (float x)
Calculate the arc hyperbolic sine of the input argument.
- `__device__ float atan2f` (float x, float y)
Calculate the arc tangent of the ratio of first and second input arguments.
- `__device__ float atanf` (float x)
Calculate the arc tangent of the input argument.
- `__device__ float atanhf` (float x)
Calculate the arc hyperbolic tangent of the input argument.
- `__device__ float cbrtf` (float x)
Calculate the cube root of the input argument.
- `__device__ float ceilf` (float x)
Calculate ceiling of the input argument.
- `__device__ float copysignf` (float x, float y)
Create value with given magnitude, copying sign of second value.
- `__device__ float cosf` (float x)
Calculate the cosine of the input argument.
- `__device__ float coshf` (float x)
Calculate the hyperbolic cosine of the input argument.
- `__device__ float cospif` (float x)
Calculate the cosine of the input argument $\times \pi$.
- `__device__ float erfcf` (float x)
Calculate the complementary error function of the input argument.
- `__device__ float erfcinvf` (float y)
Calculate the inverse complementary error function of the input argument.
- `__device__ float erfcxf` (float x)

Calculate the scaled complementary error function of the input argument.

- `__device__ float erff (float x)`
Calculate the error function of the input argument.
- `__device__ float erfinvf (float y)`
Calculate the inverse error function of the input argument.
- `__device__ float exp10f (float x)`
Calculate the base 10 exponential of the input argument.
- `__device__ float exp2f (float x)`
Calculate the base 2 exponential of the input argument.
- `__device__ float expf (float x)`
Calculate the base e exponential of the input argument.
- `__device__ float expm1f (float x)`
Calculate the base e exponential of the input argument, minus 1.
- `__device__ float fabsf (float x)`
Calculate the absolute value of its argument.
- `__device__ float fdimf (float x, float y)`
Compute the positive difference between x and y.
- `__device__ float fdividef (float x, float y)`
Divide two floating point values.
- `__device__ float floorf (float x)`
Calculate the largest integer less than or equal to x.
- `__device__ float fmaf (float x, float y, float z)`
Compute $x \times y + z$ as a single operation.
- `__device__ float fmaxf (float x, float y)`
Determine the maximum numeric value of the arguments.
- `__device__ float fminf (float x, float y)`
Determine the minimum numeric value of the arguments.
- `__device__ float fmodf (float x, float y)`
Calculate the floating-point remainder of x / y .
- `__device__ float frexpf (float x, int *nptr)`
Extract mantissa and exponent of a floating-point value.
- `__device__ float hypotf (float x, float y)`
Calculate the square root of the sum of squares of two arguments.

- `__device__ int ilogbf` (float x)
Compute the unbiased integer exponent of the argument.
- `__device__ int isfinite` (float a)
Determine whether argument is finite.
- `__device__ int isinf` (float a)
Determine whether argument is infinite.
- `__device__ int isnan` (float a)
Determine whether argument is a NaN.
- `__device__ float j0f` (float x)
Calculate the value of the Bessel function of the first kind of order 0 for the input argument.
- `__device__ float j1f` (float x)
Calculate the value of the Bessel function of the first kind of order 1 for the input argument.
- `__device__ float jnf` (int n, float x)
Calculate the value of the Bessel function of the first kind of order n for the input argument.
- `__device__ float ldexpf` (float x, int exp)
Calculate the value of $x \cdot 2^{exp}$.
- `__device__ float lgammaf` (float x)
Calculate the natural logarithm of the absolute value of the gamma function of the input argument.
- `__device__ long long int llrintf` (float x)
Round input to nearest integer value.
- `__device__ long long int llroundf` (float x)
Round to nearest integer value.
- `__device__ float log10f` (float x)
Calculate the base 10 logarithm of the input argument.
- `__device__ float log1pf` (float x)
Calculate the value of $\log_e(1 + x)$ $[x]$.
- `__device__ float log2f` (float x)
Calculate the base 2 logarithm of the input argument.
- `__device__ float logbf` (float x)
Calculate the floating point representation of the exponent of the input argument.
- `__device__ float logf` (float x)
Calculate the natural logarithm of the input argument.
- `__device__ long int lrintf` (float x)
Round input to nearest integer value.

- `__device__ long int lroundf` (float x)
Round to nearest integer value.
- `__device__ float modff` (float x, float *iptr)
Break down the input argument into fractional and integral parts.
- `__device__ float nanf` (const char *tagp)
Returns "Not a Number" value.
- `__device__ float nearbyintf` (float x)
Round the input argument to the nearest integer.
- `__device__ float nextafterf` (float x, float y)
Return next representable single-precision floating-point value after argument.
- `__device__ float normcdf` (float y)
Calculate the standard normal cumulative distribution function.
- `__device__ float normcdfinv` (float y)
Calculate the inverse of the standard normal cumulative distribution function.
- `__device__ float powf` (float x, float y)
Calculate the value of first argument to the power of second argument.
- `__device__ float rcbtrf` (float x)
Calculate reciprocal cube root function.
- `__device__ float remainderf` (float x, float y)
Compute single-precision floating-point remainder.
- `__device__ float remquof` (float x, float y, int *quo)
Compute single-precision floating-point remainder and part of quotient.
- `__device__ float rintf` (float x)
Round input to nearest integer value in floating-point.
- `__device__ float roundf` (float x)
Round to nearest integer value in floating-point.
- `__device__ float rsqrtf` (float x)
Calculate the reciprocal of the square root of the input argument.
- `__device__ float scalblnf` (float x, long int n)
Scale floating-point input by integer power of two.
- `__device__ float scalbnf` (float x, int n)
Scale floating-point input by integer power of two.
- `__device__ int signbit` (float a)

Return the sign bit of the input.

- `__device__ void sincosf` (float x, float *sptr, float *cptr)
Calculate the sine and cosine of the first input argument.
- `__device__ void sincospif` (float x, float *sptr, float *cptr)
Calculate the sine and cosine of the first input argument $\times \pi$.
- `__device__ float sinf` (float x)
Calculate the sine of the input argument.
- `__device__ float sinhf` (float x)
Calculate the hyperbolic sine of the input argument.
- `__device__ float sinpif` (float x)
Calculate the sine of the input argument $\times \pi$.
- `__device__ float sqrtf` (float x)
Calculate the square root of the input argument.
- `__device__ float tanf` (float x)
Calculate the tangent of the input argument.
- `__device__ float tanhf` (float x)
Calculate the hyperbolic tangent of the input argument.
- `__device__ float tgammaf` (float x)
Calculate the gamma function of the input argument.
- `__device__ float truncf` (float x)
Truncate input argument to the integral part.
- `__device__ float y0f` (float x)
Calculate the value of the Bessel function of the second kind of order 0 for the input argument.
- `__device__ float y1f` (float x)
Calculate the value of the Bessel function of the second kind of order 1 for the input argument.
- `__device__ float ynf` (int n, float x)
Calculate the value of the Bessel function of the second kind of order n for the input argument.

5.61.1 Detailed Description

This section describes single precision mathematical functions.

5.61.2 Function Documentation

5.61.2.1 `__device__ float acosf (float x)`

Calculate the principal value of the arc cosine of the input argument x .

Returns:

Result will be in radians, in the interval $[0, \pi]$ for x inside $[-1, +1]$.

- `acosf(1)` returns $+0$.
- `acosf(x)` returns NaN for x outside $[-1, +1]$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.2 `__device__ float acoshf (float x)`

Calculate the nonnegative arc hyperbolic cosine of the input argument x (measured in radians).

Returns:

Result will be in the interval $[0, +\infty]$.

- `acoshf(1)` returns 0 .
- `acoshf(x)` returns NaN for x in the interval $[-\infty, 1)$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.3 `__device__ float asinf (float x)`

Calculate the principal value of the arc sine of the input argument x .

Returns:

Result will be in radians, in the interval $[-\pi/2, +\pi/2]$ for x inside $[-1, +1]$.

- `asinf(0)` returns $+0$.
- `asinf(x)` returns NaN for x outside $[-1, +1]$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.4 `__device__ float asinhf (float x)`

Calculate the arc hyperbolic sine of the input argument x (measured in radians).

Returns:

- `asinhf(0)` returns 1 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.5 `__device__ float atan2f (float x, float y)`

Calculate the principal value of the arc tangent of the ratio of first and second input arguments x / y . The quadrant of the result is determined by the signs of inputs x and y .

Returns:

Result will be in radians, in the interval $[-\pi, +\pi]$.

- `atan2f(0, 1)` returns `+0`.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.6 `__device__ float atanf (float x)`

Calculate the principal value of the arc tangent of the input argument x .

Returns:

Result will be in radians, in the interval $[-\pi/2, +\pi/2]$.

- `atanf(0)` returns `+0`.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.7 `__device__ float atanhf (float x)`

Calculate the arc hyperbolic tangent of the input argument x (measured in radians).

Returns:

- `atanhf(±0)` returns ± 0 .
- `atanhf(±1)` returns $\pm\infty$.
- `atanhf(x)` returns NaN for x outside interval $[-1, 1]$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.8 `__device__ float cbrtf (float x)`

Calculate the cube root of x , $x^{1/3}$.

Returns:

Returns $x^{1/3}$.

- `cbrtf(±0)` returns ± 0 .
- `cbrtf(±∞)` returns $\pm\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.9 `__device__ float ceilf (float x)`

Compute the smallest integer value not less than x .

Returns:

Returns $\lceil x \rceil$ expressed as a floating-point number.

- `ceilf(±0)` returns ± 0 .
- `ceilf(±∞)` returns $\pm\infty$.

5.61.2.10 `__device__ float copysignf (float x, float y)`

Create a floating-point value with the magnitude x and the sign of y .

Returns:

Returns a value with the magnitude of x and the sign of y .

5.61.2.11 `__device__ float cosf (float x)`

Calculate the cosine of the input argument x (measured in radians).

Returns:

- `cosf(0)` returns 1.
- `cosf(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This function is affected by the `-use_fast_math` compiler flag. See the CUDA C Programming Guide, Appendix C, Table C-3 for a complete list of functions affected.

5.61.2.12 `__device__ float coshf (float x)`

Calculate the hyperbolic cosine of the input argument x (measured in radians).

Returns:

- `coshf(0)` returns 1.
- `coshf(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.13 `__device__ float cospif (float x)`

Calculate the cosine of $x \times \pi$ (measured in radians), where x is the input argument.

Returns:

- `cospif(±0)` returns 1.
- `cospif(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.14 `__device__ float erfcf (float x)`

Calculate the complementary error function of the input argument x , $1 - \text{erf}(x)$.

Returns:

- `erfcf(-∞)` returns 2.
- `erfcf(+∞)` returns +0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.15 `__device__ float erfcinvf (float y)`

Calculate the inverse complementary error function of the input argument y , for y in the interval $[0, 2]$. The inverse complementary error function find the value x that satisfies the equation $y = \text{erfc}(x)$, for $0 \leq y \leq 2$, and $-\infty \leq x \leq \infty$.

Returns:

- `erfcinvf(0)` returns $+\infty$.
- `erfcinvf(2)` returns $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.16 `__device__ float erfcxf (float x)`

Calculate the scaled complementary error function of the input argument x , $e^{x^2} \cdot \text{erfc}(x)$.

Returns:

- `erfcxf(-∞)` returns $+\infty$
- `erfcxf(+∞)` returns +0
- `erfcxf(x)` returns $+\infty$ if the correctly calculated value is outside the single floating point range.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.17 `__device__ float erff (float x)`

Calculate the value of the error function for the input argument x , $\frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

Returns:

- `erff(±0)` returns ± 0 .
- `erff(±∞)` returns ± 1 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.18 `__device__ float erfinvf (float y)`

Calculate the inverse error function of the input argument y , for y in the interval $[-1, 1]$. The inverse error function finds the value x that satisfies the equation $y = \text{erf}(x)$, for $-1 \leq y \leq 1$, and $-\infty \leq x \leq \infty$.

Returns:

- `erfinvf(1)` returns $+\infty$.
- `erfinvf(-1)` returns $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.19 `__device__ float exp10f (float x)`

Calculate the base 10 exponential of the input argument x .

Returns:

Returns 10^x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This function is affected by the `-use_fast_math` compiler flag. See the CUDA C Programming Guide, Appendix C, Table C-3 for a complete list of functions affected.

5.61.2.20 `__device__ float exp2f (float x)`

Calculate the base 2 exponential of the input argument x .

Returns:

Returns 2^x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.21 `__device__ float expf (float x)`

Calculate the base e exponential of the input argument x , e^x .

Returns:

Returns e^x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This function is affected by the `-use_fast_math` compiler flag. See the CUDA C Programming Guide, Appendix C, Table C-3 for a complete list of functions affected.

5.61.2.22 `__device__ float expm1f (float x)`

Calculate the base e exponential of the input argument x , minus 1.

Returns:

Returns $e^x - 1$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.23 `__device__ float fabsf (float x)`

Calculate the absolute value of the input argument x .

Returns:

Returns the absolute value of its argument.

- `fabs(±∞)` returns $+\infty$.
- `fabs(±0)` returns 0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.24 `__device__ float fdimf (float x, float y)`

Compute the positive difference between x and y . The positive difference is $x - y$ when $x > y$ and $+0$ otherwise.

Returns:

Returns the positive difference between x and y .

- `fdimf(x, y)` returns $x - y$ if $x > y$.
- `fdimf(x, y)` returns $+0$ if $x \leq y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.25 `__device__ float fdividef (float x, float y)`

Compute x divided by y . If `-use_fast_math` is specified, use `__fdividef()` for higher performance, otherwise use normal division.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This function is affected by the `-use_fast_math` compiler flag. See the CUDA C Programming Guide, Appendix C, Table C-3 for a complete list of functions affected.

5.61.2.26 `__device__ float floorf (float x)`

Calculate the largest integer value which is less than or equal to x .

Returns:

Returns $\log_e(1 + x) \lfloor x \rfloor$ expressed as a floating-point number.

- `floorf(±∞)` returns $±∞$.
- `floorf(±0)` returns $±0$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.27 `__device__ float fmaf (float x, float y, float z)`

Compute the value of $x \times y + z$ as a single ternary operation. After computing the value to infinite precision, the value is rounded once.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+∞$.
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-∞$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.28 `__device__ float fmaxf (float x, float y)`

Determines the maximum numeric value of the arguments x and y . Treats NaN arguments as missing data. If one argument is a NaN and the other is legitimate numeric value, the numeric value is chosen.

Returns:

Returns the maximum numeric values of the arguments x and y .

- If both arguments are NaN, returns NaN.
- If one argument is NaN, returns the numeric argument.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.29 `__device__ float fminf (float x, float y)`

Determines the minimum numeric value of the arguments x and y . Treats NaN arguments as missing data. If one argument is a NaN and the other is legitimate numeric value, the numeric value is chosen.

Returns:

Returns the minimum numeric values of the arguments x and y .

- If both arguments are NaN, returns NaN.
- If one argument is NaN, returns the numeric argument.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.30 `__device__ float fmodf (float x, float y)`

Calculate the floating-point remainder of x / y . The absolute value of the computed value is always less than y 's absolute value and will have the same sign as x .

Returns:

- Returns the floating point remainder of x / y .
- `fmodf(±0, y)` returns ± 0 if y is not zero.
- `fmodf(x, y)` returns NaN and raised an invalid floating point exception if x is $\pm\infty$ or y is zero.
- `fmodf(x, y)` returns zero if y is zero or the result would overflow.
- `fmodf(x, ±∞)` returns x if x is finite.
- `fmodf(x, 0)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.31 `__device__ float frexpf (float x, int * nptr)`

Decomposes the floating-point value x into a component m for the normalized fraction element and another term n for the exponent. The absolute value of m will be greater than or equal to 0.5 and less than 1.0 or it will be equal to 0; $x = m \cdot 2^n$. The integer exponent n will be stored in the location to which `nptr` points.

Returns:

Returns the fractional component m .

- `frexpf(0, nptr)` returns 0 for the fractional component and zero for the integer component.
- `frexpf(± 0 , nptr)` returns ± 0 and stores zero in the location pointed to by `nptr`.
- `frexpf($\pm\infty$, nptr)` returns $\pm\infty$ and stores an unspecified value in the location to which `nptr` points.
- `frexpf(NaN, y)` returns a NaN and stores an unspecified value in the location to which `nptr` points.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.32 `__device__ float hypotf (float x, float y)`

Calculates the length of the hypotenuse of a right triangle whose two sides have lengths x and y without undue overflow or underflow.

Returns:

Returns the length of the hypotenuse $\sqrt{x^2 + y^2}$. If the correct value would overflow, returns $+\infty$. If the correct value would underflow, returns 0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.33 `__device__ int ilogbf (float x)`

Calculates the unbiased integer exponent of the input argument x .

Returns:

- If successful, returns the unbiased exponent of the argument.
- `ilogbf(0)` returns `INT_MIN`.
- `ilogbf(NaN)` returns NaN.
- `ilogbf(x)` returns `INT_MAX` if x is ∞ or the correct value is greater than `INT_MAX`.
- `ilogbf(x)` return `INT_MIN` if the correct value is less than `INT_MIN`.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.34 `__device__ int isfinite (float a)`

Determine whether the floating-point value `a` is a finite value (zero, subnormal, or normal and not infinity or NaN).

Returns:

Returns a nonzero value if and only if `a` is a finite value.

5.61.2.35 `__device__ int isinf (float a)`

Determine whether the floating-point value `a` is an infinite value (positive or negative).

Returns:

Returns a nonzero value if and only if `a` is a infinite value.

5.61.2.36 `__device__ int isnan (float a)`

Determine whether the floating-point value `a` is a NaN.

Returns:

Returns a nonzero value if and only if `a` is a NaN value.

5.61.2.37 `__device__ float j0f (float x)`

Calculate the value of the Bessel function of the first kind of order 0 for the input argument `x`, $J_0(x)$.

Returns:

Returns the value of the Bessel function of the first kind of order 0.

- `j0f(±∞)` returns +0.
- `j0f(NaN)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.38 `__device__ float j1f (float x)`

Calculate the value of the Bessel function of the first kind of order 1 for the input argument `x`, $J_1(x)$.

Returns:

Returns the value of the Bessel function of the first kind of order 1.

- `j1f(±0)` returns ±0.
- `j1f(±∞)` returns +0.
- `j1f(NaN)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.39 `__device__ float jnf (int n, float x)`

Calculate the value of the Bessel function of the first kind of order n for the input argument x , $J_n(x)$.

Returns:

Returns the value of the Bessel function of the first kind of order n .

- `jnf(n, NaN)` returns NaN.
- `jnf(n, x)` returns NaN for $n < 0$.
- `jnf(n, +∞)` returns +0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.40 `__device__ float ldexpf (float x, int exp)`

Calculate the value of $x \cdot 2^{exp}$ of the input arguments x and `exp`.

Returns:

- `ldexpf(x)` returns $\pm\infty$ if the correctly calculated value is outside the single floating point range.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.41 `__device__ float lgammaf (float x)`

Calculate the natural logarithm of the absolute value of the gamma function of the input argument x , namely the value of $\log_e |\int_0^\infty e^{-t} t^{x-1} dt|$.

Returns:

- `lgammaf(1)` returns +0.
- `lgammaf(2)` returns +0.
- `lgammaf(x)` returns $\pm\infty$ if the correctly calculated value is outside the single floating point range.
- `lgammaf(x)` returns $+\infty$ if $x \leq 0$.
- `lgammaf(-∞)` returns $-\infty$.
- `lgammaf(+∞)` returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.42 `__device__ long long int llrintf (float x)`

Round x to the nearest integer value, with halfway cases rounded towards zero. If the result is outside the range of the return type, the result is undefined.

Returns:

Returns rounded integer value.

5.61.2.43 `__device__ long long int llroundf (float x)`

Round x to the nearest integer value, with halfway cases rounded away from zero. If the result is outside the range of the return type, the result is undefined.

Returns:

Returns rounded integer value.

Note:

This function may be slower than alternate rounding methods. See [llrintf\(\)](#).

5.61.2.44 `__device__ float log10f (float x)`

Calculate the base 10 logarithm of the input argument x .

Returns:

- $\log_{10}f(\pm 0)$ returns $-\infty$.
- $\log_{10}f(1)$ returns $+0$.
- $\log_{10}f(x)$ returns NaN for $x < 0$.
- $\log_{10}f(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.45 `__device__ float log1pf (float x)`

Calculate the value of $\log_e(1+x) \lfloor x \rfloor$ of the input argument x .

Returns:

- $\log_{1p}f(\pm 0)$ returns $-\infty$.
- $\log_{1p}f(-1)$ returns $+0$.
- $\log_{1p}f(x)$ returns NaN for $x < -1$.
- $\log_{1p}f(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.46 `__device__ float log2f (float x)`

Calculate the base 2 logarithm of the input argument x .

Returns:

- $\log_{2}f(\pm 0)$ returns $-\infty$.

- $\log_2 f(1)$ returns $+0$.
- $\log_2 f(x)$ returns NaN for $x < 0$.
- $\log_2 f(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.47 __device__ float logbf (float x)

Calculate the floating point representation of the exponent of the input argument x .

Returns:

- $\logbf \pm 0$ returns $-\infty$
- $\logbf +\infty$ returns $+\infty$

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.48 __device__ float logf (float x)

Calculate the natural logarithm of the input argument x .

Returns:

- $\logf(\pm 0)$ returns $-\infty$.
- $\logf(1)$ returns $+0$.
- $\logf(x)$ returns NaN for $x < 0$.
- $\logf(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.49 __device__ long int lrintf (float x)

Round x to the nearest integer value, with halfway cases rounded towards zero. If the result is outside the range of the return type, the result is undefined.

Returns:

Returns rounded integer value.

5.61.2.50 `__device__ long int lroundf (float x)`

Round x to the nearest integer value, with halfway cases rounded away from zero. If the result is outside the range of the return type, the result is undefined.

Returns:

Returns rounded integer value.

Note:

This function may be slower than alternate rounding methods. See [lrintf\(\)](#).

5.61.2.51 `__device__ float modff (float x, float * iptr)`

Break down the argument x into fractional and integral parts. The integral part is stored in the argument `iptr`. Fractional and integral parts are given the same sign as the argument x .

Returns:

- `modff(±x, iptr)` returns a result with the same sign as x .
- `modff(±∞, iptr)` returns $±0$ and stores $±∞$ in the object pointed to by `iptr`.
- `modff(NaN, iptr)` stores a NaN in the object pointed to by `iptr` and returns a NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.52 `__device__ float nanf (const char * tagp)`

Return a representation of a quiet NaN. Argument `tagp` selects one of the possible representations.

Returns:

- `nanf(tagp)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.53 `__device__ float nearbyintf (float x)`

Round argument x to an integer value in single precision floating-point format.

Returns:

- `nearbyintf(±0)` returns $±0$.
- `nearbyintf(±∞)` returns $±∞$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.54 `__device__ float nextafterf (float x, float y)`

Calculate the next representable single-precision floating-point value following x in the direction of y . For example, if y is greater than x , `nextafterf()` returns the smallest representable number greater than x .

Returns:

- `nextafterf($\pm\infty$, y)` returns $\pm\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.55 `__device__ float normcdf (float y)`

Calculate the cumulative distribution function of the standard normal distribution for input argument y , $\Phi(y)$.

Returns:

- `normcdf($+\infty$)` returns 1
- `normcdf($-\infty$)` returns +0

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.56 `__device__ float normcdfinv (float y)`

Calculate the inverse of the standard normal cumulative distribution function for input argument y , $\Phi^{-1}(y)$. The function is defined for input values in the interval $(0, 1)$.

Returns:

- `normcdfinv(0)` returns $-\infty$.
- `normcdfinv(1)` returns $+\infty$.
- `normcdfinv(x)` returns NaN if x is not in the interval $[0,1]$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.57 `__device__ float powf (float x, float y)`

Calculate the value of x to the power of y .

Returns:

- `powf(± 0 , y)` returns $\pm\infty$ for y an integer less than 0.
- `powf(± 0 , y)` returns ± 0 for y an odd integer greater than 0.
- `powf(± 0 , y)` returns +0 for $y > 0$ and not an odd integer.
- `powf(-1, $\pm\infty$)` returns 1.

- `powf(+1, y)` returns 1 for any y , even a NaN.
- `powf(x, ± 0)` returns 1 for any x , even a NaN.
- `powf(x, y)` returns a NaN for finite $x < 0$ and finite non-integer y .
- `powf(x, $-\infty$)` returns $+\infty$ for $|x| < 1$.
- `powf(x, $-\infty$)` returns +0 for $|x| > 1$.
- `powf(x, $+\infty$)` returns +0 for $|x| < 1$.
- `powf(x, $+\infty$)` returns $+\infty$ for $|x| > 1$.
- `powf($-\infty$, y)` returns -0 for y an odd integer less than 0.
- `powf($-\infty$, y)` returns +0 for $y < 0$ and not an odd integer.
- `powf($-\infty$, y)` returns $-\infty$ for y an odd integer greater than 0.
- `powf($-\infty$, y)` returns $+\infty$ for $y > 0$ and not an odd integer.
- `powf($+\infty$, y)` returns +0 for $y < 0$.
- `powf($+\infty$, y)` returns $+\infty$ for $y > 0$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.58 `__device__ float rcbrtf (float x)`

Calculate reciprocal cube root function of x

Returns:

- `rcbrt(± 0)` returns $\pm\infty$.
- `rcbrt($\pm\infty$)` returns ± 0 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.59 `__device__ float remainderf (float x, float y)`

Compute single-precision floating-point remainder r of dividing x by y for nonzero y . Thus $r = x - ny$. The value n is the integer value nearest $\frac{x}{y}$. In the case when $|n - \frac{x}{y}| = \frac{1}{2}$, the even n value is chosen.

Returns:

- `remainderf(x, 0)` returns NaN.
- `remainderf($\pm\infty$, y)` returns NaN.
- `remainderf(x, $\pm\infty$)` returns x for finite x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.60 `__device__ float remquof (float x, float y, int * quo)`

Compute a double-precision floating-point remainder in the same way as the `remainderf()` function. Argument `quo` returns part of quotient upon division of x by y . Value `quo` has the same sign as $\frac{x}{y}$ and may not be the exact quotient but agrees with the exact quotient in the low order 3 bits.

Returns:

Returns the remainder.

- `remquof(x, 0, quo)` returns NaN.
- `remquof($\pm\infty$, y, quo)` returns NaN.
- `remquof(x, $\pm\infty$, quo)` returns x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.61 `__device__ float rintf (float x)`

Round x to the nearest integer value in floating-point format, with halfway cases rounded towards zero.

Returns:

Returns rounded integer value.

5.61.2.62 `__device__ float roundf (float x)`

Round x to the nearest integer value in floating-point format, with halfway cases rounded away from zero.

Returns:

Returns rounded integer value.

Note:

This function may be slower than alternate rounding methods. See `rintf()`.

5.61.2.63 `__device__ float rsqrtf (float x)`

Calculate the reciprocal of the nonnegative square root of x , $1/\sqrt{x}$.

Returns:

Returns $1/\sqrt{x}$.

- `rsqrtf(+ ∞)` returns $+0$.
- `rsqrtf(± 0)` returns $\pm\infty$.
- `rsqrtf(x)` returns NaN if x is less than 0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.64 `__device__ float scalblnf (float x, long int n)`

Scale x by 2^n by efficient manipulation of the floating-point exponent.

Returns:

Returns $x * 2^n$.

- `scalblnf(±0, n)` returns $±0$.
- `scalblnf(x, 0)` returns x .
- `scalblnf(±∞, n)` returns $±∞$.

5.61.2.65 `__device__ float scalbnf (float x, int n)`

Scale x by 2^n by efficient manipulation of the floating-point exponent.

Returns:

Returns $x * 2^n$.

- `scalbnf(±0, n)` returns $±0$.
- `scalbnf(x, 0)` returns x .
- `scalbnf(±∞, n)` returns $±∞$.

5.61.2.66 `__device__ int signbit (float a)`

Determine whether the floating-point value a is negative.

Returns:

Returns a nonzero value if and only if a is negative. Reports the sign bit of all values including infinities, zeros, and NaNs.

5.61.2.67 `__device__ void sincosf (float x, float * sptr, float * cptr)`

Calculate the sine and cosine of the first input argument x (measured in radians). The results for sine and cosine are written into the second argument, `sptr`, and, respectively, third argument, `cptr`.

Returns:

- none

See also:

[sinf\(\)](#) and [cosf\(\)](#).

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This function is affected by the `-use_fast_math` compiler flag. See the CUDA C Programming Guide, Appendix C, Table C-3 for a complete list of functions affected.

5.61.2.68 `__device__ void sincospif (float x, float * sptr, float * cptr)`

Calculate the sine and cosine of the first input argument, x (measured in radians), $x\pi$. The results for sine and cosine are written into the second argument, `sptr`, and, respectively, third argument, `cptr`.

Returns:

- none

See also:

[sinpif\(\)](#) and [cospif\(\)](#).

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.69 `__device__ float sinf (float x)`

Calculate the sine of the input argument x (measured in radians).

Returns:

- `sinf(±0)` returns ± 0 .
- `sinf(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This function is affected by the `-use_fast_math` compiler flag. See the CUDA C Programming Guide, Appendix C, Table C-3 for a complete list of functions affected.

5.61.2.70 `__device__ float sinhf (float x)`

Calculate the hyperbolic sine of the input argument x (measured in radians).

Returns:

- `sinhf(±0)` returns ± 0 .
- `sinhf(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.71 `__device__ float sinpif (float x)`

Calculate the sine of $x \times \pi$ (measured in radians), where x is the input argument.

Returns:

- `sinpif(±0)` returns ± 0 .
- `sinpif(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.72 `__device__ float sqrtf (float x)`

Calculate the nonnegative square root of x , \sqrt{x} .

Returns:

Returns \sqrt{x} .

- `sqrtf(±0)` returns ± 0 .
- `sqrtf(+∞)` returns $+\infty$.
- `sqrtf(x)` returns NaN if x is less than 0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.73 `__device__ float tanf (float x)`

Calculate the tangent of the input argument x (measured in radians).

Returns:

- `tanf(±0)` returns ± 0 .
- `tanf(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This function is affected by the `-use_fast_math` compiler flag. See the CUDA C Programming Guide, Appendix C, Table C-3 for a complete list of functions affected.

5.61.2.74 `__device__ float tanhf (float x)`

Calculate the hyperbolic tangent of the input argument x (measured in radians).

Returns:

- `tanhf(±0)` returns ± 0 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.75 `__device__ float tgammaf (float x)`

Calculate the gamma function of the input argument x , namely the value of $\int_0^\infty e^{-t} t^{x-1} dt$.

Returns:

- `tgammaf(±0)` returns $\pm\infty$.
- `tgammaf(2)` returns $+0$.

- $\text{tgammaf}(x)$ returns $\pm\infty$ if the correctly calculated value is outside the single floating point range.
- $\text{tgammaf}(x)$ returns NaN if $x < 0$.
- $\text{tgammaf}(-\infty)$ returns NaN.
- $\text{tgammaf}(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.76 __device__ float truncf (float x)

Round x to the nearest integer value that does not exceed x in magnitude.

Returns:

Returns truncated integer value.

5.61.2.77 __device__ float y0f (float x)

Calculate the value of the Bessel function of the second kind of order 0 for the input argument x , $Y_0(x)$.

Returns:

Returns the value of the Bessel function of the second kind of order 0.

- $y0f(0)$ returns $-\infty$.
- $y0f(x)$ returns NaN for $x < 0$.
- $y0f(+\infty)$ returns +0.
- $y0f(\text{NaN})$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.78 __device__ float y1f (float x)

Calculate the value of the Bessel function of the second kind of order 1 for the input argument x , $Y_1(x)$.

Returns:

Returns the value of the Bessel function of the second kind of order 1.

- $y1f(0)$ returns $-\infty$.
- $y1f(x)$ returns NaN for $x < 0$.
- $y1f(+\infty)$ returns +0.
- $y1f(\text{NaN})$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.61.2.79 `__device__ float ynf (int n, float x)`

Calculate the value of the Bessel function of the second kind of order n for the input argument x , $Y_n(x)$.

Returns:

Returns the value of the Bessel function of the second kind of order n .

- `ynf(n, x)` returns NaN for $n < 0$.
- `ynf(n, 0)` returns $-\infty$.
- `ynf(n, x)` returns NaN for $x < 0$.
- `ynf(n, +∞)` returns +0.
- `ynf(n, NaN)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.62 Double Precision Mathematical Functions

Functions

- `__device__ double acos (double x)`
Calculate the arc cosine of the input argument.
- `__device__ double acosh (double x)`
Calculate the nonnegative arc hyperbolic cosine of the input argument.
- `__device__ double asin (double x)`
Calculate the arc sine of the input argument.
- `__device__ double asinh (double x)`
Calculate the arc hyperbolic sine of the input argument.
- `__device__ double atan (double x)`
Calculate the arc tangent of the input argument.
- `__device__ double atan2 (double x, double y)`
Calculate the arc tangent of the ratio of first and second input arguments.
- `__device__ double atanh (double x)`
Calculate the arc hyperbolic tangent of the input argument.
- `__device__ double cbrt (double x)`
Calculate the cube root of the input argument.
- `__device__ double ceil (double x)`
Calculate ceiling of the input argument.
- `__device__ double copysign (double x, double y)`
Create value with given magnitude, copying sign of second value.
- `__device__ double cos (double x)`
Calculate the cosine of the input argument.
- `__device__ double cosh (double x)`
Calculate the hyperbolic cosine of the input argument.
- `__device__ double cospi (double x)`
Calculate the cosine of the input argument $\times \pi$.
- `__device__ double erf (double x)`
Calculate the error function of the input argument.
- `__device__ double erfc (double x)`
Calculate the complementary error function of the input argument.
- `__device__ double erfcinv (double y)`

Calculate the inverse complementary error function of the input argument.

- `__device__ double erfcx (double x)`

Calculate the scaled complementary error function of the input argument.

- `__device__ double erfinv (double y)`

Calculate the inverse error function of the input argument.

- `__device__ double exp (double x)`

Calculate the base e exponential of the input argument.

- `__device__ double exp10 (double x)`

Calculate the base 10 exponential of the input argument.

- `__device__ double exp2 (double x)`

Calculate the base 2 exponential of the input argument.

- `__device__ double expm1 (double x)`

Calculate the base e exponential of the input argument, minus 1.

- `__device__ double fabs (double x)`

Calculate the absolute value of the input argument.

- `__device__ double fdim (double x, double y)`

Compute the positive difference between x and y .

- `__device__ double floor (double x)`

Calculate the largest integer less than or equal to x .

- `__device__ double fma (double x, double y, double z)`

Compute $x \times y + z$ as a single operation.

- `__device__ double fmax (double, double)`

Determine the maximum numeric value of the arguments.

- `__device__ double fmin (double x, double y)`

Determine the minimum numeric value of the arguments.

- `__device__ double fmod (double x, double y)`

Calculate the floating-point remainder of x / y .

- `__device__ double frexp (double x, int *nptr)`

Extract mantissa and exponent of a floating-point value.

- `__device__ double hypot (double x, double y)`

Calculate the square root of the sum of squares of two arguments.

- `__device__ int ilogb (double x)`

Compute the unbiased integer exponent of the argument.

- `__device__ int isfinite` (double a)
Determine whether argument is finite.
- `__device__ int isinf` (double a)
Determine whether argument is infinite.
- `__device__ int isnan` (double a)
Determine whether argument is a NaN.
- `__device__ double j0` (double x)
Calculate the value of the Bessel function of the first kind of order 0 for the input argument.
- `__device__ double j1` (double x)
Calculate the value of the Bessel function of the first kind of order 1 for the input argument.
- `__device__ double jn` (int n, double x)
Calculate the value of the Bessel function of the first kind of order n for the input argument.
- `__device__ double ldexp` (double x, int exp)
Calculate the value of $x \cdot 2^{\text{exp}}$.
- `__device__ double lgamma` (double x)
Calculate the natural logarithm of the absolute value of the gamma function of the input argument.
- `__device__ long long int llrint` (double x)
Round input to nearest integer value.
- `__device__ long long int llround` (double x)
Round to nearest integer value.
- `__device__ double log` (double x)
Calculate the base e logarithm of the input argument.
- `__device__ double log10` (double x)
Calculate the base 10 logarithm of the input argument.
- `__device__ double log1p` (double x)
Calculate the value of $\log_e(1 + x)$ [x].
- `__device__ double log2` (double x)
Calculate the base 2 logarithm of the input argument.
- `__device__ double logb` (double x)
Calculate the floating point representation of the exponent of the input argument.
- `__device__ long int lrint` (double x)
Round input to nearest integer value.
- `__device__ long int lround` (double x)
Round to nearest integer value.

- `__device__ double modf (double x, double *iptr)`
Break down the input argument into fractional and integral parts.
- `__device__ double nan (const char *tagp)`
Returns "Not a Number" value.
- `__device__ double nearbyint (double x)`
Round the input argument to the nearest integer.
- `__device__ double nextafter (double x, double y)`
Return next representable double-precision floating-point value after argument.
- `__device__ double normcdf (double y)`
Calculate the standard normal cumulative distribution function.
- `__device__ double normcdfinv (double y)`
Calculate the inverse of the standard normal cumulative distribution function.
- `__device__ double pow (double x, double y)`
Calculate the value of first argument to the power of second argument.
- `__device__ double rcbt (double x)`
Calculate reciprocal cube root function.
- `__device__ double remainder (double x, double y)`
Compute double-precision floating-point remainder.
- `__device__ double remquo (double x, double y, int *quo)`
Compute double-precision floating-point remainder and part of quotient.
- `__device__ double rint (double x)`
Round to nearest integer value in floating-point.
- `__device__ double round (double x)`
Round to nearest integer value in floating-point.
- `__device__ double rsqrt (double x)`
Calculate the reciprocal of the square root of the input argument.
- `__device__ double scalbln (double x, long int n)`
Scale floating-point input by integer power of two.
- `__device__ double scalbn (double x, int n)`
Scale floating-point input by integer power of two.
- `__device__ int signbit (double a)`
Return the sign bit of the input.
- `__device__ double sin (double x)`

Calculate the sine of the input argument.

- `__device__ void sincos` (double x, double *sptr, double *cptr)

Calculate the sine and cosine of the first input argument.

- `__device__ void sincospi` (double x, double *sptr, double *cptr)

Calculate the sine and cosine of the first input argument $\times \pi$.

- `__device__ double sinh` (double x)

Calculate the hyperbolic sine of the input argument.

- `__device__ double sinpi` (double x)

Calculate the sine of the input argument $\times \pi$.

- `__device__ double sqrt` (double x)

Calculate the square root of the input argument.

- `__device__ double tan` (double x)

Calculate the tangent of the input argument.

- `__device__ double tanh` (double x)

Calculate the hyperbolic tangent of the input argument.

- `__device__ double tgamma` (double x)

Calculate the gamma function of the input argument.

- `__device__ double trunc` (double x)

Truncate input argument to the integral part.

- `__device__ double y0` (double x)

Calculate the value of the Bessel function of the second kind of order 0 for the input argument.

- `__device__ double y1` (double x)

Calculate the value of the Bessel function of the second kind of order 1 for the input argument.

- `__device__ double yn` (int n, double x)

Calculate the value of the Bessel function of the second kind of order n for the input argument.

5.62.1 Detailed Description

This section describes double precision mathematical functions.

5.62.2 Function Documentation

5.62.2.1 `__device__ double acos` (double x)

Calculate the principal value of the arc cosine of the input argument x.

Returns:

Result will be in radians, in the interval $[0, \pi]$ for x inside $[-1, +1]$.

- `acos(1)` returns `+0`.
- `acos(x)` returns NaN for x outside $[-1, +1]$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.2 `__device__ double acosh (double x)`

Calculate the nonnegative arc hyperbolic cosine of the input argument x (measured in radians).

Returns:

Result will be in the interval $[0, +\infty]$.

- `acosh(1)` returns `0`.
- `acosh(x)` returns NaN for x in the interval $[-\infty, 1)$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.3 `__device__ double asin (double x)`

Calculate the principal value of the arc sine of the input argument x .

Returns:

Result will be in radians, in the interval $[-\pi/2, +\pi/2]$ for x inside $[-1, +1]$.

- `asin(0)` returns `+0`.
- `asin(x)` returns NaN for x outside $[-1, +1]$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.4 `__device__ double asinh (double x)`

Calculate the arc hyperbolic sine of the input argument x (measured in radians).

Returns:

- `asinh(0)` returns `1`.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.5 __device__ double atan (double x)

Calculate the principal value of the arc tangent of the input argument x .

Returns:

Result will be in radians, in the interval $[-\pi/2, +\pi/2]$.

- `atan(0)` returns `+0`.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.6 __device__ double atan2 (double x, double y)

Calculate the principal value of the arc tangent of the ratio of first and second input arguments x/y . The quadrant of the result is determined by the signs of inputs x and y .

Returns:

Result will be in radians, in the interval $[-\pi, +\pi]$.

- `atan2(0, 1)` returns `+0`.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.7 __device__ double atanh (double x)

Calculate the arc hyperbolic tangent of the input argument x (measured in radians).

Returns:

- `atanh(± 0)` returns ± 0 .
- `atanh(± 1)` returns $\pm\infty$.
- `atanh(x)` returns NaN for x outside interval $[-1, 1]$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.8 __device__ double cbrt (double x)

Calculate the cube root of x , $x^{1/3}$.

Returns:

Returns $x^{1/3}$.

- `cbrt(± 0)` returns ± 0 .
- `cbrt($\pm\infty$)` returns $\pm\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.9 `__device__ double ceil (double x)`

Compute the smallest integer value not less than x .

Returns:

Returns $\lceil x \rceil$ expressed as a floating-point number.

- `ceil(±0)` returns ± 0 .
- `ceil(±∞)` returns $\pm\infty$.

5.62.2.10 `__device__ double copysign (double x, double y)`

Create a floating-point value with the magnitude x and the sign of y .

Returns:

Returns a value with the magnitude of x and the sign of y .

5.62.2.11 `__device__ double cos (double x)`

Calculate the cosine of the input argument x (measured in radians).

Returns:

- `cos(±0)` returns 1.
- `cos(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.12 `__device__ double cosh (double x)`

Calculate the hyperbolic cosine of the input argument x (measured in radians).

Returns:

- `cosh(0)` returns 1.
- `cosh(±∞)` returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.13 `__device__ double cospi (double x)`

Calculate the cosine of $x \times \pi$ (measured in radians), where x is the input argument.

Returns:

- `cospi(±0)` returns 1.

- `cospi(±∞)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.14 __device__ double erf (double x)

Calculate the value of the error function for the input argument x , $\frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

Returns:

- `erf(±0)` returns ± 0 .
- `erf(±∞)` returns ± 1 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.15 __device__ double erfc (double x)

Calculate the complementary error function of the input argument x , $1 - \text{erf}(x)$.

Returns:

- `erfc(-∞)` returns 2.
- `erfc(+∞)` returns +0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.16 __device__ double erfcinv (double y)

Calculate the inverse complementary error function of the input argument y , for y in the interval $[0, 2]$. The inverse complementary error function find the value x that satisfies the equation $y = \text{erfc}(x)$, for $0 \leq y \leq 2$, and $-\infty \leq x \leq \infty$.

Returns:

- `erfcinv(0)` returns $+\infty$.
- `erfcinv(2)` returns $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.17 `__device__ double erfcx (double x)`

Calculate the scaled complementary error function of the input argument x , $e^{x^2} \cdot \text{erfc}(x)$.

Returns:

- `erfcx(-∞)` returns $+\infty$
- `erfcx(+∞)` returns $+0$
- `erfcx(x)` returns $+\infty$ if the correctly calculated value is outside the double floating point range.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.18 `__device__ double erfinv (double y)`

Calculate the inverse error function of the input argument y , for y in the interval $[-1, 1]$. The inverse error function finds the value x that satisfies the equation $y = \text{erf}(x)$, for $-1 \leq y \leq 1$, and $-\infty \leq x \leq \infty$.

Returns:

- `erfinv(1)` returns $+\infty$.
- `erfinv(-1)` returns $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.19 `__device__ double exp (double x)`

Calculate the base e exponential of the input argument x .

Returns:

Returns e^x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.20 `__device__ double exp10 (double x)`

Calculate the base 10 exponential of the input argument x .

Returns:

Returns 10^x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.21 `__device__ double exp2 (double x)`

Calculate the base 2 exponential of the input argument x .

Returns:

Returns 2^x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.22 `__device__ double expm1 (double x)`

Calculate the base e exponential of the input argument x , minus 1.

Returns:

Returns $e^x - 1$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.23 `__device__ double fabs (double x)`

Calculate the absolute value of the input argument x .

Returns:

Returns the absolute value of the input argument.

- `fabs($\pm\infty$)` returns $+\infty$.
- `fabs(± 0)` returns 0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.24 `__device__ double fdim (double x, double y)`

Compute the positive difference between x and y . The positive difference is $x - y$ when $x > y$ and +0 otherwise.

Returns:

Returns the positive difference between x and y .

- `fdim(x, y)` returns $x - y$ if $x > y$.
- `fdim(x, y)` returns +0 if $x \leq y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.62.2.25 `__device__ double floor (double x)`

Calculates the largest integer value which is less than or equal to x .

Returns:

Returns $\log_e(1 + x) \lfloor x \rfloor$ expressed as a floating-point number.

- `floor($\pm\infty$)` returns $\pm\infty$.
- `floor(± 0)` returns ± 0 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.26 `__device__ double fma (double x, double y, double z)`

Compute the value of $x \times y + z$ as a single ternary operation. After computing the value to infinite precision, the value is rounded once.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fma($\pm\infty$, ± 0 , z)` returns NaN.
- `fma(± 0 , $\pm\infty$, z)` returns NaN.
- `fma(x , y , $-\infty$)` returns NaN if $x \times y$ is an exact $+\infty$.
- `fma(x , y , $+\infty$)` returns NaN if $x \times y$ is an exact $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.27 `__device__ double fmax (double, double)`

Determines the maximum numeric value of the arguments x and y . Treats NaN arguments as missing data. If one argument is a NaN and the other is legitimate numeric value, the numeric value is chosen.

Returns:

Returns the maximum numeric values of the arguments x and y .

- If both arguments are NaN, returns NaN.
- If one argument is NaN, returns the numeric argument.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.28 `__device__ double fmin (double x, double y)`

Determines the minimum numeric value of the arguments x and y . Treats NaN arguments as missing data. If one argument is a NaN and the other is legitimate numeric value, the numeric value is chosen.

Returns:

Returns the minimum numeric values of the arguments x and y .

- If both arguments are NaN, returns NaN.
- If one argument is NaN, returns the numeric argument.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.29 `__device__ double fmod (double x, double y)`

Calculate the floating-point remainder of x / y . The absolute value of the computed value is always less than y 's absolute value and will have the same sign as x .

Returns:

- Returns the floating point remainder of x / y .
- `fmod(±0, y)` returns $±0$ if y is not zero.
- `fmod(x, y)` returns NaN and raised an invalid floating point exception if x is $±∞$ or y is zero.
- `fmod(x, y)` returns zero if y is zero or the result would overflow.
- `fmod(x, ±∞)` returns x if x is finite.
- `fmod(x, 0)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.30 `__device__ double frexp (double x, int * nptr)`

Decompose the floating-point value x into a component m for the normalized fraction element and another term n for the exponent. The absolute value of m will be greater than or equal to 0.5 and less than 1.0 or it will be equal to 0; $x = m \cdot 2^n$. The integer exponent n will be stored in the location to which `nptr` points.

Returns:

Returns the fractional component m .

- `frexp(0, nptr)` returns 0 for the fractional component and zero for the integer component.
- `frexp(±0, nptr)` returns $±0$ and stores zero in the location pointed to by `nptr`.
- `frexp(±∞, nptr)` returns $±∞$ and stores an unspecified value in the location to which `nptr` points.
- `frexp(NaN, y)` returns a NaN and stores an unspecified value in the location to which `nptr` points.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.31 `__device__ double hypot (double x, double y)`

Calculate the length of the hypotenuse of a right triangle whose two sides have lengths x and y without undue overflow or underflow.

Returns:

Returns the length of the hypotenuse $\sqrt{x^2 + y^2}$. If the correct value would overflow, returns $+\infty$. If the correct value would underflow, returns 0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.32 `__device__ int ilogb (double x)`

Calculates the unbiased integer exponent of the input argument x .

Returns:

- If successful, returns the unbiased exponent of the argument.
- `ilogb(0)` returns `INT_MIN`.
- `ilogb(NaN)` returns `NaN`.
- `ilogb(x)` returns `INT_MAX` if x is ∞ or the correct value is greater than `INT_MAX`.
- `ilogb(x)` return `INT_MIN` if the correct value is less than `INT_MIN`.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.33 `__device__ int isfinite (double a)`

Determine whether the floating-point value a is a finite value (zero, subnormal, or normal and not infinity or NaN).

Returns:

Returns a nonzero value if and only if a is a finite value.

5.62.2.34 `__device__ int isinf (double a)`

Determine whether the floating-point value a is an infinite value (positive or negative).

Returns:

Returns a nonzero value if and only if a is a infinite value.

5.62.2.35 `__device__ int isnan (double a)`

Determine whether the floating-point value a is a NaN.

Returns:

Returns a nonzero value if and only if a is a NaN value.

5.62.2.36 `__device__ double j0 (double x)`

Calculate the value of the Bessel function of the first kind of order 0 for the input argument x , $J_0(x)$.

Returns:

Returns the value of the Bessel function of the first kind of order 0.

- $j0(\pm\infty)$ returns +0.
- $j0(\text{NaN})$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.37 `__device__ double j1 (double x)`

Calculate the value of the Bessel function of the first kind of order 1 for the input argument x , $J_1(x)$.

Returns:

Returns the value of the Bessel function of the first kind of order 1.

- $j1(\pm 0)$ returns ± 0 .
- $j1(\pm\infty)$ returns +0.
- $j1(\text{NaN})$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.38 `__device__ double jn (int n, double x)`

Calculate the value of the Bessel function of the first kind of order n for the input argument x , $J_n(x)$.

Returns:

Returns the value of the Bessel function of the first kind of order n .

- $jn(n, \text{NaN})$ returns NaN.
- $jn(n, x)$ returns NaN for $n < 0$.
- $jn(n, +\infty)$ returns +0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.39 `__device__ double ldexp (double x, int exp)`

Calculate the value of $x \cdot 2^{\text{exp}}$ of the input arguments x and exp .

Returns:

- $\text{ldexp}(x)$ returns $\pm\infty$ if the correctly calculated value is outside the double floating point range.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.40 __device__ double lgamma (double x)

Calculate the natural logarithm of the absolute value of the gamma function of the input argument x , namely the value of $\log_e \left| \int_0^\infty e^{-t} t^{x-1} dt \right|$

Returns:

- `lgamma(1)` returns `+0`.
- `lgamma(2)` returns `+0`.
- `lgamma(x)` returns $\pm\infty$ if the correctly calculated value is outside the double floating point range.
- `lgamma(x)` returns $+\infty$ if $x \leq 0$.
- `lgamma(-∞)` returns $-\infty$.
- `lgamma(+∞)` returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.41 __device__ long long int llrint (double x)

Round x to the nearest integer value, with halfway cases rounded towards zero. If the result is outside the range of the return type, the result is undefined.

Returns:

Returns rounded integer value.

5.62.2.42 __device__ long long int llround (double x)

Round x to the nearest integer value, with halfway cases rounded away from zero. If the result is outside the range of the return type, the result is undefined.

Returns:

Returns rounded integer value.

Note:

This function may be slower than alternate rounding methods. See [llrint\(\)](#).

5.62.2.43 __device__ double log (double x)

Calculate the base e logarithm of the input argument x .

Returns:

- `log(±0)` returns $-\infty$.

- $\log(1)$ returns +0.
- $\log(x)$ returns NaN for $x < 0$.
- $\log(+\infty)$ returns $+\infty$

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.44 __device__ double log10 (double x)

Calculate the base 10 logarithm of the input argument x .

Returns:

- $\log_{10}(\pm 0)$ returns $-\infty$.
- $\log_{10}(1)$ returns +0.
- $\log_{10}(x)$ returns NaN for $x < 0$.
- $\log_{10}(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.45 __device__ double log1p (double x)

Calculate the value of $\log_e(1 + x)$ of the input argument x .

Returns:

- $\log_{1p}(\pm 0)$ returns $-\infty$.
- $\log_{1p}(-1)$ returns +0.
- $\log_{1p}(x)$ returns NaN for $x < -1$.
- $\log_{1p}(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.46 __device__ double log2 (double x)

Calculate the base 2 logarithm of the input argument x .

Returns:

- $\log_2(\pm 0)$ returns $-\infty$.
- $\log_2(1)$ returns +0.
- $\log_2(x)$ returns NaN for $x < 0$.
- $\log_2(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.47 `__device__ double logb (double x)`

Calculate the floating point representation of the exponent of the input argument x .

Returns:

- $\text{logb } \pm 0$ returns $-\infty$
- $\text{logb } \pm \infty$ returns $+\infty$

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.48 `__device__ long int lrint (double x)`

Round x to the nearest integer value, with halfway cases rounded towards zero. If the result is outside the range of the return type, the result is undefined.

Returns:

Returns rounded integer value.

5.62.2.49 `__device__ long int lround (double x)`

Round x to the nearest integer value, with halfway cases rounded away from zero. If the result is outside the range of the return type, the result is undefined.

Returns:

Returns rounded integer value.

Note:

This function may be slower than alternate rounding methods. See [lrint\(\)](#).

5.62.2.50 `__device__ double modf (double x, double * iptr)`

Break down the argument x into fractional and integral parts. The integral part is stored in the argument `iptr`. Fractional and integral parts are given the same sign as the argument x .

Returns:

- $\text{modf}(\pm x, \text{iptr})$ returns a result with the same sign as x .
- $\text{modf}(\pm \infty, \text{iptr})$ returns ± 0 and stores $\pm \infty$ in the object pointed to by `iptr`.
- $\text{modf}(\text{NaN}, \text{iptr})$ stores a NaN in the object pointed to by `iptr` and returns a NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.51 `__device__ double nan (const char * tagp)`

Return a representation of a quiet NaN. Argument `tagp` selects one of the possible representations.

Returns:

- `nan(tagp)` returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.52 `__device__ double nearbyint (double x)`

Round argument `x` to an integer value in double precision floating-point format.

Returns:

- `nearbyint(±0)` returns ± 0 .
- `nearbyint(±∞)` returns $\pm\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.53 `__device__ double nextafter (double x, double y)`

Calculate the next representable double-precision floating-point value following `x` in the direction of `y`. For example, if `y` is greater than `x`, `nextafter()` returns the smallest representable number greater than `x`.

Returns:

- `nextafter(±∞, y)` returns $\pm\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.54 `__device__ double normcdf (double y)`

Calculate the cumulative distribution function of the standard normal distribution for input argument `y`, $\Phi(y)$.

Returns:

- `normcdf(+∞)` returns 1
- `normcdf(-∞)` returns +0

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.55 `__device__ double normcdfinv (double y)`

Calculate the inverse of the standard normal cumulative distribution function for input argument y , $\Phi^{-1}(y)$. The function is defined for input values in the interval $(0, 1)$.

Returns:

- `normcdfinv(0)` returns $-\infty$.
- `normcdfinv(1)` returns $+\infty$.
- `normcdfinv(x)` returns NaN if x is not in the interval $[0,1]$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.56 `__device__ double pow (double x, double y)`

Calculate the value of x to the power of y

Returns:

- `pow(± 0 , y)` returns $\pm\infty$ for y an integer less than 0.
- `pow(± 0 , y)` returns ± 0 for y an odd integer greater than 0.
- `pow(± 0 , y)` returns +0 for $y > 0$ and not an odd integer.
- `pow(-1, $\pm\infty$)` returns 1.
- `pow(+1, y)` returns 1 for any y , even a NaN.
- `pow(x , ± 0)` returns 1 for any x , even a NaN.
- `pow(x , y)` returns a NaN for finite $x < 0$ and finite non-integer y .
- `pow(x , $-\infty$)` returns $+\infty$ for $|x| < 1$.
- `pow(x , $-\infty$)` returns +0 for $|x| > 1$.
- `pow(x , $+\infty$)` returns +0 for $|x| < 1$.
- `pow(x , $+\infty$)` returns $+\infty$ for $|x| > 1$.
- `pow($-\infty$, y)` returns -0 for y an odd integer less than 0.
- `pow($-\infty$, y)` returns +0 for $y < 0$ and not an odd integer.
- `pow($-\infty$, y)` returns $-\infty$ for y an odd integer greater than 0.
- `pow($-\infty$, y)` returns $+\infty$ for $y > 0$ and not an odd integer.
- `pow($+\infty$, y)` returns +0 for $y < 0$.
- `pow($+\infty$, y)` returns $+\infty$ for $y > 0$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.57 `__device__ double rcbrt (double x)`

Calculate reciprocal cube root function of x

Returns:

- `rcbrt(±0)` returns $\pm\infty$.
- `rcbrt(±∞)` returns ± 0 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.58 `__device__ double remainder (double x, double y)`

Compute double-precision floating-point remainder r of dividing x by y for nonzero y . Thus $r = x - ny$. The value n is the integer value nearest $\frac{x}{y}$. In the case when $|n - \frac{x}{y}| = \frac{1}{2}$, the even n value is chosen.

Returns:

- `remainder(x, 0)` returns NaN.
- `remainder(±∞, y)` returns NaN.
- `remainder(x, ±∞)` returns x for finite x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.59 `__device__ double remquo (double x, double y, int * quo)`

Compute a double-precision floating-point remainder in the same way as the `remainder()` function. Argument `quo` returns part of quotient upon division of x by y . Value `quo` has the same sign as $\frac{x}{y}$ and may not be the exact quotient but agrees with the exact quotient in the low order 3 bits.

Returns:

Returns the remainder.

- `remquo(x, 0, quo)` returns NaN.
- `remquo(±∞, y, quo)` returns NaN.
- `remquo(x, ±∞, quo)` returns x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.60 `__device__ double rint (double x)`

Round x to the nearest integer value in floating-point format, with halfway cases rounded towards zero.

Returns:

Returns rounded integer value.

5.62.2.61 `__device__ double round (double x)`

Round x to the nearest integer value in floating-point format, with halfway cases rounded away from zero.

Returns:

Returns rounded integer value.

Note:

This function may be slower than alternate rounding methods. See [rint\(\)](#).

5.62.2.62 `__device__ double rsqrt (double x)`

Calculate the reciprocal of the nonnegative square root of x , $1/\sqrt{x}$.

Returns:

Returns $1/\sqrt{x}$.

- `rsqrt(+∞)` returns `+0`.
- `rsqrt(±0)` returns $\pm\infty$.
- `rsqrt(x)` returns NaN if x is less than 0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.63 `__device__ double scalbn (double x, long int n)`

Scale x by 2^n by efficient manipulation of the floating-point exponent.

Returns:

Returns $x * 2^n$.

- `scalbn(±0, n)` returns ± 0 .
- `scalbn(x, 0)` returns x .
- `scalbn(±∞, n)` returns $\pm\infty$.

5.62.2.64 `__device__ double scalbn (double x, int n)`

Scale x by 2^n by efficient manipulation of the floating-point exponent.

Returns:

Returns $x * 2^n$.

- `scalbn(±0, n)` returns ± 0 .
- `scalbn(x, 0)` returns x .
- `scalbn(±∞, n)` returns $\pm\infty$.

5.62.2.65 `__device__ int signbit (double a)`

Determine whether the floating-point value a is negative.

Returns:

Returns a nonzero value if and only if a is negative. Reports the sign bit of all values including infinities, zeros, and NaNs.

5.62.2.66 `__device__ double sin (double x)`

Calculate the sine of the input argument x (measured in radians).

Returns:

- $\sin(\pm 0)$ returns ± 0 .
- $\sin(\pm\infty)$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.67 `__device__ void sincos (double x, double * sptr, double * cptr)`

Calculate the sine and cosine of the first input argument x (measured in radians). The results for sine and cosine are written into the second argument, `sptr`, and, respectively, third argument, `cptr`.

Returns:

- none

See also:

[sin\(\)](#) and [cos\(\)](#).

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.68 `__device__ void sincospi (double x, double * sptr, double * cptr)`

Calculate the sine and cosine of the first input argument, x (measured in radians), $\times \pi$. The results for sine and cosine are written into the second argument, `sptr`, and, respectively, third argument, `cptr`.

Returns:

- none

See also:

[sinpi\(\)](#) and [cospi\(\)](#).

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.69 `__device__ double sinh (double x)`

Calculate the hyperbolic sine of the input argument x (measured in radians).

Returns:

- $\sinh(\pm 0)$ returns ± 0 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.70 `__device__ double sinpi (double x)`

Calculate the sine of $x \times \pi$ (measured in radians), where x is the input argument.

Returns:

- $\sinpi(\pm 0)$ returns ± 0 .
- $\sinpi(\pm \infty)$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.71 `__device__ double sqrt (double x)`

Calculate the nonnegative square root of x , \sqrt{x} .

Returns:

Returns \sqrt{x} .

- $\text{sqrt}(\pm 0)$ returns ± 0 .
- $\text{sqrt}(+\infty)$ returns $+\infty$.
- $\text{sqrt}(x)$ returns NaN if x is less than 0.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.72 `__device__ double tan (double x)`

Calculate the tangent of the input argument x (measured in radians).

Returns:

- $\tan(\pm 0)$ returns ± 0 .
- $\tan(\pm \infty)$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.73 `__device__ double tanh (double x)`

Calculate the hyperbolic tangent of the input argument x (measured in radians).

Returns:

- $\tanh(\pm 0)$ returns ± 0 .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.74 `__device__ double tgamma (double x)`

Calculate the gamma function of the input argument x , namely the value of $\int_0^\infty e^{-t} t^{x-1} dt$.

Returns:

- $\text{tgamma}(\pm 0)$ returns $\pm\infty$.
- $\text{tgamma}(2)$ returns $+0$.
- $\text{tgamma}(x)$ returns $\pm\infty$ if the correctly calculated value is outside the double floating point range.
- $\text{tgamma}(x)$ returns NaN if $x < 0$.
- $\text{tgamma}(-\infty)$ returns NaN.
- $\text{tgamma}(+\infty)$ returns $+\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.75 `__device__ double trunc (double x)`

Round x to the nearest integer value that does not exceed x in magnitude.

Returns:

Returns truncated integer value.

5.62.2.76 `__device__ double y0 (double x)`

Calculate the value of the Bessel function of the second kind of order 0 for the input argument x , $Y_0(x)$.

Returns:

Returns the value of the Bessel function of the second kind of order 0.

- $y_0(0)$ returns $-\infty$.
- $y_0(x)$ returns NaN for $x < 0$.
- $y_0(+\infty)$ returns $+0$.
- $y_0(\text{NaN})$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.77 `__device__ double y1 (double x)`

Calculate the value of the Bessel function of the second kind of order 1 for the input argument x , $Y_1(x)$.

Returns:

Returns the value of the Bessel function of the second kind of order 1.

- $y1(0)$ returns $-\infty$.
- $y1(x)$ returns NaN for $x < 0$.
- $y1(+\infty)$ returns +0.
- $y1(\text{NaN})$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.62.2.78 `__device__ double yn (int n, double x)`

Calculate the value of the Bessel function of the second kind of order n for the input argument x , $Y_n(x)$.

Returns:

Returns the value of the Bessel function of the second kind of order n .

- $yn(n, x)$ returns NaN for $n < 0$.
- $yn(n, 0)$ returns $-\infty$.
- $yn(n, x)$ returns NaN for $x < 0$.
- $yn(n, +\infty)$ returns +0.
- $yn(n, \text{NaN})$ returns NaN.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.63 Single Precision Ininsics

Functions

- `__device__ __cudart_builtin__ float __cosf (float x)`
Calculate the fast approximate cosine of the input argument.
- `__device__ __cudart_builtin__ float __exp10f (float x)`
Calculate the fast approximate base 10 exponential of the input argument.
- `__device__ __cudart_builtin__ float __expf (float x)`
Calculate the fast approximate base e exponential of the input argument.
- `__device__ float __fadd_rd (float x, float y)`
Add two floating point values in round-down mode.
- `__device__ float __fadd_rn (float x, float y)`
Add two floating point values in round-to-nearest-even mode.
- `__device__ float __fadd_ru (float x, float y)`
Add two floating point values in round-up mode.
- `__device__ float __fadd_rz (float x, float y)`
Add two floating point values in round-towards-zero mode.
- `__device__ float __fdiv_rd (float x, float y)`
Divide two floating point values in round-down mode.
- `__device__ float __fdiv_rn (float x, float y)`
Divide two floating point values in round-to-nearest-even mode.
- `__device__ float __fdiv_ru (float x, float y)`
Divide two floating point values in round-up mode.
- `__device__ float __fdiv_rz (float x, float y)`
Divide two floating point values in round-towards-zero mode.
- `__device__ float __fdividef (float x, float y)`
Calculate the fast approximate division of the input arguments.
- `__device__ float __fmaf_rd (float x, float y, float z)`
Compute $x \times y + z$ as a single operation, in round-down mode.
- `__device__ float __fmaf_rn (float x, float y, float z)`
Compute $x \times y + z$ as a single operation, in round-to-nearest-even mode.
- `__device__ float __fmaf_ru (float x, float y, float z)`
Compute $x \times y + z$ as a single operation, in round-up mode.
- `__device__ float __fmaf_rz (float x, float y, float z)`

Compute $x \times y + z$ as a single operation, in round-towards-zero mode.

- `__device__ float __fmul_rd` (float x, float y)
Multiply two floating point values in round-down mode.
- `__device__ float __fmul_rn` (float x, float y)
Multiply two floating point values in round-to-nearest-even mode.
- `__device__ float __fmul_ru` (float x, float y)
Multiply two floating point values in round-up mode.
- `__device__ float __fmul_rz` (float x, float y)
Multiply two floating point values in round-towards-zero mode.
- `__device__ float __frcp_rd` (float x)
Compute $\frac{1}{x}$ in round-down mode.
- `__device__ float __frcp_rn` (float x)
Compute $\frac{1}{x}$ in round-to-nearest-even mode.
- `__device__ float __frcp_ru` (float x)
Compute $\frac{1}{x}$ in round-up mode.
- `__device__ float __frcp_rz` (float x)
Compute $\frac{1}{x}$ in round-towards-zero mode.
- `__device__ float __frsqrt_rn` (float x)
Compute $1/\sqrt{x}$ in round-to-nearest-even mode.
- `__device__ float __fsqrt_rd` (float x)
Compute \sqrt{x} in round-down mode.
- `__device__ float __fsqrt_rn` (float x)
Compute \sqrt{x} in round-to-nearest-even mode.
- `__device__ float __fsqrt_ru` (float x)
Compute \sqrt{x} in round-up mode.
- `__device__ float __fsqrt_rz` (float x)
Compute \sqrt{x} in round-towards-zero mode.
- `__device__ __cudart_builtin__ float __log10f` (float x)
Calculate the fast approximate base 10 logarithm of the input argument.
- `__device__ __cudart_builtin__ float __log2f` (float x)
Calculate the fast approximate base 2 logarithm of the input argument.
- `__device__ __cudart_builtin__ float __logf` (float x)
Calculate the fast approximate base e logarithm of the input argument.

- `__device__ __cudart_builtin__ float __powf (float x, float y)`
Calculate the fast approximate of x^y .
- `__device__ float __saturatef (float x)`
Clamp the input argument to $[+0.0, 1.0]$.
- `__device__ __cudart_builtin__ void __sincosf (float x, float *sptr, float *cptr)`
Calculate the fast approximate of sine and cosine of the first input argument.
- `__device__ __cudart_builtin__ float __sinf (float x)`
Calculate the fast approximate sine of the input argument.
- `__device__ __cudart_builtin__ float __tanf (float x)`
Calculate the fast approximate tangent of the input argument.

5.63.1 Detailed Description

This section describes single precision intrinsic functions that are only supported in device code.

5.63.2 Function Documentation

5.63.2.1 `__device__ __cudart_builtin__ float __cosf (float x)`

Calculate the fast approximate cosine of the input argument x , measured in radians.

Returns:

Returns the approximate cosine of x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Input and output in the denormal range is flushed to sign preserving 0.0.

5.63.2.2 `__device__ __cudart_builtin__ float __exp10f (float x)`

Calculate the fast approximate base 10 exponential of the input argument x , 10^x .

Returns:

Returns an approximation to 10^x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Most input and output values around denormal range are flushed to sign preserving 0.0.

5.63.2.3 `__device__ __cudart_builtin__ float __expf (float x)`

Calculate the fast approximate base e exponential of the input argument x , e^x .

Returns:

Returns an approximation to e^x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Most input and output values around denormal range are flushed to sign preserving 0.0.

5.63.2.4 `__device__ float __fadd_rd (float x, float y)`

Compute the sum of x and y in round-down (to negative infinity) mode.

Returns:

Returns $x + y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This operation will never be merged into a single multiply-add instruction.

5.63.2.5 `__device__ float __fadd_rn (float x, float y)`

Compute the sum of x and y in round-to-nearest-even rounding mode.

Returns:

Returns $x + y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This operation will never be merged into a single multiply-add instruction.

5.63.2.6 `__device__ float __fadd_ru (float x, float y)`

Compute the sum of x and y in round-up (to positive infinity) mode.

Returns:

Returns $x + y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This operation will never be merged into a single multiply-add instruction.

5.63.2.7 `__device__ float __fadd_rz (float x, float y)`

Compute the sum of x and y in round-towards-zero mode.

Returns:

Returns $x + y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This operation will never be merged into a single multiply-add instruction.

5.63.2.8 `__device__ float __fdiv_rd (float x, float y)`

Divide two floating point values x by y in round-down (to negative infinity) mode.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.9 `__device__ float __fdiv_rn (float x, float y)`

Divide two floating point values x by y in round-to-nearest-even mode.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.10 `__device__ float __fdiv_ru (float x, float y)`

Divide two floating point values x by y in round-up (to positive infinity) mode.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.11 `__device__ float __fdiv_rz (float x, float y)`

Divide two floating point values x by y in round-towards-zero mode.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.12 `__device__ float __fdivdef (float x, float y)`

Calculate the fast approximate division of x by y .

Returns:

Returns x / y .

- `__fdivdef(∞, y)` returns NaN for $2^{126} < y < 2^{128}$.
- `__fdivdef(x, y)` returns 0 for $2^{126} < y < 2^{128}$ and $x \neq \infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4.

5.63.2.13 `__device__ float __fmaf_rd (float x, float y, float z)`

Computes the value of $x \times y + z$ as a single ternary operation, rounding the result once in round-down (to negative infinity) mode.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+\infty$.
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.14 `__device__ float __fmaf_rn (float x, float y, float z)`

Computes the value of $x \times y + z$ as a single ternary operation, rounding the result once in round-to-nearest-even mode.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+\infty$.
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.15 `__device__ float __fmaf_ru(float x, float y, float z)`

Computes the value of $x \times y + z$ as a single ternary operation, rounding the result once in round-up (to positive infinity) mode.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+\infty$.
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.16 `__device__ float __fmaf_rz(float x, float y, float z)`

Computes the value of $x \times y + z$ as a single ternary operation, rounding the result once in round-towards-zero mode.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+\infty$.
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-\infty$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.17 `__device__ float __fmul_rd(float x, float y)`

Compute the product of x and y in round-down (to negative infinity) mode.

Returns:

Returns $x * y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This operation will never be merged into a single multiply-add instruction.

5.63.2.18 `__device__ float __fmul_rn (float x, float y)`

Compute the product of x and y in round-to-nearest-even mode.

Returns:

Returns $x * y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This operation will never be merged into a single multiply-add instruction.

5.63.2.19 `__device__ float __fmul_ru (float x, float y)`

Compute the product of x and y in round-up (to positive infinity) mode.

Returns:

Returns $x * y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This operation will never be merged into a single multiply-add instruction.

5.63.2.20 `__device__ float __fmul_rz (float x, float y)`

Compute the product of x and y in round-towards-zero mode.

Returns:

Returns $x * y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1. This operation will never be merged into a single multiply-add instruction.

5.63.2.21 `__device__ float __frcp_rd (float x)`

Compute the reciprocal of x in round-down (to negative infinity) mode.

Returns:

Returns $\frac{1}{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.22 `__device__ float __frcp_rn (float x)`

Compute the reciprocal of x in round-to-nearest-even mode.

Returns:

Returns $\frac{1}{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.23 `__device__ float __frcp_ru (float x)`

Compute the reciprocal of x in round-up (to positive infinity) mode.

Returns:

Returns $\frac{1}{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.24 `__device__ float __frcp_rz (float x)`

Compute the reciprocal of x in round-towards-zero mode.

Returns:

Returns $\frac{1}{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.25 `__device__ float __frsqrt_rn (float x)`

Compute the reciprocal square root of x in round-to-nearest-even mode.

Returns:

Returns $1/\sqrt{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.26 `__device__ float __fsqrt_rd (float x)`

Compute the square root of x in round-down (to negative infinity) mode.

Returns:

Returns \sqrt{x} .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.27 `__device__ float __fsqrt_rn (float x)`

Compute the square root of x in round-to-nearest-even mode.

Returns:

Returns \sqrt{x} .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.28 `__device__ float __fsqrt_ru (float x)`

Compute the square root of x in round-up (to positive infinity) mode.

Returns:

Returns \sqrt{x} .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.29 `__device__ float __fsqrt_rz (float x)`

Compute the square root of x in round-towards-zero mode.

Returns:

Returns \sqrt{x} .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-1.

5.63.2.30 `__device__ __cudart_builtin__ float __log10f (float x)`

Calculate the fast approximate base 10 logarithm of the input argument x .

Returns:

Returns an approximation to $\log_{10}(x)$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Most input and output values around denormal range are flushed to sign preserving 0.0.

5.63.2.31 `__device__ __cudart_builtin__ float __log2f (float x)`

Calculate the fast approximate base 2 logarithm of the input argument x .

Returns:

Returns an approximation to $\log_2(x)$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Input and output in the denormal range is flushed to sign preserving 0.0.

5.63.2.32 `__device__ __cudart_builtin__ float __logf (float x)`

Calculate the fast approximate base e logarithm of the input argument x .

Returns:

Returns an approximation to $\log_e(x)$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Most input and output values around denormal range are flushed to sign preserving 0.0.

5.63.2.33 `__device__ __cudart_builtin__ float __powf (float x, float y)`

Calculate the fast approximate of x , the first input argument, raised to the power of y , the second input argument, x^y .

Returns:

Returns an approximation to x^y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Most input and output values around denormal range are flushed to sign preserving 0.0.

5.63.2.34 `__device__ float __saturatef(float x)`

Clamp the input argument x to be within the interval $[+0.0, 1.0]$.

Returns:

- `__saturatef(x)` returns 0 if $x < 0$.
- `__saturatef(x)` returns 1 if $x > 1$.
- `__saturatef(x)` returns x if $0 \leq x \leq 1$.
- `__saturatef(NaN)` returns 0.

5.63.2.35 `__device__ __cudart_builtin__ void __sincosf(float x, float * sptr, float * cptr)`

Calculate the fast approximate of sine and cosine of the first input argument x (measured in radians). The results for sine and cosine are written into the second argument, `sptr`, and, respectively, third argument, `cptr`.

Returns:

- none

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Denorm input/output is flushed to sign preserving 0.0.

5.63.2.36 `__device__ __cudart_builtin__ float __sinf(float x)`

Calculate the fast approximate sine of the input argument x , measured in radians.

Returns:

Returns the approximate sine of x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. Input and output in the denormal range is flushed to sign preserving 0.0.

5.63.2.37 `__device__ __cudart_builtin__ float __tanf(float x)`

Calculate the fast approximate tangent of the input argument x , measured in radians.

Returns:

Returns the approximate tangent of x .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-4. The result is computed as the fast divide of `__sinf()` by `__cosf()`. Denormal input and output are flushed to sign-preserving 0.0 at each step of the computation.

5.64 Double Precision Ininsics

Functions

- `__device__ double __dadd_rd (double x, double y)`
Add two floating point values in round-down mode.
- `__device__ double __dadd_rn (double x, double y)`
Add two floating point values in round-to-nearest-even mode.
- `__device__ double __dadd_ru (double x, double y)`
Add two floating point values in round-up mode.
- `__device__ double __dadd_rz (double x, double y)`
Add two floating point values in round-towards-zero mode.
- `__device__ double __ddiv_rd (double x, double y)`
Divide two floating point values in round-down mode.
- `__device__ double __ddiv_rn (double x, double y)`
Divide two floating point values in round-to-nearest-even mode.
- `__device__ double __ddiv_ru (double x, double y)`
Divide two floating point values in round-up mode.
- `__device__ double __ddiv_rz (double x, double y)`
Divide two floating point values in round-towards-zero mode.
- `__device__ double __dmul_rd (double x, double y)`
Multiply two floating point values in round-down mode.
- `__device__ double __dmul_rn (double x, double y)`
Multiply two floating point values in round-to-nearest-even mode.
- `__device__ double __dmul_ru (double x, double y)`
Multiply two floating point values in round-up mode.
- `__device__ double __dmul_rz (double x, double y)`
Multiply two floating point values in round-towards-zero mode.
- `__device__ double __drcp_rd (double x)`
Compute $\frac{1}{x}$ in round-down mode.
- `__device__ double __drcp_rn (double x)`
Compute $\frac{1}{x}$ in round-to-nearest-even mode.
- `__device__ double __drcp_ru (double x)`
Compute $\frac{1}{x}$ in round-up mode.
- `__device__ double __drcp_rz (double x)`

Compute $\frac{1}{x}$ in round-towards-zero mode.

- `__device__ double __dsqrt_rd (double x)`
Compute \sqrt{x} in round-down mode.
- `__device__ double __dsqrt_rn (double x)`
Compute \sqrt{x} in round-to-nearest-even mode.
- `__device__ double __dsqrt_ru (double x)`
Compute \sqrt{x} in round-up mode.
- `__device__ double __dsqrt_rz (double x)`
Compute \sqrt{x} in round-towards-zero mode.
- `__device__ double __fma_rd (double x, double y, double z)`
Compute $x \times y + z$ as a single operation in round-down mode.
- `__device__ double __fma_rn (double x, double y, double z)`
Compute $x \times y + z$ as a single operation in round-to-nearest-even mode.
- `__device__ double __fma_ru (double x, double y, double z)`
Compute $x \times y + z$ as a single operation in round-up mode.
- `__device__ double __fma_rz (double x, double y, double z)`
Compute $x \times y + z$ as a single operation in round-towards-zero mode.

5.64.1 Detailed Description

This section describes double precision intrinsic functions that are only supported in device code.

5.64.2 Function Documentation

5.64.2.1 `__device__ double __dadd_rd (double x, double y)`

Adds two floating point values x and y in round-down (to negative infinity) mode.

Returns:

Returns $x + y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. This operation will never be merged into a single multiply-add instruction.

5.64.2.2 `__device__ double __dadd_rn (double x, double y)`

Adds two floating point values x and y in round-to-nearest-even mode.

Returns:

Returns $x + y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. This operation will never be merged into a single multiply-add instruction.

5.64.2.3 `__device__ double __dadd_ru (double x, double y)`

Adds two floating point values x and y in round-up (to positive infinity) mode.

Returns:

Returns $x + y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. This operation will never be merged into a single multiply-add instruction.

5.64.2.4 `__device__ double __dadd_rz (double x, double y)`

Adds two floating point values x and y in round-towards-zero mode.

Returns:

Returns $x + y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. This operation will never be merged into a single multiply-add instruction.

5.64.2.5 `__device__ double __ddiv_rd (double x, double y)`

Divides two floating point values x by y in round-down (to negative infinity) mode.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. Requires compute capability ≥ 2.0 .

5.64.2.6 `__device__ double __ddiv_rn (double x, double y)`

Divides two floating point values x by y in round-to-nearest-even mode.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.7 `__device__ double __ddiv_ru (double x, double y)`

Divides two floating point values x by y in round-up (to positive infinity) mode.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.8 `__device__ double __ddiv_rz (double x, double y)`

Divides two floating point values x by y in round-towards-zero mode.

Returns:

Returns x / y .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.9 `__device__ double __dmul_rd (double x, double y)`

Multiplies two floating point values x and y in round-down (to negative infinity) mode.

Returns:

Returns $x * y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
This operation will never be merged into a single multiply-add instruction.

5.64.2.10 `__device__ double __dmul_rn (double x, double y)`

Multiplies two floating point values x and y in round-to-nearest-even mode.

Returns:

Returns $x * y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. This operation will never be merged into a single multiply-add instruction.

5.64.2.11 `__device__ double __dmul_ru (double x, double y)`

Multiplies two floating point values x and y in round-up (to positive infinity) mode.

Returns:

Returns $x * y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. This operation will never be merged into a single multiply-add instruction.

5.64.2.12 `__device__ double __dmul_rz (double x, double y)`

Multiplies two floating point values x and y in round-towards-zero mode.

Returns:

Returns $x * y$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. This operation will never be merged into a single multiply-add instruction.

5.64.2.13 `__device__ double __drcp_rd (double x)`

Compute the reciprocal of x in round-down (to negative infinity) mode.

Returns:

Returns $\frac{1}{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2. Requires compute capability ≥ 2.0 .

5.64.2.14 `__device__ double __drcp_rn (double x)`

Compute the reciprocal of x in round-to-nearest-even mode.

Returns:

Returns $\frac{1}{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.15 `__device__ double __drcp_ru (double x)`

Compute the reciprocal of x in round-up (to positive infinity) mode.

Returns:

Returns $\frac{1}{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.16 `__device__ double __drcp_rz (double x)`

Compute the reciprocal of x in round-towards-zero mode.

Returns:

Returns $\frac{1}{x}$.

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.17 `__device__ double __dsqrt_rd (double x)`

Compute the square root of x in round-down (to negative infinity) mode.

Returns:

Returns \sqrt{x} .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.18 `__device__ double __dsqrt_rn (double x)`

Compute the square root of x in round-to-nearest-even mode.

Returns:

Returns \sqrt{x} .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.19 `__device__ double __dsqrt_ru (double x)`

Compute the square root of x in round-up (to positive infinity) mode.

Returns:

Returns \sqrt{x} .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.20 `__device__ double __dsqrt_rz (double x)`

Compute the square root of x in round-towards-zero mode.

Returns:

Returns \sqrt{x} .

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.
Requires compute capability ≥ 2.0 .

5.64.2.21 `__device__ double __fma_rd (double x, double y, double z)`

Computes the value of $x \times y + z$ as a single ternary operation, rounding the result once in round-down (to negative infinity) mode.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+\infty$
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-\infty$

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.64.2.22 `__device__ double __fma_rn (double x, double y, double z)`

Computes the value of $x \times y + z$ as a single ternary operation, rounding the result once in round-to-nearest-even mode.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+\infty$
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-\infty$

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.64.2.23 `__device__ double __fma_ru (double x, double y, double z)`

Computes the value of $x \times y + z$ as a single ternary operation, rounding the result once in round-up (to positive infinity) mode.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+\infty$
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-\infty$

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.64.2.24 `__device__ double __fma_rz (double x, double y, double z)`

Computes the value of $x \times y + z$ as a single ternary operation, rounding the result once in round-towards-zero mode.

Returns:

Returns the rounded value of $x \times y + z$ as a single operation.

- `fmaf(±∞, ±0, z)` returns NaN.
- `fmaf(±0, ±∞, z)` returns NaN.
- `fmaf(x, y, -∞)` returns NaN if $x \times y$ is an exact $+\infty$
- `fmaf(x, y, +∞)` returns NaN if $x \times y$ is an exact $-\infty$

Note:

For accuracy information for this function see the CUDA C Programming Guide, Appendix C, Table C-2.

5.65 Integer Ininsics

Functions

- `__device__ unsigned int __brev` (unsigned int x)
Reverse the bit order of a 32 bit unsigned integer.
- `__device__ unsigned long long int __brevll` (unsigned long long int x)
Reverse the bit order of a 64 bit unsigned integer.
- `__device__ unsigned int __byte_perm` (unsigned int x, unsigned int y, unsigned int s)
Return selected bytes from two 32 bit unsigned integers.
- `__device__ int __clz` (int x)
Return the number of consecutive high-order zero bits in a 32 bit integer.
- `__device__ int __clzll` (long long int x)
Count the number of consecutive high-order zero bits in a 64 bit integer.
- `__device__ int __ffs` (int x)
Find the position of the least significant bit set to 1 in a 32 bit integer.
- `__device__ int __ffsll` (long long int x)
Find the position of the least significant bit set to 1 in a 64 bit integer.
- `__device__ int __mul24` (int x, int y)
Calculate the least significant 32 bits of the product of the least significant 24 bits of two integers.
- `__device__ long long int __mul64hi` (long long int x, long long int y)
Calculate the most significant 64 bits of the product of the two 64 bit integers.
- `__device__ int __mulhi` (int x, int y)
Calculate the most significant 32 bits of the product of the two 32 bit integers.
- `__device__ int __popc` (unsigned int x)
Count the number of bits that are set to 1 in a 32 bit integer.
- `__device__ int __popc11` (unsigned long long int x)
Count the number of bits that are set to 1 in a 64 bit integer.
- `__device__ unsigned int __sad` (int x, int y, unsigned int z)
Calculate $|x - y| + z$, the sum of absolute difference.
- `__device__ unsigned int __umul24` (unsigned int x, unsigned int y)
Calculate the least significant 32 bits of the product of the least significant 24 bits of two unsigned integers.
- `__device__ unsigned long long int __umul64hi` (unsigned long long int x, unsigned long long int y)
Calculate the most significant 64 bits of the product of the two 64 unsigned bit integers.
- `__device__ unsigned int __umulhi` (unsigned int x, unsigned int y)

Calculate the most significant 32 bits of the product of the two 32 bit unsigned integers.

- `__device__ unsigned int __usad (unsigned int x, unsigned int y, unsigned int z)`

Calculate $|x - y| + z$, the sum of absolute difference.

5.65.1 Detailed Description

This section describes integer intrinsic functions that are only supported in device code.

5.65.2 Function Documentation

5.65.2.1 `__device__ unsigned int __brev (unsigned int x)`

Reverses the bit order of the 32 bit unsigned integer x .

Returns:

Returns the bit-reversed value of x . i.e. bit N of the return value corresponds to bit $31-N$ of x .

5.65.2.2 `__device__ unsigned long long int __brevll (unsigned long long int x)`

Reverses the bit order of the 64 bit unsigned integer x .

Returns:

Returns the bit-reversed value of x . i.e. bit N of the return value corresponds to bit $63-N$ of x .

5.65.2.3 `__device__ unsigned int __byte_perm (unsigned int x, unsigned int y, unsigned int s)`

`byte_perm(x,y,s)` returns a 32-bit integer consisting of four bytes from eight input bytes provided in the two input integers x and y , as specified by a selector, s .

The input bytes are indexed as follows:

```
input [0] = x<0:7>   input [1] = x<8:15>
input [2] = x<16:23> input [3] = x<24:31>
input [4] = y<0:7>   input [5] = y<8:15>
input [6] = y<16:23> input [7] = y<24:31>
```

The selector indices are stored in 4-bit nibbles (with the upper 16-bits of the selector not being used):

```
selector[0] = s<0:3>   selector[1] = s<4:7>
selector[2] = s<8:11> selector[3] = s<12:15>
```

Returns:

The returned value r is computed to be: `result [n] := input [selector [n]]` where `result [n]` is the n th byte of r .

5.65.2.4 `__device__ int __clz (int x)`

Count the number of consecutive leading zero bits, starting at the most significant bit (bit 31) of x .

Returns:

Returns a value between 0 and 32 inclusive representing the number of zero bits.

5.65.2.5 `__device__ int __clzll (long long int x)`

Count the number of consecutive leading zero bits, starting at the most significant bit (bit 63) of x .

Returns:

Returns a value between 0 and 64 inclusive representing the number of zero bits.

5.65.2.6 `__device__ int __ffs (int x)`

Find the position of the first (least significant) bit set to 1 in x , where the least significant bit position is 1.

Returns:

Returns a value between 0 and 32 inclusive representing the position of the first bit set.

- `__ffs(0)` returns 0.

5.65.2.7 `__device__ int __ffsll (long long int x)`

Find the position of the first (least significant) bit set to 1 in x , where the least significant bit position is 1.

Returns:

Returns a value between 0 and 64 inclusive representing the position of the first bit set.

- `__ffsll(0)` returns 0.

5.65.2.8 `__device__ int __mul24 (int x, int y)`

Calculate the least significant 32 bits of the product of the least significant 24 bits of x and y . The high order 8 bits of x and y are ignored.

Returns:

Returns the least significant 32 bits of the product $x * y$.

5.65.2.9 `__device__ long long int __mul64hi (long long int x, long long int y)`

Calculate the most significant 64 bits of the 128-bit product $x * y$, where x and y are 64-bit integers.

Returns:

Returns the most significant 64 bits of the product $x * y$.

5.65.2.10 `__device__ int __mulhi (int x, int y)`

Calculate the most significant 32 bits of the 64-bit product $x * y$, where x and y are 32-bit integers.

Returns:

Returns the most significant 32 bits of the product $x * y$.

5.65.2.11 `__device__ int __popc (unsigned int x)`

Count the number of bits that are set to 1 in x .

Returns:

Returns a value between 0 and 32 inclusive representing the number of set bits.

5.65.2.12 `__device__ int __popcll (unsigned long long int x)`

Count the number of bits that are set to 1 in x .

Returns:

Returns a value between 0 and 64 inclusive representing the number of set bits.

5.65.2.13 `__device__ unsigned int __sad (int x, int y, unsigned int z)`

Calculate $|x - y| + z$, the 32-bit sum of the third argument z plus and the absolute value of the difference between the first argument, x , and second argument, y .

Inputs x and y are signed 32-bit integers, input z is a 32-bit unsigned integer.

Returns:

Returns $|x - y| + z$.

5.65.2.14 `__device__ unsigned int __umul24 (unsigned int x, unsigned int y)`

Calculate the least significant 32 bits of the product of the least significant 24 bits of x and y . The high order 8 bits of x and y are ignored.

Returns:

Returns the least significant 32 bits of the product $x * y$.

5.65.2.15 `__device__ unsigned long long int __umul64hi (unsigned long long int x, unsigned long long int y)`

Calculate the most significant 64 bits of the 128-bit product $x * y$, where x and y are 64-bit unsigned integers.

Returns:

Returns the most significant 64 bits of the product $x * y$.

5.65.2.16 `__device__ unsigned int __umulhi (unsigned int x, unsigned int y)`

Calculate the most significant 32 bits of the 64-bit product $x * y$, where x and y are 32-bit unsigned integers.

Returns:

Returns the most significant 32 bits of the product $x * y$.

5.65.2.17 `__device__ unsigned int __usad (unsigned int x, unsigned int y, unsigned int z)`

Calculate $|x - y| + z$, the 32-bit sum of the third argument z plus and the absolute value of the difference between the first argument, x , and second argument, y .

Inputs x , y , and z are unsigned 32-bit integers.

Returns:

Returns $|x - y| + z$.

5.66 Type Casting Ininsics

Functions

- `__device__ float __double2float_rd (double x)`
Convert a double to a float in round-down mode.
- `__device__ float __double2float_rn (double x)`
Convert a double to a float in round-to-nearest-even mode.
- `__device__ float __double2float_ru (double x)`
Convert a double to a float in round-up mode.
- `__device__ float __double2float_rz (double x)`
Convert a double to a float in round-towards-zero mode.
- `__device__ int __double2hiint (double x)`
Reinterpret high 32 bits in a double as a signed integer.
- `__device__ int __double2int_rd (double x)`
Convert a double to a signed int in round-down mode.
- `__device__ int __double2int_rn (double x)`
Convert a double to a signed int in round-to-nearest-even mode.
- `__device__ int __double2int_ru (double x)`
Convert a double to a signed int in round-up mode.
- `__device__ int __double2int_rz (double)`
Convert a double to a signed int in round-towards-zero mode.
- `__device__ long long int __double2ll_rd (double x)`
Convert a double to a signed 64-bit int in round-down mode.
- `__device__ long long int __double2ll_rn (double x)`
Convert a double to a signed 64-bit int in round-to-nearest-even mode.
- `__device__ long long int __double2ll_ru (double x)`
Convert a double to a signed 64-bit int in round-up mode.
- `__device__ long long int __double2ll_rz (double)`
Convert a double to a signed 64-bit int in round-towards-zero mode.
- `__device__ int __double2loint (double x)`
Reinterpret low 32 bits in a double as a signed integer.
- `__device__ unsigned int __double2uint_rd (double x)`
Convert a double to an unsigned int in round-down mode.
- `__device__ unsigned int __double2uint_rn (double x)`

Convert a double to an unsigned int in round-to-nearest-even mode.

- `__device__ unsigned int __double2uint_ru` (double x)
Convert a double to an unsigned int in round-up mode.
- `__device__ unsigned int __double2uint_rz` (double)
Convert a double to an unsigned int in round-towards-zero mode.
- `__device__ unsigned long long int __double2ull_rd` (double x)
Convert a double to an unsigned 64-bit int in round-down mode.
- `__device__ unsigned long long int __double2ull_rn` (double x)
Convert a double to an unsigned 64-bit int in round-to-nearest-even mode.
- `__device__ unsigned long long int __double2ull_ru` (double x)
Convert a double to an unsigned 64-bit int in round-up mode.
- `__device__ unsigned long long int __double2ull_rz` (double)
Convert a double to an unsigned 64-bit int in round-towards-zero mode.
- `__device__ long long int __double_as_longlong` (double x)
Reinterpret bits in a double as a 64-bit signed integer.
- `__device__ unsigned short __float2half_rn` (float x)
Convert a single-precision float to a half-precision float in round-to-nearest-even mode.
- `__device__ int __float2int_rd` (float x)
Convert a float to a signed integer in round-down mode.
- `__device__ int __float2int_rn` (float x)
Convert a float to a signed integer in round-to-nearest-even mode.
- `__device__ int __float2int_ru` (float)
Convert a float to a signed integer in round-up mode.
- `__device__ int __float2int_rz` (float x)
Convert a float to a signed integer in round-towards-zero mode.
- `__device__ long long int __float2ll_rd` (float x)
Convert a float to a signed 64-bit integer in round-down mode.
- `__device__ long long int __float2ll_rn` (float x)
Convert a float to a signed 64-bit integer in round-to-nearest-even mode.
- `__device__ long long int __float2ll_ru` (float x)
Convert a float to a signed 64-bit integer in round-up mode.
- `__device__ long long int __float2ll_rz` (float x)
Convert a float to a signed 64-bit integer in round-towards-zero mode.

- `__device__ unsigned int __float2uint_rd (float x)`
Convert a float to an unsigned integer in round-down mode.
- `__device__ unsigned int __float2uint_rn (float x)`
Convert a float to an unsigned integer in round-to-nearest-even mode.
- `__device__ unsigned int __float2uint_ru (float x)`
Convert a float to an unsigned integer in round-up mode.
- `__device__ unsigned int __float2uint_rz (float x)`
Convert a float to an unsigned integer in round-towards-zero mode.
- `__device__ unsigned long long int __float2ull_rd (float x)`
Convert a float to an unsigned 64-bit integer in round-down mode.
- `__device__ unsigned long long int __float2ull_rn (float x)`
Convert a float to an unsigned 64-bit integer in round-to-nearest-even mode.
- `__device__ unsigned long long int __float2ull_ru (float x)`
Convert a float to an unsigned 64-bit integer in round-up mode.
- `__device__ unsigned long long int __float2ull_rz (float x)`
Convert a float to an unsigned 64-bit integer in round-towards-zero mode.
- `__device__ int __float_as_int (float x)`
Reinterpret bits in a float as a signed integer.
- `__device__ float __half2float (unsigned short x)`
Convert a half-precision float to a single-precision float in round-to-nearest-even mode.
- `__device__ double __hiloInt2double (int hi, int lo)`
Reinterpret high and low 32-bit integer values as a double.
- `__device__ double __int2double_rn (int x)`
Convert a signed int to a double.
- `__device__ float __int2float_rd (int x)`
Convert a signed integer to a float in round-down mode.
- `__device__ float __int2float_rn (int x)`
Convert a signed integer to a float in round-to-nearest-even mode.
- `__device__ float __int2float_ru (int x)`
Convert a signed integer to a float in round-up mode.
- `__device__ float __int2float_rz (int x)`
Convert a signed integer to a float in round-towards-zero mode.
- `__device__ float __int_as_float (int x)`
Reinterpret bits in an integer as a float.

- `__device__ double __ll2double_rd` (long long int x)
Convert a signed 64-bit int to a double in round-down mode.
- `__device__ double __ll2double_rn` (long long int x)
Convert a signed 64-bit int to a double in round-to-nearest-even mode.
- `__device__ double __ll2double_ru` (long long int x)
Convert a signed 64-bit int to a double in round-up mode.
- `__device__ double __ll2double_rz` (long long int x)
Convert a signed 64-bit int to a double in round-towards-zero mode.
- `__device__ float __ll2float_rd` (long long int x)
Convert a signed integer to a float in round-down mode.
- `__device__ float __ll2float_rn` (long long int x)
Convert a signed 64-bit integer to a float in round-to-nearest-even mode.
- `__device__ float __ll2float_ru` (long long int x)
Convert a signed integer to a float in round-up mode.
- `__device__ float __ll2float_rz` (long long int x)
Convert a signed integer to a float in round-towards-zero mode.
- `__device__ double __longlong_as_double` (long long int x)
Reinterpret bits in a 64-bit signed integer as a double.
- `__device__ double __uint2double_rn` (unsigned int x)
Convert an unsigned int to a double.
- `__device__ float __uint2float_rd` (unsigned int x)
Convert an unsigned integer to a float in round-down mode.
- `__device__ float __uint2float_rn` (unsigned int x)
Convert an unsigned integer to a float in round-to-nearest-even mode.
- `__device__ float __uint2float_ru` (unsigned int x)
Convert an unsigned integer to a float in round-up mode.
- `__device__ float __uint2float_rz` (unsigned int x)
Convert an unsigned integer to a float in round-towards-zero mode.
- `__device__ double __ull2double_rd` (unsigned long long int x)
Convert an unsigned 64-bit int to a double in round-down mode.
- `__device__ double __ull2double_rn` (unsigned long long int x)
Convert an unsigned 64-bit int to a double in round-to-nearest-even mode.
- `__device__ double __ull2double_ru` (unsigned long long int x)

Convert an unsigned 64-bit int to a double in round-up mode.

- `__device__ double __ull2double_rz` (unsigned long long int x)
Convert an unsigned 64-bit int to a double in round-towards-zero mode.
- `__device__ float __ull2float_rd` (unsigned long long int x)
Convert an unsigned integer to a float in round-down mode.
- `__device__ float __ull2float_rn` (unsigned long long int x)
Convert an unsigned integer to a float in round-to-nearest-even mode.
- `__device__ float __ull2float_ru` (unsigned long long int x)
Convert an unsigned integer to a float in round-up mode.
- `__device__ float __ull2float_rz` (unsigned long long int x)
Convert an unsigned integer to a float in round-towards-zero mode.

5.66.1 Detailed Description

This section describes type casting intrinsic functions that are only supported in device code.

5.66.2 Function Documentation

5.66.2.1 `__device__ float __double2float_rd` (double x)

Convert the double-precision floating point value x to a single-precision floating point value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.2 `__device__ float __double2float_rn` (double x)

Convert the double-precision floating point value x to a single-precision floating point value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.3 `__device__ float __double2float_ru` (double x)

Convert the double-precision floating point value x to a single-precision floating point value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.4 `__device__ float __double2float_rz (double x)`

Convert the double-precision floating point value x to a single-precision floating point value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.5 `__device__ int __double2hiint (double x)`

Reinterpret the high 32 bits in the double-precision floating point value x as a signed integer.

Returns:

Returns reinterpreted value.

5.66.2.6 `__device__ int __double2int_rd (double x)`

Convert the double-precision floating point value x to a signed integer value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.7 `__device__ int __double2int_rn (double x)`

Convert the double-precision floating point value x to a signed integer value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.8 `__device__ int __double2int_ru (double x)`

Convert the double-precision floating point value x to a signed integer value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.9 `__device__ int __double2int_rz (double)`

Convert the double-precision floating point value x to a signed integer value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.10 `__device__ long long int __double2ll_rd (double x)`

Convert the double-precision floating point value x to a signed 64-bit integer value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.11 `__device__ long long int __double2ll_rn (double x)`

Convert the double-precision floating point value x to a signed 64-bit integer value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.12 `__device__ long long int __double2ll_ru (double x)`

Convert the double-precision floating point value x to a signed 64-bit integer value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.13 `__device__ long long int __double2ll_rz (double)`

Convert the double-precision floating point value x to a signed 64-bit integer value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.14 `__device__ int __double2loint (double x)`

Reinterpret the low 32 bits in the double-precision floating point value x as a signed integer.

Returns:

Returns reinterpreted value.

5.66.2.15 `__device__ unsigned int __double2uint_rd (double x)`

Convert the double-precision floating point value x to an unsigned integer value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.16 `__device__ unsigned int __double2uint_rn (double x)`

Convert the double-precision floating point value x to an unsigned integer value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.17 `__device__ unsigned int __double2uint_ru (double x)`

Convert the double-precision floating point value x to an unsigned integer value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.18 `__device__ unsigned int __double2uint_rz (double)`

Convert the double-precision floating point value x to an unsigned integer value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.19 `__device__ unsigned long long int __double2ull_rd (double x)`

Convert the double-precision floating point value x to an unsigned 64-bit integer value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.20 `__device__ unsigned long long int __double2ull_rn (double x)`

Convert the double-precision floating point value x to an unsigned 64-bit integer value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.21 `__device__ unsigned long long int __double2ull_ru (double x)`

Convert the double-precision floating point value x to an unsigned 64-bit integer value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.22 `__device__ unsigned long long int __double2ull_rz (double)`

Convert the double-precision floating point value x to an unsigned 64-bit integer value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.23 `__device__ long long int __double_as_longlong (double x)`

Reinterpret the bits in the double-precision floating point value x as a signed 64-bit integer.

Returns:

Returns reinterpreted value.

5.66.2.24 `__device__ unsigned short __float2half_rn (float x)`

Convert the single-precision float value x to a half-precision floating point value represented in `unsigned short` format, in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.25 `__device__ int __float2int_rd (float x)`

Convert the single-precision floating point value x to a signed integer in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.26 `__device__ int __float2int_rn (float x)`

Convert the single-precision floating point value x to a signed integer in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.27 `__device__ int __float2int_ru (float)`

Convert the single-precision floating point value x to a signed integer in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.28 `__device__ int __float2int_rz (float x)`

Convert the single-precision floating point value x to a signed integer in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.29 `__device__ long long int __float2ll_rd (float x)`

Convert the single-precision floating point value x to a signed 64-bit integer in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.30 `__device__ long long int __float2ll_rn (float x)`

Convert the single-precision floating point value x to a signed 64-bit integer in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.31 `__device__ long long int __float2ll_ru (float x)`

Convert the single-precision floating point value x to a signed 64-bit integer in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.32 `__device__ long long int __float2ll_rz (float x)`

Convert the single-precision floating point value x to a signed 64-bit integer in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.33 `__device__ unsigned int __float2uint_rd (float x)`

Convert the single-precision floating point value x to an unsigned integer in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.34 `__device__ unsigned int __float2uint_rn(float x)`

Convert the single-precision floating point value x to an unsigned integer in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.35 `__device__ unsigned int __float2uint_ru(float x)`

Convert the single-precision floating point value x to an unsigned integer in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.36 `__device__ unsigned int __float2uint_rz(float x)`

Convert the single-precision floating point value x to an unsigned integer in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.37 `__device__ unsigned long long int __float2ull_rd(float x)`

Convert the single-precision floating point value x to an unsigned 64-bit integer in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.38 `__device__ unsigned long long int __float2ull_rn(float x)`

Convert the single-precision floating point value x to an unsigned 64-bit integer in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.39 `__device__ unsigned long long int __float2ull_ru(float x)`

Convert the single-precision floating point value x to an unsigned 64-bit integer in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.40 `__device__ unsigned long long int __float2ull_rz (float x)`

Convert the single-precision floating point value x to an unsigned 64-bit integer in round-towards_zero mode.

Returns:

Returns converted value.

5.66.2.41 `__device__ int __float_as_int (float x)`

Reinterpret the bits in the single-precision floating point value x as a signed integer.

Returns:

Returns reinterpreted value.

5.66.2.42 `__device__ float __half2float (unsigned short x)`

Convert the half-precision floating point value x represented in `unsigned short` format to a single-precision floating point value.

Returns:

Returns converted value.

5.66.2.43 `__device__ double __hiloint2double (int hi, int lo)`

Reinterpret the integer value of `hi` as the high 32 bits of a double-precision floating point value and the integer value of `lo` as the low 32 bits of the same double-precision floating point value.

Returns:

Returns reinterpreted value.

5.66.2.44 `__device__ double __int2double_rn (int x)`

Convert the signed integer value x to a double-precision floating point value.

Returns:

Returns converted value.

5.66.2.45 `__device__ float __int2float_rd (int x)`

Convert the signed integer value x to a single-precision floating point value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.46 `__device__ float __int2float_rn (int x)`

Convert the signed integer value x to a single-precision floating point value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.47 `__device__ float __int2float_ru (int x)`

Convert the signed integer value x to a single-precision floating point value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.48 `__device__ float __int2float_rz (int x)`

Convert the signed integer value x to a single-precision floating point value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.49 `__device__ float __int_as_float (int x)`

Reinterpret the bits in the signed integer value x as a single-precision floating point value.

Returns:

Returns reinterpreted value.

5.66.2.50 `__device__ double __ll2double_rd (long long int x)`

Convert the signed 64-bit integer value x to a double-precision floating point value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.51 `__device__ double __ll2double_rn (long long int x)`

Convert the signed 64-bit integer value x to a double-precision floating point value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.52 `__device__ double __ll2double_ru (long long int x)`

Convert the signed 64-bit integer value x to a double-precision floating point value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.53 `__device__ double __ll2double_rz (long long int x)`

Convert the signed 64-bit integer value x to a double-precision floating point value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.54 `__device__ float __ll2float_rd (long long int x)`

Convert the signed integer value x to a single-precision floating point value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.55 `__device__ float __ll2float_rn (long long int x)`

Convert the signed 64-bit integer value x to a single-precision floating point value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.56 `__device__ float __ll2float_ru (long long int x)`

Convert the signed integer value x to a single-precision floating point value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.57 `__device__ float __ll2float_rz (long long int x)`

Convert the signed integer value x to a single-precision floating point value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.58 `__device__ double __longlong_as_double (long long int x)`

Reinterpret the bits in the 64-bit signed integer value x as a double-precision floating point value.

Returns:

Returns reinterpreted value.

5.66.2.59 `__device__ double __uint2double_rn (unsigned int x)`

Convert the unsigned integer value x to a double-precision floating point value.

Returns:

Returns converted value.

5.66.2.60 `__device__ float __uint2float_rd (unsigned int x)`

Convert the unsigned integer value x to a single-precision floating point value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.61 `__device__ float __uint2float_rn (unsigned int x)`

Convert the unsigned integer value x to a single-precision floating point value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.62 `__device__ float __uint2float_ru (unsigned int x)`

Convert the unsigned integer value x to a single-precision floating point value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.63 `__device__ float __uint2float_rz (unsigned int x)`

Convert the unsigned integer value x to a single-precision floating point value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.64 `__device__ double __ull2double_rd (unsigned long long int x)`

Convert the unsigned 64-bit integer value x to a double-precision floating point value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.65 `__device__ double __ull2double_rn (unsigned long long int x)`

Convert the unsigned 64-bit integer value x to a double-precision floating point value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.66 `__device__ double __ull2double_ru (unsigned long long int x)`

Convert the unsigned 64-bit integer value x to a double-precision floating point value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.67 `__device__ double __ull2double_rz (unsigned long long int x)`

Convert the unsigned 64-bit integer value x to a double-precision floating point value in round-towards-zero mode.

Returns:

Returns converted value.

5.66.2.68 `__device__ float __ull2float_rd (unsigned long long int x)`

Convert the unsigned integer value x to a single-precision floating point value in round-down (to negative infinity) mode.

Returns:

Returns converted value.

5.66.2.69 `__device__ float __ull2float_rn (unsigned long long int x)`

Convert the unsigned integer value x to a single-precision floating point value in round-to-nearest-even mode.

Returns:

Returns converted value.

5.66.2.70 `__device__ float __ull2float_ru (unsigned long long int x)`

Convert the unsigned integer value x to a single-precision floating point value in round-up (to positive infinity) mode.

Returns:

Returns converted value.

5.66.2.71 `__device__ float __ull2float_rz (unsigned long long int x)`

Convert the unsigned integer value x to a single-precision floating point value in round-towards-zero mode.

Returns:

Returns converted value.

Chapter 6

Data Structure Documentation

6.1 CUDA_ARRAY3D_DESCRIPTOR_st Struct Reference

Data Fields

- [size_t Depth](#)
- [unsigned int Flags](#)
- [CUarray_format Format](#)
- [size_t Height](#)
- [unsigned int NumChannels](#)
- [size_t Width](#)

6.1.1 Detailed Description

3D array descriptor

6.1.2 Field Documentation

6.1.2.1 `size_t CUDA_ARRAY3D_DESCRIPTOR_st::Depth`

Depth of 3D array

6.1.2.2 `unsigned int CUDA_ARRAY3D_DESCRIPTOR_st::Flags`

Flags

6.1.2.3 `CUarray_format CUDA_ARRAY3D_DESCRIPTOR_st::Format`

Array format

6.1.2.4 `size_t CUDA_ARRAY3D_DESCRIPTOR_st::Height`

Height of 3D array

6.1.2.5 unsigned int CUDA_ARRAY3D_DESCRIPTOR_st::NumChannels

Channels per array element

6.1.2.6 size_t CUDA_ARRAY3D_DESCRIPTOR_st::Width

Width of 3D array

6.2 CUDA_ARRAY_DESCRIPTOR_st Struct Reference

Data Fields

- [CUarray_format](#) Format
- [size_t](#) Height
- [unsigned int](#) NumChannels
- [size_t](#) Width

6.2.1 Detailed Description

Array descriptor

6.2.2 Field Documentation

6.2.2.1 [CUarray_format](#) CUDA_ARRAY_DESCRIPTOR_st::Format

Array format

6.2.2.2 [size_t](#) CUDA_ARRAY_DESCRIPTOR_st::Height

Height of array

6.2.2.3 [unsigned int](#) CUDA_ARRAY_DESCRIPTOR_st::NumChannels

Channels per array element

6.2.2.4 [size_t](#) CUDA_ARRAY_DESCRIPTOR_st::Width

Width of array

6.3 CUDA_MEMCPY2D_st Struct Reference

Data Fields

- [CUarray dstArray](#)
- [CUdeviceptr dstDevice](#)
- [void * dstHost](#)
- [CUmemorytype dstMemoryType](#)
- [size_t dstPitch](#)
- [size_t dstXInBytes](#)
- [size_t dstY](#)
- [size_t Height](#)
- [CUarray srcArray](#)
- [CUdeviceptr srcDevice](#)
- [const void * srcHost](#)
- [CUmemorytype srcMemoryType](#)
- [size_t srcPitch](#)
- [size_t srcXInBytes](#)
- [size_t srcY](#)
- [size_t WidthInBytes](#)

6.3.1 Detailed Description

2D memory copy parameters

6.3.2 Field Documentation

6.3.2.1 CUarray CUDA_MEMCPY2D_st::dstArray

Destination array reference

6.3.2.2 CUdeviceptr CUDA_MEMCPY2D_st::dstDevice

Destination device pointer

6.3.2.3 void* CUDA_MEMCPY2D_st::dstHost

Destination host pointer

6.3.2.4 CUmemorytype CUDA_MEMCPY2D_st::dstMemoryType

Destination memory type (host, device, array)

6.3.2.5 size_t CUDA_MEMCPY2D_st::dstPitch

Destination pitch (ignored when dst is array)

6.3.2.6 size_t CUDA_MEMCPY2D_st::dstXInBytes

Destination X in bytes

6.3.2.7 size_t CUDA_MEMCPY2D_st::dstY

Destination Y

6.3.2.8 size_t CUDA_MEMCPY2D_st::Height

Height of 2D memory copy

6.3.2.9 CUarray CUDA_MEMCPY2D_st::srcArray

Source array reference

6.3.2.10 CUdeviceptr CUDA_MEMCPY2D_st::srcDevice

Source device pointer

6.3.2.11 const void* CUDA_MEMCPY2D_st::srcHost

Source host pointer

6.3.2.12 CUmemorytype CUDA_MEMCPY2D_st::srcMemoryType

Source memory type (host, device, array)

6.3.2.13 size_t CUDA_MEMCPY2D_st::srcPitch

Source pitch (ignored when src is array)

6.3.2.14 size_t CUDA_MEMCPY2D_st::srcXInBytes

Source X in bytes

6.3.2.15 size_t CUDA_MEMCPY2D_st::srcY

Source Y

6.3.2.16 size_t CUDA_MEMCPY2D_st::WidthInBytes

Width of 2D memory copy in bytes

6.4 CUDA_MEMCPY3D_PEER_st Struct Reference

Data Fields

- [size_t Depth](#)
- [CUarray dstArray](#)
- [CUcontext dstContext](#)
- [CUdeviceptr dstDevice](#)
- [size_t dstHeight](#)
- [void * dstHost](#)
- [size_t dstLOD](#)
- [CUmemorytype dstMemoryType](#)
- [size_t dstPitch](#)
- [size_t dstXInBytes](#)
- [size_t dstY](#)
- [size_t dstZ](#)
- [size_t Height](#)
- [CUarray srcArray](#)
- [CUcontext srcContext](#)
- [CUdeviceptr srcDevice](#)
- [size_t srcHeight](#)
- [const void * srcHost](#)
- [size_t srcLOD](#)
- [CUmemorytype srcMemoryType](#)
- [size_t srcPitch](#)
- [size_t srcXInBytes](#)
- [size_t srcY](#)
- [size_t srcZ](#)
- [size_t WidthInBytes](#)

6.4.1 Detailed Description

3D memory cross-context copy parameters

6.4.2 Field Documentation

6.4.2.1 [size_t CUDA_MEMCPY3D_PEER_st::Depth](#)

Depth of 3D memory copy

6.4.2.2 [CUarray CUDA_MEMCPY3D_PEER_st::dstArray](#)

Destination array reference

6.4.2.3 [CUcontext CUDA_MEMCPY3D_PEER_st::dstContext](#)

Destination context (ignored with [dstMemoryType](#) is [CU_MEMORYTYPE_ARRAY](#))

6.4.2.4 CUdeviceptr CUDA_MEMCPY3D_PEER_st::dstDevice

Destination device pointer

6.4.2.5 size_t CUDA_MEMCPY3D_PEER_st::dstHeight

Destination height (ignored when dst is array; may be 0 if Depth==1)

6.4.2.6 void* CUDA_MEMCPY3D_PEER_st::dstHost

Destination host pointer

6.4.2.7 size_t CUDA_MEMCPY3D_PEER_st::dstLOD

Destination LOD

6.4.2.8 CUmemorytype CUDA_MEMCPY3D_PEER_st::dstMemoryType

Destination memory type (host, device, array)

6.4.2.9 size_t CUDA_MEMCPY3D_PEER_st::dstPitch

Destination pitch (ignored when dst is array)

6.4.2.10 size_t CUDA_MEMCPY3D_PEER_st::dstXInBytes

Destination X in bytes

6.4.2.11 size_t CUDA_MEMCPY3D_PEER_st::dstY

Destination Y

6.4.2.12 size_t CUDA_MEMCPY3D_PEER_st::dstZ

Destination Z

6.4.2.13 size_t CUDA_MEMCPY3D_PEER_st::Height

Height of 3D memory copy

6.4.2.14 CUarray CUDA_MEMCPY3D_PEER_st::srcArray

Source array reference

6.4.2.15 CUcontext CUDA_MEMCPY3D_PEER_st::srcContext

Source context (ignored with srcMemoryType is [CU_MEMORYTYPE_ARRAY](#))

6.4.2.16 CUdeviceptr CUDA_MEMCPY3D_PEER_st::srcDevice

Source device pointer

6.4.2.17 size_t CUDA_MEMCPY3D_PEER_st::srcHeight

Source height (ignored when src is array; may be 0 if Depth==1)

6.4.2.18 const void* CUDA_MEMCPY3D_PEER_st::srcHost

Source host pointer

6.4.2.19 size_t CUDA_MEMCPY3D_PEER_st::srcLOD

Source LOD

6.4.2.20 CUmemorytype CUDA_MEMCPY3D_PEER_st::srcMemoryType

Source memory type (host, device, array)

6.4.2.21 size_t CUDA_MEMCPY3D_PEER_st::srcPitch

Source pitch (ignored when src is array)

6.4.2.22 size_t CUDA_MEMCPY3D_PEER_st::srcXInBytes

Source X in bytes

6.4.2.23 size_t CUDA_MEMCPY3D_PEER_st::srcY

Source Y

6.4.2.24 size_t CUDA_MEMCPY3D_PEER_st::srcZ

Source Z

6.4.2.25 size_t CUDA_MEMCPY3D_PEER_st::WidthInBytes

Width of 3D memory copy in bytes

6.5 CUDA_MEMCPY3D_st Struct Reference

Data Fields

- [size_t Depth](#)
- [CUarray dstArray](#)
- [CUdeviceptr dstDevice](#)
- [size_t dstHeight](#)
- [void * dstHost](#)
- [size_t dstLOD](#)
- [CUmemorytype dstMemoryType](#)
- [size_t dstPitch](#)
- [size_t dstXInBytes](#)
- [size_t dstY](#)
- [size_t dstZ](#)
- [size_t Height](#)
- [void * reserved0](#)
- [void * reserved1](#)
- [CUarray srcArray](#)
- [CUdeviceptr srcDevice](#)
- [size_t srcHeight](#)
- [const void * srcHost](#)
- [size_t srcLOD](#)
- [CUmemorytype srcMemoryType](#)
- [size_t srcPitch](#)
- [size_t srcXInBytes](#)
- [size_t srcY](#)
- [size_t srcZ](#)
- [size_t WidthInBytes](#)

6.5.1 Detailed Description

3D memory copy parameters

6.5.2 Field Documentation

6.5.2.1 [size_t CUDA_MEMCPY3D_st::Depth](#)

Depth of 3D memory copy

6.5.2.2 [CUarray CUDA_MEMCPY3D_st::dstArray](#)

Destination array reference

6.5.2.3 [CUdeviceptr CUDA_MEMCPY3D_st::dstDevice](#)

Destination device pointer

6.5.2.4 size_t CUDA_MEMCPY3D_st::dstHeight

Destination height (ignored when dst is array; may be 0 if Depth==1)

6.5.2.5 void* CUDA_MEMCPY3D_st::dstHost

Destination host pointer

6.5.2.6 size_t CUDA_MEMCPY3D_st::dstLOD

Destination LOD

6.5.2.7 CUmemorytype CUDA_MEMCPY3D_st::dstMemoryType

Destination memory type (host, device, array)

6.5.2.8 size_t CUDA_MEMCPY3D_st::dstPitch

Destination pitch (ignored when dst is array)

6.5.2.9 size_t CUDA_MEMCPY3D_st::dstXInBytes

Destination X in bytes

6.5.2.10 size_t CUDA_MEMCPY3D_st::dstY

Destination Y

6.5.2.11 size_t CUDA_MEMCPY3D_st::dstZ

Destination Z

6.5.2.12 size_t CUDA_MEMCPY3D_st::Height

Height of 3D memory copy

6.5.2.13 void* CUDA_MEMCPY3D_st::reserved0

Must be NULL

6.5.2.14 void* CUDA_MEMCPY3D_st::reserved1

Must be NULL

6.5.2.15 CUarray CUDA_MEMCPY3D_st::srcArray

Source array reference

6.5.2.16 CUdeviceptr CUDA_MEMCPY3D_st::srcDevice

Source device pointer

6.5.2.17 size_t CUDA_MEMCPY3D_st::srcHeight

Source height (ignored when src is array; may be 0 if Depth==1)

6.5.2.18 const void* CUDA_MEMCPY3D_st::srcHost

Source host pointer

6.5.2.19 size_t CUDA_MEMCPY3D_st::srcLOD

Source LOD

6.5.2.20 CUmemorytype CUDA_MEMCPY3D_st::srcMemoryType

Source memory type (host, device, array)

6.5.2.21 size_t CUDA_MEMCPY3D_st::srcPitch

Source pitch (ignored when src is array)

6.5.2.22 size_t CUDA_MEMCPY3D_st::srcXInBytes

Source X in bytes

6.5.2.23 size_t CUDA_MEMCPY3D_st::srcY

Source Y

6.5.2.24 size_t CUDA_MEMCPY3D_st::srcZ

Source Z

6.5.2.25 size_t CUDA_MEMCPY3D_st::WidthInBytes

Width of 3D memory copy in bytes

6.6 CUDA_POINTER_ATTRIBUTE_P2P_TOKENS_st Struct Reference

6.6.1 Detailed Description

GPU Direct v3 tokens

6.7 CUDA_RESOURCE_DESC_st Struct Reference

Data Fields

- unsigned int `flags`
- `CUresourcetype` `resType`
- `CUdeviceptr` `devPtr`
- `CUarray_format` `format`
- `CUarray` `hArray`
- `size_t` `height`
- `CUmipmappedArray` `hMipmappedArray`
- unsigned int `numChannels`
- `size_t` `pitchInBytes`
- `size_t` `sizeInBytes`
- `size_t` `width`

6.7.1 Detailed Description

CUDA Resource descriptor

6.7.2 Field Documentation

6.7.2.1 `CUdeviceptr` `CUDA_RESOURCE_DESC_st::devPtr`

Device pointer

6.7.2.2 `unsigned int` `CUDA_RESOURCE_DESC_st::flags`

Flags (must be zero)

6.7.2.3 `CUarray_format` `CUDA_RESOURCE_DESC_st::format`

Array format

6.7.2.4 `CUarray` `CUDA_RESOURCE_DESC_st::hArray`

CUDA array

6.7.2.5 `size_t` `CUDA_RESOURCE_DESC_st::height`

Height of the array in elements

6.7.2.6 `CUmipmappedArray` `CUDA_RESOURCE_DESC_st::hMipmappedArray`

CUDA mipmapped array

6.7.2.7 unsigned int CUDA_RESOURCE_DESC_st::numChannels

Channels per array element

6.7.2.8 size_t CUDA_RESOURCE_DESC_st::pitchInBytes

Pitch between two rows in bytes

6.7.2.9 CUresourcetype CUDA_RESOURCE_DESC_st::resType

Resource type

6.7.2.10 size_t CUDA_RESOURCE_DESC_st::sizeInBytes

Size in bytes

6.7.2.11 size_t CUDA_RESOURCE_DESC_st::width

Width of the array in elements

6.8 CUDA_RESOURCE_VIEW_DESC_st Struct Reference

Data Fields

- size_t [depth](#)
- unsigned int [firstLayer](#)
- unsigned int [firstMipmapLevel](#)
- [CUresourceViewFormat](#) format
- size_t [height](#)
- unsigned int [lastLayer](#)
- unsigned int [lastMipmapLevel](#)
- size_t [width](#)

6.8.1 Detailed Description

Resource view descriptor

6.8.2 Field Documentation

6.8.2.1 size_t CUDA_RESOURCE_VIEW_DESC_st::depth

Depth of the resource view

6.8.2.2 unsigned int CUDA_RESOURCE_VIEW_DESC_st::firstLayer

First layer index

6.8.2.3 unsigned int CUDA_RESOURCE_VIEW_DESC_st::firstMipmapLevel

First defined mipmap level

6.8.2.4 CUresourceViewFormat CUDA_RESOURCE_VIEW_DESC_st::format

Resource view format

6.8.2.5 size_t CUDA_RESOURCE_VIEW_DESC_st::height

Height of the resource view

6.8.2.6 unsigned int CUDA_RESOURCE_VIEW_DESC_st::lastLayer

Last layer index

6.8.2.7 unsigned int CUDA_RESOURCE_VIEW_DESC_st::lastMipmapLevel

Last defined mipmap level

6.8.2.8 `size_t CUDA_RESOURCE_VIEW_DESC_st::width`

Width of the resource view

6.9 CUDA_TEXTURE_DESC_st Struct Reference

Data Fields

- [CUaddress_mode](#) addressMode [3]
- [CUfilter_mode](#) filterMode
- unsigned int flags
- unsigned int maxAnisotropy
- float maxMipmapLevelClamp
- float minMipmapLevelClamp
- [CUfilter_mode](#) mipmapFilterMode
- float mipmapLevelBias

6.9.1 Detailed Description

Texture descriptor

6.9.2 Field Documentation

6.9.2.1 CUaddress_mode CUDA_TEXTURE_DESC_st::addressMode[3]

Address modes

6.9.2.2 CUfilter_mode CUDA_TEXTURE_DESC_st::filterMode

Filter mode

6.9.2.3 unsigned int CUDA_TEXTURE_DESC_st::flags

Flags

6.9.2.4 unsigned int CUDA_TEXTURE_DESC_st::maxAnisotropy

Maximum anistropy ratio

6.9.2.5 float CUDA_TEXTURE_DESC_st::maxMipmapLevelClamp

Mipmap maximum level clamp

6.9.2.6 float CUDA_TEXTURE_DESC_st::minMipmapLevelClamp

Mipmap minimum level clamp

6.9.2.7 CUfilter_mode CUDA_TEXTURE_DESC_st::mipmapFilterMode

Mipmap filter mode

6.9.2.8 float CUDA_TEXTURE_DESC_st::mipmapLevelBias

Mipmap level bias

6.10 cudaChannelFormatDesc Struct Reference

Data Fields

- enum [cudaChannelFormatKind](#) f
- int w
- int x
- int y
- int z

6.10.1 Detailed Description

CUDA Channel format descriptor

6.10.2 Field Documentation

6.10.2.1 enum [cudaChannelFormatKind](#) [cudaChannelFormatDesc::f](#)

Channel format kind

6.10.2.2 int [cudaChannelFormatDesc::w](#)

w

6.10.2.3 int [cudaChannelFormatDesc::x](#)

x

6.10.2.4 int [cudaChannelFormatDesc::y](#)

y

6.10.2.5 int [cudaChannelFormatDesc::z](#)

z

6.11 cudaDeviceProp Struct Reference

Data Fields

- int [asyncEngineCount](#)
- int [canMapHostMemory](#)
- int [clockRate](#)
- int [computeMode](#)
- int [concurrentKernels](#)
- int [deviceOverlap](#)
- int [ECCEEnabled](#)
- int [integrated](#)
- int [kernelExecTimeoutEnabled](#)
- int [l2CacheSize](#)
- int [major](#)
- int [maxGridSize](#) [3]
- int [maxSurface1D](#)
- int [maxSurface1DLayered](#) [2]
- int [maxSurface2D](#) [2]
- int [maxSurface2DLayered](#) [3]
- int [maxSurface3D](#) [3]
- int [maxSurfaceCubemap](#)
- int [maxSurfaceCubemapLayered](#) [2]
- int [maxTexture1D](#)
- int [maxTexture1DLayered](#) [2]
- int [maxTexture1DLinear](#)
- int [maxTexture1DMipmap](#)
- int [maxTexture2D](#) [2]
- int [maxTexture2DGather](#) [2]
- int [maxTexture2DLayered](#) [3]
- int [maxTexture2DLinear](#) [3]
- int [maxTexture2DMipmap](#) [2]
- int [maxTexture3D](#) [3]
- int [maxTextureCubemap](#)
- int [maxTextureCubemapLayered](#) [2]
- int [maxThreadsDim](#) [3]
- int [maxThreadsPerBlock](#)
- int [maxThreadsPerMultiProcessor](#)
- int [memoryBusWidth](#)
- int [memoryClockRate](#)
- size_t [memPitch](#)
- int [minor](#)
- int [multiProcessorCount](#)
- char [name](#) [256]
- int [pciBusID](#)
- int [pciDeviceID](#)
- int [pciDomainID](#)
- int [regsPerBlock](#)
- size_t [sharedMemPerBlock](#)
- size_t [surfaceAlignment](#)

- int [tccDriver](#)
- size_t [textureAlignment](#)
- size_t [texturePitchAlignment](#)
- size_t [totalConstMem](#)
- size_t [totalGlobalMem](#)
- int [unifiedAddressing](#)
- int [warpSize](#)

6.11.1 Detailed Description

CUDA device properties

6.11.2 Field Documentation

6.11.2.1 int cudaDeviceProp::asyncEngineCount

Number of asynchronous engines

6.11.2.2 int cudaDeviceProp::canMapHostMemory

Device can map host memory with [cudaHostAlloc/cudaHostGetDevicePointer](#)

6.11.2.3 int cudaDeviceProp::clockRate

Clock frequency in kilohertz

6.11.2.4 int cudaDeviceProp::computeMode

Compute mode (See [cudaComputeMode](#))

6.11.2.5 int cudaDeviceProp::concurrentKernels

Device can possibly execute multiple kernels concurrently

6.11.2.6 int cudaDeviceProp::deviceOverlap

Device can concurrently copy memory and execute a kernel. Deprecated. Use instead [asyncEngineCount](#).

6.11.2.7 int cudaDeviceProp::ECCEnabled

Device has ECC support enabled

6.11.2.8 int cudaDeviceProp::integrated

Device is integrated as opposed to discrete

6.11.2.9 int cudaDeviceProp::kernelExecTimeoutEnabled

Specified whether there is a run time limit on kernels

6.11.2.10 int cudaDeviceProp::l2CacheSize

Size of L2 cache in bytes

6.11.2.11 int cudaDeviceProp::major

Major compute capability

6.11.2.12 int cudaDeviceProp::maxGridSize[3]

Maximum size of each dimension of a grid

6.11.2.13 int cudaDeviceProp::maxSurface1D

Maximum 1D surface size

6.11.2.14 int cudaDeviceProp::maxSurface1DLayered[2]

Maximum 1D layered surface dimensions

6.11.2.15 int cudaDeviceProp::maxSurface2D[2]

Maximum 2D surface dimensions

6.11.2.16 int cudaDeviceProp::maxSurface2DLayered[3]

Maximum 2D layered surface dimensions

6.11.2.17 int cudaDeviceProp::maxSurface3D[3]

Maximum 3D surface dimensions

6.11.2.18 int cudaDeviceProp::maxSurfaceCubemap

Maximum Cubemap surface dimensions

6.11.2.19 int cudaDeviceProp::maxSurfaceCubemapLayered[2]

Maximum Cubemap layered surface dimensions

6.11.2.20 int cudaDeviceProp::maxTexture1D

Maximum 1D texture size

6.11.2.21 int cudaDeviceProp::maxTexture1DLayered[2]

Maximum 1D layered texture dimensions

6.11.2.22 int cudaDeviceProp::maxTexture1DLinear

Maximum size for 1D textures bound to linear memory

6.11.2.23 int cudaDeviceProp::maxTexture1DMipmap

Maximum 1D mipmapped texture size

6.11.2.24 int cudaDeviceProp::maxTexture2D[2]

Maximum 2D texture dimensions

6.11.2.25 int cudaDeviceProp::maxTexture2DGather[2]

Maximum 2D texture dimensions if texture gather operations have to be performed

6.11.2.26 int cudaDeviceProp::maxTexture2DLayered[3]

Maximum 2D layered texture dimensions

6.11.2.27 int cudaDeviceProp::maxTexture2DLinear[3]

Maximum dimensions (width, height, pitch) for 2D textures bound to pitched memory

6.11.2.28 int cudaDeviceProp::maxTexture2DMipmap[2]

Maximum 2D mipmapped texture dimensions

6.11.2.29 int cudaDeviceProp::maxTexture3D[3]

Maximum 3D texture dimensions

6.11.2.30 int cudaDeviceProp::maxTextureCubemap

Maximum Cubemap texture dimensions

6.11.2.31 int cudaDeviceProp::maxTextureCubemapLayered[2]

Maximum Cubemap layered texture dimensions

6.11.2.32 int cudaDeviceProp::maxThreadsDim[3]

Maximum size of each dimension of a block

6.11.2.33 int cudaDeviceProp::maxThreadsPerBlock

Maximum number of threads per block

6.11.2.34 int cudaDeviceProp::maxThreadsPerMultiProcessor

Maximum resident threads per multiprocessor

6.11.2.35 int cudaDeviceProp::memoryBusWidth

Global memory bus width in bits

6.11.2.36 int cudaDeviceProp::memoryClockRate

Peak memory clock frequency in kilohertz

6.11.2.37 size_t cudaDeviceProp::memPitch

Maximum pitch in bytes allowed by memory copies

6.11.2.38 int cudaDeviceProp::minor

Minor compute capability

6.11.2.39 int cudaDeviceProp::multiProcessorCount

Number of multiprocessors on device

6.11.2.40 char cudaDeviceProp::name[256]

ASCII string identifying device

6.11.2.41 int cudaDeviceProp::pciBusID

PCI bus ID of the device

6.11.2.42 int cudaDeviceProp::pciDeviceID

PCI device ID of the device

6.11.2.43 int cudaDeviceProp::pciDomainID

PCI domain ID of the device

6.11.2.44 int cudaDeviceProp::regsPerBlock

32-bit registers available per block

6.11.2.45 size_t cudaDeviceProp::sharedMemPerBlock

Shared memory available per block in bytes

6.11.2.46 size_t cudaDeviceProp::surfaceAlignment

Alignment requirements for surfaces

6.11.2.47 int cudaDeviceProp::tccDriver

1 if device is a Tesla device using TCC driver, 0 otherwise

6.11.2.48 size_t cudaDeviceProp::textureAlignment

Alignment requirement for textures

6.11.2.49 size_t cudaDeviceProp::texturePitchAlignment

Pitch alignment requirement for texture references bound to pitched memory

6.11.2.50 size_t cudaDeviceProp::totalConstMem

Constant memory available on device in bytes

6.11.2.51 size_t cudaDeviceProp::totalGlobalMem

Global memory available on device in bytes

6.11.2.52 int cudaDeviceProp::unifiedAddressing

Device shares a unified address space with the host

6.11.2.53 int cudaDeviceProp::warpSize

Warp size in threads

6.12 cudaExtent Struct Reference

Data Fields

- [size_t depth](#)
- [size_t height](#)
- [size_t width](#)

6.12.1 Detailed Description

CUDA extent

See also:

[make_cudaExtent](#)

6.12.2 Field Documentation

6.12.2.1 `size_t cudaExtent::depth`

Depth in elements

6.12.2.2 `size_t cudaExtent::height`

Height in elements

6.12.2.3 `size_t cudaExtent::width`

Width in elements when referring to array memory, in bytes when referring to linear memory

6.13 cudaFuncAttributes Struct Reference

Data Fields

- int [binaryVersion](#)
- size_t [constSizeBytes](#)
- size_t [localSizeBytes](#)
- int [maxThreadsPerBlock](#)
- int [numRegs](#)
- int [ptxVersion](#)
- size_t [sharedSizeBytes](#)

6.13.1 Detailed Description

CUDA function attributes

6.13.2 Field Documentation

6.13.2.1 int cudaFuncAttributes::binaryVersion

The binary architecture version for which the function was compiled. This value is the major binary version * 10 + the minor binary version, so a binary version 1.3 function would return the value 13.

6.13.2.2 size_t cudaFuncAttributes::constSizeBytes

The size in bytes of user-allocated constant memory required by this function.

6.13.2.3 size_t cudaFuncAttributes::localSizeBytes

The size in bytes of local memory used by each thread of this function.

6.13.2.4 int cudaFuncAttributes::maxThreadsPerBlock

The maximum number of threads per block, beyond which a launch of the function would fail. This number depends on both the function and the device on which the function is currently loaded.

6.13.2.5 int cudaFuncAttributes::numRegs

The number of registers used by each thread of this function.

6.13.2.6 int cudaFuncAttributes::ptxVersion

The PTX virtual architecture version for which the function was compiled. This value is the major PTX version * 10 + the minor PTX version, so a PTX version 1.3 function would return the value 13.

6.13.2.7 `size_t cudaFuncAttributes::sharedSizeBytes`

The size in bytes of statically-allocated shared memory per block required by this function. This does not include dynamically-allocated shared memory requested by the user at runtime.

6.14 cudaIpcEventHandle_st Struct Reference

6.14.1 Detailed Description

CUDA IPC event handle

6.15 `cudaIpcMemHandle_st` Struct Reference

6.15.1 Detailed Description

CUDA IPC memory handle

6.16 cudaMemcpy3DParms Struct Reference

Data Fields

- [cudaArray_t dstArray](#)
- struct [cudaPos dstPos](#)
- struct [cudaPitchedPtr dstPtr](#)
- struct [cudaExtent extent](#)
- enum [cudaMemcpyKind kind](#)
- [cudaArray_t srcArray](#)
- struct [cudaPos srcPos](#)
- struct [cudaPitchedPtr srcPtr](#)

6.16.1 Detailed Description

CUDA 3D memory copying parameters

6.16.2 Field Documentation

6.16.2.1 `cudaArray_t cudaMemcpy3DParms::dstArray`

Destination memory address

6.16.2.2 `struct cudaPos cudaMemcpy3DParms::dstPos` [read]

Destination position offset

6.16.2.3 `struct cudaPitchedPtr cudaMemcpy3DParms::dstPtr` [read]

Pitched destination memory address

6.16.2.4 `struct cudaExtent cudaMemcpy3DParms::extent` [read]

Requested memory copy size

6.16.2.5 `enum cudaMemcpyKind cudaMemcpy3DParms::kind`

Type of transfer

6.16.2.6 `cudaArray_t cudaMemcpy3DParms::srcArray`

Source memory address

6.16.2.7 `struct cudaPos cudaMemcpy3DParms::srcPos` [read]

Source position offset

6.16.2.8 struct cudaPitchedPtr cudaMemcpy3DParms::srcPtr [read]

Pitched source memory address

6.17 cudaMemcpy3DPeerParms Struct Reference

Data Fields

- [cudaArray_t dstArray](#)
- [int dstDevice](#)
- [struct cudaPos dstPos](#)
- [struct cudaPitchedPtr dstPtr](#)
- [struct cudaExtent extent](#)
- [cudaArray_t srcArray](#)
- [int srcDevice](#)
- [struct cudaPos srcPos](#)
- [struct cudaPitchedPtr srcPtr](#)

6.17.1 Detailed Description

CUDA 3D cross-device memory copying parameters

6.17.2 Field Documentation

6.17.2.1 [cudaArray_t cudaMemcpy3DPeerParms::dstArray](#)

Destination memory address

6.17.2.2 [int cudaMemcpy3DPeerParms::dstDevice](#)

Destination device

6.17.2.3 [struct cudaPos cudaMemcpy3DPeerParms::dstPos](#) [read]

Destination position offset

6.17.2.4 [struct cudaPitchedPtr cudaMemcpy3DPeerParms::dstPtr](#) [read]

Pitched destination memory address

6.17.2.5 [struct cudaExtent cudaMemcpy3DPeerParms::extent](#) [read]

Requested memory copy size

6.17.2.6 [cudaArray_t cudaMemcpy3DPeerParms::srcArray](#)

Source memory address

6.17.2.7 [int cudaMemcpy3DPeerParms::srcDevice](#)

Source device

6.17.2.8 `struct cudaPos cudaMemcpy3DPeerParms::srcPos` [read]

Source position offset

6.17.2.9 `struct cudaPitchedPtr cudaMemcpy3DPeerParms::srcPtr` [read]

Pitched source memory address

6.18 cudaPitchedPtr Struct Reference

Data Fields

- [size_t pitch](#)
- [void * ptr](#)
- [size_t xsize](#)
- [size_t ysize](#)

6.18.1 Detailed Description

CUDA Pitched memory pointer

See also:

[make_cudaPitchedPtr](#)

6.18.2 Field Documentation

6.18.2.1 `size_t cudaPitchedPtr::pitch`

Pitch of allocated memory in bytes

6.18.2.2 `void* cudaPitchedPtr::ptr`

Pointer to allocated memory

6.18.2.3 `size_t cudaPitchedPtr::xsize`

Logical width of allocation in elements

6.18.2.4 `size_t cudaPitchedPtr::ysize`

Logical height of allocation in elements

6.19 cudaPointerAttributes Struct Reference

Data Fields

- int [device](#)
- void * [devicePointer](#)
- void * [hostPointer](#)
- enum [cudaMemoryType](#) [memoryType](#)

6.19.1 Detailed Description

CUDA pointer attributes

6.19.2 Field Documentation

6.19.2.1 int cudaPointerAttributes::device

The device against which the memory was allocated or registered. If the memory type is [cudaMemoryTypeDevice](#) then this identifies the device on which the memory referred physically resides. If the memory type is [cudaMemoryTypeHost](#) then this identifies the device which was current when the memory was allocated or registered (and if that device is deinitialized then this allocation will vanish with that device's state).

6.19.2.2 void* cudaPointerAttributes::devicePointer

The address which may be dereferenced on the current device to access the memory or NULL if no such address exists.

6.19.2.3 void* cudaPointerAttributes::hostPointer

The address which may be dereferenced on the host to access the memory or NULL if no such address exists.

6.19.2.4 enum cudaMemoryType cudaPointerAttributes::memoryType

The physical location of the memory, [cudaMemoryTypeHost](#) or [cudaMemoryTypeDevice](#).

6.20 cudaPos Struct Reference

Data Fields

- [size_t x](#)
- [size_t y](#)
- [size_t z](#)

6.20.1 Detailed Description

CUDA 3D position

See also:

[make_cudaPos](#)

6.20.2 Field Documentation

6.20.2.1 size_t cudaPos::x

x

6.20.2.2 size_t cudaPos::y

y

6.20.2.3 size_t cudaPos::z

z

6.21 cudaResourceDesc Struct Reference

Data Fields

- enum [cudaResourceType](#) resType
- [cudaArray_t](#) array
- struct [cudaChannelFormatDesc](#) desc
- void * devPtr
- size_t height
- [cudaMipmappedArray_t](#) mipmap
- size_t pitchInBytes
- size_t sizeInBytes
- size_t width

6.21.1 Detailed Description

CUDA resource descriptor

6.21.2 Field Documentation

6.21.2.1 [cudaArray_t](#) [cudaResourceDesc::array](#)

CUDA array

6.21.2.2 [struct cudaChannelFormatDesc](#) [cudaResourceDesc::desc](#) [read]

Channel descriptor

6.21.2.3 [void*](#) [cudaResourceDesc::devPtr](#)

Device pointer

6.21.2.4 [size_t](#) [cudaResourceDesc::height](#)

Height of the array in elements

6.21.2.5 [cudaMipmappedArray_t](#) [cudaResourceDesc::mipmap](#)

CUDA mipmapped array

6.21.2.6 [size_t](#) [cudaResourceDesc::pitchInBytes](#)

Pitch between two rows in bytes

6.21.2.7 [enum cudaResourceType](#) [cudaResourceDesc::resType](#)

Resource type

6.21.2.8 `size_t cudaResourceDesc::sizeInBytes`

Size in bytes

6.21.2.9 `size_t cudaResourceDesc::width`

Width of the array in elements

6.22 cudaResourceViewDesc Struct Reference

Data Fields

- size_t [depth](#)
- unsigned int [firstLayer](#)
- unsigned int [firstMipmapLevel](#)
- enum [cudaResourceViewFormat](#) [format](#)
- size_t [height](#)
- unsigned int [lastLayer](#)
- unsigned int [lastMipmapLevel](#)
- size_t [width](#)

6.22.1 Detailed Description

CUDA resource view descriptor

6.22.2 Field Documentation

6.22.2.1 size_t cudaResourceViewDesc::depth

Depth of the resource view

6.22.2.2 unsigned int cudaResourceViewDesc::firstLayer

First layer index

6.22.2.3 unsigned int cudaResourceViewDesc::firstMipmapLevel

First defined mipmap level

6.22.2.4 enum cudaResourceViewFormat cudaResourceViewDesc::format

Resource view format

6.22.2.5 size_t cudaResourceViewDesc::height

Height of the resource view

6.22.2.6 unsigned int cudaResourceViewDesc::lastLayer

Last layer index

6.22.2.7 unsigned int cudaResourceViewDesc::lastMipmapLevel

Last defined mipmap level

6.22.2.8 size_t cudaResourceViewDesc::width

Width of the resource view

6.23 cudaTextureDesc Struct Reference

Data Fields

- enum [cudaTextureAddressMode](#) `addressMode` [3]
- enum [cudaTextureFilterMode](#) `filterMode`
- unsigned int `maxAnisotropy`
- float `maxMipmapLevelClamp`
- float `minMipmapLevelClamp`
- enum [cudaTextureFilterMode](#) `mipmapFilterMode`
- float `mipmapLevelBias`
- int `normalizedCoords`
- enum [cudaTextureReadMode](#) `readMode`
- int `sRGB`

6.23.1 Detailed Description

CUDA texture descriptor

6.23.2 Field Documentation

6.23.2.1 enum `cudaTextureAddressMode` `cudaTextureDesc::addressMode`[3]

Texture address mode for up to 3 dimensions

6.23.2.2 enum `cudaTextureFilterMode` `cudaTextureDesc::filterMode`

Texture filter mode

6.23.2.3 unsigned int `cudaTextureDesc::maxAnisotropy`

Limit to the anisotropy ratio

6.23.2.4 float `cudaTextureDesc::maxMipmapLevelClamp`

Upper end of the mipmap level range to clamp access to

6.23.2.5 float `cudaTextureDesc::minMipmapLevelClamp`

Lower end of the mipmap level range to clamp access to

6.23.2.6 enum `cudaTextureFilterMode` `cudaTextureDesc::mipmapFilterMode`

Mipmap filter mode

6.23.2.7 float `cudaTextureDesc::mipmapLevelBias`

Offset applied to the supplied mipmap level

6.23.2.8 int cudaTextureDesc::normalizedCoords

Indicates whether texture reads are normalized or not

6.23.2.9 enum cudaTextureReadMode cudaTextureDesc::readMode

Texture read mode

6.23.2.10 int cudaTextureDesc::sRGB

Perform sRGB->linear conversion during texture read

6.24 CUdevprop_st Struct Reference

Data Fields

- int [clockRate](#)
- int [maxGridSize](#) [3]
- int [maxThreadsDim](#) [3]
- int [maxThreadsPerBlock](#)
- int [memPitch](#)
- int [regsPerBlock](#)
- int [sharedMemPerBlock](#)
- int [SIMDWidth](#)
- int [textureAlign](#)
- int [totalConstantMemory](#)

6.24.1 Detailed Description

Legacy device properties

6.24.2 Field Documentation

6.24.2.1 int CUdevprop_st::clockRate

Clock frequency in kilohertz

6.24.2.2 int CUdevprop_st::maxGridSize[3]

Maximum size of each dimension of a grid

6.24.2.3 int CUdevprop_st::maxThreadsDim[3]

Maximum size of each dimension of a block

6.24.2.4 int CUdevprop_st::maxThreadsPerBlock

Maximum number of threads per block

6.24.2.5 int CUdevprop_st::memPitch

Maximum pitch in bytes allowed by memory copies

6.24.2.6 int CUdevprop_st::regsPerBlock

32-bit registers available per block

6.24.2.7 int CUdevprop_st::sharedMemPerBlock

Shared memory available per block in bytes

6.24.2.8 int CUdevprop_st::SIMDWidth

Warp size in threads

6.24.2.9 int CUdevprop_st::textureAlign

Alignment requirement for textures

6.24.2.10 int CUdevprop_st::totalConstantMemory

Constant memory available on device in bytes

6.25 CUipcEventHandle_st Struct Reference

6.25.1 Detailed Description

CUDA IPC event handle

6.26 CUipcMemHandle_st Struct Reference

6.26.1 Detailed Description

CUDA IPC mem handle

6.27 surfaceReference Struct Reference

Data Fields

- struct [cudaChannelFormatDesc](#) `channelDesc`

6.27.1 Detailed Description

CUDA Surface reference

6.27.2 Field Documentation

6.27.2.1 struct `cudaChannelFormatDesc` `surfaceReference::channelDesc` [read]

Channel descriptor for surface reference

6.28 textureReference Struct Reference

Data Fields

- enum [cudaTextureAddressMode](#) `addressMode` [3]
- struct [cudaChannelFormatDesc](#) `channelDesc`
- enum [cudaTextureFilterMode](#) `filterMode`
- unsigned int [maxAnisotropy](#)
- float [maxMipmapLevelClamp](#)
- float [minMipmapLevelClamp](#)
- enum [cudaTextureFilterMode](#) `mipmapFilterMode`
- float [mipmapLevelBias](#)
- int [normalized](#)
- int [sRGB](#)

6.28.1 Detailed Description

CUDA texture reference

6.28.2 Field Documentation

6.28.2.1 enum [cudaTextureAddressMode](#) `textureReference::addressMode`[3]

Texture address mode for up to 3 dimensions

6.28.2.2 struct [cudaChannelFormatDesc](#) `textureReference::channelDesc` [read]

Channel descriptor for the texture reference

6.28.2.3 enum [cudaTextureFilterMode](#) `textureReference::filterMode`

Texture filter mode

6.28.2.4 unsigned int `textureReference::maxAnisotropy`

Limit to the anisotropy ratio

6.28.2.5 float `textureReference::maxMipmapLevelClamp`

Upper end of the mipmap level range to clamp access to

6.28.2.6 float `textureReference::minMipmapLevelClamp`

Lower end of the mipmap level range to clamp access to

6.28.2.7 enum [cudaTextureFilterMode](#) `textureReference::mipmapFilterMode`

Mipmap filter mode

6.28.2.8 float textureReference::mipmapLevelBias

Offset applied to the supplied mipmap level

6.28.2.9 int textureReference::normalized

Indicates whether texture reads are normalized or not

6.28.2.10 int textureReference::sRGB

Perform sRGB->linear conversion during texture read

Index

__brev
 CUDA_MATH_INTRINSIC_INT, 514

__brevll
 CUDA_MATH_INTRINSIC_INT, 514

__byte_perm
 CUDA_MATH_INTRINSIC_INT, 514

__clz
 CUDA_MATH_INTRINSIC_INT, 514

__clzll
 CUDA_MATH_INTRINSIC_INT, 515

__cosf
 CUDA_MATH_INTRINSIC_SINGLE, 495

__dadd_rd
 CUDA_MATH_INTRINSIC_DOUBLE, 506

__dadd_rn
 CUDA_MATH_INTRINSIC_DOUBLE, 506

__dadd_ru
 CUDA_MATH_INTRINSIC_DOUBLE, 507

__dadd_rz
 CUDA_MATH_INTRINSIC_DOUBLE, 507

__ddiv_rd
 CUDA_MATH_INTRINSIC_DOUBLE, 507

__ddiv_rn
 CUDA_MATH_INTRINSIC_DOUBLE, 507

__ddiv_ru
 CUDA_MATH_INTRINSIC_DOUBLE, 508

__ddiv_rz
 CUDA_MATH_INTRINSIC_DOUBLE, 508

__dmul_rd
 CUDA_MATH_INTRINSIC_DOUBLE, 508

__dmul_rn
 CUDA_MATH_INTRINSIC_DOUBLE, 508

__dmul_ru
 CUDA_MATH_INTRINSIC_DOUBLE, 509

__dmul_rz
 CUDA_MATH_INTRINSIC_DOUBLE, 509

__double2float_rd
 CUDA_MATH_INTRINSIC_CAST, 522

__double2float_rn
 CUDA_MATH_INTRINSIC_CAST, 522

__double2float_ru
 CUDA_MATH_INTRINSIC_CAST, 522

__double2float_rz
 CUDA_MATH_INTRINSIC_CAST, 522

__double2hiint
 CUDA_MATH_INTRINSIC_CAST, 523

__double2int_rd
 CUDA_MATH_INTRINSIC_CAST, 523

__double2int_rn
 CUDA_MATH_INTRINSIC_CAST, 523

__double2int_ru
 CUDA_MATH_INTRINSIC_CAST, 523

__double2int_rz
 CUDA_MATH_INTRINSIC_CAST, 523

__double2ll_rd
 CUDA_MATH_INTRINSIC_CAST, 523

__double2ll_rn
 CUDA_MATH_INTRINSIC_CAST, 524

__double2ll_ru
 CUDA_MATH_INTRINSIC_CAST, 524

__double2ll_rz
 CUDA_MATH_INTRINSIC_CAST, 524

__double2loint
 CUDA_MATH_INTRINSIC_CAST, 524

__double2uint_rd
 CUDA_MATH_INTRINSIC_CAST, 524

__double2uint_rn
 CUDA_MATH_INTRINSIC_CAST, 524

__double2uint_ru
 CUDA_MATH_INTRINSIC_CAST, 525

__double2uint_rz
 CUDA_MATH_INTRINSIC_CAST, 525

__double2ull_rd
 CUDA_MATH_INTRINSIC_CAST, 525

__double2ull_rn
 CUDA_MATH_INTRINSIC_CAST, 525

__double2ull_ru
 CUDA_MATH_INTRINSIC_CAST, 525

__double2ull_rz
 CUDA_MATH_INTRINSIC_CAST, 525

__double_as_longlong
 CUDA_MATH_INTRINSIC_CAST, 526

__drcep_rd
 CUDA_MATH_INTRINSIC_DOUBLE, 509

__drcep_rn
 CUDA_MATH_INTRINSIC_DOUBLE, 509

__drcep_ru
 CUDA_MATH_INTRINSIC_DOUBLE, 510

__drcep_rz
 CUDA_MATH_INTRINSIC_DOUBLE, 510

- __dsqrt_rd
 - CUDA_MATH_INTRINSIC_DOUBLE, 510
- __dsqrt_rn
 - CUDA_MATH_INTRINSIC_DOUBLE, 510
- __dsqrt_ru
 - CUDA_MATH_INTRINSIC_DOUBLE, 511
- __dsqrt_rz
 - CUDA_MATH_INTRINSIC_DOUBLE, 511
- __exp10f
 - CUDA_MATH_INTRINSIC_SINGLE, 495
- __expf
 - CUDA_MATH_INTRINSIC_SINGLE, 495
- __fadd_rd
 - CUDA_MATH_INTRINSIC_SINGLE, 496
- __fadd_rn
 - CUDA_MATH_INTRINSIC_SINGLE, 496
- __fadd_ru
 - CUDA_MATH_INTRINSIC_SINGLE, 496
- __fadd_rz
 - CUDA_MATH_INTRINSIC_SINGLE, 496
- __fdiv_rd
 - CUDA_MATH_INTRINSIC_SINGLE, 497
- __fdiv_rn
 - CUDA_MATH_INTRINSIC_SINGLE, 497
- __fdiv_ru
 - CUDA_MATH_INTRINSIC_SINGLE, 497
- __fdiv_rz
 - CUDA_MATH_INTRINSIC_SINGLE, 497
- __fdividef
 - CUDA_MATH_INTRINSIC_SINGLE, 498
- __ffs
 - CUDA_MATH_INTRINSIC_INT, 515
- __ffsll
 - CUDA_MATH_INTRINSIC_INT, 515
- __float2half_rn
 - CUDA_MATH_INTRINSIC_CAST, 526
- __float2int_rd
 - CUDA_MATH_INTRINSIC_CAST, 526
- __float2int_rn
 - CUDA_MATH_INTRINSIC_CAST, 526
- __float2int_ru
 - CUDA_MATH_INTRINSIC_CAST, 526
- __float2int_rz
 - CUDA_MATH_INTRINSIC_CAST, 526
- __float2ll_rd
 - CUDA_MATH_INTRINSIC_CAST, 527
- __float2ll_rn
 - CUDA_MATH_INTRINSIC_CAST, 527
- __float2ll_ru
 - CUDA_MATH_INTRINSIC_CAST, 527
- __float2ll_rz
 - CUDA_MATH_INTRINSIC_CAST, 527
- __float2uint_rd
 - CUDA_MATH_INTRINSIC_CAST, 527
- __float2uint_rn
 - CUDA_MATH_INTRINSIC_CAST, 527
- __float2uint_ru
 - CUDA_MATH_INTRINSIC_CAST, 528
- __float2uint_rz
 - CUDA_MATH_INTRINSIC_CAST, 528
- __float2ull_rd
 - CUDA_MATH_INTRINSIC_CAST, 528
- __float2ull_rn
 - CUDA_MATH_INTRINSIC_CAST, 528
- __float2ull_ru
 - CUDA_MATH_INTRINSIC_CAST, 528
- __float2ull_rz
 - CUDA_MATH_INTRINSIC_CAST, 528
- __float_as_int
 - CUDA_MATH_INTRINSIC_CAST, 529
- __fma_rd
 - CUDA_MATH_INTRINSIC_DOUBLE, 511
- __fma_rn
 - CUDA_MATH_INTRINSIC_DOUBLE, 511
- __fma_ru
 - CUDA_MATH_INTRINSIC_DOUBLE, 512
- __fma_rz
 - CUDA_MATH_INTRINSIC_DOUBLE, 512
- __fmaf_rd
 - CUDA_MATH_INTRINSIC_SINGLE, 498
- __fmaf_rn
 - CUDA_MATH_INTRINSIC_SINGLE, 498
- __fmaf_ru
 - CUDA_MATH_INTRINSIC_SINGLE, 499
- __fmaf_rz
 - CUDA_MATH_INTRINSIC_SINGLE, 499
- __fmul_rd
 - CUDA_MATH_INTRINSIC_SINGLE, 499
- __fmul_rn
 - CUDA_MATH_INTRINSIC_SINGLE, 499
- __fmul_ru
 - CUDA_MATH_INTRINSIC_SINGLE, 500
- __fmul_rz
 - CUDA_MATH_INTRINSIC_SINGLE, 500
- __frcp_rd
 - CUDA_MATH_INTRINSIC_SINGLE, 500
- __frcp_rn
 - CUDA_MATH_INTRINSIC_SINGLE, 500
- __frcp_ru
 - CUDA_MATH_INTRINSIC_SINGLE, 501
- __frcp_rz
 - CUDA_MATH_INTRINSIC_SINGLE, 501
- __frsqrt_rn
 - CUDA_MATH_INTRINSIC_SINGLE, 501
- __fsqrt_rd
 - CUDA_MATH_INTRINSIC_SINGLE, 501
- __fsqrt_rn
 - CUDA_MATH_INTRINSIC_SINGLE, 502

- __fsqrt_ru
 - CUDA_MATH_INTRINSIC_SINGLE, 502
- __fsqrt_rz
 - CUDA_MATH_INTRINSIC_SINGLE, 502
- __half2float
 - CUDA_MATH_INTRINSIC_CAST, 529
- __hiloint2double
 - CUDA_MATH_INTRINSIC_CAST, 529
- __int2double_rn
 - CUDA_MATH_INTRINSIC_CAST, 529
- __int2float_rd
 - CUDA_MATH_INTRINSIC_CAST, 529
- __int2float_rn
 - CUDA_MATH_INTRINSIC_CAST, 529
- __int2float_ru
 - CUDA_MATH_INTRINSIC_CAST, 530
- __int2float_rz
 - CUDA_MATH_INTRINSIC_CAST, 530
- __int_as_float
 - CUDA_MATH_INTRINSIC_CAST, 530
- __ll2double_rd
 - CUDA_MATH_INTRINSIC_CAST, 530
- __ll2double_rn
 - CUDA_MATH_INTRINSIC_CAST, 530
- __ll2double_ru
 - CUDA_MATH_INTRINSIC_CAST, 530
- __ll2double_rz
 - CUDA_MATH_INTRINSIC_CAST, 531
- __ll2float_rd
 - CUDA_MATH_INTRINSIC_CAST, 531
- __ll2float_rn
 - CUDA_MATH_INTRINSIC_CAST, 531
- __ll2float_ru
 - CUDA_MATH_INTRINSIC_CAST, 531
- __ll2float_rz
 - CUDA_MATH_INTRINSIC_CAST, 531
- __log10f
 - CUDA_MATH_INTRINSIC_SINGLE, 502
- __log2f
 - CUDA_MATH_INTRINSIC_SINGLE, 503
- __logf
 - CUDA_MATH_INTRINSIC_SINGLE, 503
- __longlong_as_double
 - CUDA_MATH_INTRINSIC_CAST, 531
- __mul24
 - CUDA_MATH_INTRINSIC_INT, 515
- __mul64hi
 - CUDA_MATH_INTRINSIC_INT, 515
- __mulhi
 - CUDA_MATH_INTRINSIC_INT, 515
- __popc
 - CUDA_MATH_INTRINSIC_INT, 516
- __popcll
 - CUDA_MATH_INTRINSIC_INT, 516
- __powf
 - CUDA_MATH_INTRINSIC_SINGLE, 503
- __sad
 - CUDA_MATH_INTRINSIC_INT, 516
- __saturatef
 - CUDA_MATH_INTRINSIC_SINGLE, 503
- __sincosf
 - CUDA_MATH_INTRINSIC_SINGLE, 504
- __sinf
 - CUDA_MATH_INTRINSIC_SINGLE, 504
- __tanf
 - CUDA_MATH_INTRINSIC_SINGLE, 504
- __uint2double_rn
 - CUDA_MATH_INTRINSIC_CAST, 532
- __uint2float_rd
 - CUDA_MATH_INTRINSIC_CAST, 532
- __uint2float_rn
 - CUDA_MATH_INTRINSIC_CAST, 532
- __uint2float_ru
 - CUDA_MATH_INTRINSIC_CAST, 532
- __uint2float_rz
 - CUDA_MATH_INTRINSIC_CAST, 532
- __ull2double_rd
 - CUDA_MATH_INTRINSIC_CAST, 532
- __ull2double_rn
 - CUDA_MATH_INTRINSIC_CAST, 533
- __ull2double_ru
 - CUDA_MATH_INTRINSIC_CAST, 533
- __ull2double_rz
 - CUDA_MATH_INTRINSIC_CAST, 533
- __ull2float_rd
 - CUDA_MATH_INTRINSIC_CAST, 533
- __ull2float_rn
 - CUDA_MATH_INTRINSIC_CAST, 533
- __ull2float_ru
 - CUDA_MATH_INTRINSIC_CAST, 533
- __ull2float_rz
 - CUDA_MATH_INTRINSIC_CAST, 534
- __umul24
 - CUDA_MATH_INTRINSIC_INT, 516
- __umul64hi
 - CUDA_MATH_INTRINSIC_INT, 516
- __umulhi
 - CUDA_MATH_INTRINSIC_INT, 516
- __usad
 - CUDA_MATH_INTRINSIC_INT, 517
- acos
 - CUDA_MATH_DOUBLE, 471
- acosf
 - CUDA_MATH_SINGLE, 445
- acosh
 - CUDA_MATH_DOUBLE, 472
- acoshf
 -

- CUDA_MATH_SINGLE, 445
- addressMode
 - CUDA_TEXTURE_DESC_st, 551
 - cudaTextureDesc, 576
 - textureReference, 583
- array
 - cudaResourceDesc, 572
- asin
 - CUDA_MATH_DOUBLE, 472
- asinf
 - CUDA_MATH_SINGLE, 445
- asinh
 - CUDA_MATH_DOUBLE, 472
- asinhf
 - CUDA_MATH_SINGLE, 445
- asyncEngineCount
 - cudaDeviceProp, 555
- atan
 - CUDA_MATH_DOUBLE, 472
- atan2
 - CUDA_MATH_DOUBLE, 473
- atan2f
 - CUDA_MATH_SINGLE, 445
- atanf
 - CUDA_MATH_SINGLE, 446
- atanh
 - CUDA_MATH_DOUBLE, 473
- atanhf
 - CUDA_MATH_SINGLE, 446
- binaryVersion
 - cudaFuncAttributes, 561
- C++ API Routines, 170
- canMapHostMemory
 - cudaDeviceProp, 555
- cbrt
 - CUDA_MATH_DOUBLE, 473
- cbrtf
 - CUDA_MATH_SINGLE, 446
- ceil
 - CUDA_MATH_DOUBLE, 473
- ceilf
 - CUDA_MATH_SINGLE, 446
- channelDesc
 - surfaceReference, 582
 - textureReference, 583
- clockRate
 - cudaDeviceProp, 555
 - CUdevprop_st, 578
- computeMode
 - cudaDeviceProp, 555
- concurrentKernels
 - cudaDeviceProp, 555
- constSizeBytes
 - cudaFuncAttributes, 561
- Context Management, 252
- copysign
 - CUDA_MATH_DOUBLE, 474
- copysignf
 - CUDA_MATH_SINGLE, 447
- cos
 - CUDA_MATH_DOUBLE, 474
- cosf
 - CUDA_MATH_SINGLE, 447
- cosh
 - CUDA_MATH_DOUBLE, 474
- coshf
 - CUDA_MATH_SINGLE, 447
- cospi
 - CUDA_MATH_DOUBLE, 474
- cospif
 - CUDA_MATH_SINGLE, 447
- CU_AD_FORMAT_FLOAT
 - CUDA_TYPES, 230
- CU_AD_FORMAT_HALF
 - CUDA_TYPES, 230
- CU_AD_FORMAT_SIGNED_INT16
 - CUDA_TYPES, 230
- CU_AD_FORMAT_SIGNED_INT32
 - CUDA_TYPES, 230
- CU_AD_FORMAT_SIGNED_INT8
 - CUDA_TYPES, 230
- CU_AD_FORMAT_UNSIGNED_INT16
 - CUDA_TYPES, 230
- CU_AD_FORMAT_UNSIGNED_INT32
 - CUDA_TYPES, 230
- CU_AD_FORMAT_UNSIGNED_INT8
 - CUDA_TYPES, 230
- CU_COMPUTEMODE_DEFAULT
 - CUDA_TYPES, 230
- CU_COMPUTEMODE_EXCLUSIVE
 - CUDA_TYPES, 230
- CU_COMPUTEMODE_EXCLUSIVE_PROCESS
 - CUDA_TYPES, 230
- CU_COMPUTEMODE_PROHIBITED
 - CUDA_TYPES, 230
- CU_CTX_BLOCKING_SYNC
 - CUDA_TYPES, 230
- CU_CTX_LMEM_RESIZE_TO_MAX
 - CUDA_TYPES, 230
- CU_CTX_MAP_HOST
 - CUDA_TYPES, 230
- CU_CTX_SCHED_AUTO
 - CUDA_TYPES, 230
- CU_CTX_SCHED_BLOCKING_SYNC
 - CUDA_TYPES, 230
- CU_CTX_SCHED_SPIN

- CUDA_TYPES, [230](#)
- CU_CTX_SCHED_YIELD
 - CUDA_TYPES, [230](#)
- CU_CUBEMAP_FACE_NEGATIVE_X
 - CUDA_TYPES, [229](#)
- CU_CUBEMAP_FACE_NEGATIVE_Y
 - CUDA_TYPES, [229](#)
- CU_CUBEMAP_FACE_NEGATIVE_Z
 - CUDA_TYPES, [229](#)
- CU_CUBEMAP_FACE_POSITIVE_X
 - CUDA_TYPES, [229](#)
- CU_CUBEMAP_FACE_POSITIVE_Y
 - CUDA_TYPES, [229](#)
- CU_CUBEMAP_FACE_POSITIVE_Z
 - CUDA_TYPES, [229](#)
- CU_D3D10_DEVICE_LIST_ALL
 - CUDA_D3D10, [415](#)
- CU_D3D10_DEVICE_LIST_CURRENT_FRAME
 - CUDA_D3D10, [415](#)
- CU_D3D10_DEVICE_LIST_NEXT_FRAME
 - CUDA_D3D10, [415](#)
- CU_D3D11_DEVICE_LIST_ALL
 - CUDA_D3D11, [430](#)
- CU_D3D11_DEVICE_LIST_CURRENT_FRAME
 - CUDA_D3D11, [430](#)
- CU_D3D11_DEVICE_LIST_NEXT_FRAME
 - CUDA_D3D11, [430](#)
- CU_D3D9_DEVICE_LIST_ALL
 - CUDA_D3D9, [400](#)
- CU_D3D9_DEVICE_LIST_CURRENT_FRAME
 - CUDA_D3D9, [400](#)
- CU_D3D9_DEVICE_LIST_NEXT_FRAME
 - CUDA_D3D9, [400](#)
- CU_DEVICE_ATTRIBUTE_ASYNC_ENGINE_-
COUNT
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_CAN_MAP_HOST_-
MEMORY
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_CAN_TEX2D_GATHER
 - CUDA_TYPES, [235](#)
- CU_DEVICE_ATTRIBUTE_CLOCK_RATE
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_COMPUTE_-
CAPABILITY_MAJOR
 - CUDA_TYPES, [236](#)
- CU_DEVICE_ATTRIBUTE_COMPUTE_-
CAPABILITY_MINOR
 - CUDA_TYPES, [236](#)
- CU_DEVICE_ATTRIBUTE_COMPUTE_MODE
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_CONCURRENT_-
KERNELS
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_ECC_ENABLED
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_GLOBAL_MEMORY_-
BUS_WIDTH
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_GPU_OVERLAP
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_INTEGRATED
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_KERNEL_EXEC_-
TIMEOUT
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_L2_CACHE_SIZE
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_X
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Y
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Z
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_X
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Y
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Z
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_PITCH
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_REGISTERS_-
PER_BLOCK
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_SHARED_-
MEMORY_PER_BLOCK
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_THREADS_PER_-
BLOCK
 - CUDA_TYPES, [233](#)
- CU_DEVICE_ATTRIBUTE_MAX_THREADS_PER_-
MULTIPROCESSOR
 - CUDA_TYPES, [234](#)
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE1D_LAYERED_LAYERS
 - CUDA_TYPES, [235](#)
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE1D_LAYERED_WIDTH
 - CUDA_TYPES, [235](#)
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE1D_WIDTH
 - CUDA_TYPES, [235](#)
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_HEIGHT
 - CUDA_TYPES, [235](#)
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_LAYERED_HEIGHT

- CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_LAYERED_LAYERS
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_LAYERED_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE3D_DEPTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE3D_HEIGHT
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE3D_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACECUBEMAP_LAYERED_LAYERS
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACECUBEMAP_LAYERED_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACECUBEMAP_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_LAYERED_LAYERS
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_LAYERED_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_LINEAR_WIDTH
CUDA_TYPES, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_MIPMAPPED_WIDTH
CUDA_TYPES, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_WIDTH
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_ARRAY_HEIGHT
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_ARRAY_NUMSLICES
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_ARRAY_WIDTH
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_GATHER_HEIGHT
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_GATHER_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_HEIGHT
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LAYERED_HEIGHT
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LAYERED_LAYERS
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LAYERED_WIDTH
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LINEAR_HEIGHT
CUDA_TYPES, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LINEAR_PITCH
CUDA_TYPES, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LINEAR_WIDTH
CUDA_TYPES, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_MIPMAPPED_HEIGHT
CUDA_TYPES, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_MIPMAPPED_WIDTH
CUDA_TYPES, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_WIDTH
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_DEPTH
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_DEPTH_ALTERNATE
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_HEIGHT
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_HEIGHT_ALTERNATE
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_WIDTH
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_WIDTH_ALTERNATE
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURECUBEMAP_LAYERED_LAYERS
CUDA_TYPES, 235

- CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_LAYERED_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_TEXTURECUBEMAP_WIDTH
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_MEMORY_CLOCK_RATE
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_MULTIPROCESSOR_COUNT
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_PCI_BUS_ID
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_PCI_DEVICE_ID
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_PCI_DOMAIN_ID
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_REGISTERS_PER_BLOCK
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_SHARED_MEMORY_PER_BLOCK
CUDA_TYPES, 233
- CU_DEVICE_ATTRIBUTE_SURFACE_ALIGNMENT
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_TCC_DRIVER
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_TEXTURE_ALIGNMENT
CUDA_TYPES, 234
- CU_DEVICE_ATTRIBUTE_TEXTURE_PITCH_ALIGNMENT
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_TOTAL_CONSTANT_MEMORY
CUDA_TYPES, 233
- CU_DEVICE_ATTRIBUTE_UNIFIED_ADDRESSING
CUDA_TYPES, 235
- CU_DEVICE_ATTRIBUTE_WARP_SIZE
CUDA_TYPES, 233
- CU_EVENT_BLOCKING_SYNC
CUDA_TYPES, 236
- CU_EVENT_DEFAULT
CUDA_TYPES, 236
- CU_EVENT_DISABLE_TIMING
CUDA_TYPES, 236
- CU_EVENT_INTERPROCESS
CUDA_TYPES, 236
- CU_FUNC_ATTRIBUTE_BINARY_VERSION
CUDA_TYPES, 237
- CU_FUNC_ATTRIBUTE_CONST_SIZE_BYTES
CUDA_TYPES, 237
- CU_FUNC_ATTRIBUTE_LOCAL_SIZE_BYTES
CUDA_TYPES, 237
- CU_FUNC_ATTRIBUTE_MAX_THREADS_PER_BLOCK
CUDA_TYPES, 237
- CU_FUNC_ATTRIBUTE_NUM_REGS
CUDA_TYPES, 237
- CU_FUNC_ATTRIBUTE_PTX_VERSION
CUDA_TYPES, 237
- CU_FUNC_ATTRIBUTE_SHARED_SIZE_BYTES
CUDA_TYPES, 237
- CU_FUNC_CACHE_PREFER_EQUAL
CUDA_TYPES, 237
- CU_FUNC_CACHE_PREFER_L1
CUDA_TYPES, 237
- CU_FUNC_CACHE_PREFER_NONE
CUDA_TYPES, 236
- CU_FUNC_CACHE_PREFER_SHARED
CUDA_TYPES, 236
- CU_GL_DEVICE_LIST_ALL
CUDA_GL, 388
- CU_GL_DEVICE_LIST_CURRENT_FRAME
CUDA_GL, 388
- CU_GL_DEVICE_LIST_NEXT_FRAME
CUDA_GL, 388
- CU_IPC_MEM_LAZY_ENABLE_PEER_ACCESS
CUDA_TYPES, 237
- CU_JIT_ERROR_LOG_BUFFER
CUDA_TYPES, 238
- CU_JIT_ERROR_LOG_BUFFER_SIZE_BYTES
CUDA_TYPES, 238
- CU_JIT_FALLBACK_STRATEGY
CUDA_TYPES, 239
- CU_JIT_INFO_LOG_BUFFER
CUDA_TYPES, 238
- CU_JIT_INFO_LOG_BUFFER_SIZE_BYTES
CUDA_TYPES, 238
- CU_JIT_MAX_REGISTERS
CUDA_TYPES, 238
- CU_JIT_OPTIMIZATION_LEVEL
CUDA_TYPES, 238
- CU_JIT_TARGET
CUDA_TYPES, 238
- CU_JIT_TARGET_FROM_CUCONTEXT
CUDA_TYPES, 238
- CU_JIT_THREADS_PER_BLOCK
CUDA_TYPES, 238
- CU_JIT_WALL_TIME
CUDA_TYPES, 238
- CU_LIMIT_DEV_RUNTIME_PENDING_LAUNCH_COUNT
CUDA_TYPES, 239
- CU_LIMIT_DEV_RUNTIME_SYNC_DEPTH
CUDA_TYPES, 239
- CU_LIMIT_MALLOC_HEAP_SIZE

- CUDA_TYPES, 239
- CU_LIMIT_PRINTF_FIFO_SIZE
 - CUDA_TYPES, 239
- CU_LIMIT_STACK_SIZE
 - CUDA_TYPES, 239
- CU_MEMORYTYPE_ARRAY
 - CUDA_TYPES, 239
- CU_MEMORYTYPE_DEVICE
 - CUDA_TYPES, 239
- CU_MEMORYTYPE_HOST
 - CUDA_TYPES, 239
- CU_MEMORYTYPE_UNIFIED
 - CUDA_TYPES, 239
- CU_POINTER_ATTRIBUTE_CONTEXT
 - CUDA_TYPES, 240
- CU_POINTER_ATTRIBUTE_DEVICE_POINTER
 - CUDA_TYPES, 240
- CU_POINTER_ATTRIBUTE_HOST_POINTER
 - CUDA_TYPES, 240
- CU_POINTER_ATTRIBUTE_MEMORY_TYPE
 - CUDA_TYPES, 240
- CU_POINTER_ATTRIBUTE_P2P_TOKENS
 - CUDA_TYPES, 240
- CU_PREFER_BINARY
 - CUDA_TYPES, 238
- CU_PREFER_PTX
 - CUDA_TYPES, 238
- CU_RES_VIEW_FORMAT_FLOAT_1X16
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_FLOAT_1X32
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_FLOAT_2X16
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_FLOAT_2X32
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_FLOAT_4X16
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_FLOAT_4X32
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_NONE
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_SIGNED_BC4
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_SIGNED_BC5
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_SIGNED_BC6H
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_SINT_1X16
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_SINT_1X32
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_SINT_1X8
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_SINT_2X16
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_SINT_2X32
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_SINT_2X8
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_SINT_4X16
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_SINT_4X32
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_SINT_4X8
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_1X16
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_1X32
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_1X8
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_2X16
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_2X32
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_2X8
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_4X16
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_4X32
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UINT_4X8
 - CUDA_TYPES, 240
- CU_RES_VIEW_FORMAT_UNSIGNED_BC1
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC2
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC3
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC4
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC5
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC6H
 - CUDA_TYPES, 241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC7
 - CUDA_TYPES, 241
- CU_RESOURCE_TYPE_ARRAY
 - CUDA_TYPES, 240
- CU_RESOURCE_TYPE_LINEAR
 - CUDA_TYPES, 240
- CU_RESOURCE_TYPE_MIPMAPPED_ARRAY
 - CUDA_TYPES, 240
- CU_RESOURCE_TYPE_PITCH2D
 - CUDA_TYPES, 240
- CU_SHARED_MEM_CONFIG_DEFAULT_BANK_SIZE
 - CUDA_TYPES, 241

- CU_SHARED_MEM_CONFIG_EIGHT_BYTE_-
BANK_SIZE
CUDA_TYPES, 241
- CU_SHARED_MEM_CONFIG_FOUR_BYTE_-
BANK_SIZE
CUDA_TYPES, 241
- CU_STREAM_DEFAULT
CUDA_TYPES, 241
- CU_STREAM_NON_BLOCKING
CUDA_TYPES, 241
- CU_TARGET_COMPUTE_10
CUDA_TYPES, 239
- CU_TARGET_COMPUTE_11
CUDA_TYPES, 239
- CU_TARGET_COMPUTE_12
CUDA_TYPES, 239
- CU_TARGET_COMPUTE_13
CUDA_TYPES, 239
- CU_TARGET_COMPUTE_20
CUDA_TYPES, 239
- CU_TARGET_COMPUTE_21
CUDA_TYPES, 239
- CU_TARGET_COMPUTE_30
CUDA_TYPES, 239
- CU_TARGET_COMPUTE_35
CUDA_TYPES, 239
- CU_TR_ADDRESS_MODE_BORDER
CUDA_TYPES, 229
- CU_TR_ADDRESS_MODE_CLAMP
CUDA_TYPES, 229
- CU_TR_ADDRESS_MODE_MIRROR
CUDA_TYPES, 229
- CU_TR_ADDRESS_MODE_WRAP
CUDA_TYPES, 229
- CU_TR_FILTER_MODE_LINEAR
CUDA_TYPES, 236
- CU_TR_FILTER_MODE_POINT
CUDA_TYPES, 236
- CU_IPC_HANDLE_SIZE
CUDA_TYPES, 222
- CU_LAUNCH_PARAM_BUFFER_POINTER
CUDA_TYPES, 222
- CU_LAUNCH_PARAM_BUFFER_SIZE
CUDA_TYPES, 222
- CU_LAUNCH_PARAM_END
CUDA_TYPES, 223
- CU_MEMHOSTALLOC_DEVICEMAP
CUDA_TYPES, 223
- CU_MEMHOSTALLOC_PORTABLE
CUDA_TYPES, 223
- CU_MEMHOSTALLOC_WRITECOMBINED
CUDA_TYPES, 223
- CU_MEMHOSTREGISTER_DEVICEMAP
CUDA_TYPES, 223
- CU_MEMHOSTREGISTER_PORTABLE
CUDA_TYPES, 223
- CU_PARAM_TR_DEFAULT
CUDA_TYPES, 223
- CU_TRSA_OVERRIDE_FORMAT
CUDA_TYPES, 223
- CU_TRSF_NORMALIZED_COORDINATES
CUDA_TYPES, 223
- CU_TRSF_READ_AS_INTEGER
CUDA_TYPES, 223
- CU_TRSF_SRGB
CUDA_TYPES, 224
- CUaddress_mode
CUDA_TYPES, 224
- CUaddress_mode_enum
CUDA_TYPES, 229
- CUarray
CUDA_TYPES, 224
- cuArray3DCreate
CUDA_MEM, 276
- cuArray3DGetDescriptor
CUDA_MEM, 279
- CUarray_cubemap_face
CUDA_TYPES, 224
- CUarray_cubemap_face_enum
CUDA_TYPES, 229
- CUarray_format
CUDA_TYPES, 225
- CUarray_format_enum
CUDA_TYPES, 229
- cuArrayCreate
CUDA_MEM, 279
- cuArrayDestroy
CUDA_MEM, 281
- cuArrayGetDescriptor
CUDA_MEM, 281
- CUcomputemode
CUDA_TYPES, 225
- CUcomputemode_enum
CUDA_TYPES, 230
- CUcontext
CUDA_TYPES, 225
- CUctx_flags
CUDA_TYPES, 225
- CUctx_flags_enum
CUDA_TYPES, 230
- cuCtxAttach
CUDA_CTX_DEPRECATED, 263
- cuCtxCreate
CUDA_CTX, 253
- cuCtxDestroy
CUDA_CTX, 254
- cuCtxDetach
CUDA_CTX_DEPRECATED, 263

- cuCtxDisablePeerAccess
 - CUDA_PEER_ACCESS, 378
- cuCtxEnablePeerAccess
 - CUDA_PEER_ACCESS, 378
- cuCtxGetApiVersion
 - CUDA_CTX, 254
- cuCtxGetCacheConfig
 - CUDA_CTX, 255
- cuCtxGetCurrent
 - CUDA_CTX, 255
- cuCtxGetDevice
 - CUDA_CTX, 256
- cuCtxGetLimit
 - CUDA_CTX, 256
- cuCtxGetSharedMemConfig
 - CUDA_CTX, 257
- cuCtxPopCurrent
 - CUDA_CTX, 257
- cuCtxPushCurrent
 - CUDA_CTX, 258
- cuCtxSetCacheConfig
 - CUDA_CTX, 258
- cuCtxSetCurrent
 - CUDA_CTX, 259
- cuCtxSetLimit
 - CUDA_CTX, 260
- cuCtxSetSharedMemConfig
 - CUDA_CTX, 261
- cuCtxSynchronize
 - CUDA_CTX, 261
- cuD3D10CtxCreate
 - CUDA_D3D10_DEPRECATED, 419
- cuD3D10CtxCreateOnDevice
 - CUDA_D3D10_DEPRECATED, 420
- CUD3d10DeviceList
 - CUDA_D3D10, 414
- CUD3d10DeviceList_enum
 - CUDA_D3D10, 414
- cuD3D10GetDevice
 - CUDA_D3D10, 415
- cuD3D10GetDevices
 - CUDA_D3D10, 415
- cuD3D10GetDirect3DDevice
 - CUDA_D3D10_DEPRECATED, 420
- CUD3D10map_flags
 - CUDA_D3D10_DEPRECATED, 419
- CUD3D10map_flags_enum
 - CUDA_D3D10_DEPRECATED, 419
- cuD3D10MapResources
 - CUDA_D3D10_DEPRECATED, 421
- CUD3D10register_flags
 - CUDA_D3D10_DEPRECATED, 419
- CUD3D10register_flags_enum
 - CUDA_D3D10_DEPRECATED, 419
- cuD3D10RegisterResource
 - CUDA_D3D10_DEPRECATED, 422
- cuD3D10ResourceGetMappedArray
 - CUDA_D3D10_DEPRECATED, 423
- cuD3D10ResourceGetMappedPitch
 - CUDA_D3D10_DEPRECATED, 423
- cuD3D10ResourceGetMappedPointer
 - CUDA_D3D10_DEPRECATED, 424
- cuD3D10ResourceGetMappedSize
 - CUDA_D3D10_DEPRECATED, 425
- cuD3D10ResourceGetSurfaceDimensions
 - CUDA_D3D10_DEPRECATED, 426
- cuD3D10ResourceSetMapFlags
 - CUDA_D3D10_DEPRECATED, 426
- cuD3D10UnmapResources
 - CUDA_D3D10_DEPRECATED, 427
- cuD3D10UnregisterResource
 - CUDA_D3D10_DEPRECATED, 428
- cuD3D11CtxCreate
 - CUDA_D3D11_DEPRECATED, 433
- cuD3D11CtxCreateOnDevice
 - CUDA_D3D11_DEPRECATED, 433
- CUD3d11DeviceList
 - CUDA_D3D11, 429
- CUD3d11DeviceList_enum
 - CUDA_D3D11, 429
- cuD3D11GetDevice
 - CUDA_D3D11, 430
- cuD3D11GetDevices
 - CUDA_D3D11, 430
- cuD3D11GetDirect3DDevice
 - CUDA_D3D11_DEPRECATED, 434
- cuD3D9CtxCreate
 - CUDA_D3D9, 400
- cuD3D9CtxCreateOnDevice
 - CUDA_D3D9, 401
- CUD3d9DeviceList
 - CUDA_D3D9, 400
- CUD3d9DeviceList_enum
 - CUDA_D3D9, 400
- cuD3D9GetDevice
 - CUDA_D3D9, 401
- cuD3D9GetDevices
 - CUDA_D3D9, 402
- cuD3D9GetDirect3DDevice
 - CUDA_D3D9, 402
- CUD3d9map_flags
 - CUDA_D3D9_DEPRECATED, 406
- CUD3d9map_flags_enum
 - CUDA_D3D9_DEPRECATED, 406
- cuD3D9MapResources
 - CUDA_D3D9_DEPRECATED, 406
- CUD3d9register_flags
 - CUDA_D3D9_DEPRECATED, 406

- CuD3d9register_flags_enum
 - CUDA_D3D9_DEPRECATED, 406
- cuD3D9RegisterResource
 - CUDA_D3D9_DEPRECATED, 407
- cuD3D9ResourceGetMappedArray
 - CUDA_D3D9_DEPRECATED, 408
- cuD3D9ResourceGetMappedPitch
 - CUDA_D3D9_DEPRECATED, 409
- cuD3D9ResourceGetMappedPointer
 - CUDA_D3D9_DEPRECATED, 410
- cuD3D9ResourceGetMappedSize
 - CUDA_D3D9_DEPRECATED, 410
- cuD3D9ResourceGetSurfaceDimensions
 - CUDA_D3D9_DEPRECATED, 411
- cuD3D9ResourceSetMapFlags
 - CUDA_D3D9_DEPRECATED, 412
- cuD3D9UnmapResources
 - CUDA_D3D9_DEPRECATED, 412
- cuD3D9UnregisterResource
 - CUDA_D3D9_DEPRECATED, 413
- CUDA_D3D10
 - CU_D3D10_DEVICE_LIST_ALL, 415
 - CU_D3D10_DEVICE_LIST_CURRENT_FRAME, 415
 - CU_D3D10_DEVICE_LIST_NEXT_FRAME, 415
- CUDA_D3D11
 - CU_D3D11_DEVICE_LIST_ALL, 430
 - CU_D3D11_DEVICE_LIST_CURRENT_FRAME, 430
 - CU_D3D11_DEVICE_LIST_NEXT_FRAME, 430
- CUDA_D3D9
 - CU_D3D9_DEVICE_LIST_ALL, 400
 - CU_D3D9_DEVICE_LIST_CURRENT_FRAME, 400
 - CU_D3D9_DEVICE_LIST_NEXT_FRAME, 400
- CUDA_ERROR_ALREADY_ACQUIRED
 - CUDA_TYPES, 232
- CUDA_ERROR_ALREADY_MAPPED
 - CUDA_TYPES, 231
- CUDA_ERROR_ARRAY_IS_MAPPED
 - CUDA_TYPES, 231
- CUDA_ERROR_ASSERT
 - CUDA_TYPES, 233
- CUDA_ERROR_CONTEXT_ALREADY_CURRENT
 - CUDA_TYPES, 231
- CUDA_ERROR_CONTEXT_ALREADY_IN_USE
 - CUDA_TYPES, 232
- CUDA_ERROR_CONTEXT_IS_DESTROYED
 - CUDA_TYPES, 233
- CUDA_ERROR_DEINITIALIZED
 - CUDA_TYPES, 231
- CUDA_ERROR_ECC_UNCORRECTABLE
 - CUDA_TYPES, 232
- CUDA_ERROR_FILE_NOT_FOUND
 - CUDA_TYPES, 231
- CUDA_ERROR_HOST_MEMORY_ALREADY_REGISTERED
 - CUDA_TYPES, 233
- CUDA_ERROR_HOST_MEMORY_NOT_REGISTERED
 - CUDA_TYPES, 233
- CUDA_ERROR_INVALID_CONTEXT
 - CUDA_TYPES, 231
- CUDA_ERROR_INVALID_DEVICE
 - CUDA_TYPES, 231
- CUDA_ERROR_INVALID_HANDLE
 - CUDA_TYPES, 232
- CUDA_ERROR_INVALID_IMAGE
 - CUDA_TYPES, 231
- CUDA_ERROR_INVALID_SOURCE
 - CUDA_TYPES, 232
- CUDA_ERROR_INVALID_VALUE
 - CUDA_TYPES, 231
- CUDA_ERROR_LAUNCH_FAILED
 - CUDA_TYPES, 232
- CUDA_ERROR_LAUNCH_INCOMPATIBLE_TEXTURING
 - CUDA_TYPES, 232
- CUDA_ERROR_LAUNCH_OUT_OF_RESOURCES
 - CUDA_TYPES, 232
- CUDA_ERROR_LAUNCH_TIMEOUT
 - CUDA_TYPES, 232
- CUDA_ERROR_MAP_FAILED
 - CUDA_TYPES, 231
- CUDA_ERROR_NO_BINARY_FOR_GPU
 - CUDA_TYPES, 231
- CUDA_ERROR_NO_DEVICE
 - CUDA_TYPES, 231
- CUDA_ERROR_NOT_FOUND
 - CUDA_TYPES, 232
- CUDA_ERROR_NOT_INITIALIZED
 - CUDA_TYPES, 231
- CUDA_ERROR_NOT_MAPPED
 - CUDA_TYPES, 232
- CUDA_ERROR_NOT_MAPPED_AS_ARRAY
 - CUDA_TYPES, 232
- CUDA_ERROR_NOT_MAPPED_AS_POINTER
 - CUDA_TYPES, 232
- CUDA_ERROR_NOT_PERMITTED
 - CUDA_TYPES, 233
- CUDA_ERROR_NOT_READY
 - CUDA_TYPES, 232
- CUDA_ERROR_NOT_SUPPORTED
 - CUDA_TYPES, 233
- CUDA_ERROR_OPERATING_SYSTEM
 - CUDA_TYPES, 232
- CUDA_ERROR_OUT_OF_MEMORY
 - CUDA_TYPES, 231

- CUDA_ERROR_PEER_ACCESS_ALREADY_ENABLED
 - CUDA_TYPES, 233
- CUDA_ERROR_PEER_ACCESS_NOT_ENABLED
 - CUDA_TYPES, 233
- CUDA_ERROR_PEER_ACCESS_UNSUPPORTED
 - CUDA_TYPES, 232
- CUDA_ERROR_PRIMARY_CONTEXT_ACTIVE
 - CUDA_TYPES, 233
- CUDA_ERROR_PROFILER_ALREADY_STARTED
 - CUDA_TYPES, 231
- CUDA_ERROR_PROFILER_ALREADY_STOPPED
 - CUDA_TYPES, 231
- CUDA_ERROR_PROFILER_DISABLED
 - CUDA_TYPES, 231
- CUDA_ERROR_PROFILER_NOT_INITIALIZED
 - CUDA_TYPES, 231
- CUDA_ERROR_SHARED_OBJECT_INIT_FAILED
 - CUDA_TYPES, 232
- CUDA_ERROR_SHARED_OBJECT_SYMBOL_NOT_FOUND
 - CUDA_TYPES, 232
- CUDA_ERROR_TOO_MANY_PEERS
 - CUDA_TYPES, 233
- CUDA_ERROR_UNKNOWN
 - CUDA_TYPES, 233
- CUDA_ERROR_UNMAP_FAILED
 - CUDA_TYPES, 231
- CUDA_ERROR_UNSUPPORTED_LIMIT
 - CUDA_TYPES, 232
- CUDA_GL
 - CU_GL_DEVICE_LIST_ALL, 388
 - CU_GL_DEVICE_LIST_CURRENT_FRAME, 388
 - CU_GL_DEVICE_LIST_NEXT_FRAME, 388
- CUDA_SUCCESS
 - CUDA_TYPES, 231
- CUDA_TYPES
 - CU_AD_FORMAT_FLOAT, 230
 - CU_AD_FORMAT_HALF, 230
 - CU_AD_FORMAT_SIGNED_INT16, 230
 - CU_AD_FORMAT_SIGNED_INT32, 230
 - CU_AD_FORMAT_SIGNED_INT8, 230
 - CU_AD_FORMAT_UNSIGNED_INT16, 230
 - CU_AD_FORMAT_UNSIGNED_INT32, 230
 - CU_AD_FORMAT_UNSIGNED_INT8, 230
 - CU_COMPUTEMODE_DEFAULT, 230
 - CU_COMPUTEMODE_EXCLUSIVE, 230
 - CU_COMPUTEMODE_EXCLUSIVE_PROCESS, 230
 - CU_COMPUTEMODE_PROHIBITED, 230
 - CU_CTX_BLOCKING_SYNC, 230
 - CU_CTX_LMEM_RESIZE_TO_MAX, 230
 - CU_CTX_MAP_HOST, 230
 - CU_CTX_SCHED_AUTO, 230
 - CU_CTX_SCHED_BLOCKING_SYNC, 230
 - CU_CTX_SCHED_SPIN, 230
 - CU_CTX_SCHED_YIELD, 230
 - CU_CUBEMAP_FACE_NEGATIVE_X, 229
 - CU_CUBEMAP_FACE_NEGATIVE_Y, 229
 - CU_CUBEMAP_FACE_NEGATIVE_Z, 229
 - CU_CUBEMAP_FACE_POSITIVE_X, 229
 - CU_CUBEMAP_FACE_POSITIVE_Y, 229
 - CU_CUBEMAP_FACE_POSITIVE_Z, 229
 - CU_DEVICE_ATTRIBUTE_ASYNC_ENGINE_COUNT, 234
 - CU_DEVICE_ATTRIBUTE_CAN_MAP_HOST_MEMORY, 234
 - CU_DEVICE_ATTRIBUTE_CAN_TEX2D_GATHER, 235
 - CU_DEVICE_ATTRIBUTE_CLOCK_RATE, 234
 - CU_DEVICE_ATTRIBUTE_COMPUTE_CAPABILITY_MAJOR, 236
 - CU_DEVICE_ATTRIBUTE_COMPUTE_CAPABILITY_MINOR, 236
 - CU_DEVICE_ATTRIBUTE_COMPUTE_MODE, 234
 - CU_DEVICE_ATTRIBUTE_CONCURRENT_KERNELS, 234
 - CU_DEVICE_ATTRIBUTE_ECC_ENABLED, 234
 - CU_DEVICE_ATTRIBUTE_GLOBAL_MEMORY_BUS_WIDTH, 234
 - CU_DEVICE_ATTRIBUTE_GPU_OVERLAP, 234
 - CU_DEVICE_ATTRIBUTE_INTEGRATED, 234
 - CU_DEVICE_ATTRIBUTE_KERNEL_EXEC_TIMEOUT, 234
 - CU_DEVICE_ATTRIBUTE_L2_CACHE_SIZE, 234
 - CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_X, 233
 - CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Y, 233
 - CU_DEVICE_ATTRIBUTE_MAX_BLOCK_DIM_Z, 233
 - CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_X, 233
 - CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Y, 233
 - CU_DEVICE_ATTRIBUTE_MAX_GRID_DIM_Z, 233
 - CU_DEVICE_ATTRIBUTE_MAX_PITCH, 233
 - CU_DEVICE_ATTRIBUTE_MAX_REGISTERS_PER_BLOCK, 233
 - CU_DEVICE_ATTRIBUTE_MAX_SHARED_MEMORY_PER_BLOCK, 233
 - CU_DEVICE_ATTRIBUTE_MAX_THREADS_

- PER_BLOCK, 233
- CU_DEVICE_ATTRIBUTE_MAX_THREADS_-
PER_MULTIPROCESSOR, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE1D_LAYERED_LAYERS, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE1D_LAYERED_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE1D_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_HEIGHT, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_LAYERED_HEIGHT, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_LAYERED_LAYERS, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_LAYERED_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE2D_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE3D_DEPTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE3D_HEIGHT, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACE3D_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACECUBEMAP_LAYERED_LAYERS,
235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACECUBEMAP_LAYERED_WIDTH,
235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
SURFACECUBEMAP_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_LAYERED_LAYERS, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_LAYERED_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_LINEAR_WIDTH, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_MIPMAPPED_WIDTH, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE1D_WIDTH, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_ARRAY_HEIGHT, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_ARRAY_NUMSLICES, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_ARRAY_WIDTH, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_GATHER_HEIGHT, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_GATHER_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_HEIGHT, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LAYERED_HEIGHT, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LAYERED_LAYERS, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LAYERED_WIDTH, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LINEAR_HEIGHT, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LINEAR_PITCH, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_LINEAR_WIDTH, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_MIPMAPPED_HEIGHT,
236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_MIPMAPPED_WIDTH, 236
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE2D_WIDTH, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_DEPTH, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_DEPTH_ALTERNATE, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_HEIGHT, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_HEIGHT_ALTERNATE, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_WIDTH, 234
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURE3D_WIDTH_ALTERNATE, 235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURECUBEMAP_LAYERED_LAYERS,
235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURECUBEMAP_LAYERED_WIDTH,
235
- CU_DEVICE_ATTRIBUTE_MAXIMUM_-
TEXTURECUBEMAP_WIDTH, 235
- CU_DEVICE_ATTRIBUTE_MEMORY_-
CLOCK_RATE, 234
- CU_DEVICE_ATTRIBUTE_-
MULTIPROCESSOR_COUNT, 234
- CU_DEVICE_ATTRIBUTE_PCI_BUS_ID, 234
- CU_DEVICE_ATTRIBUTE_PCI_DEVICE_ID,
234
- CU_DEVICE_ATTRIBUTE_PCI_DOMAIN_ID,
235
- CU_DEVICE_ATTRIBUTE_REGISTERS_PER_-
BLOCK, 234
- CU_DEVICE_ATTRIBUTE_SHARED_-
MEMORY_PER_BLOCK, 233
- CU_DEVICE_ATTRIBUTE_SURFACE_-

- ALIGNMENT, 234
- CU_DEVICE_ATTRIBUTE_TCC_DRIVER, 234
- CU_DEVICE_ATTRIBUTE_TEXTURE_-
ALIGNMENT, 234
- CU_DEVICE_ATTRIBUTE_TEXTURE_PITCH_-
ALIGNMENT, 235
- CU_DEVICE_ATTRIBUTE_TOTAL_-
CONSTANT_MEMORY, 233
- CU_DEVICE_ATTRIBUTE_UNIFIED_-
ADDRESSING, 235
- CU_DEVICE_ATTRIBUTE_WARP_SIZE, 233
- CU_EVENT_BLOCKING_SYNC, 236
- CU_EVENT_DEFAULT, 236
- CU_EVENT_DISABLE_TIMING, 236
- CU_EVENT_INTERPROCESS, 236
- CU_FUNC_ATTRIBUTE_BINARY_VERSION,
237
- CU_FUNC_ATTRIBUTE_CONST_SIZE_BYTES,
237
- CU_FUNC_ATTRIBUTE_LOCAL_SIZE_BYTES,
237
- CU_FUNC_ATTRIBUTE_MAX_THREADS_-
PER_BLOCK, 237
- CU_FUNC_ATTRIBUTE_NUM_REGS, 237
- CU_FUNC_ATTRIBUTE_PTX_VERSION, 237
- CU_FUNC_ATTRIBUTE_SHARED_SIZE_-
BYTES, 237
- CU_FUNC_CACHE_PREFER_EQUAL, 237
- CU_FUNC_CACHE_PREFER_L1, 237
- CU_FUNC_CACHE_PREFER_NONE, 236
- CU_FUNC_CACHE_PREFER_SHARED, 236
- CU_IPC_MEM_LAZY_ENABLE_PEER_-
ACCESS, 237
- CU_JIT_ERROR_LOG_BUFFER, 238
- CU_JIT_ERROR_LOG_BUFFER_SIZE_BYTES,
238
- CU_JIT_FALLBACK_STRATEGY, 239
- CU_JIT_INFO_LOG_BUFFER, 238
- CU_JIT_INFO_LOG_BUFFER_SIZE_BYTES,
238
- CU_JIT_MAX_REGISTERS, 238
- CU_JIT_OPTIMIZATION_LEVEL, 238
- CU_JIT_TARGET, 238
- CU_JIT_TARGET_FROM_CUCONTEXT, 238
- CU_JIT_THREADS_PER_BLOCK, 238
- CU_JIT_WALL_TIME, 238
- CU_LIMIT_DEV_RUNTIME_PENDING_-
LAUNCH_COUNT, 239
- CU_LIMIT_DEV_RUNTIME_SYNC_DEPTH,
239
- CU_LIMIT_MALLOC_HEAP_SIZE, 239
- CU_LIMIT_PRINTF_FIFO_SIZE, 239
- CU_LIMIT_STACK_SIZE, 239
- CU_MEMORYTYPE_ARRAY, 239
- CU_MEMORYTYPE_DEVICE, 239
- CU_MEMORYTYPE_HOST, 239
- CU_MEMORYTYPE_UNIFIED, 239
- CU_POINTER_ATTRIBUTE_CONTEXT, 240
- CU_POINTER_ATTRIBUTE_DEVICE_-
POINTER, 240
- CU_POINTER_ATTRIBUTE_HOST_POINTER,
240
- CU_POINTER_ATTRIBUTE_MEMORY_TYPE,
240
- CU_POINTER_ATTRIBUTE_P2P_TOKENS, 240
- CU_PREFER_BINARY, 238
- CU_PREFER_PTX, 238
- CU_RES_VIEW_FORMAT_FLOAT_1X16, 241
- CU_RES_VIEW_FORMAT_FLOAT_1X32, 241
- CU_RES_VIEW_FORMAT_FLOAT_2X16, 241
- CU_RES_VIEW_FORMAT_FLOAT_2X32, 241
- CU_RES_VIEW_FORMAT_FLOAT_4X16, 241
- CU_RES_VIEW_FORMAT_FLOAT_4X32, 241
- CU_RES_VIEW_FORMAT_NONE, 240
- CU_RES_VIEW_FORMAT_SIGNED_BC4, 241
- CU_RES_VIEW_FORMAT_SIGNED_BC5, 241
- CU_RES_VIEW_FORMAT_SIGNED_BC6H, 241
- CU_RES_VIEW_FORMAT_SINT_1X16, 240
- CU_RES_VIEW_FORMAT_SINT_1X32, 241
- CU_RES_VIEW_FORMAT_SINT_1X8, 240
- CU_RES_VIEW_FORMAT_SINT_2X16, 240
- CU_RES_VIEW_FORMAT_SINT_2X32, 241
- CU_RES_VIEW_FORMAT_SINT_2X8, 240
- CU_RES_VIEW_FORMAT_SINT_4X16, 240
- CU_RES_VIEW_FORMAT_SINT_4X32, 241
- CU_RES_VIEW_FORMAT_SINT_4X8, 240
- CU_RES_VIEW_FORMAT_UINT_1X16, 240
- CU_RES_VIEW_FORMAT_UINT_1X32, 240
- CU_RES_VIEW_FORMAT_UINT_1X8, 240
- CU_RES_VIEW_FORMAT_UINT_2X16, 240
- CU_RES_VIEW_FORMAT_UINT_2X32, 240
- CU_RES_VIEW_FORMAT_UINT_2X8, 240
- CU_RES_VIEW_FORMAT_UINT_4X16, 240
- CU_RES_VIEW_FORMAT_UINT_4X32, 240
- CU_RES_VIEW_FORMAT_UINT_4X8, 240
- CU_RES_VIEW_FORMAT_UNSIGNED_BC1,
241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC2,
241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC3,
241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC4,
241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC5,
241
- CU_RES_VIEW_FORMAT_UNSIGNED_BC6H,
241

- CU_RES_VIEW_FORMAT_UNSIGNED_BC7, 241
- CU_RESOURCE_TYPE_ARRAY, 240
- CU_RESOURCE_TYPE_LINEAR, 240
- CU_RESOURCE_TYPE_MIPMAPPED_ARRAY, 240
- CU_RESOURCE_TYPE_PITCH2D, 240
- CU_SHARED_MEM_CONFIG_DEFAULT_-
BANK_SIZE, 241
- CU_SHARED_MEM_CONFIG_EIGHT_BYTE_-
BANK_SIZE, 241
- CU_SHARED_MEM_CONFIG_FOUR_BYTE_-
BANK_SIZE, 241
- CU_STREAM_DEFAULT, 241
- CU_STREAM_NON_BLOCKING, 241
- CU_TARGET_COMPUTE_10, 239
- CU_TARGET_COMPUTE_11, 239
- CU_TARGET_COMPUTE_12, 239
- CU_TARGET_COMPUTE_13, 239
- CU_TARGET_COMPUTE_20, 239
- CU_TARGET_COMPUTE_21, 239
- CU_TARGET_COMPUTE_30, 239
- CU_TARGET_COMPUTE_35, 239
- CU_TR_ADDRESS_MODE_BORDER, 229
- CU_TR_ADDRESS_MODE_CLAMP, 229
- CU_TR_ADDRESS_MODE_MIRROR, 229
- CU_TR_ADDRESS_MODE_WRAP, 229
- CU_TR_FILTER_MODE_LINEAR, 236
- CU_TR_FILTER_MODE_POINT, 236
- CUDA_ERROR_ALREADY_ACQUIRED, 232
- CUDA_ERROR_ALREADY_MAPPED, 231
- CUDA_ERROR_ARRAY_IS_MAPPED, 231
- CUDA_ERROR_ASSERT, 233
- CUDA_ERROR_CONTEXT_ALREADY_-
CURRENT, 231
- CUDA_ERROR_CONTEXT_ALREADY_IN_-
USE, 232
- CUDA_ERROR_CONTEXT_IS_DESTROYED, 233
- CUDA_ERROR_DEINITIALIZED, 231
- CUDA_ERROR_ECC_UNCORRECTABLE, 232
- CUDA_ERROR_FILE_NOT_FOUND, 232
- CUDA_ERROR_HOST_MEMORY_ALREADY_-
REGISTERED, 233
- CUDA_ERROR_HOST_MEMORY_NOT_-
REGISTERED, 233
- CUDA_ERROR_INVALID_CONTEXT, 231
- CUDA_ERROR_INVALID_DEVICE, 231
- CUDA_ERROR_INVALID_HANDLE, 232
- CUDA_ERROR_INVALID_IMAGE, 231
- CUDA_ERROR_INVALID_SOURCE, 232
- CUDA_ERROR_INVALID_VALUE, 231
- CUDA_ERROR_LAUNCH_FAILED, 232
- CUDA_ERROR_LAUNCH_INCOMPATIBLE_-
TEXTURING, 232
- CUDA_ERROR_LAUNCH_OUT_OF_-
RESOURCES, 232
- CUDA_ERROR_LAUNCH_TIMEOUT, 232
- CUDA_ERROR_MAP_FAILED, 231
- CUDA_ERROR_NO_BINARY_FOR_GPU, 231
- CUDA_ERROR_NO_DEVICE, 231
- CUDA_ERROR_NOT_FOUND, 232
- CUDA_ERROR_NOT_INITIALIZED, 231
- CUDA_ERROR_NOT_MAPPED, 232
- CUDA_ERROR_NOT_MAPPED_AS_ARRAY, 232
- CUDA_ERROR_NOT_MAPPED_AS_POINTER, 232
- CUDA_ERROR_NOT_PERMITTED, 233
- CUDA_ERROR_NOT_READY, 232
- CUDA_ERROR_NOT_SUPPORTED, 233
- CUDA_ERROR_OPERATING_SYSTEM, 232
- CUDA_ERROR_OUT_OF_MEMORY, 231
- CUDA_ERROR_PEER_ACCESS_ALREADY_-
ENABLED, 233
- CUDA_ERROR_PEER_ACCESS_NOT_-
ENABLED, 233
- CUDA_ERROR_PEER_ACCESS_-
UNSUPPORTED, 232
- CUDA_ERROR_PRIMARY_CONTEXT_-
ACTIVE, 233
- CUDA_ERROR_PROFILER_ALREADY_-
STARTED, 231
- CUDA_ERROR_PROFILER_ALREADY_-
STOPPED, 231
- CUDA_ERROR_PROFILER_DISABLED, 231
- CUDA_ERROR_PROFILER_NOT_INITIALIZED, 231
- CUDA_ERROR_SHARED_OBJECT_INIT_-
FAILED, 232
- CUDA_ERROR_SHARED_OBJECT_SYMBOL_-
NOT_FOUND, 232
- CUDA_ERROR_TOO_MANY_PEERS, 233
- CUDA_ERROR_UNKNOWN, 233
- CUDA_ERROR_UNMAP_FAILED, 231
- CUDA_ERROR_UNSUPPORTED_LIMIT, 232
- CUDA_SUCCESS, 231
- CUDA_ARRAY3D_2DARRAY
CUDA_TYPES, 224
- CUDA_ARRAY3D_CUBEMAP
CUDA_TYPES, 224
- CUDA_ARRAY3D_DESCRIPTOR
CUDA_TYPES, 225
- CUDA_ARRAY3D_DESCRIPTOR_st, 535
 - Depth, 535
 - Flags, 535
 - Format, 535

- Height, [535](#)
- NumChannels, [535](#)
- Width, [536](#)
- CUDA_ARRAY3D_LAYERED
 - CUDA_TYPES, [224](#)
- CUDA_ARRAY3D_SURFACE_LDST
 - CUDA_TYPES, [224](#)
- CUDA_ARRAY3D_TEXTURE_GATHER
 - CUDA_TYPES, [224](#)
- CUDA_ARRAY_DESCRIPTOR
 - CUDA_TYPES, [225](#)
- CUDA_ARRAY_DESCRIPTOR_st, [537](#)
 - Format, [537](#)
 - Height, [537](#)
 - NumChannels, [537](#)
 - Width, [537](#)
- CUDA_CTX
 - cuCtxCreate, [253](#)
 - cuCtxDestroy, [254](#)
 - cuCtxGetApiVersion, [254](#)
 - cuCtxGetCacheConfig, [255](#)
 - cuCtxGetCurrent, [255](#)
 - cuCtxGetDevice, [256](#)
 - cuCtxGetLimit, [256](#)
 - cuCtxGetSharedMemConfig, [257](#)
 - cuCtxPopCurrent, [257](#)
 - cuCtxPushCurrent, [258](#)
 - cuCtxSetCacheConfig, [258](#)
 - cuCtxSetCurrent, [259](#)
 - cuCtxSetLimit, [260](#)
 - cuCtxSetSharedMemConfig, [261](#)
 - cuCtxSynchronize, [261](#)
- CUDA_CTX_DEPRECATED
 - cuCtxAttach, [263](#)
 - cuCtxDetach, [263](#)
- CUDA_D3D10
 - CUd3d10DeviceList, [414](#)
 - CUd3d10DeviceList_enum, [414](#)
 - cuD3D10GetDevice, [415](#)
 - cuD3D10GetDevices, [415](#)
 - cuGraphicsD3D10RegisterResource, [416](#)
- CUDA_D3D10_DEPRECATED
 - cuD3D10CtxCreate, [419](#)
 - cuD3D10CtxCreateOnDevice, [420](#)
 - cuD3D10GetDirect3DDevice, [420](#)
 - CUD3D10map_flags, [419](#)
 - CUD3D10map_flags_enum, [419](#)
 - cuD3D10MapResources, [421](#)
 - CUD3D10register_flags, [419](#)
 - CUD3D10register_flags_enum, [419](#)
 - cuD3D10RegisterResource, [422](#)
 - cuD3D10ResourceGetMappedArray, [423](#)
 - cuD3D10ResourceGetMappedPitch, [423](#)
 - cuD3D10ResourceGetMappedPointer, [424](#)
 - cuD3D10ResourceGetMappedSize, [425](#)
 - cuD3D10ResourceGetSurfaceDimensions, [426](#)
 - cuD3D10ResourceSetMapFlags, [426](#)
 - cuD3D10UnmapResources, [427](#)
 - cuD3D10UnregisterResource, [428](#)
- CUDA_D3D11
 - CUd3d11DeviceList, [429](#)
 - CUd3d11DeviceList_enum, [429](#)
 - cuD3D11GetDevice, [430](#)
 - cuD3D11GetDevices, [430](#)
 - cuGraphicsD3D11RegisterResource, [431](#)
- CUDA_D3D11_DEPRECATED
 - cuD3D11CtxCreate, [433](#)
 - cuD3D11CtxCreateOnDevice, [433](#)
 - cuD3D11GetDirect3DDevice, [434](#)
- CUDA_D3D9
 - cuD3D9CtxCreate, [400](#)
 - cuD3D9CtxCreateOnDevice, [401](#)
 - CUd3d9DeviceList, [400](#)
 - CUd3d9DeviceList_enum, [400](#)
 - cuD3D9GetDevice, [401](#)
 - cuD3D9GetDevices, [402](#)
 - cuD3D9GetDirect3DDevice, [402](#)
 - cuGraphicsD3D9RegisterResource, [403](#)
- CUDA_D3D9_DEPRECATED
 - CUd3d9map_flags, [406](#)
 - CUd3d9map_flags_enum, [406](#)
 - cuD3D9MapResources, [406](#)
 - CUd3d9register_flags, [406](#)
 - CUd3d9register_flags_enum, [406](#)
 - cuD3D9RegisterResource, [407](#)
 - cuD3D9ResourceGetMappedArray, [408](#)
 - cuD3D9ResourceGetMappedPitch, [409](#)
 - cuD3D9ResourceGetMappedPointer, [410](#)
 - cuD3D9ResourceGetMappedSize, [410](#)
 - cuD3D9ResourceGetSurfaceDimensions, [411](#)
 - cuD3D9ResourceSetMapFlags, [412](#)
 - cuD3D9UnmapResources, [412](#)
 - cuD3D9UnregisterResource, [413](#)
- CUDA_DEVICE
 - cuDeviceGet, [244](#)
 - cuDeviceGetAttribute, [244](#)
 - cuDeviceGetCount, [248](#)
 - cuDeviceGetName, [248](#)
 - cuDeviceTotalMem, [249](#)
- CUDA_DEVICE_DEPRECATED
 - cuDeviceComputeCapability, [250](#)
 - cuDeviceGetProperties, [250](#)
- CUDA_EVENT
 - cuEventCreate, [337](#)
 - cuEventDestroy, [338](#)
 - cuEventElapsedTime, [338](#)
 - cuEventQuery, [339](#)
 - cuEventRecord, [339](#)

- cuEventSynchronize, 340
- CUDA_EXEC
 - cuFuncGetAttribute, 341
 - cuFuncSetCacheConfig, 342
 - cuFuncSetSharedMemConfig, 343
 - cuLaunchKernel, 343
- CUDA_EXEC_DEPRECATED
 - cuFuncSetBlockShape, 346
 - cuFuncSetSharedSize, 347
 - cuLaunch, 347
 - cuLaunchGrid, 348
 - cuLaunchGridAsync, 348
 - cuParamSetf, 349
 - cuParamSeti, 350
 - cuParamSetSize, 350
 - cuParamSetTexRef, 351
 - cuParamSetv, 351
- CUDA_GL
 - CUGLDeviceList, 387
 - CUGLDeviceList_enum, 388
 - cuGLGetDevices, 388
 - cuGraphicsGLRegisterBuffer, 388
 - cuGraphicsGLRegisterImage, 389
 - cuWGLGetDevice, 390
- CUDA_GL_DEPRECATED
 - cuGLCtxCreate, 393
 - cuGLInit, 393
 - CUGLmap_flags, 392
 - CUGLmap_flags_enum, 393
 - cuGLMapBufferObject, 394
 - cuGLMapBufferObjectAsync, 394
 - cuGLRegisterBufferObject, 395
 - cuGLSetBufferObjectMapFlags, 395
 - cuGLUnmapBufferObject, 396
 - cuGLUnmapBufferObjectAsync, 397
 - cuGLUnregisterBufferObject, 397
- CUDA_GRAPHICS
 - cuGraphicsMapResources, 380
 - cuGraphicsResourceGetMappedMipmappedArray, 381
 - cuGraphicsResourceGetMappedPointer, 381
 - cuGraphicsResourceSetMapFlags, 382
 - cuGraphicsSubResourceGetMappedArray, 382
 - cuGraphicsUnmapResources, 383
 - cuGraphicsUnregisterResource, 384
- CUDA_INITIALIZE
 - cuInit, 242
- CUDA_IPC_HANDLE_SIZE
 - CUDART_TYPES, 198
- CUDA_MATH_DOUBLE
 - acos, 471
 - acosh, 472
 - asin, 472
 - asinh, 472
 - atan, 472
 - atan2, 473
 - atanh, 473
 - cbrt, 473
 - ceil, 473
 - copysign, 474
 - cos, 474
 - cosh, 474
 - cospi, 474
 - erf, 475
 - erfc, 475
 - erfcinv, 475
 - erfcx, 475
 - erfinv, 476
 - exp, 476
 - exp10, 476
 - exp2, 476
 - expm1, 477
 - fabs, 477
 - fdim, 477
 - floor, 477
 - fma, 478
 - fmax, 478
 - fmin, 478
 - fmod, 479
 - frexp, 479
 - hypot, 479
 - ilogb, 480
 - isfinite, 480
 - isinf, 480
 - isnan, 480
 - j0, 480
 - j1, 481
 - jn, 481
 - ldexp, 481
 - lgamma, 482
 - llrint, 482
 - llround, 482
 - log, 482
 - log10, 483
 - log1p, 483
 - log2, 483
 - logb, 483
 - lrint, 484
 - lround, 484
 - modf, 484
 - nan, 484
 - nearbyint, 485
 - nextafter, 485
 - normcdf, 485
 - normcdfinv, 485
 - pow, 486
 - rcbrt, 486
 - remainder, 487

- remquo, 487
 - rint, 487
 - round, 487
 - rsqrt, 488
 - scalbln, 488
 - scalbn, 488
 - signbit, 488
 - sin, 489
 - sincos, 489
 - sincospi, 489
 - sinh, 489
 - sinpi, 490
 - sqrt, 490
 - tan, 490
 - tanh, 490
 - tgamma, 491
 - trunc, 491
 - y0, 491
 - y1, 491
 - yn, 492
- CUDA_MATH_INTRINSIC_CAST
- __double2float_rd, 522
 - __double2float_rn, 522
 - __double2float_ru, 522
 - __double2float_rz, 522
 - __double2hiint, 523
 - __double2int_rd, 523
 - __double2int_rn, 523
 - __double2int_ru, 523
 - __double2int_rz, 523
 - __double2ll_rd, 523
 - __double2ll_rn, 524
 - __double2ll_ru, 524
 - __double2ll_rz, 524
 - __double2loint, 524
 - __double2uint_rd, 524
 - __double2uint_rn, 524
 - __double2uint_ru, 525
 - __double2uint_rz, 525
 - __double2ull_rd, 525
 - __double2ull_rn, 525
 - __double2ull_ru, 525
 - __double2ull_rz, 525
 - __double_as_longlong, 526
 - __float2half_rn, 526
 - __float2int_rd, 526
 - __float2int_rn, 526
 - __float2int_ru, 526
 - __float2int_rz, 526
 - __float2ll_rd, 527
 - __float2ll_rn, 527
 - __float2ll_ru, 527
 - __float2ll_rz, 527
 - __float2uint_rd, 527
 - __float2uint_rn, 527
 - __float2uint_ru, 528
 - __float2uint_rz, 528
 - __float2ull_rd, 528
 - __float2ull_rn, 528
 - __float2ull_ru, 528
 - __float2ull_rz, 528
 - __float_as_int, 529
 - __half2float, 529
 - __hiloint2double, 529
 - __int2double_rn, 529
 - __int2float_rd, 529
 - __int2float_rn, 529
 - __int2float_ru, 530
 - __int2float_rz, 530
 - __int_as_float, 530
 - __ll2double_rd, 530
 - __ll2double_rn, 530
 - __ll2double_ru, 530
 - __ll2double_rz, 531
 - __ll2float_rd, 531
 - __ll2float_rn, 531
 - __ll2float_ru, 531
 - __ll2float_rz, 531
 - __longlong_as_double, 531
 - __uint2double_rn, 532
 - __uint2float_rd, 532
 - __uint2float_rn, 532
 - __uint2float_ru, 532
 - __uint2float_rz, 532
 - __ull2double_rd, 532
 - __ull2double_rn, 533
 - __ull2double_ru, 533
 - __ull2double_rz, 533
 - __ull2float_rd, 533
 - __ull2float_rn, 533
 - __ull2float_ru, 533
 - __ull2float_rz, 534
- CUDA_MATH_INTRINSIC_DOUBLE
- __dadd_rd, 506
 - __dadd_rn, 506
 - __dadd_ru, 507
 - __dadd_rz, 507
 - __ddiv_rd, 507
 - __ddiv_rn, 507
 - __ddiv_ru, 508
 - __ddiv_rz, 508
 - __dmul_rd, 508
 - __dmul_rn, 508
 - __dmul_ru, 509
 - __dmul_rz, 509
 - __drcp_rd, 509
 - __drcp_rn, 509
 - __drcp_ru, 510

- __drcp_rz, 510
- __dsqrt_rd, 510
- __dsqrt_rn, 510
- __dsqrt_ru, 511
- __dsqrt_rz, 511
- __fma_rd, 511
- __fma_rn, 511
- __fma_ru, 512
- __fma_rz, 512
- CUDA_MATH_INTRINSIC_INT
 - __brev, 514
 - __brevll, 514
 - __byte_perm, 514
 - __clz, 514
 - __clzll, 515
 - __ffs, 515
 - __ffsll, 515
 - __mul24, 515
 - __mul64hi, 515
 - __mulhi, 515
 - __popc, 516
 - __popc11, 516
 - __sad, 516
 - __umul24, 516
 - __umul64hi, 516
 - __umulhi, 516
 - __usad, 517
- CUDA_MATH_INTRINSIC_SINGLE
 - __cosf, 495
 - __exp10f, 495
 - __expf, 495
 - __fadd_rd, 496
 - __fadd_rn, 496
 - __fadd_ru, 496
 - __fadd_rz, 496
 - __fdiv_rd, 497
 - __fdiv_rn, 497
 - __fdiv_ru, 497
 - __fdiv_rz, 497
 - __fdividef, 498
 - __fmaf_rd, 498
 - __fmaf_rn, 498
 - __fmaf_ru, 499
 - __fmaf_rz, 499
 - __fmul_rd, 499
 - __fmul_rn, 499
 - __fmul_ru, 500
 - __fmul_rz, 500
 - __frcp_rd, 500
 - __frcp_rn, 500
 - __frcp_ru, 501
 - __frcp_rz, 501
 - __frsqrt_rn, 501
 - __fsqrt_rd, 501
 - __fsqrt_rn, 502
 - __fsqrt_ru, 502
 - __fsqrt_rz, 502
 - __log10f, 502
 - __log2f, 503
 - __logf, 503
 - __powf, 503
 - __saturatef, 503
 - __sincosf, 504
 - __sinf, 504
 - __tanf, 504
- CUDA_MATH_SINGLE
 - acosf, 445
 - acoshf, 445
 - asinf, 445
 - asinhf, 445
 - atan2f, 445
 - atanf, 446
 - atanhf, 446
 - cbrtf, 446
 - ceilf, 446
 - copysignf, 447
 - cosf, 447
 - coshf, 447
 - cospif, 447
 - erfcf, 448
 - erfcinvf, 448
 - erfcxf, 448
 - erff, 448
 - erfinvf, 449
 - exp10f, 449
 - exp2f, 449
 - expf, 449
 - expm1f, 450
 - fabsf, 450
 - fdimf, 450
 - fdividef, 450
 - floorf, 451
 - fmaf, 451
 - fmaxf, 451
 - fminf, 452
 - fmodf, 452
 - frexpf, 452
 - hypotf, 453
 - ilogbf, 453
 - isfinite, 453
 - isinf, 454
 - isnan, 454
 - j0f, 454
 - j1f, 454
 - jnf, 454
 - ldexpf, 455
 - lgammaf, 455
 - llrintf, 455

- llroundf, 455
- log10f, 456
- log1pf, 456
- log2f, 456
- logbf, 457
- logf, 457
- lrintf, 457
- lroundf, 457
- modff, 458
- nanf, 458
- nearbyintf, 458
- nextafterf, 458
- normcdf, 459
- normcdfinvf, 459
- powf, 459
- rcbrtf, 460
- remainderf, 460
- remquof, 460
- rintf, 461
- roundf, 461
- rsqrtf, 461
- scalblnf, 461
- scalbnf, 462
- signbit, 462
- sincosf, 462
- sincospif, 462
- sinf, 463
- sinhf, 463
- sinpif, 463
- sqrtf, 463
- tanf, 464
- tanhf, 464
- tgammaf, 464
- truncf, 465
- y0f, 465
- y1f, 465
- ynf, 465
- CUDA_MEM
 - cuArray3DCreate, 276
 - cuArray3DGetDescriptor, 279
 - cuArrayCreate, 279
 - cuArrayDestroy, 281
 - cuArrayGetDescriptor, 281
 - cuDeviceGetByPCIBusId, 282
 - cuDeviceGetPCIBusId, 282
 - cuIpcCloseMemHandle, 283
 - cuIpcGetEventHandle, 283
 - cuIpcGetMemHandle, 284
 - cuIpcOpenEventHandle, 284
 - cuIpcOpenMemHandle, 285
 - cuMemAlloc, 285
 - cuMemAllocHost, 286
 - cuMemAllocPitch, 287
 - cuMemcpy, 287
 - cuMemcpy2D, 288
 - cuMemcpy2DAsync, 290
 - cuMemcpy2DUnaligned, 293
 - cuMemcpy3D, 295
 - cuMemcpy3DAsync, 298
 - cuMemcpy3DPeer, 300
 - cuMemcpy3DPeerAsync, 301
 - cuMemcpyAsync, 301
 - cuMemcpyAtoA, 302
 - cuMemcpyAtoD, 303
 - cuMemcpyAtoH, 303
 - cuMemcpyAtoHAsync, 304
 - cuMemcpyDtoA, 305
 - cuMemcpyDtoD, 305
 - cuMemcpyDtoDAsync, 306
 - cuMemcpyDtoH, 306
 - cuMemcpyDtoHAsync, 307
 - cuMemcpyHtoA, 308
 - cuMemcpyHtoAAsync, 308
 - cuMemcpyHtoD, 309
 - cuMemcpyHtoDAsync, 309
 - cuMemcpyPeer, 310
 - cuMemcpyPeerAsync, 311
 - cuMemFree, 311
 - cuMemFreeHost, 312
 - cuMemGetAddressRange, 312
 - cuMemGetInfo, 313
 - cuMemHostAlloc, 313
 - cuMemHostGetDevicePointer, 315
 - cuMemHostGetFlags, 315
 - cuMemHostRegister, 316
 - cuMemHostUnregister, 317
 - cuMemsetD16, 317
 - cuMemsetD16Async, 318
 - cuMemsetD2D16, 318
 - cuMemsetD2D16Async, 319
 - cuMemsetD2D32, 320
 - cuMemsetD2D32Async, 321
 - cuMemsetD2D8, 321
 - cuMemsetD2D8Async, 322
 - cuMemsetD32, 323
 - cuMemsetD32Async, 323
 - cuMemsetD8, 324
 - cuMemsetD8Async, 324
 - cuMipmappedArrayCreate, 325
 - cuMipmappedArrayDestroy, 327
 - cuMipmappedArrayGetLevel, 328
- CUDA_MEMCPY2D
 - CUDA_TYPES, 225
- CUDA_MEMCPY2D_st, 538
 - dstArray, 538
 - dstDevice, 538
 - dstHost, 538
 - dstMemoryType, 538

- dstPitch, 538
- dstXInBytes, 538
- dstY, 539
- Height, 539
- srcArray, 539
- srcDevice, 539
- srcHost, 539
- srcMemoryType, 539
- srcPitch, 539
- srcXInBytes, 539
- srcY, 539
- WidthInBytes, 539
- CUDA_MEMCPY3D
 - CUDA_TYPES, 225
- CUDA_MEMCPY3D_PEER
 - CUDA_TYPES, 225
- CUDA_MEMCPY3D_PEER_st, 540
 - Depth, 540
 - dstArray, 540
 - dstContext, 540
 - dstDevice, 540
 - dstHeight, 541
 - dstHost, 541
 - dstLOD, 541
 - dstMemoryType, 541
 - dstPitch, 541
 - dstXInBytes, 541
 - dstY, 541
 - dstZ, 541
 - Height, 541
 - srcArray, 541
 - srcContext, 541
 - srcDevice, 541
 - srcHeight, 542
 - srcHost, 542
 - srcLOD, 542
 - srcMemoryType, 542
 - srcPitch, 542
 - srcXInBytes, 542
 - srcY, 542
 - srcZ, 542
 - WidthInBytes, 542
- CUDA_MEMCPY3D_st, 543
 - Depth, 543
 - dstArray, 543
 - dstDevice, 543
 - dstHeight, 543
 - dstHost, 544
 - dstLOD, 544
 - dstMemoryType, 544
 - dstPitch, 544
 - dstXInBytes, 544
 - dstY, 544
 - dstZ, 544
 - Height, 544
 - reserved0, 544
 - reserved1, 544
 - srcArray, 544
 - srcDevice, 544
 - srcHeight, 545
 - srcHost, 545
 - srcLOD, 545
 - srcMemoryType, 545
 - srcPitch, 545
 - srcXInBytes, 545
 - srcY, 545
 - srcZ, 545
 - WidthInBytes, 545
- CUDA_MODULE
 - cuModuleGetFunction, 265
 - cuModuleGetGlobal, 266
 - cuModuleGetSurfRef, 266
 - cuModuleGetTexRef, 267
 - cuModuleLoad, 267
 - cuModuleLoadData, 268
 - cuModuleLoadDataEx, 268
 - cuModuleLoadFatBinary, 270
 - cuModuleUnload, 270
- CUDA_PEER_ACCESS
 - cuCtxDisablePeerAccess, 378
 - cuCtxEnablePeerAccess, 378
 - cuDeviceCanAccessPeer, 379
- CUDA_POINTER_ATTRIBUTE_P2P_TOKENS
 - CUDA_TYPES, 225
- CUDA_POINTER_ATTRIBUTE_P2P_TOKENS_st, 546
- CUDA_PROFILER
 - cuProfilerInitialize, 385
 - cuProfilerStart, 386
 - cuProfilerStop, 386
- CUDA_RESOURCE_DESC
 - CUDA_TYPES, 225
- CUDA_RESOURCE_DESC_st, 547
 - devPtr, 547
 - flags, 547
 - format, 547
 - hArray, 547
 - height, 547
 - hMipmappedArray, 547
 - numChannels, 547
 - pitchInBytes, 548
 - resType, 548
 - sizeInBytes, 548
 - width, 548
- CUDA_RESOURCE_VIEW_DESC
 - CUDA_TYPES, 226
- CUDA_RESOURCE_VIEW_DESC_st, 549
 - depth, 549

- firstLayer, [549](#)
- firstMipmapLevel, [549](#)
- format, [549](#)
- height, [549](#)
- lastLayer, [549](#)
- lastMipmapLevel, [549](#)
- width, [549](#)
- CUDA_STREAM
 - cuStreamAddCallback, [333](#)
 - cuStreamCreate, [334](#)
 - cuStreamDestroy, [334](#)
 - cuStreamQuery, [335](#)
 - cuStreamSynchronize, [335](#)
 - cuStreamWaitEvent, [335](#)
- CUDA_SURF_OBJECT
 - cuSurfObjectCreate, [376](#)
 - cuSurfObjectDestroy, [376](#)
 - cuSurfObjectGetResourceDesc, [377](#)
- CUDA_SURFREF
 - cuSurfRefGetArray, [368](#)
 - cuSurfRefSetArray, [368](#)
- CUDA_TEXOBJECT
 - cuTexObjectCreate, [370](#)
 - cuTexObjectDestroy, [374](#)
 - cuTexObjectGetResourceDesc, [374](#)
 - cuTexObjectGetResourceViewDesc, [374](#)
 - cuTexObjectGetTextureDesc, [375](#)
- CUDA_TEXREF
 - cuTexRefGetAddress, [354](#)
 - cuTexRefGetAddressMode, [355](#)
 - cuTexRefGetArray, [355](#)
 - cuTexRefGetFilterMode, [355](#)
 - cuTexRefGetFlags, [356](#)
 - cuTexRefGetFormat, [356](#)
 - cuTexRefGetMaxAnisotropy, [357](#)
 - cuTexRefGetMipmapFilterMode, [357](#)
 - cuTexRefGetMipmapLevelBias, [357](#)
 - cuTexRefGetMipmapLevelClamp, [358](#)
 - cuTexRefGetMipmappedArray, [358](#)
 - cuTexRefSetAddress, [359](#)
 - cuTexRefSetAddress2D, [359](#)
 - cuTexRefSetAddressMode, [360](#)
 - cuTexRefSetArray, [361](#)
 - cuTexRefSetFilterMode, [361](#)
 - cuTexRefSetFlags, [362](#)
 - cuTexRefSetFormat, [362](#)
 - cuTexRefSetMaxAnisotropy, [363](#)
 - cuTexRefSetMipmapFilterMode, [363](#)
 - cuTexRefSetMipmapLevelBias, [364](#)
 - cuTexRefSetMipmapLevelClamp, [364](#)
 - cuTexRefSetMipmappedArray, [364](#)
- CUDA_TEXREF_DEPRECATED
 - cuTexRefCreate, [366](#)
 - cuTexRefDestroy, [366](#)
- CUDA_TEXTURE_DESC
 - CUDA_TYPES, [226](#)
- CUDA_TEXTURE_DESC_st, [551](#)
 - addressMode, [551](#)
 - filterMode, [551](#)
 - flags, [551](#)
 - maxAnisotropy, [551](#)
 - maxMipmapLevelClamp, [551](#)
 - minMipmapLevelClamp, [551](#)
 - mipmapFilterMode, [551](#)
 - mipmapLevelBias, [551](#)
- CUDA_TYPES
 - CU_IPC_HANDLE_SIZE, [222](#)
 - CU_LAUNCH_PARAM_BUFFER_POINTER, [222](#)
 - CU_LAUNCH_PARAM_BUFFER_SIZE, [222](#)
 - CU_LAUNCH_PARAM_END, [223](#)
 - CU_MEMHOSTALLOC_DEVICEMAP, [223](#)
 - CU_MEMHOSTALLOC_PORTABLE, [223](#)
 - CU_MEMHOSTALLOC_WRITECOMBINED, [223](#)
 - CU_MEMHOSTREGISTER_DEVICEMAP, [223](#)
 - CU_MEMHOSTREGISTER_PORTABLE, [223](#)
 - CU_PARAM_TR_DEFAULT, [223](#)
 - CU_TRSA_OVERRIDE_FORMAT, [223](#)
 - CU_TRSF_NORMALIZED_COORDINATES, [223](#)
 - CU_TRSF_READ_AS_INTEGER, [223](#)
 - CU_TRSF_SRGB, [224](#)
 - CUaddress_mode, [224](#)
 - CUaddress_mode_enum, [229](#)
 - CUarray, [224](#)
 - CUarray_cubemap_face, [224](#)
 - CUarray_cubemap_face_enum, [229](#)
 - CUarray_format, [225](#)
 - CUarray_format_enum, [229](#)
 - CUcomputemode, [225](#)
 - CUcomputemode_enum, [230](#)
 - CUcontext, [225](#)
 - CUctx_flags, [225](#)
 - CUctx_flags_enum, [230](#)
 - CUDA_ARRAY3D_2DARRAY, [224](#)
 - CUDA_ARRAY3D_CUBEMAP, [224](#)
 - CUDA_ARRAY3D_DESCRIPTOR, [225](#)
 - CUDA_ARRAY3D_LAYERED, [224](#)
 - CUDA_ARRAY3D_SURFACE_LDST, [224](#)
 - CUDA_ARRAY3D_TEXTURE_GATHER, [224](#)
 - CUDA_ARRAY_DESCRIPTOR, [225](#)
 - CUDA_MEMCPY2D, [225](#)
 - CUDA_MEMCPY3D, [225](#)
 - CUDA_MEMCPY3D_PEER, [225](#)
 - CUDA_POINTER_ATTRIBUTE_P2P_TOKENS, [225](#)
 - CUDA_RESOURCE_DESC, [225](#)
 - CUDA_RESOURCE_VIEW_DESC, [226](#)
 - CUDA_TEXTURE_DESC, [226](#)

- CUDA_VERSION, 224
- cudaError_enum, 230
- CUdevice, 226
- CUdevice_attribute, 226
- CUdevice_attribute_enum, 233
- CUdeviceptr, 226
- CUdevprop, 226
- CUEvent, 226
- CUEvent_flags, 226
- CUEvent_flags_enum, 236
- CUfilter_mode, 226
- CUfilter_mode_enum, 236
- CUfunc_cache, 226
- CUfunc_cache_enum, 236
- CUfunction, 226
- CUfunction_attribute, 226
- CUfunction_attribute_enum, 237
- CUgraphicsMapResourceFlags, 227
- CUgraphicsMapResourceFlags_enum, 237
- CUgraphicsRegisterFlags, 227
- CUgraphicsRegisterFlags_enum, 237
- CUgraphicsResource, 227
- CUipcEventHandle, 227
- CUipcMem_flags, 227
- CUipcMem_flags_enum, 237
- CUipcMemHandle, 227
- CUjit_fallback, 227
- CUjit_fallback_enum, 237
- CUjit_option, 227
- CUjit_option_enum, 238
- CUjit_target, 227
- CUjit_target_enum, 239
- CULimit, 227
- CULimit_enum, 239
- CUMemorytype, 227
- CUMemorytype_enum, 239
- CUmipmappedArray, 227
- CUmodule, 228
- CUpointer_attribute, 228
- CUpointer_attribute_enum, 239
- CUresourcetype, 228
- CUresourcetype_enum, 240
- CUresourceViewFormat, 228
- CUresourceViewFormat_enum, 240
- CUresult, 228
- CUsharedconfig, 228
- CUsharedconfig_enum, 241
- CUstream, 228
- CUstream_flags, 228
- CUstream_flags_enum, 241
- CUstreamCallback, 228
- CUsurfObject, 228
- CUsurfref, 229
- CUtexObject, 229
- CUtexref, 229
- CUDA_UNIFIED
 - cuPointerGetAttribute, 330
- CUDA_VDPAU
 - cuGraphicsVDPAURegisterOutputSurface, 435
 - cuGraphicsVDPAURegisterVideoSurface, 436
 - cuVDPAUCtxCreate, 437
 - cuVDPAUGetDevice, 437
- CUDA_VERSION
 - CUDA_TYPES, 224
 - cuDriverGetVersion, 243
- cudaAddressModeBorder
 - CUDART_TYPES, 213
- cudaAddressModeClamp
 - CUDART_TYPES, 213
- cudaAddressModeMirror
 - CUDART_TYPES, 213
- cudaAddressModeWrap
 - CUDART_TYPES, 213
- cudaArray_const_t
 - CUDART_TYPES, 201
- cudaArray_t
 - CUDART_TYPES, 201
- cudaArrayCubemap
 - CUDART_TYPES, 198
- cudaArrayDefault
 - CUDART_TYPES, 198
- cudaArrayGetInfo
 - CUDART_MEMORY, 60
- cudaArrayLayered
 - CUDART_TYPES, 198
- cudaArraySurfaceLoadStore
 - CUDART_TYPES, 199
- cudaArrayTextureGather
 - CUDART_TYPES, 199
- cudaBindSurfaceToArray
 - CUDART_HIGHLEVEL, 172
 - CUDART_SURFACE, 159
- cudaBindTexture
 - CUDART_HIGHLEVEL, 173
 - CUDART_TEXTURE, 153
- cudaBindTexture2D
 - CUDART_HIGHLEVEL, 174, 175
 - CUDART_TEXTURE, 154
- cudaBindTextureToArray
 - CUDART_HIGHLEVEL, 175, 176
 - CUDART_TEXTURE, 155
- cudaBindTextureToMipmappedArray
 - CUDART_HIGHLEVEL, 176, 177
 - CUDART_TEXTURE, 155
- cudaBoundaryModeClamp
 - CUDART_TYPES, 212
- cudaBoundaryModeTrap
 - CUDART_TYPES, 212

- cudaBoundaryModeZero
 - CUDART_TYPES, [212](#)
- cudaChannelFormatDesc, [553](#)
 - f, [553](#)
 - w, [553](#)
 - x, [553](#)
 - y, [553](#)
 - z, [553](#)
- cudaChannelFormatKind
 - CUDART_TYPES, [202](#)
- cudaChannelFormatKindFloat
 - CUDART_TYPES, [202](#)
- cudaChannelFormatKindNone
 - CUDART_TYPES, [202](#)
- cudaChannelFormatKindSigned
 - CUDART_TYPES, [202](#)
- cudaChannelFormatKindUnsigned
 - CUDART_TYPES, [202](#)
- cudaChooseDevice
 - CUDART_DEVICE, [16](#)
- cudaComputeMode
 - CUDART_TYPES, [202](#)
- cudaComputeModeDefault
 - CUDART_TYPES, [203](#)
- cudaComputeModeExclusive
 - CUDART_TYPES, [203](#)
- cudaComputeModeExclusiveProcess
 - CUDART_TYPES, [203](#)
- cudaComputeModeProhibited
 - CUDART_TYPES, [203](#)
- cudaConfigureCall
 - CUDART_EXECUTION, [51](#)
- cudaCreateChannelDesc
 - CUDART_HIGHLEVEL, [177](#)
 - CUDART_TEXTURE, [156](#)
- cudaCreateSurfaceObject
 - CUDART_SURFACE_OBJECT, [167](#)
- cudaCreateTextureObject
 - CUDART_TEXTURE_OBJECT, [161](#)
- cudaCSV
 - CUDART_TYPES, [211](#)
- cudaD3D10DeviceList
 - CUDART_D3D10, [125](#)
- cudaD3D10DeviceListAll
 - CUDART_D3D10, [125](#)
- cudaD3D10DeviceListCurrentFrame
 - CUDART_D3D10, [125](#)
- cudaD3D10DeviceListNextFrame
 - CUDART_D3D10, [125](#)
- cudaD3D10GetDevice
 - CUDART_D3D10, [125](#)
- cudaD3D10GetDevices
 - CUDART_D3D10, [126](#)
- cudaD3D10GetDirect3DDevice
 - CUDART_D3D10_DEPRECATED, [130](#)
- cudaD3D10MapFlags
 - CUDART_D3D10_DEPRECATED, [130](#)
- cudaD3D10MapFlagsNone
 - CUDART_D3D10_DEPRECATED, [130](#)
- cudaD3D10MapFlagsReadOnly
 - CUDART_D3D10_DEPRECATED, [130](#)
- cudaD3D10MapFlagsWriteDiscard
 - CUDART_D3D10_DEPRECATED, [130](#)
- cudaD3D10MapResources
 - CUDART_D3D10_DEPRECATED, [131](#)
- cudaD3D10RegisterFlags
 - CUDART_D3D10_DEPRECATED, [130](#)
- cudaD3D10RegisterFlagsArray
 - CUDART_D3D10_DEPRECATED, [130](#)
- cudaD3D10RegisterFlagsNone
 - CUDART_D3D10_DEPRECATED, [130](#)
- cudaD3D10RegisterResource
 - CUDART_D3D10_DEPRECATED, [131](#)
- cudaD3D10ResourceGetMappedArray
 - CUDART_D3D10_DEPRECATED, [132](#)
- cudaD3D10ResourceGetMappedPitch
 - CUDART_D3D10_DEPRECATED, [133](#)
- cudaD3D10ResourceGetMappedPointer
 - CUDART_D3D10_DEPRECATED, [134](#)
- cudaD3D10ResourceGetMappedSize
 - CUDART_D3D10_DEPRECATED, [134](#)
- cudaD3D10ResourceGetSurfaceDimensions
 - CUDART_D3D10_DEPRECATED, [135](#)
- cudaD3D10ResourceSetMapFlags
 - CUDART_D3D10_DEPRECATED, [136](#)
- cudaD3D10SetDirect3DDevice
 - CUDART_D3D10_DEPRECATED, [136](#)
- cudaD3D10UnmapResources
 - CUDART_D3D10_DEPRECATED, [137](#)
- cudaD3D10UnregisterResource
 - CUDART_D3D10_DEPRECATED, [137](#)
- cudaD3D11DeviceList
 - CUDART_D3D11, [139](#)
- cudaD3D11DeviceListAll
 - CUDART_D3D11, [139](#)
- cudaD3D11DeviceListCurrentFrame
 - CUDART_D3D11, [139](#)
- cudaD3D11DeviceListNextFrame
 - CUDART_D3D11, [139](#)
- cudaD3D11GetDevice
 - CUDART_D3D11, [139](#)
- cudaD3D11GetDevices
 - CUDART_D3D11, [140](#)
- cudaD3D11GetDirect3DDevice
 - CUDART_D3D11_DEPRECATED, [143](#)
- cudaD3D11SetDirect3DDevice
 - CUDART_D3D11_DEPRECATED, [143](#)
- cudaD3D9DeviceList

- CUDART_D3D9, [111](#)
- cudaD3D9DeviceListAll
 - CUDART_D3D9, [111](#)
- cudaD3D9DeviceListCurrentFrame
 - CUDART_D3D9, [111](#)
- cudaD3D9DeviceListNextFrame
 - CUDART_D3D9, [111](#)
- cudaD3D9GetDevice
 - CUDART_D3D9, [112](#)
- cudaD3D9GetDevices
 - CUDART_D3D9, [112](#)
- cudaD3D9GetDirect3DDevice
 - CUDART_D3D9, [112](#)
- cudaD3D9MapFlags
 - CUDART_D3D9_DEPRECATED, [117](#)
- cudaD3D9MapFlagsNone
 - CUDART_D3D9_DEPRECATED, [117](#)
- cudaD3D9MapFlagsReadOnly
 - CUDART_D3D9_DEPRECATED, [117](#)
- cudaD3D9MapFlagsWriteDiscard
 - CUDART_D3D9_DEPRECATED, [117](#)
- cudaD3D9MapResources
 - CUDART_D3D9_DEPRECATED, [117](#)
- cudaD3D9RegisterFlags
 - CUDART_D3D9_DEPRECATED, [117](#)
- cudaD3D9RegisterFlagsArray
 - CUDART_D3D9_DEPRECATED, [117](#)
- cudaD3D9RegisterFlagsNone
 - CUDART_D3D9_DEPRECATED, [117](#)
- cudaD3D9RegisterResource
 - CUDART_D3D9_DEPRECATED, [118](#)
- cudaD3D9ResourceGetMappedArray
 - CUDART_D3D9_DEPRECATED, [119](#)
- cudaD3D9ResourceGetMappedPitch
 - CUDART_D3D9_DEPRECATED, [119](#)
- cudaD3D9ResourceGetMappedPointer
 - CUDART_D3D9_DEPRECATED, [120](#)
- cudaD3D9ResourceGetMappedSize
 - CUDART_D3D9_DEPRECATED, [121](#)
- cudaD3D9ResourceGetSurfaceDimensions
 - CUDART_D3D9_DEPRECATED, [122](#)
- cudaD3D9ResourceSetMapFlags
 - CUDART_D3D9_DEPRECATED, [122](#)
- cudaD3D9SetDirect3DDevice
 - CUDART_D3D9, [113](#)
- cudaD3D9UnmapResources
 - CUDART_D3D9_DEPRECATED, [123](#)
- cudaD3D9UnregisterResource
 - CUDART_D3D9_DEPRECATED, [124](#)
- cudaDestroySurfaceObject
 - CUDART_SURFACE_OBJECT, [167](#)
- cudaDestroyTextureObject
 - CUDART_TEXTURE_OBJECT, [165](#)
- cudaDevAttrAsyncEngineCount
 - CUDART_TYPES, [204](#)
- cudaDevAttrCanMapHostMemory
 - CUDART_TYPES, [203](#)
- cudaDevAttrClockRate
 - CUDART_TYPES, [203](#)
- cudaDevAttrComputeCapabilityMajor
 - CUDART_TYPES, [205](#)
- cudaDevAttrComputeCapabilityMinor
 - CUDART_TYPES, [205](#)
- cudaDevAttrComputeMode
 - CUDART_TYPES, [203](#)
- cudaDevAttrConcurrentKernels
 - CUDART_TYPES, [204](#)
- cudaDevAttrEccEnabled
 - CUDART_TYPES, [204](#)
- cudaDevAttrGlobalMemoryBusWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrGpuOverlap
 - CUDART_TYPES, [203](#)
- cudaDevAttrIntegrated
 - CUDART_TYPES, [203](#)
- cudaDevAttrKernelExecTimeout
 - CUDART_TYPES, [203](#)
- cudaDevAttrL2CacheSize
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxBlockDimX
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxBlockDimY
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxBlockDimZ
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxGridDimX
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxGridDimY
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxGridDimZ
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxPitch
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxRegistersPerBlock
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxSharedMemoryPerBlock
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxSurface1DLayeredLayers
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface1DLayeredWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface1DWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface2DHeight
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface2DLayeredHeight
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface2DLayeredLayers

- CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface2DLayeredWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface2DWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface3DDepth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface3DHeight
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurface3DWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurfaceCubemapLayeredLayers
 - CUDART_TYPES, [205](#)
- cudaDevAttrMaxSurfaceCubemapLayeredWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxSurfaceCubemapWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture1DLayeredLayers
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture1DLayeredWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture1DLinearWidth
 - CUDART_TYPES, [205](#)
- cudaDevAttrMaxTexture1DMipmappedWidth
 - CUDART_TYPES, [205](#)
- cudaDevAttrMaxTexture1DWidth
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxTexture2DGatherHeight
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture2DGatherWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture2DHeight
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxTexture2DLayeredHeight
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture2DLayeredLayers
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture2DLayeredWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture2DLinearHeight
 - CUDART_TYPES, [205](#)
- cudaDevAttrMaxTexture2DLinearPitch
 - CUDART_TYPES, [205](#)
- cudaDevAttrMaxTexture2DLinearWidth
 - CUDART_TYPES, [205](#)
- cudaDevAttrMaxTexture2DMipmappedHeight
 - CUDART_TYPES, [205](#)
- cudaDevAttrMaxTexture2DMipmappedWidth
 - CUDART_TYPES, [205](#)
- cudaDevAttrMaxTexture2DWidth
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxTexture3DDepth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture3DDepthAlt
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture3DHeight
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxTexture3DHeightAlt
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTexture3DWidth
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxTexture3DWidthAlt
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTextureCubemapLayeredLayers
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTextureCubemapLayeredWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxTextureCubemapWidth
 - CUDART_TYPES, [204](#)
- cudaDevAttrMaxThreadsPerBlock
 - CUDART_TYPES, [203](#)
- cudaDevAttrMaxThreadsPerMultiProcessor
 - CUDART_TYPES, [204](#)
- cudaDevAttrMemoryClockRate
 - CUDART_TYPES, [204](#)
- cudaDevAttrMultiProcessorCount
 - CUDART_TYPES, [203](#)
- cudaDevAttrPciBusId
 - CUDART_TYPES, [204](#)
- cudaDevAttrPciDeviceId
 - CUDART_TYPES, [204](#)
- cudaDevAttrPciDomainId
 - CUDART_TYPES, [204](#)
- cudaDevAttrSurfaceAlignment
 - CUDART_TYPES, [204](#)
- cudaDevAttrTccDriver
 - CUDART_TYPES, [204](#)
- cudaDevAttrTextureAlignment
 - CUDART_TYPES, [203](#)
- cudaDevAttrTexturePitchAlignment
 - CUDART_TYPES, [204](#)
- cudaDevAttrTotalConstantMemory
 - CUDART_TYPES, [203](#)
- cudaDevAttrUnifiedAddressing
 - CUDART_TYPES, [204](#)
- cudaDevAttrWarpSize
 - CUDART_TYPES, [203](#)
- cudaDeviceAttr
 - CUDART_TYPES, [203](#)
- cudaDeviceBlockingSync
 - CUDART_TYPES, [199](#)
- cudaDeviceCanAccessPeer
 - CUDART_PEER, [99](#)
- cudaDeviceDisablePeerAccess
 - CUDART_PEER, [99](#)
- cudaDeviceEnablePeerAccess
 - CUDART_PEER, [100](#)
- cudaDeviceGetAttribute

- CUDART_DEVICE, 17
- cudaDeviceGetByPCIBusId
 - CUDART_DEVICE, 20
- cudaDeviceGetCacheConfig
 - CUDART_DEVICE, 20
- cudaDeviceGetLimit
 - CUDART_DEVICE, 21
- cudaDeviceGetPCIBusId
 - CUDART_DEVICE, 21
- cudaDeviceGetSharedMemConfig
 - CUDART_DEVICE, 22
- cudaDeviceLmemResizeToMax
 - CUDART_TYPES, 199
- cudaDeviceMapHost
 - CUDART_TYPES, 199
- cudaDeviceMask
 - CUDART_TYPES, 199
- cudaDeviceProp, 554
 - asyncEngineCount, 555
 - canMapHostMemory, 555
 - clockRate, 555
 - computeMode, 555
 - concurrentKernels, 555
 - deviceOverlap, 555
 - ECCEntered, 555
 - integrated, 555
 - kernelExecTimeoutEnabled, 555
 - l2CacheSize, 556
 - major, 556
 - maxGridSize, 556
 - maxSurface1D, 556
 - maxSurface1DLayered, 556
 - maxSurface2D, 556
 - maxSurface2DLayered, 556
 - maxSurface3D, 556
 - maxSurfaceCubemap, 556
 - maxSurfaceCubemapLayered, 556
 - maxTexture1D, 556
 - maxTexture1DLayered, 556
 - maxTexture1DLinear, 557
 - maxTexture1DMipmap, 557
 - maxTexture2D, 557
 - maxTexture2DGather, 557
 - maxTexture2DLayered, 557
 - maxTexture2DLinear, 557
 - maxTexture2DMipmap, 557
 - maxTexture3D, 557
 - maxTextureCubemap, 557
 - maxTextureCubemapLayered, 557
 - maxThreadsDim, 557
 - maxThreadsPerBlock, 557
 - maxThreadsPerMultiProcessor, 558
 - memoryBusWidth, 558
 - memoryClockRate, 558
 - memPitch, 558
 - minor, 558
 - multiProcessorCount, 558
 - name, 558
 - pciBusID, 558
 - pciDeviceID, 558
 - pciDomainID, 558
 - regsPerBlock, 558
 - sharedMemPerBlock, 558
 - surfaceAlignment, 559
 - tccDriver, 559
 - textureAlignment, 559
 - texturePitchAlignment, 559
 - totalConstMem, 559
 - totalGlobalMem, 559
 - unifiedAddressing, 559
 - warpSize, 559
- cudaDevicePropDontCare
 - CUDART_TYPES, 199
- cudaDeviceReset
 - CUDART_DEVICE, 22
- cudaDeviceScheduleAuto
 - CUDART_TYPES, 199
- cudaDeviceScheduleBlockingSync
 - CUDART_TYPES, 199
- cudaDeviceScheduleMask
 - CUDART_TYPES, 199
- cudaDeviceScheduleSpin
 - CUDART_TYPES, 200
- cudaDeviceScheduleYield
 - CUDART_TYPES, 200
- cudaDeviceSetCacheConfig
 - CUDART_DEVICE, 23
- cudaDeviceSetLimit
 - CUDART_DEVICE, 23
- cudaDeviceSetSharedMemConfig
 - CUDART_DEVICE, 24
- cudaDeviceSynchronize
 - CUDART_DEVICE, 25
- cudaDriverGetVersion
 - CUDART__VERSION, 169
- cudaError
 - CUDART_TYPES, 205
- cudaError_enum
 - CUDA_TYPES, 230
- cudaError_t
 - CUDART_TYPES, 201
- cudaErrorAddressOfConstant
 - CUDART_TYPES, 206
- cudaErrorApiFailureBase
 - CUDART_TYPES, 209
- cudaErrorAssert
 - CUDART_TYPES, 208
- cudaErrorCudartUnloading

- CUDART_TYPES, [207](#)
- cudaErrorDeviceAlreadyInUse
 - CUDART_TYPES, [208](#)
- cudaErrorDevicesUnavailable
 - CUDART_TYPES, [208](#)
- cudaErrorDuplicateSurfaceName
 - CUDART_TYPES, [207](#)
- cudaErrorDuplicateTextureName
 - CUDART_TYPES, [207](#)
- cudaErrorDuplicateVariableName
 - CUDART_TYPES, [207](#)
- cudaErrorECCUncorrectable
 - CUDART_TYPES, [207](#)
- cudaErrorHostMemoryAlreadyRegistered
 - CUDART_TYPES, [208](#)
- cudaErrorHostMemoryNotRegistered
 - CUDART_TYPES, [208](#)
- cudaErrorIncompatibleDriverContext
 - CUDART_TYPES, [208](#)
- cudaErrorInitializationError
 - CUDART_TYPES, [205](#)
- cudaErrorInsufficientDriver
 - CUDART_TYPES, [207](#)
- cudaErrorInvalidChannelDescriptor
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidConfiguration
 - CUDART_TYPES, [205](#)
- cudaErrorInvalidDevice
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidDeviceFunction
 - CUDART_TYPES, [205](#)
- cudaErrorInvalidDevicePointer
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidFilterSetting
 - CUDART_TYPES, [207](#)
- cudaErrorInvalidHostPointer
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidKernelImage
 - CUDART_TYPES, [208](#)
- cudaErrorInvalidMemcpyDirection
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidNormSetting
 - CUDART_TYPES, [207](#)
- cudaErrorInvalidPitchValue
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidResourceHandle
 - CUDART_TYPES, [207](#)
- cudaErrorInvalidSurface
 - CUDART_TYPES, [207](#)
- cudaErrorInvalidSymbol
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidTexture
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidTextureBinding
 - CUDART_TYPES, [206](#)
- cudaErrorInvalidValue
 - CUDART_TYPES, [206](#)
- cudaErrorLaunchFailure
 - CUDART_TYPES, [205](#)
- cudaErrorLaunchFileScopedSurf
 - CUDART_TYPES, [209](#)
- cudaErrorLaunchFileScopedTex
 - CUDART_TYPES, [208](#)
- cudaErrorLaunchMaxDepthExceeded
 - CUDART_TYPES, [208](#)
- cudaErrorLaunchOutOfResources
 - CUDART_TYPES, [205](#)
- cudaErrorLaunchPendingCountExceeded
 - CUDART_TYPES, [209](#)
- cudaErrorLaunchTimeout
 - CUDART_TYPES, [205](#)
- cudaErrorMapBufferObjectFailed
 - CUDART_TYPES, [206](#)
- cudaErrorMemoryAllocation
 - CUDART_TYPES, [205](#)
- cudaErrorMemoryValueTooLarge
 - CUDART_TYPES, [207](#)
- cudaErrorMissingConfiguration
 - CUDART_TYPES, [205](#)
- cudaErrorMixedDeviceExecution
 - CUDART_TYPES, [207](#)
- cudaErrorNoDevice
 - CUDART_TYPES, [207](#)
- cudaErrorNoKernelImageForDevice
 - CUDART_TYPES, [208](#)
- cudaErrorNotPermitted
 - CUDART_TYPES, [209](#)
- cudaErrorNotReady
 - CUDART_TYPES, [207](#)
- cudaErrorNotSupported
 - CUDART_TYPES, [209](#)
- cudaErrorNotYetImplemented
 - CUDART_TYPES, [207](#)
- cudaErrorOperatingSystem
 - CUDART_TYPES, [208](#)
- cudaErrorPeerAccessAlreadyEnabled
 - CUDART_TYPES, [208](#)
- cudaErrorPeerAccessNotEnabled
 - CUDART_TYPES, [208](#)
- cudaErrorPeerAccessUnsupported
 - CUDART_TYPES, [208](#)
- cudaErrorPriorLaunchFailure
 - CUDART_TYPES, [205](#)
- cudaErrorProfilerAlreadyStarted
 - CUDART_TYPES, [208](#)
- cudaErrorProfilerAlreadyStopped
 - CUDART_TYPES, [208](#)
- cudaErrorProfilerDisabled

- CUDART_TYPES, 208
- cudaErrorProfilerNotInitialized
 - CUDART_TYPES, 208
- cudaErrorSetOnActiveProcess
 - CUDART_TYPES, 207
- cudaErrorSharedObjectInitFailed
 - CUDART_TYPES, 207
- cudaErrorSharedObjectSymbolNotFound
 - CUDART_TYPES, 207
- cudaErrorStartupFailure
 - CUDART_TYPES, 209
- cudaErrorSyncDepthExceeded
 - CUDART_TYPES, 209
- cudaErrorSynchronizationError
 - CUDART_TYPES, 206
- cudaErrorTextureFetchFailed
 - CUDART_TYPES, 206
- cudaErrorTextureNotBound
 - CUDART_TYPES, 206
- cudaErrorTooManyPeers
 - CUDART_TYPES, 208
- cudaErrorUnknown
 - CUDART_TYPES, 207
- cudaErrorUnmapBufferObjectFailed
 - CUDART_TYPES, 206
- cudaErrorUnsupportedLimit
 - CUDART_TYPES, 207
- cudaEvent_t
 - CUDART_TYPES, 201
- cudaEventBlockingSync
 - CUDART_TYPES, 200
- cudaEventCreate
 - CUDART_EVENT, 47
 - CUDART_HIGHLEVEL, 178
- cudaEventCreateWithFlags
 - CUDART_EVENT, 47
- cudaEventDefault
 - CUDART_TYPES, 200
- cudaEventDestroy
 - CUDART_EVENT, 48
- cudaEventDisableTiming
 - CUDART_TYPES, 200
- cudaEventElapsedTime
 - CUDART_EVENT, 48
- cudaEventInterprocess
 - CUDART_TYPES, 200
- cudaEventQuery
 - CUDART_EVENT, 49
- cudaEventRecord
 - CUDART_EVENT, 49
- cudaEventSynchronize
 - CUDART_EVENT, 50
- cudaExtent, 560
 - depth, 560
 - height, 560
 - width, 560
- cudaFilterModeLinear
 - CUDART_TYPES, 213
- cudaFilterModePoint
 - CUDART_TYPES, 213
- cudaFormatModeAuto
 - CUDART_TYPES, 212
- cudaFormatModeForced
 - CUDART_TYPES, 212
- cudaFree
 - CUDART_MEMORY, 60
- cudaFreeArray
 - CUDART_MEMORY, 60
- cudaFreeHost
 - CUDART_MEMORY, 61
- cudaFreeMipmappedArray
 - CUDART_MEMORY, 61
- cudaFuncAttributes, 561
 - binaryVersion, 561
 - constSizeBytes, 561
 - localSizeBytes, 561
 - maxThreadsPerBlock, 561
 - numRegs, 561
 - ptxVersion, 561
 - sharedSizeBytes, 561
- cudaFuncCache
 - CUDART_TYPES, 209
- cudaFuncCachePreferEqual
 - CUDART_TYPES, 209
- cudaFuncCachePreferL1
 - CUDART_TYPES, 209
- cudaFuncCachePreferNone
 - CUDART_TYPES, 209
- cudaFuncCachePreferShared
 - CUDART_TYPES, 209
- cudaFuncGetAttributes
 - CUDART_EXECUTION, 52
 - CUDART_HIGHLEVEL, 178
- cudaFuncSetCacheConfig
 - CUDART_EXECUTION, 52
 - CUDART_HIGHLEVEL, 179
- cudaFuncSetSharedMemConfig
 - CUDART_EXECUTION, 53
- cudaGetChannelDesc
 - CUDART_TEXTURE, 156
- cudaGetDevice
 - CUDART_DEVICE, 25
- cudaGetDeviceCount
 - CUDART_DEVICE, 26
- cudaGetDeviceProperties
 - CUDART_DEVICE, 26
- cudaGetErrorString
 - CUDART_ERROR, 40

- cudaGetLastError
 - CUDART_ERROR, 40
- cudaGetMipmappedArrayLevel
 - CUDART_MEMORY, 61
- cudaGetSurfaceObjectResourceDesc
 - CUDART_SURFACE_OBJECT, 168
- cudaGetSurfaceReference
 - CUDART_SURFACE, 159
- cudaGetSymbolAddress
 - CUDART_HIGHLEVEL, 180
 - CUDART_MEMORY, 62
- cudaGetSymbolSize
 - CUDART_HIGHLEVEL, 180
 - CUDART_MEMORY, 62
- cudaGetTextureAlignmentOffset
 - CUDART_HIGHLEVEL, 180
 - CUDART_TEXTURE, 157
- cudaGetTextureObjectResourceDesc
 - CUDART_TEXTURE_OBJECT, 165
- cudaGetTextureObjectResourceViewDesc
 - CUDART_TEXTURE_OBJECT, 165
- cudaGetTextureObjectTextureDesc
 - CUDART_TEXTURE_OBJECT, 166
- cudaGetTextureReference
 - CUDART_TEXTURE, 157
- cudaGLDeviceList
 - CUDART_OPENGL, 101
- cudaGLDeviceListAll
 - CUDART_OPENGL, 101
- cudaGLDeviceListCurrentFrame
 - CUDART_OPENGL, 101
- cudaGLDeviceListNextFrame
 - CUDART_OPENGL, 101
- cudaGLGetDevices
 - CUDART_OPENGL, 102
- cudaGLMapBufferObject
 - CUDART_OPENGL_DEPRECATED, 106
- cudaGLMapBufferObjectAsync
 - CUDART_OPENGL_DEPRECATED, 106
- cudaGLMapFlags
 - CUDART_OPENGL_DEPRECATED, 105
- cudaGLMapFlagsNone
 - CUDART_OPENGL_DEPRECATED, 105
- cudaGLMapFlagsReadOnly
 - CUDART_OPENGL_DEPRECATED, 105
- cudaGLMapFlagsWriteDiscard
 - CUDART_OPENGL_DEPRECATED, 106
- cudaGLRegisterBufferObject
 - CUDART_OPENGL_DEPRECATED, 107
- cudaGLSetBufferObjectMapFlags
 - CUDART_OPENGL_DEPRECATED, 107
- cudaGLSetGLDevice
 - CUDART_OPENGL_DEPRECATED, 108
- cudaGLUnmapBufferObject
 - CUDART_OPENGL_DEPRECATED, 108
- cudaGLUnmapBufferObjectAsync
 - CUDART_OPENGL_DEPRECATED, 109
- cudaGLUnregisterBufferObject
 - CUDART_OPENGL_DEPRECATED, 109
- cudaGraphicsCubeFace
 - CUDART_TYPES, 209
- cudaGraphicsCubeFaceNegativeX
 - CUDART_TYPES, 209
- cudaGraphicsCubeFaceNegativeY
 - CUDART_TYPES, 209
- cudaGraphicsCubeFaceNegativeZ
 - CUDART_TYPES, 210
- cudaGraphicsCubeFacePositiveX
 - CUDART_TYPES, 209
- cudaGraphicsCubeFacePositiveY
 - CUDART_TYPES, 209
- cudaGraphicsCubeFacePositiveZ
 - CUDART_TYPES, 210
- cudaGraphicsD3D10RegisterResource
 - CUDART_D3D10, 126
- cudaGraphicsD3D11RegisterResource
 - CUDART_D3D11, 140
- cudaGraphicsD3D9RegisterResource
 - CUDART_D3D9, 113
- cudaGraphicsGLRegisterBuffer
 - CUDART_OPENGL, 102
- cudaGraphicsGLRegisterImage
 - CUDART_OPENGL, 103
- cudaGraphicsMapFlags
 - CUDART_TYPES, 210
- cudaGraphicsMapFlagsNone
 - CUDART_TYPES, 210
- cudaGraphicsMapFlagsReadOnly
 - CUDART_TYPES, 210
- cudaGraphicsMapFlagsWriteDiscard
 - CUDART_TYPES, 210
- cudaGraphicsMapResources
 - CUDART_INTEROP, 148
- cudaGraphicsRegisterFlags
 - CUDART_TYPES, 210
- cudaGraphicsRegisterFlagsNone
 - CUDART_TYPES, 210
- cudaGraphicsRegisterFlagsReadOnly
 - CUDART_TYPES, 210
- cudaGraphicsRegisterFlagsSurfaceLoadStore
 - CUDART_TYPES, 210
- cudaGraphicsRegisterFlagsTextureGather
 - CUDART_TYPES, 210
- cudaGraphicsRegisterFlagsWriteDiscard
 - CUDART_TYPES, 210
- cudaGraphicsResource_t
 - CUDART_TYPES, 201
- cudaGraphicsResourceGetMappedMipmappedArray

- CUDART_INTEROP, [149](#)
- cudaGraphicsResourceGetMappedPointer
 - CUDART_INTEROP, [149](#)
- cudaGraphicsResourceSetMapFlags
 - CUDART_INTEROP, [150](#)
- cudaGraphicsSubResourceGetMappedArray
 - CUDART_INTEROP, [150](#)
- cudaGraphicsUnmapResources
 - CUDART_INTEROP, [151](#)
- cudaGraphicsUnregisterResource
 - CUDART_INTEROP, [152](#)
- cudaGraphicsVDPAURegisterOutputSurface
 - CUDART_VDPAU, [145](#)
- cudaGraphicsVDPAURegisterVideoSurface
 - CUDART_VDPAU, [146](#)
- cudaHostAlloc
 - CUDART_MEMORY, [63](#)
- cudaHostAllocDefault
 - CUDART_TYPES, [200](#)
- cudaHostAllocMapped
 - CUDART_TYPES, [200](#)
- cudaHostAllocPortable
 - CUDART_TYPES, [200](#)
- cudaHostAllocWriteCombined
 - CUDART_TYPES, [200](#)
- cudaHostGetDevicePointer
 - CUDART_MEMORY, [64](#)
- cudaHostGetFlags
 - CUDART_MEMORY, [64](#)
- cudaHostRegister
 - CUDART_MEMORY, [65](#)
- cudaHostRegisterDefault
 - CUDART_TYPES, [200](#)
- cudaHostRegisterMapped
 - CUDART_TYPES, [200](#)
- cudaHostRegisterPortable
 - CUDART_TYPES, [201](#)
- cudaHostUnregister
 - CUDART_MEMORY, [66](#)
- cudaIpcCloseMemHandle
 - CUDART_DEVICE, [29](#)
- cudaIpcEventHandle_st, [563](#)
- cudaIpcEventHandle_t
 - CUDART_TYPES, [201](#)
- cudaIpcGetEventHandle
 - CUDART_DEVICE, [30](#)
- cudaIpcGetMemHandle
 - CUDART_DEVICE, [30](#)
- cudaIpcMemHandle_st, [564](#)
- cudaIpcMemHandle_t
 - CUDART_TYPES, [202](#)
- cudaIpcMemLazyEnablePeerAccess
 - CUDART_TYPES, [201](#)
- cudaIpcOpenEventHandle
 - CUDART_DEVICE, [31](#)
- cudaIpcOpenMemHandle
 - CUDART_DEVICE, [31](#)
- cudaKeyValuePair
 - CUDART_TYPES, [211](#)
- cudaLaunch
 - CUDART_EXECUTION, [54](#)
 - CUDART_HIGHLEVEL, [181](#)
- cudaLimit
 - CUDART_TYPES, [210](#)
- cudaLimitDevRuntimePendingLaunchCount
 - CUDART_TYPES, [210](#)
- cudaLimitDevRuntimeSyncDepth
 - CUDART_TYPES, [210](#)
- cudaLimitMallocHeapSize
 - CUDART_TYPES, [210](#)
- cudaLimitPrintfFifoSize
 - CUDART_TYPES, [210](#)
- cudaLimitStackSize
 - CUDART_TYPES, [210](#)
- cudaMalloc
 - CUDART_MEMORY, [66](#)
- cudaMalloc3D
 - CUDART_MEMORY, [66](#)
- cudaMalloc3DArray
 - CUDART_MEMORY, [67](#)
- cudaMallocArray
 - CUDART_MEMORY, [69](#)
- cudaMallocHost
 - CUDART_HIGHLEVEL, [181](#)
 - CUDART_MEMORY, [70](#)
- cudaMallocMipmappedArray
 - CUDART_MEMORY, [70](#)
- cudaMallocPitch
 - CUDART_MEMORY, [72](#)
- cudaMemcpy
 - CUDART_MEMORY, [72](#)
- cudaMemcpy2D
 - CUDART_MEMORY, [73](#)
- cudaMemcpy2DArrayToArray
 - CUDART_MEMORY, [74](#)
- cudaMemcpy2DAsync
 - CUDART_MEMORY, [74](#)
- cudaMemcpy2DFromArray
 - CUDART_MEMORY, [75](#)
- cudaMemcpy2DFromArrayAsync
 - CUDART_MEMORY, [76](#)
- cudaMemcpy2DToArray
 - CUDART_MEMORY, [77](#)
- cudaMemcpy2DToArrayAsync
 - CUDART_MEMORY, [78](#)
- cudaMemcpy3D
 - CUDART_MEMORY, [78](#)
- cudaMemcpy3DAsync

- CUDART_MEMORY, 80
- cudaMemcpy3DParms, 565
 - dstArray, 565
 - dstPos, 565
 - dstPtr, 565
 - extent, 565
 - kind, 565
 - srcArray, 565
 - srcPos, 565
 - srcPtr, 565
- cudaMemcpy3DPeer
 - CUDART_MEMORY, 81
- cudaMemcpy3DPeerAsync
 - CUDART_MEMORY, 82
- cudaMemcpy3DPeerParms, 567
 - dstArray, 567
 - dstDevice, 567
 - dstPos, 567
 - dstPtr, 567
 - extent, 567
 - srcArray, 567
 - srcDevice, 567
 - srcPos, 567
 - srcPtr, 568
- cudaMemcpyArrayToArray
 - CUDART_MEMORY, 82
- cudaMemcpyAsync
 - CUDART_MEMORY, 83
- cudaMemcpyDefault
 - CUDART_TYPES, 210
- cudaMemcpyDeviceToDevice
 - CUDART_TYPES, 210
- cudaMemcpyDeviceToHost
 - CUDART_TYPES, 210
- cudaMemcpyFromArray
 - CUDART_MEMORY, 84
- cudaMemcpyFromArrayAsync
 - CUDART_MEMORY, 84
- cudaMemcpyFromSymbol
 - CUDART_HIGHLEVEL, 182
 - CUDART_MEMORY, 85
- cudaMemcpyFromSymbolAsync
 - CUDART_HIGHLEVEL, 183
 - CUDART_MEMORY, 86
- cudaMemcpyHostToDevice
 - CUDART_TYPES, 210
- cudaMemcpyHostToHost
 - CUDART_TYPES, 210
- cudaMemcpyKind
 - CUDART_TYPES, 210
- cudaMemcpyPeer
 - CUDART_MEMORY, 86
- cudaMemcpyPeerAsync
 - CUDART_MEMORY, 87
- cudaMemcpyToArray
 - CUDART_MEMORY, 87
- cudaMemcpyToArrayAsync
 - CUDART_MEMORY, 88
- cudaMemcpyToSymbol
 - CUDART_HIGHLEVEL, 184
 - CUDART_MEMORY, 89
- cudaMemcpyToSymbolAsync
 - CUDART_HIGHLEVEL, 184
 - CUDART_MEMORY, 89
- cudaMemGetInfo
 - CUDART_MEMORY, 90
- cudaMemoryType
 - CUDART_TYPES, 210
- cudaMemoryTypeDevice
 - CUDART_TYPES, 211
- cudaMemoryTypeHost
 - CUDART_TYPES, 211
- cudaMemset
 - CUDART_MEMORY, 90
- cudaMemset2D
 - CUDART_MEMORY, 91
- cudaMemset2DAsync
 - CUDART_MEMORY, 91
- cudaMemset3D
 - CUDART_MEMORY, 92
- cudaMemset3DAsync
 - CUDART_MEMORY, 93
- cudaMemsetAsync
 - CUDART_MEMORY, 93
- cudaMipmappedArray_const_t
 - CUDART_TYPES, 202
- cudaMipmappedArray_t
 - CUDART_TYPES, 202
- cudaOutputMode
 - CUDART_TYPES, 211
- cudaOutputMode_t
 - CUDART_TYPES, 202
- cudaPeekAtLastError
 - CUDART_ERROR, 41
- cudaPeerAccessDefault
 - CUDART_TYPES, 201
- cudaPitchedPtr, 569
 - pitch, 569
 - ptr, 569
 - xsize, 569
 - ysize, 569
- cudaPointerAttributes, 570
 - device, 570
 - devicePointer, 570
 - hostPointer, 570
 - memoryType, 570
- cudaPointerGetAttributes
 - CUDART_UNIFIED, 97

cudaPos, 571
 x, 571
 y, 571
 z, 571
cudaProfilerInitialize
 CUDA_RT_PROFILER, 189
cudaProfilerStart
 CUDA_RT_PROFILER, 190
cudaProfilerStop
 CUDA_RT_PROFILER, 190
cudaReadModeElementType
 CUDA_RT_TYPES, 213
cudaReadModeNormalizedFloat
 CUDA_RT_TYPES, 213
cudaResourceDesc, 572
 array, 572
 desc, 572
 devPtr, 572
 height, 572
 mipmap, 572
 pitchInBytes, 572
 resType, 572
 sizeInBytes, 572
 width, 573
cudaResourceType
 CUDA_RT_TYPES, 211
cudaResourceTypeArray
 CUDA_RT_TYPES, 211
cudaResourceTypeLinear
 CUDA_RT_TYPES, 211
cudaResourceTypeMipmappedArray
 CUDA_RT_TYPES, 211
cudaResourceTypePitch2D
 CUDA_RT_TYPES, 211
cudaResourceViewDesc, 574
 depth, 574
 firstLayer, 574
 firstMipmapLevel, 574
 format, 574
 height, 574
 lastLayer, 574
 lastMipmapLevel, 574
 width, 574
cudaResourceViewFormat
 CUDA_RT_TYPES, 211
cudaResViewFormatFloat1
 CUDA_RT_TYPES, 212
cudaResViewFormatFloat2
 CUDA_RT_TYPES, 212
cudaResViewFormatFloat4
 CUDA_RT_TYPES, 212
cudaResViewFormatHalf1
 CUDA_RT_TYPES, 212
cudaResViewFormatHalf2
 CUDA_RT_TYPES, 212
 CUDA_RT_TYPES, 212
cudaResViewFormatHalf4
 CUDA_RT_TYPES, 212
cudaResViewFormatNone
 CUDA_RT_TYPES, 211
cudaResViewFormatSignedBlockCompressed4
 CUDA_RT_TYPES, 212
cudaResViewFormatSignedBlockCompressed5
 CUDA_RT_TYPES, 212
cudaResViewFormatSignedBlockCompressed6H
 CUDA_RT_TYPES, 212
cudaResViewFormatSignedChar1
 CUDA_RT_TYPES, 211
cudaResViewFormatSignedChar2
 CUDA_RT_TYPES, 211
cudaResViewFormatSignedChar4
 CUDA_RT_TYPES, 211
cudaResViewFormatSignedInt1
 CUDA_RT_TYPES, 212
cudaResViewFormatSignedInt2
 CUDA_RT_TYPES, 212
cudaResViewFormatSignedInt4
 CUDA_RT_TYPES, 212
cudaResViewFormatSignedShort1
 CUDA_RT_TYPES, 211
cudaResViewFormatSignedShort2
 CUDA_RT_TYPES, 211
cudaResViewFormatSignedShort4
 CUDA_RT_TYPES, 211
cudaResViewFormatUnsignedBlockCompressed1
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedBlockCompressed2
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedBlockCompressed3
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedBlockCompressed4
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedBlockCompressed5
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedBlockCompressed6H
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedBlockCompressed7
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedChar1
 CUDA_RT_TYPES, 211
cudaResViewFormatUnsignedChar2
 CUDA_RT_TYPES, 211
cudaResViewFormatUnsignedChar4
 CUDA_RT_TYPES, 211
cudaResViewFormatUnsignedInt1
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedInt2
 CUDA_RT_TYPES, 212
cudaResViewFormatUnsignedInt4
 CUDA_RT_TYPES, 212

- CUDART_TYPES, 212
- cudaResViewFormatUnsignedShort1
 - CUDART_TYPES, 211
- cudaResViewFormatUnsignedShort2
 - CUDART_TYPES, 211
- cudaResViewFormatUnsignedShort4
 - CUDART_TYPES, 211
- CUDART_D3D10
 - cudaD3D10DeviceListAll, 125
 - cudaD3D10DeviceListCurrentFrame, 125
 - cudaD3D10DeviceListNextFrame, 125
- CUDART_D3D10_DEPRECATED
 - cudaD3D10MapFlagsNone, 130
 - cudaD3D10MapFlagsReadOnly, 130
 - cudaD3D10MapFlagsWriteDiscard, 130
 - cudaD3D10RegisterFlagsArray, 130
 - cudaD3D10RegisterFlagsNone, 130
- CUDART_D3D11
 - cudaD3D11DeviceListAll, 139
 - cudaD3D11DeviceListCurrentFrame, 139
 - cudaD3D11DeviceListNextFrame, 139
- CUDART_D3D9
 - cudaD3D9DeviceListAll, 111
 - cudaD3D9DeviceListCurrentFrame, 111
 - cudaD3D9DeviceListNextFrame, 111
- CUDART_D3D9_DEPRECATED
 - cudaD3D9MapFlagsNone, 117
 - cudaD3D9MapFlagsReadOnly, 117
 - cudaD3D9MapFlagsWriteDiscard, 117
 - cudaD3D9RegisterFlagsArray, 117
 - cudaD3D9RegisterFlagsNone, 117
- CUDART_OPENGL
 - cudaGLDeviceListAll, 101
 - cudaGLDeviceListCurrentFrame, 101
 - cudaGLDeviceListNextFrame, 101
- CUDART_OPENGL_DEPRECATED
 - cudaGLMapFlagsNone, 105
 - cudaGLMapFlagsReadOnly, 105
 - cudaGLMapFlagsWriteDiscard, 106
- CUDART_TYPES
 - cudaAddressModeBorder, 213
 - cudaAddressModeClamp, 213
 - cudaAddressModeMirror, 213
 - cudaAddressModeWrap, 213
 - cudaBoundaryModeClamp, 212
 - cudaBoundaryModeTrap, 212
 - cudaBoundaryModeZero, 212
 - cudaChannelFormatKindFloat, 202
 - cudaChannelFormatKindNone, 202
 - cudaChannelFormatKindSigned, 202
 - cudaChannelFormatKindUnsigned, 202
 - cudaComputeModeDefault, 203
 - cudaComputeModeExclusive, 203
 - cudaComputeModeExclusiveProcess, 203
 - cudaComputeModeProhibited, 203
 - cudaCSV, 211
 - cudaDevAttrAsyncEngineCount, 204
 - cudaDevAttrCanMapHostMemory, 203
 - cudaDevAttrClockRate, 203
 - cudaDevAttrComputeCapabilityMajor, 205
 - cudaDevAttrComputeCapabilityMinor, 205
 - cudaDevAttrComputeMode, 203
 - cudaDevAttrConcurrentKernels, 204
 - cudaDevAttrEccEnabled, 204
 - cudaDevAttrGlobalMemoryBusWidth, 204
 - cudaDevAttrGpuOverlap, 203
 - cudaDevAttrIntegrated, 203
 - cudaDevAttrKernelExecTimeout, 203
 - cudaDevAttrL2CacheSize, 204
 - cudaDevAttrMaxBlockDimX, 203
 - cudaDevAttrMaxBlockDimY, 203
 - cudaDevAttrMaxBlockDimZ, 203
 - cudaDevAttrMaxGridDimX, 203
 - cudaDevAttrMaxGridDimY, 203
 - cudaDevAttrMaxGridDimZ, 203
 - cudaDevAttrMaxPitch, 203
 - cudaDevAttrMaxRegistersPerBlock, 203
 - cudaDevAttrMaxSharedMemoryPerBlock, 203
 - cudaDevAttrMaxSurface1DLayeredLayers, 204
 - cudaDevAttrMaxSurface1DLayeredWidth, 204
 - cudaDevAttrMaxSurface1DWidth, 204
 - cudaDevAttrMaxSurface2DHeight, 204
 - cudaDevAttrMaxSurface2DLayeredHeight, 204
 - cudaDevAttrMaxSurface2DLayeredLayers, 204
 - cudaDevAttrMaxSurface2DLayeredWidth, 204
 - cudaDevAttrMaxSurface2DWidth, 204
 - cudaDevAttrMaxSurface3DDepth, 204
 - cudaDevAttrMaxSurface3DHeight, 204
 - cudaDevAttrMaxSurface3DWidth, 204
 - cudaDevAttrMaxSurfaceCubemapLayeredLayers, 205
 - cudaDevAttrMaxSurfaceCubemapLayeredWidth, 204
 - cudaDevAttrMaxSurfaceCubemapWidth, 204
 - cudaDevAttrMaxTexture1DLayeredLayers, 204
 - cudaDevAttrMaxTexture1DLayeredWidth, 204
 - cudaDevAttrMaxTexture1DLinearWidth, 205
 - cudaDevAttrMaxTexture1DMipmappedWidth, 205
 - cudaDevAttrMaxTexture1DWidth, 203
 - cudaDevAttrMaxTexture2DGatherHeight, 204
 - cudaDevAttrMaxTexture2DGatherWidth, 204
 - cudaDevAttrMaxTexture2DHeight, 203
 - cudaDevAttrMaxTexture2DLayeredHeight, 204
 - cudaDevAttrMaxTexture2DLayeredLayers, 204
 - cudaDevAttrMaxTexture2DLayeredWidth, 204
 - cudaDevAttrMaxTexture2DLinearHeight, 205
 - cudaDevAttrMaxTexture2DLinearPitch, 205
 - cudaDevAttrMaxTexture2DLinearWidth, 205

- cudaDevAttrMaxTexture2DMipmappedHeight, 205
- cudaDevAttrMaxTexture2DMipmappedWidth, 205
- cudaDevAttrMaxTexture2DWidth, 203
- cudaDevAttrMaxTexture3DDepth, 204
- cudaDevAttrMaxTexture3DDepthAlt, 204
- cudaDevAttrMaxTexture3DHeight, 203
- cudaDevAttrMaxTexture3DHeightAlt, 204
- cudaDevAttrMaxTexture3DWidth, 203
- cudaDevAttrMaxTexture3DWidthAlt, 204
- cudaDevAttrMaxTextureCubemapLayeredLayers, 204
- cudaDevAttrMaxTextureCubemapLayeredWidth, 204
- cudaDevAttrMaxTextureCubemapWidth, 204
- cudaDevAttrMaxThreadsPerBlock, 203
- cudaDevAttrMaxThreadsPerMultiProcessor, 204
- cudaDevAttrMemoryClockRate, 204
- cudaDevAttrMultiProcessorCount, 203
- cudaDevAttrPciBusId, 204
- cudaDevAttrPciDeviceId, 204
- cudaDevAttrPciDomainId, 204
- cudaDevAttrSurfaceAlignment, 204
- cudaDevAttrTccDriver, 204
- cudaDevAttrTextureAlignment, 203
- cudaDevAttrTexturePitchAlignment, 204
- cudaDevAttrTotalConstantMemory, 203
- cudaDevAttrUnifiedAddressing, 204
- cudaDevAttrWarpSize, 203
- cudaErrorAddressOfConstant, 206
- cudaErrorApiFailureBase, 209
- cudaErrorAssert, 208
- cudaErrorCudartUnloading, 207
- cudaErrorDeviceAlreadyInUse, 208
- cudaErrorDevicesUnavailable, 208
- cudaErrorDuplicateSurfaceName, 207
- cudaErrorDuplicateTextureName, 207
- cudaErrorDuplicateVariableName, 207
- cudaErrorECCUncorrectable, 207
- cudaErrorHostMemoryAlreadyRegistered, 208
- cudaErrorHostMemoryNotRegistered, 208
- cudaErrorIncompatibleDriverContext, 208
- cudaErrorInitializationError, 205
- cudaErrorInsufficientDriver, 207
- cudaErrorInvalidChannelDescriptor, 206
- cudaErrorInvalidConfiguration, 205
- cudaErrorInvalidDevice, 206
- cudaErrorInvalidDeviceFunction, 205
- cudaErrorInvalidDevicePointer, 206
- cudaErrorInvalidFilterSetting, 207
- cudaErrorInvalidHostPointer, 206
- cudaErrorInvalidKernelImage, 208
- cudaErrorInvalidMemcpyDirection, 206
- cudaErrorInvalidNormSetting, 207
- cudaErrorInvalidPitchValue, 206
- cudaErrorInvalidResourceHandle, 207
- cudaErrorInvalidSurface, 207
- cudaErrorInvalidSymbol, 206
- cudaErrorInvalidTexture, 206
- cudaErrorInvalidTextureBinding, 206
- cudaErrorInvalidValue, 206
- cudaErrorLaunchFailure, 205
- cudaErrorLaunchFileScopedSurf, 209
- cudaErrorLaunchFileScopedTex, 208
- cudaErrorLaunchMaxDepthExceeded, 208
- cudaErrorLaunchOutOfResources, 205
- cudaErrorLaunchPendingCountExceeded, 209
- cudaErrorLaunchTimeout, 205
- cudaErrorMapBufferObjectFailed, 206
- cudaErrorMemoryAllocation, 205
- cudaErrorMemoryValueTooLarge, 207
- cudaErrorMissingConfiguration, 205
- cudaErrorMixedDeviceExecution, 207
- cudaErrorNoDevice, 207
- cudaErrorNoKernelImageForDevice, 208
- cudaErrorNotPermitted, 209
- cudaErrorNotReady, 207
- cudaErrorNotSupported, 209
- cudaErrorNotYetImplemented, 207
- cudaErrorOperatingSystem, 208
- cudaErrorPeerAccessAlreadyEnabled, 208
- cudaErrorPeerAccessNotEnabled, 208
- cudaErrorPeerAccessUnsupported, 208
- cudaErrorPriorLaunchFailure, 205
- cudaErrorProfilerAlreadyStarted, 208
- cudaErrorProfilerAlreadyStopped, 208
- cudaErrorProfilerDisabled, 208
- cudaErrorProfilerNotInitialized, 208
- cudaErrorSetOnActiveProcess, 207
- cudaErrorSharedObjectInitFailed, 207
- cudaErrorSharedObjectSymbolNotFound, 207
- cudaErrorStartupFailure, 209
- cudaErrorSyncDepthExceeded, 209
- cudaErrorSynchronizationError, 206
- cudaErrorTextureFetchFailed, 206
- cudaErrorTextureNotBound, 206
- cudaErrorTooManyPeers, 208
- cudaErrorUnknown, 207
- cudaErrorUnmapBufferObjectFailed, 206
- cudaErrorUnsupportedLimit, 207
- cudaFilterModeLinear, 213
- cudaFilterModePoint, 213
- cudaFormatModeAuto, 212
- cudaFormatModeForced, 212
- cudaFuncCachePreferEqual, 209
- cudaFuncCachePreferL1, 209
- cudaFuncCachePreferNone, 209
- cudaFuncCachePreferShared, 209
- cudaGraphicsCubeFaceNegativeX, 209

- cudaGraphicsCubeFaceNegativeY, 209
- cudaGraphicsCubeFaceNegativeZ, 210
- cudaGraphicsCubeFacePositiveX, 209
- cudaGraphicsCubeFacePositiveY, 209
- cudaGraphicsCubeFacePositiveZ, 210
- cudaGraphicsMapFlagsNone, 210
- cudaGraphicsMapFlagsReadOnly, 210
- cudaGraphicsMapFlagsWriteDiscard, 210
- cudaGraphicsRegisterFlagsNone, 210
- cudaGraphicsRegisterFlagsReadOnly, 210
- cudaGraphicsRegisterFlagsSurfaceLoadStore, 210
- cudaGraphicsRegisterFlagsTextureGather, 210
- cudaGraphicsRegisterFlagsWriteDiscard, 210
- cudaKeyValuePair, 211
- cudaLimitDevRuntimePendingLaunchCount, 210
- cudaLimitDevRuntimeSyncDepth, 210
- cudaLimitMallocHeapSize, 210
- cudaLimitPrintfFifoSize, 210
- cudaLimitStackSize, 210
- cudaMemcpyDefault, 210
- cudaMemcpyDeviceToDevice, 210
- cudaMemcpyDeviceToHost, 210
- cudaMemcpyHostToDevice, 210
- cudaMemcpyHostToHost, 210
- cudaMemoryTypeDevice, 211
- cudaMemoryTypeHost, 211
- cudaReadModeElementType, 213
- cudaReadModeNormalizedFloat, 213
- cudaResourceTypeArray, 211
- cudaResourceTypeLinear, 211
- cudaResourceTypeMipmappedArray, 211
- cudaResourceTypePitch2D, 211
- cudaResViewFormatFloat1, 212
- cudaResViewFormatFloat2, 212
- cudaResViewFormatFloat4, 212
- cudaResViewFormatHalf1, 212
- cudaResViewFormatHalf2, 212
- cudaResViewFormatHalf4, 212
- cudaResViewFormatNone, 211
- cudaResViewFormatSignedBlockCompressed4, 212
- cudaResViewFormatSignedBlockCompressed5, 212
- cudaResViewFormatSignedBlockCompressed6H, 212
- cudaResViewFormatSignedChar1, 211
- cudaResViewFormatSignedChar2, 211
- cudaResViewFormatSignedChar4, 211
- cudaResViewFormatSignedInt1, 212
- cudaResViewFormatSignedInt2, 212
- cudaResViewFormatSignedInt4, 212
- cudaResViewFormatSignedShort1, 211
- cudaResViewFormatSignedShort2, 211
- cudaResViewFormatSignedShort4, 211
- cudaResViewFormatUnsignedBlockCompressed1, 212
- cudaResViewFormatUnsignedBlockCompressed2, 212
- cudaResViewFormatUnsignedBlockCompressed3, 212
- cudaResViewFormatUnsignedBlockCompressed4, 212
- cudaResViewFormatUnsignedBlockCompressed5, 212
- cudaResViewFormatUnsignedBlockCompressed6H, 212
- cudaResViewFormatUnsignedBlockCompressed7, 212
- cudaResViewFormatUnsignedChar1, 211
- cudaResViewFormatUnsignedChar2, 211
- cudaResViewFormatUnsignedChar4, 211
- cudaResViewFormatUnsignedInt1, 212
- cudaResViewFormatUnsignedInt2, 212
- cudaResViewFormatUnsignedInt4, 212
- cudaResViewFormatUnsignedShort1, 211
- cudaResViewFormatUnsignedShort2, 211
- cudaResViewFormatUnsignedShort4, 211
- cudaSuccess, 205
- CUDART__VERSION
 - cudaDriverGetVersion, 169
 - cudaRuntimeGetVersion, 169
- CUDART_D3D10
 - cudaD3D10DeviceList, 125
 - cudaD3D10GetDevice, 125
 - cudaD3D10GetDevices, 126
 - cudaGraphicsD3D10RegisterResource, 126
- CUDART_D3D10_DEPRECATED
 - cudaD3D10GetDirect3DDevice, 130
 - cudaD3D10MapFlags, 130
 - cudaD3D10MapResources, 131
 - cudaD3D10RegisterFlags, 130
 - cudaD3D10RegisterResource, 131
 - cudaD3D10ResourceGetMappedArray, 132
 - cudaD3D10ResourceGetMappedPitch, 133
 - cudaD3D10ResourceGetMappedPointer, 134
 - cudaD3D10ResourceGetMappedSize, 134
 - cudaD3D10ResourceGetSurfaceDimensions, 135
 - cudaD3D10ResourceSetMapFlags, 136
 - cudaD3D10SetDirect3DDevice, 136
 - cudaD3D10UnmapResources, 137
 - cudaD3D10UnregisterResource, 137
- CUDART_D3D11
 - cudaD3D11DeviceList, 139
 - cudaD3D11GetDevice, 139
 - cudaD3D11GetDevices, 140
 - cudaGraphicsD3D11RegisterResource, 140
- CUDART_D3D11_DEPRECATED
 - cudaD3D11GetDirect3DDevice, 143
 - cudaD3D11SetDirect3DDevice, 143
- CUDART_D3D9

- cudaD3D9DeviceList, 111
- cudaD3D9GetDevice, 112
- cudaD3D9GetDevices, 112
- cudaD3D9GetDirect3DDevice, 112
- cudaD3D9SetDirect3DDevice, 113
- cudaGraphicsD3D9RegisterResource, 113
- CUDART_D3D9_DEPRECATED
 - cudaD3D9MapFlags, 117
 - cudaD3D9MapResources, 117
 - cudaD3D9RegisterFlags, 117
 - cudaD3D9RegisterResource, 118
 - cudaD3D9ResourceGetMappedArray, 119
 - cudaD3D9ResourceGetMappedPitch, 119
 - cudaD3D9ResourceGetMappedPointer, 120
 - cudaD3D9ResourceGetMappedSize, 121
 - cudaD3D9ResourceGetSurfaceDimensions, 122
 - cudaD3D9ResourceSetMapFlags, 122
 - cudaD3D9UnmapResources, 123
 - cudaD3D9UnregisterResource, 124
- CUDART_DEVICE
 - cudaChooseDevice, 16
 - cudaDeviceGetAttribute, 17
 - cudaDeviceGetByPCIBusId, 20
 - cudaDeviceGetCacheConfig, 20
 - cudaDeviceGetLimit, 21
 - cudaDeviceGetPCIBusId, 21
 - cudaDeviceGetSharedMemConfig, 22
 - cudaDeviceReset, 22
 - cudaDeviceSetCacheConfig, 23
 - cudaDeviceSetLimit, 23
 - cudaDeviceSetSharedMemConfig, 24
 - cudaDeviceSynchronize, 25
 - cudaGetDevice, 25
 - cudaGetDeviceCount, 26
 - cudaGetDeviceProperties, 26
 - cudaIpcCloseMemHandle, 29
 - cudaIpcGetEventHandle, 30
 - cudaIpcGetMemHandle, 30
 - cudaIpcOpenEventHandle, 31
 - cudaIpcOpenMemHandle, 31
 - cudaSetDevice, 32
 - cudaSetDeviceFlags, 32
 - cudaSetValidDevices, 33
- CUDART_ERROR
 - cudaGetErrorString, 40
 - cudaGetLastError, 40
 - cudaPeekAtLastError, 41
- CUDART_EVENT
 - cudaEventCreate, 47
 - cudaEventCreateWithFlags, 47
 - cudaEventDestroy, 48
 - cudaEventElapsedTime, 48
 - cudaEventQuery, 49
 - cudaEventRecord, 49
 - cudaEventSynchronize, 50
- CUDART_EXECUTION
 - cudaConfigureCall, 51
 - cudaFuncGetAttributes, 52
 - cudaFuncSetCacheConfig, 52
 - cudaFuncSetSharedMemConfig, 53
 - cudaLaunch, 54
 - cudaSetDoubleForDevice, 54
 - cudaSetDoubleForHost, 55
 - cudaSetupArgument, 55
- CUDART_HIGHLEVEL
 - cudaBindSurfaceToArray, 172
 - cudaBindTexture, 173
 - cudaBindTexture2D, 174, 175
 - cudaBindTextureToArray, 175, 176
 - cudaBindTextureToMipmappedArray, 176, 177
 - cudaCreateChannelDesc, 177
 - cudaEventCreate, 178
 - cudaFuncGetAttributes, 178
 - cudaFuncSetCacheConfig, 179
 - cudaGetSymbolAddress, 180
 - cudaGetSymbolSize, 180
 - cudaGetTextureAlignmentOffset, 180
 - cudaLaunch, 181
 - cudaMallocHost, 181
 - cudaMemcpyFromSymbol, 182
 - cudaMemcpyFromSymbolAsync, 183
 - cudaMemcpyToSymbol, 184
 - cudaMemcpyToSymbolAsync, 184
 - cudaSetupArgument, 185
 - cudaUnbindTexture, 186
- CUDART_INTEROP
 - cudaGraphicsMapResources, 148
 - cudaGraphicsResourceGetMappedMipmappedArray, 149
 - cudaGraphicsResourceGetMappedPointer, 149
 - cudaGraphicsResourceSetMapFlags, 150
 - cudaGraphicsSubResourceGetMappedArray, 150
 - cudaGraphicsUnmapResources, 151
 - cudaGraphicsUnregisterResource, 152
- CUDART_MEMORY
 - cudaArrayGetInfo, 60
 - cudaFree, 60
 - cudaFreeArray, 60
 - cudaFreeHost, 61
 - cudaFreeMipmappedArray, 61
 - cudaGetMipmappedArrayLevel, 61
 - cudaGetSymbolAddress, 62
 - cudaGetSymbolSize, 62
 - cudaHostAlloc, 63
 - cudaHostGetDevicePointer, 64
 - cudaHostGetFlags, 64
 - cudaHostRegister, 65
 - cudaHostUnregister, 66

- cudaMalloc, 66
- cudaMalloc3D, 66
- cudaMalloc3DArray, 67
- cudaMallocArray, 69
- cudaMallocHost, 70
- cudaMallocMipmappedArray, 70
- cudaMallocPitch, 72
- cudaMemcpy, 72
- cudaMemcpy2D, 73
- cudaMemcpy2DArrayToArray, 74
- cudaMemcpy2DAsync, 74
- cudaMemcpy2DFromArray, 75
- cudaMemcpy2DFromArrayAsync, 76
- cudaMemcpy2DToArray, 77
- cudaMemcpy2DToArrayAsync, 78
- cudaMemcpy3D, 78
- cudaMemcpy3DAsync, 80
- cudaMemcpy3DPeer, 81
- cudaMemcpy3DPeerAsync, 82
- cudaMemcpyArrayToArray, 82
- cudaMemcpyAsync, 83
- cudaMemcpyFromArray, 84
- cudaMemcpyFromArrayAsync, 84
- cudaMemcpyFromSymbol, 85
- cudaMemcpyFromSymbolAsync, 86
- cudaMemcpyPeer, 86
- cudaMemcpyPeerAsync, 87
- cudaMemcpyToArray, 87
- cudaMemcpyToArrayAsync, 88
- cudaMemcpyToSymbol, 89
- cudaMemcpyToSymbolAsync, 89
- cudaMemGetInfo, 90
- cudaMemset, 90
- cudaMemset2D, 91
- cudaMemset2DAsync, 91
- cudaMemset3D, 92
- cudaMemset3DAsync, 93
- cudaMemsetAsync, 93
- make_cudaExtent, 94
- make_cudaPitchedPtr, 94
- make_cudaPos, 95
- CUDART_OPENGL
 - cudaGLDeviceList, 101
 - cudaGLGetDevices, 102
 - cudaGraphicsGLRegisterBuffer, 102
 - cudaGraphicsGLRegisterImage, 103
 - cudaWGLGetDevice, 104
- CUDART_OPENGL_DEPRECATED
 - cudaGLMapBufferObject, 106
 - cudaGLMapBufferObjectAsync, 106
 - cudaGLMapFlags, 105
 - cudaGLRegisterBufferObject, 107
 - cudaGLSetBufferObjectMapFlags, 107
 - cudaGLSetGLDevice, 108
 - cudaGLUnmapBufferObject, 108
 - cudaGLUnmapBufferObjectAsync, 109
 - cudaGLUnregisterBufferObject, 109
- CUDART_PEER
 - cudaDeviceCanAccessPeer, 99
 - cudaDeviceDisablePeerAccess, 99
 - cudaDeviceEnablePeerAccess, 100
- CUDART_PROFILER
 - cudaProfilerInitialize, 189
 - cudaProfilerStart, 190
 - cudaProfilerStop, 190
- CUDART_STREAM
 - cudaStreamAddCallback, 43
 - cudaStreamCallback_t, 42
 - cudaStreamCreate, 43
 - cudaStreamCreateWithFlags, 44
 - cudaStreamDestroy, 44
 - cudaStreamQuery, 44
 - cudaStreamSynchronize, 45
 - cudaStreamWaitEvent, 45
- CUDART_SURFACE
 - cudaBindSurfaceToArray, 159
 - cudaGetSurfaceReference, 159
- CUDART_SURFACE_OBJECT
 - cudaCreateSurfaceObject, 167
 - cudaDestroySurfaceObject, 167
 - cudaGetSurfaceObjectResourceDesc, 168
- CUDART_TEXTURE
 - cudaBindTexture, 153
 - cudaBindTexture2D, 154
 - cudaBindTextureToArray, 155
 - cudaBindTextureToMipmappedArray, 155
 - cudaCreateChannelDesc, 156
 - cudaGetChannelDesc, 156
 - cudaGetTextureAlignmentOffset, 157
 - cudaGetTextureReference, 157
 - cudaUnbindTexture, 158
- CUDART_TEXTURE_OBJECT
 - cudaCreateTextureObject, 161
 - cudaDestroyTextureObject, 165
 - cudaGetTextureObjectResourceDesc, 165
 - cudaGetTextureObjectResourceViewDesc, 165
 - cudaGetTextureObjectTextureDesc, 166
- CUDART_THREAD_DEPRECATED
 - cudaThreadExit, 35
 - cudaThreadGetCacheConfig, 36
 - cudaThreadGetLimit, 36
 - cudaThreadSetCacheConfig, 37
 - cudaThreadSetLimit, 38
 - cudaThreadSynchronize, 38
- CUDART_TYPES
 - CUDA_IPC_HANDLE_SIZE, 198
 - cudaArray_const_t, 201
 - cudaArray_t, 201

- cudaArrayCubemap, 198
- cudaArrayDefault, 198
- cudaArrayLayered, 198
- cudaArraySurfaceLoadStore, 199
- cudaArrayTextureGather, 199
- cudaChannelFormatKind, 202
- cudaComputeMode, 202
- cudaDeviceAttr, 203
- cudaDeviceBlockingSync, 199
- cudaDeviceLmemResizeToMax, 199
- cudaDeviceMapHost, 199
- cudaDeviceMask, 199
- cudaDevicePropDontCare, 199
- cudaDeviceScheduleAuto, 199
- cudaDeviceScheduleBlockingSync, 199
- cudaDeviceScheduleMask, 199
- cudaDeviceScheduleSpin, 200
- cudaDeviceScheduleYield, 200
- cudaError, 205
- cudaError_t, 201
- cudaEvent_t, 201
- cudaEventBlockingSync, 200
- cudaEventDefault, 200
- cudaEventDisableTiming, 200
- cudaEventInterprocess, 200
- cudaFuncCache, 209
- cudaGraphicsCubeFace, 209
- cudaGraphicsMapFlags, 210
- cudaGraphicsRegisterFlags, 210
- cudaGraphicsResource_t, 201
- cudaHostAllocDefault, 200
- cudaHostAllocMapped, 200
- cudaHostAllocPortable, 200
- cudaHostAllocWriteCombined, 200
- cudaHostRegisterDefault, 200
- cudaHostRegisterMapped, 200
- cudaHostRegisterPortable, 201
- cudaIpcEventHandle_t, 201
- cudaIpcMemHandle_t, 202
- cudaIpcMemLazyEnablePeerAccess, 201
- cudaLimit, 210
- cudaMemcpyKind, 210
- cudaMemoryType, 210
- cudaMipmappedArray_const_t, 202
- cudaMipmappedArray_t, 202
- cudaOutputMode, 211
- cudaOutputMode_t, 202
- cudaPeerAccessDefault, 201
- cudaResourceType, 211
- cudaResourceViewFormat, 211
- cudaSharedMemConfig, 212
- cudaStream_t, 202
- cudaStreamDefault, 201
- cudaStreamNonBlocking, 201
- cudaSurfaceBoundaryMode, 212
- cudaSurfaceFormatMode, 212
- cudaSurfaceObject_t, 202
- cudaTextureAddressMode, 212
- cudaTextureFilterMode, 213
- cudaTextureObject_t, 202
- cudaTextureReadMode, 213
- cudaUUID_t, 202
- CUDART_UNIFIED
 - cudaPointerGetAttributes, 97
- CUDART_VDPAU
 - cudaGraphicsVDPAURegisterOutputSurface, 145
 - cudaGraphicsVDPAURegisterVideoSurface, 146
 - cudaVDPAUGetDevice, 146
 - cudaVDPAUSetVDPAUDevice, 147
- cudaRuntimeGetVersion
 - CUDART__VERSION, 169
- cudaSetDevice
 - CUDART_DEVICE, 32
- cudaSetDeviceFlags
 - CUDART_DEVICE, 32
- cudaSetDoubleForDevice
 - CUDART_EXECUTION, 54
- cudaSetDoubleForHost
 - CUDART_EXECUTION, 55
- cudaSetupArgument
 - CUDART_EXECUTION, 55
 - CUDART_HIGHLEVEL, 185
- cudaSetValidDevices
 - CUDART_DEVICE, 33
- cudaSharedMemConfig
 - CUDART_TYPES, 212
- cudaStream_t
 - CUDART_TYPES, 202
- cudaStreamAddCallback
 - CUDART_STREAM, 43
- cudaStreamCallback_t
 - CUDART_STREAM, 42
- cudaStreamCreate
 - CUDART_STREAM, 43
- cudaStreamCreateWithFlags
 - CUDART_STREAM, 44
- cudaStreamDefault
 - CUDART_TYPES, 201
- cudaStreamDestroy
 - CUDART_STREAM, 44
- cudaStreamNonBlocking
 - CUDART_TYPES, 201
- cudaStreamQuery
 - CUDART_STREAM, 44
- cudaStreamSynchronize
 - CUDART_STREAM, 45
- cudaStreamWaitEvent
 - CUDART_STREAM, 45

- cudaSuccess
 - CUDART_TYPES, 205
- cudaSurfaceBoundaryMode
 - CUDART_TYPES, 212
- cudaSurfaceFormatMode
 - CUDART_TYPES, 212
- cudaSurfaceObject_t
 - CUDART_TYPES, 202
- cudaTextureAddressMode
 - CUDART_TYPES, 212
- cudaTextureDesc, 576
 - addressMode, 576
 - filterMode, 576
 - maxAnisotropy, 576
 - maxMipmapLevelClamp, 576
 - minMipmapLevelClamp, 576
 - mipmapFilterMode, 576
 - mipmapLevelBias, 576
 - normalizedCoords, 576
 - readMode, 577
 - sRGB, 577
- cudaTextureFilterMode
 - CUDART_TYPES, 213
- cudaTextureObject_t
 - CUDART_TYPES, 202
- cudaTextureReadMode
 - CUDART_TYPES, 213
- cudaThreadExit
 - CUDART_THREAD_DEPRECATED, 35
- cudaThreadGetCacheConfig
 - CUDART_THREAD_DEPRECATED, 36
- cudaThreadGetLimit
 - CUDART_THREAD_DEPRECATED, 36
- cudaThreadSetCacheConfig
 - CUDART_THREAD_DEPRECATED, 37
- cudaThreadSetLimit
 - CUDART_THREAD_DEPRECATED, 38
- cudaThreadSynchronize
 - CUDART_THREAD_DEPRECATED, 38
- cudaUnbindTexture
 - CUDART_HIGHLEVEL, 186
 - CUDART_TEXTURE, 158
- cudaUUID_t
 - CUDART_TYPES, 202
- cudaVDPAUGetDevice
 - CUDART_VDPAU, 146
- cudaVDPAUSetVDPAUDevice
 - CUDART_VDPAU, 147
- cudaWGLGetDevice
 - CUDART_OPENGL, 104
- CUdevice
 - CUDA_TYPES, 226
- CUdevice_attribute
 - CUDA_TYPES, 226
- CUdevice_attribute_enum
 - CUDA_TYPES, 233
- cuDeviceCanAccessPeer
 - CUDA_PEER_ACCESS, 379
- cuDeviceComputeCapability
 - CUDA_DEVICE_DEPRECATED, 250
- cuDeviceGet
 - CUDA_DEVICE, 244
- cuDeviceGetAttribute
 - CUDA_DEVICE, 244
- cuDeviceGetByPCIBusId
 - CUDA_MEM, 282
- cuDeviceGetCount
 - CUDA_DEVICE, 248
- cuDeviceGetName
 - CUDA_DEVICE, 248
- cuDeviceGetPCIBusId
 - CUDA_MEM, 282
- cuDeviceGetProperties
 - CUDA_DEVICE_DEPRECATED, 250
- CUdeviceptr
 - CUDA_TYPES, 226
- cuDeviceTotalMem
 - CUDA_DEVICE, 249
- CUdevprop
 - CUDA_TYPES, 226
- CUdevprop_st, 578
 - clockRate, 578
 - maxGridSize, 578
 - maxThreadsDim, 578
 - maxThreadsPerBlock, 578
 - memPitch, 578
 - regsPerBlock, 578
 - sharedMemPerBlock, 578
 - SIMDWidth, 578
 - textureAlign, 579
 - totalConstantMemory, 579
- cuDriverGetVersion
 - CUDA_VERSION, 243
- CUevent
 - CUDA_TYPES, 226
- CUevent_flags
 - CUDA_TYPES, 226
- CUevent_flags_enum
 - CUDA_TYPES, 236
- cuEventCreate
 - CUDA_EVENT, 337
- cuEventDestroy
 - CUDA_EVENT, 338
- cuEventElapsedTime
 - CUDA_EVENT, 338
- cuEventQuery
 - CUDA_EVENT, 339
- cuEventRecord

- CUDA_EVENT, 339
- cuEventSynchronize
 - CUDA_EVENT, 340
- CUfilter_mode
 - CUDA_TYPES, 226
- CUfilter_mode_enum
 - CUDA_TYPES, 236
- CUfunc_cache
 - CUDA_TYPES, 226
- CUfunc_cache_enum
 - CUDA_TYPES, 236
- cuFuncGetAttribute
 - CUDA_EXEC, 341
- cuFuncSetBlockShape
 - CUDA_EXEC_DEPRECATED, 346
- cuFuncSetCacheConfig
 - CUDA_EXEC, 342
- cuFuncSetSharedMemConfig
 - CUDA_EXEC, 343
- cuFuncSetSharedSize
 - CUDA_EXEC_DEPRECATED, 347
- CUfunction
 - CUDA_TYPES, 226
- CUfunction_attribute
 - CUDA_TYPES, 226
- CUfunction_attribute_enum
 - CUDA_TYPES, 237
- cuGLCtxCreate
 - CUDA_GL_DEPRECATED, 393
- CUGLDeviceList
 - CUDA_GL, 387
- CUGLDeviceList_enum
 - CUDA_GL, 388
- cuGLGetDevices
 - CUDA_GL, 388
- cuGLInit
 - CUDA_GL_DEPRECATED, 393
- CUGLmap_flags
 - CUDA_GL_DEPRECATED, 392
- CUGLmap_flags_enum
 - CUDA_GL_DEPRECATED, 393
- cuGLMapBufferObject
 - CUDA_GL_DEPRECATED, 394
- cuGLMapBufferObjectAsync
 - CUDA_GL_DEPRECATED, 394
- cuGLRegisterBufferObject
 - CUDA_GL_DEPRECATED, 395
- cuGLSetBufferObjectMapFlags
 - CUDA_GL_DEPRECATED, 395
- cuGLUnmapBufferObject
 - CUDA_GL_DEPRECATED, 396
- cuGLUnmapBufferObjectAsync
 - CUDA_GL_DEPRECATED, 397
- cuGLUnregisterBufferObject
 - CUDA_GL_DEPRECATED, 397
- CUDA_GL_DEPRECATED, 397
- cuGraphicsD3D10RegisterResource
 - CUDA_D3D10, 416
- cuGraphicsD3D11RegisterResource
 - CUDA_D3D11, 431
- cuGraphicsD3D9RegisterResource
 - CUDA_D3D9, 403
- cuGraphicsGLRegisterBuffer
 - CUDA_GL, 388
- cuGraphicsGLRegisterImage
 - CUDA_GL, 389
- CUgraphicsMapResourceFlags
 - CUDA_TYPES, 227
- CUgraphicsMapResourceFlags_enum
 - CUDA_TYPES, 237
- cuGraphicsMapResources
 - CUDA_GRAPHICS, 380
- CUgraphicsRegisterFlags
 - CUDA_TYPES, 227
- CUgraphicsRegisterFlags_enum
 - CUDA_TYPES, 237
- CUgraphicsResource
 - CUDA_TYPES, 227
- cuGraphicsResourceGetMappedMipmappedArray
 - CUDA_GRAPHICS, 381
- cuGraphicsResourceGetMappedPointer
 - CUDA_GRAPHICS, 381
- cuGraphicsResourceSetMapFlags
 - CUDA_GRAPHICS, 382
- cuGraphicsSubResourceGetMappedArray
 - CUDA_GRAPHICS, 382
- cuGraphicsUnmapResources
 - CUDA_GRAPHICS, 383
- cuGraphicsUnregisterResource
 - CUDA_GRAPHICS, 384
- cuGraphicsVDPAURegisterOutputSurface
 - CUDA_VDPAU, 435
- cuGraphicsVDPAURegisterVideoSurface
 - CUDA_VDPAU, 436
- cuInit
 - CUDA_INITIALIZE, 242
- cuIpcCloseMemHandle
 - CUDA_MEM, 283
- CUipcEventHandle
 - CUDA_TYPES, 227
- CUipcEventHandle_st, 580
- cuIpcGetEventHandle
 - CUDA_MEM, 283
- cuIpcGetMemHandle
 - CUDA_MEM, 284
- CUipcMem_flags
 - CUDA_TYPES, 227
- CUipcMem_flags_enum
 - CUDA_TYPES, 237

- CUipcMemHandle
 - CUDA_TYPES, 227
- CUipcMemHandle_st, 581
- cuIpcOpenEventHandle
 - CUDA_MEM, 284
- cuIpcOpenMemHandle
 - CUDA_MEM, 285
- CUjit_fallback
 - CUDA_TYPES, 227
- CUjit_fallback_enum
 - CUDA_TYPES, 237
- CUjit_option
 - CUDA_TYPES, 227
- CUjit_option_enum
 - CUDA_TYPES, 238
- CUjit_target
 - CUDA_TYPES, 227
- CUjit_target_enum
 - CUDA_TYPES, 239
- cuLaunch
 - CUDA_EXEC_DEPRECATED, 347
- cuLaunchGrid
 - CUDA_EXEC_DEPRECATED, 348
- cuLaunchGridAsync
 - CUDA_EXEC_DEPRECATED, 348
- cuLaunchKernel
 - CUDA_EXEC, 343
- CUlimit
 - CUDA_TYPES, 227
- CUlimit_enum
 - CUDA_TYPES, 239
- cuMemAlloc
 - CUDA_MEM, 285
- cuMemAllocHost
 - CUDA_MEM, 286
- cuMemAllocPitch
 - CUDA_MEM, 287
- cuMemcpy
 - CUDA_MEM, 287
- cuMemcpy2D
 - CUDA_MEM, 288
- cuMemcpy2DAsync
 - CUDA_MEM, 290
- cuMemcpy2DUnaligned
 - CUDA_MEM, 293
- cuMemcpy3D
 - CUDA_MEM, 295
- cuMemcpy3DAsync
 - CUDA_MEM, 298
- cuMemcpy3DPeer
 - CUDA_MEM, 300
- cuMemcpy3DPeerAsync
 - CUDA_MEM, 301
- cuMemcpyAsync
 - CUDA_MEM, 301
- cuMemcpyAtoA
 - CUDA_MEM, 302
- cuMemcpyAtoD
 - CUDA_MEM, 303
- cuMemcpyAtoH
 - CUDA_MEM, 303
- cuMemcpyAtoHAsync
 - CUDA_MEM, 304
- cuMemcpyDtoA
 - CUDA_MEM, 305
- cuMemcpyDtoD
 - CUDA_MEM, 305
- cuMemcpyDtoDAsync
 - CUDA_MEM, 306
- cuMemcpyDtoH
 - CUDA_MEM, 306
- cuMemcpyDtoHAsync
 - CUDA_MEM, 307
- cuMemcpyHtoA
 - CUDA_MEM, 308
- cuMemcpyHtoAAsync
 - CUDA_MEM, 308
- cuMemcpyHtoD
 - CUDA_MEM, 309
- cuMemcpyHtoDAsync
 - CUDA_MEM, 309
- cuMemcpyPeer
 - CUDA_MEM, 310
- cuMemcpyPeerAsync
 - CUDA_MEM, 311
- cuMemFree
 - CUDA_MEM, 311
- cuMemFreeHost
 - CUDA_MEM, 312
- cuMemGetAddressRange
 - CUDA_MEM, 312
- cuMemGetInfo
 - CUDA_MEM, 313
- cuMemHostAlloc
 - CUDA_MEM, 313
- cuMemHostGetDevicePointer
 - CUDA_MEM, 315
- cuMemHostGetFlags
 - CUDA_MEM, 315
- cuMemHostRegister
 - CUDA_MEM, 316
- cuMemHostUnregister
 - CUDA_MEM, 317
- CUmemorytype
 - CUDA_TYPES, 227
- CUmemorytype_enum
 - CUDA_TYPES, 239
- cuMemsetD16

- CUDA_MEM, 317
- cuMemsetD16Async
 - CUDA_MEM, 318
- cuMemsetD2D16
 - CUDA_MEM, 318
- cuMemsetD2D16Async
 - CUDA_MEM, 319
- cuMemsetD2D32
 - CUDA_MEM, 320
- cuMemsetD2D32Async
 - CUDA_MEM, 321
- cuMemsetD2D8
 - CUDA_MEM, 321
- cuMemsetD2D8Async
 - CUDA_MEM, 322
- cuMemsetD32
 - CUDA_MEM, 323
- cuMemsetD32Async
 - CUDA_MEM, 323
- cuMemsetD8
 - CUDA_MEM, 324
- cuMemsetD8Async
 - CUDA_MEM, 324
- CUmipmappedArray
 - CUDA_TYPES, 227
- cuMipmappedArrayCreate
 - CUDA_MEM, 325
- cuMipmappedArrayDestroy
 - CUDA_MEM, 327
- cuMipmappedArrayGetLevel
 - CUDA_MEM, 328
- CUmodule
 - CUDA_TYPES, 228
- cuModuleGetFunction
 - CUDA_MODULE, 265
- cuModuleGetGlobal
 - CUDA_MODULE, 266
- cuModuleGetSurfRef
 - CUDA_MODULE, 266
- cuModuleGetTexRef
 - CUDA_MODULE, 267
- cuModuleLoad
 - CUDA_MODULE, 267
- cuModuleLoadData
 - CUDA_MODULE, 268
- cuModuleLoadDataEx
 - CUDA_MODULE, 268
- cuModuleLoadFatBinary
 - CUDA_MODULE, 270
- cuModuleUnload
 - CUDA_MODULE, 270
- cuParamSetf
 - CUDA_EXEC_DEPRECATED, 349
- cuParamSeti
 - CUDA_EXEC_DEPRECATED, 350
- cuParamSetSize
 - CUDA_EXEC_DEPRECATED, 350
- cuParamSetTexRef
 - CUDA_EXEC_DEPRECATED, 351
- cuParamSetv
 - CUDA_EXEC_DEPRECATED, 351
- CUpointer_attribute
 - CUDA_TYPES, 228
- CUpointer_attribute_enum
 - CUDA_TYPES, 239
- cuPointerGetAttribute
 - CUDA_UNIFIED, 330
- cuProfilerInitialize
 - CUDA_PROFILER, 385
- cuProfilerStart
 - CUDA_PROFILER, 386
- cuProfilerStop
 - CUDA_PROFILER, 386
- CUresourcetype
 - CUDA_TYPES, 228
- CUresourcetype_enum
 - CUDA_TYPES, 240
- CUresourceViewFormat
 - CUDA_TYPES, 228
- CUresourceViewFormat_enum
 - CUDA_TYPES, 240
- CUresult
 - CUDA_TYPES, 228
- CUsharedconfig
 - CUDA_TYPES, 228
- CUsharedconfig_enum
 - CUDA_TYPES, 241
- CUstream
 - CUDA_TYPES, 228
- CUstream_flags
 - CUDA_TYPES, 228
- CUstream_flags_enum
 - CUDA_TYPES, 241
- cuStreamAddCallback
 - CUDA_STREAM, 333
- CUstreamCallback
 - CUDA_TYPES, 228
- cuStreamCreate
 - CUDA_STREAM, 334
- cuStreamDestroy
 - CUDA_STREAM, 334
- cuStreamQuery
 - CUDA_STREAM, 335
- cuStreamSynchronize
 - CUDA_STREAM, 335
- cuStreamWaitEvent
 - CUDA_STREAM, 335
- CUsurfObject

- CUDA_TYPES, 228
 - cuSurfObjectCreate
 - CUDA_SURFOBJECT, 376
 - cuSurfObjectDestroy
 - CUDA_SURFOBJECT, 376
 - cuSurfObjectGetResourceDesc
 - CUDA_SURFOBJECT, 377
 - CUsurfref
 - CUDA_TYPES, 229
 - cuSurfRefGetArray
 - CUDA_SURFREF, 368
 - cuSurfRefSetArray
 - CUDA_SURFREF, 368
 - CUtexObject
 - CUDA_TYPES, 229
 - cuTexObjectCreate
 - CUDA_TEXOBJECT, 370
 - cuTexObjectDestroy
 - CUDA_TEXOBJECT, 374
 - cuTexObjectGetResourceDesc
 - CUDA_TEXOBJECT, 374
 - cuTexObjectGetResourceViewDesc
 - CUDA_TEXOBJECT, 374
 - cuTexObjectGetTextureDesc
 - CUDA_TEXOBJECT, 375
 - CUtexref
 - CUDA_TYPES, 229
 - cuTexRefCreate
 - CUDA_TEXREF_DEPRECATED, 366
 - cuTexRefDestroy
 - CUDA_TEXREF_DEPRECATED, 366
 - cuTexRefGetAddress
 - CUDA_TEXREF, 354
 - cuTexRefGetAddressMode
 - CUDA_TEXREF, 355
 - cuTexRefGetArray
 - CUDA_TEXREF, 355
 - cuTexRefGetFilterMode
 - CUDA_TEXREF, 355
 - cuTexRefGetFlags
 - CUDA_TEXREF, 356
 - cuTexRefGetFormat
 - CUDA_TEXREF, 356
 - cuTexRefGetMaxAnisotropy
 - CUDA_TEXREF, 357
 - cuTexRefGetMipmapFilterMode
 - CUDA_TEXREF, 357
 - cuTexRefGetMipmapLevelBias
 - CUDA_TEXREF, 357
 - cuTexRefGetMipmapLevelClamp
 - CUDA_TEXREF, 358
 - cuTexRefGetMipmappedArray
 - CUDA_TEXREF, 358
 - cuTexRefSetAddress
 - CUDA_TEXREF, 359
 - cuTexRefSetAddress2D
 - CUDA_TEXREF, 359
 - cuTexRefSetAddressMode
 - CUDA_TEXREF, 360
 - cuTexRefSetArray
 - CUDA_TEXREF, 361
 - cuTexRefSetFilterMode
 - CUDA_TEXREF, 361
 - cuTexRefSetFlags
 - CUDA_TEXREF, 362
 - cuTexRefSetFormat
 - CUDA_TEXREF, 362
 - cuTexRefSetMaxAnisotropy
 - CUDA_TEXREF, 363
 - cuTexRefSetMipmapFilterMode
 - CUDA_TEXREF, 363
 - cuTexRefSetMipmapLevelBias
 - CUDA_TEXREF, 364
 - cuTexRefSetMipmapLevelClamp
 - CUDA_TEXREF, 364
 - cuTexRefSetMipmappedArray
 - CUDA_TEXREF, 364
 - cuVDPAUCtxCreate
 - CUDA_VDPAU, 437
 - cuVDPAUGetDevice
 - CUDA_VDPAU, 437
 - cuWGLGetDevice
 - CUDA_GL, 390
- Data types used by CUDA driver, 214
- Data types used by CUDA Runtime, 191
- Depth
 - CUDA_ARRAY3D_DESCRIPTOR_st, 535
 - CUDA_MEMCPY3D_PEER_st, 540
 - CUDA_MEMCPY3D_st, 543
- depth
 - CUDA_RESOURCE_VIEW_DESC_st, 549
 - cudaExtent, 560
 - cudaResourceViewDesc, 574
- desc
 - cudaResourceDesc, 572
- device
 - cudaPointerAttributes, 570
- Device Management, 15, 244
- deviceOverlap
 - cudaDeviceProp, 555
- devicePointer
 - cudaPointerAttributes, 570
- devPtr
 - CUDA_RESOURCE_DESC_st, 547
 - cudaResourceDesc, 572
- Direct3D 10 Interoperability, 125, 414
- Direct3D 11 Interoperability, 139, 429

- Direct3D 9 Interoperability, [111](#), [399](#)
- Double Precision Ininsics, [505](#)
- Double Precision Mathematical Functions, [467](#)
- dstArray
 - [CUDA_MEMCPY2D_st](#), [538](#)
 - [CUDA_MEMCPY3D_PEER_st](#), [540](#)
 - [CUDA_MEMCPY3D_st](#), [543](#)
 - [cudaMemcpy3DParms](#), [565](#)
 - [cudaMemcpy3DPeerParms](#), [567](#)
- dstContext
 - [CUDA_MEMCPY3D_PEER_st](#), [540](#)
- dstDevice
 - [CUDA_MEMCPY2D_st](#), [538](#)
 - [CUDA_MEMCPY3D_PEER_st](#), [540](#)
 - [CUDA_MEMCPY3D_st](#), [543](#)
 - [cudaMemcpy3DPeerParms](#), [567](#)
- dstHeight
 - [CUDA_MEMCPY3D_PEER_st](#), [541](#)
 - [CUDA_MEMCPY3D_st](#), [543](#)
- dstHost
 - [CUDA_MEMCPY2D_st](#), [538](#)
 - [CUDA_MEMCPY3D_PEER_st](#), [541](#)
 - [CUDA_MEMCPY3D_st](#), [544](#)
- dstLOD
 - [CUDA_MEMCPY3D_PEER_st](#), [541](#)
 - [CUDA_MEMCPY3D_st](#), [544](#)
- dstMemoryType
 - [CUDA_MEMCPY2D_st](#), [538](#)
 - [CUDA_MEMCPY3D_PEER_st](#), [541](#)
 - [CUDA_MEMCPY3D_st](#), [544](#)
- dstPitch
 - [CUDA_MEMCPY2D_st](#), [538](#)
 - [CUDA_MEMCPY3D_PEER_st](#), [541](#)
 - [CUDA_MEMCPY3D_st](#), [544](#)
- dstPos
 - [cudaMemcpy3DParms](#), [565](#)
 - [cudaMemcpy3DPeerParms](#), [567](#)
- dstPtr
 - [cudaMemcpy3DParms](#), [565](#)
 - [cudaMemcpy3DPeerParms](#), [567](#)
- dstXInBytes
 - [CUDA_MEMCPY2D_st](#), [538](#)
 - [CUDA_MEMCPY3D_PEER_st](#), [541](#)
 - [CUDA_MEMCPY3D_st](#), [544](#)
- dstY
 - [CUDA_MEMCPY2D_st](#), [539](#)
 - [CUDA_MEMCPY3D_PEER_st](#), [541](#)
 - [CUDA_MEMCPY3D_st](#), [544](#)
- dstZ
 - [CUDA_MEMCPY3D_PEER_st](#), [541](#)
 - [CUDA_MEMCPY3D_st](#), [544](#)
- ECCEnabled
 - [cudaDeviceProp](#), [555](#)
- erf
 - [CUDA_MATH_DOUBLE](#), [475](#)
- erfc
 - [CUDA_MATH_DOUBLE](#), [475](#)
- erfcf
 - [CUDA_MATH_SINGLE](#), [448](#)
- erfcinv
 - [CUDA_MATH_DOUBLE](#), [475](#)
- erfcinvf
 - [CUDA_MATH_SINGLE](#), [448](#)
- erfcx
 - [CUDA_MATH_DOUBLE](#), [475](#)
- erfcxf
 - [CUDA_MATH_SINGLE](#), [448](#)
- erff
 - [CUDA_MATH_SINGLE](#), [448](#)
- erfinv
 - [CUDA_MATH_DOUBLE](#), [476](#)
- erfinvf
 - [CUDA_MATH_SINGLE](#), [449](#)
- Error Handling, [40](#)
- Event Management, [47](#), [337](#)
- Execution Control, [51](#), [341](#)
- exp
 - [CUDA_MATH_DOUBLE](#), [476](#)
- exp10
 - [CUDA_MATH_DOUBLE](#), [476](#)
- exp10f
 - [CUDA_MATH_SINGLE](#), [449](#)
- exp2
 - [CUDA_MATH_DOUBLE](#), [476](#)
- exp2f
 - [CUDA_MATH_SINGLE](#), [449](#)
- expf
 - [CUDA_MATH_SINGLE](#), [449](#)
- expm1
 - [CUDA_MATH_DOUBLE](#), [477](#)
- expm1f
 - [CUDA_MATH_SINGLE](#), [450](#)
- extent
 - [cudaMemcpy3DParms](#), [565](#)
 - [cudaMemcpy3DPeerParms](#), [567](#)
- f
 - [cudaChannelFormatDesc](#), [553](#)
- fabs
 - [CUDA_MATH_DOUBLE](#), [477](#)
- fabsf
 - [CUDA_MATH_SINGLE](#), [450](#)
- fdim
 - [CUDA_MATH_DOUBLE](#), [477](#)
- fdimf
 - [CUDA_MATH_SINGLE](#), [450](#)
- fdividef

- CUDA_MATH_SINGLE, 450
- filterMode
 - CUDA_TEXTURE_DESC_st, 551
 - cudaTextureDesc, 576
 - textureReference, 583
- firstLayer
 - CUDA_RESOURCE_VIEW_DESC_st, 549
 - cudaResourceViewDesc, 574
- firstMipmapLevel
 - CUDA_RESOURCE_VIEW_DESC_st, 549
 - cudaResourceViewDesc, 574
- Flags
 - CUDA_ARRAY3D_DESCRIPTOR_st, 535
- flags
 - CUDA_RESOURCE_DESC_st, 547
 - CUDA_TEXTURE_DESC_st, 551
- floor
 - CUDA_MATH_DOUBLE, 477
- floorf
 - CUDA_MATH_SINGLE, 451
- fma
 - CUDA_MATH_DOUBLE, 478
- fmaf
 - CUDA_MATH_SINGLE, 451
- fmax
 - CUDA_MATH_DOUBLE, 478
- fmaxf
 - CUDA_MATH_SINGLE, 451
- fmin
 - CUDA_MATH_DOUBLE, 478
- fminf
 - CUDA_MATH_SINGLE, 452
- fmod
 - CUDA_MATH_DOUBLE, 479
- fmodf
 - CUDA_MATH_SINGLE, 452
- Format
 - CUDA_ARRAY3D_DESCRIPTOR_st, 535
 - CUDA_ARRAY_DESCRIPTOR_st, 537
- format
 - CUDA_RESOURCE_DESC_st, 547
 - CUDA_RESOURCE_VIEW_DESC_st, 549
 - cudaResourceViewDesc, 574
- frexp
 - CUDA_MATH_DOUBLE, 479
- frexpf
 - CUDA_MATH_SINGLE, 452
- Graphics Interoperability, 148, 380
- hArray
 - CUDA_RESOURCE_DESC_st, 547
- Height
 - CUDA_ARRAY3D_DESCRIPTOR_st, 535
- CUDA_ARRAY_DESCRIPTOR_st, 537
- CUDA_MEMCPY2D_st, 539
- CUDA_MEMCPY3D_PEER_st, 541
- CUDA_MEMCPY3D_st, 544
- height
 - CUDA_RESOURCE_DESC_st, 547
 - CUDA_RESOURCE_VIEW_DESC_st, 549
 - cudaExtent, 560
 - cudaResourceDesc, 572
 - cudaResourceViewDesc, 574
- hMipmappedArray
 - CUDA_RESOURCE_DESC_st, 547
- hostPointer
 - cudaPointerAttributes, 570
- hypot
 - CUDA_MATH_DOUBLE, 479
- hypotf
 - CUDA_MATH_SINGLE, 453
- ilogb
 - CUDA_MATH_DOUBLE, 480
- ilogbf
 - CUDA_MATH_SINGLE, 453
- Initialization, 242
- Integer Intrinsic, 513
- integrated
 - cudaDeviceProp, 555
- Interactions with the CUDA Driver API, 187
- isfinite
 - CUDA_MATH_DOUBLE, 480
 - CUDA_MATH_SINGLE, 453
- isinf
 - CUDA_MATH_DOUBLE, 480
 - CUDA_MATH_SINGLE, 454
- isnan
 - CUDA_MATH_DOUBLE, 480
 - CUDA_MATH_SINGLE, 454
- j0
 - CUDA_MATH_DOUBLE, 480
- j0f
 - CUDA_MATH_SINGLE, 454
- j1
 - CUDA_MATH_DOUBLE, 481
- j1f
 - CUDA_MATH_SINGLE, 454
- jn
 - CUDA_MATH_DOUBLE, 481
- jnf
 - CUDA_MATH_SINGLE, 454
- kernelExecTimeoutEnabled
 - cudaDeviceProp, 555
- kind
 - cudaMemcpy3DParms, 565

- l2CacheSize
 - cudaDeviceProp, 556
- lastLayer
 - CUDA_RESOURCE_VIEW_DESC_st, 549
 - cudaResourceViewDesc, 574
- lastMipmapLevel
 - CUDA_RESOURCE_VIEW_DESC_st, 549
 - cudaResourceViewDesc, 574
- ldexp
 - CUDA_MATH_DOUBLE, 481
- ldexpf
 - CUDA_MATH_SINGLE, 455
- lgamma
 - CUDA_MATH_DOUBLE, 482
- lgammaf
 - CUDA_MATH_SINGLE, 455
- llrint
 - CUDA_MATH_DOUBLE, 482
- llrintf
 - CUDA_MATH_SINGLE, 455
- llround
 - CUDA_MATH_DOUBLE, 482
- llroundf
 - CUDA_MATH_SINGLE, 455
- localSizeBytes
 - cudaFuncAttributes, 561
- log
 - CUDA_MATH_DOUBLE, 482
- log10
 - CUDA_MATH_DOUBLE, 483
- log10f
 - CUDA_MATH_SINGLE, 456
- log1p
 - CUDA_MATH_DOUBLE, 483
- log1pf
 - CUDA_MATH_SINGLE, 456
- log2
 - CUDA_MATH_DOUBLE, 483
- log2f
 - CUDA_MATH_SINGLE, 456
- logb
 - CUDA_MATH_DOUBLE, 483
- logbf
 - CUDA_MATH_SINGLE, 457
- logf
 - CUDA_MATH_SINGLE, 457
- lrint
 - CUDA_MATH_DOUBLE, 484
- lrintf
 - CUDA_MATH_SINGLE, 457
- lround
 - CUDA_MATH_DOUBLE, 484
- lroundf
 - CUDA_MATH_SINGLE, 457
- major
 - cudaDeviceProp, 556
- make_cudaExtent
 - CUDART_MEMORY, 94
- make_cudaPitchedPtr
 - CUDART_MEMORY, 94
- make_cudaPos
 - CUDART_MEMORY, 95
- Mathematical Functions, 439
- maxAnisotropy
 - CUDA_TEXTURE_DESC_st, 551
 - cudaTextureDesc, 576
 - textureReference, 583
- maxGridSize
 - cudaDeviceProp, 556
 - CUdevprop_st, 578
- maxMipmapLevelClamp
 - CUDA_TEXTURE_DESC_st, 551
 - cudaTextureDesc, 576
 - textureReference, 583
- maxSurface1D
 - cudaDeviceProp, 556
- maxSurface1DLayered
 - cudaDeviceProp, 556
- maxSurface2D
 - cudaDeviceProp, 556
- maxSurface2DLayered
 - cudaDeviceProp, 556
- maxSurface3D
 - cudaDeviceProp, 556
- maxSurfaceCubemap
 - cudaDeviceProp, 556
- maxSurfaceCubemapLayered
 - cudaDeviceProp, 556
- maxTexture1D
 - cudaDeviceProp, 556
- maxTexture1DLayered
 - cudaDeviceProp, 556
- maxTexture1DLinear
 - cudaDeviceProp, 557
- maxTexture1DMipmap
 - cudaDeviceProp, 557
- maxTexture2D
 - cudaDeviceProp, 557
- maxTexture2DGather
 - cudaDeviceProp, 557
- maxTexture2DLayered
 - cudaDeviceProp, 557
- maxTexture2DLinear
 - cudaDeviceProp, 557
- maxTexture2DMipmap
 - cudaDeviceProp, 557
- maxTexture3D
 - cudaDeviceProp, 557

- maxTextureCubemap
 - cudaDeviceProp, 557
- maxTextureCubemapLayered
 - cudaDeviceProp, 557
- maxThreadsDim
 - cudaDeviceProp, 557
 - CUdevprop_st, 578
- maxThreadsPerBlock
 - cudaDeviceProp, 557
 - cudaFuncAttributes, 561
 - CUdevprop_st, 578
- maxThreadsPerMultiProcessor
 - cudaDeviceProp, 558
- Memory Management, 56, 272
- memoryBusWidth
 - cudaDeviceProp, 558
- memoryClockRate
 - cudaDeviceProp, 558
- memoryType
 - cudaPointerAttributes, 570
- memPitch
 - cudaDeviceProp, 558
 - CUdevprop_st, 578
- minMipmapLevelClamp
 - CUDA_TEXTURE_DESC_st, 551
 - cudaTextureDesc, 576
 - textureReference, 583
- minor
 - cudaDeviceProp, 558
- mipmap
 - cudaResourceDesc, 572
- mipmapFilterMode
 - CUDA_TEXTURE_DESC_st, 551
 - cudaTextureDesc, 576
 - textureReference, 583
- mipmapLevelBias
 - CUDA_TEXTURE_DESC_st, 551
 - cudaTextureDesc, 576
 - textureReference, 583
- modf
 - CUDA_MATH_DOUBLE, 484
- modff
 - CUDA_MATH_SINGLE, 458
- Module Management, 265
- multiProcessorCount
 - cudaDeviceProp, 558
- name
 - cudaDeviceProp, 558
- nan
 - CUDA_MATH_DOUBLE, 484
- nanf
 - CUDA_MATH_SINGLE, 458
- nearbyint
 - CUDA_MATH_DOUBLE, 485
- nearbyintf
 - CUDA_MATH_SINGLE, 458
- nextafter
 - CUDA_MATH_DOUBLE, 485
- nextafterf
 - CUDA_MATH_SINGLE, 458
- normalized
 - textureReference, 584
- normalizedCoords
 - cudaTextureDesc, 576
- normcdf
 - CUDA_MATH_DOUBLE, 485
- normcdfi
 - CUDA_MATH_SINGLE, 459
- normcdfinv
 - CUDA_MATH_DOUBLE, 485
- normcdfinvi
 - CUDA_MATH_SINGLE, 459
- NumChannels
 - CUDA_ARRAY3D_DESCRIPTOR_st, 535
 - CUDA_ARRAY_DESCRIPTOR_st, 537
- numChannels
 - CUDA_RESOURCE_DESC_st, 547
- numRegs
 - cudaFuncAttributes, 561
- OpenGL Interoperability, 101, 387
- pciBusID
 - cudaDeviceProp, 558
- pciDeviceID
 - cudaDeviceProp, 558
- pciDomainID
 - cudaDeviceProp, 558
- Peer Context Memory Access, 378
- Peer Device Memory Access, 99
- pitch
 - cudaPitchedPtr, 569
- pitchInBytes
 - CUDA_RESOURCE_DESC_st, 548
 - cudaResourceDesc, 572
- pow
 - CUDA_MATH_DOUBLE, 486
- powf
 - CUDA_MATH_SINGLE, 459
- Profiler Control, 189, 385
- ptr
 - cudaPitchedPtr, 569
- ptxVersion
 - cudaFuncAttributes, 561
- rcbrt
 - CUDA_MATH_DOUBLE, 486
- rcbrtf

- CUDA_MATH_SINGLE, 460
- readMode
 - cudaTextureDesc, 577
- regsPerBlock
 - cudaDeviceProp, 558
 - CUdevprop_st, 578
- remainder
 - CUDA_MATH_DOUBLE, 487
- remainderf
 - CUDA_MATH_SINGLE, 460
- remquo
 - CUDA_MATH_DOUBLE, 487
- remquof
 - CUDA_MATH_SINGLE, 460
- reserved0
 - CUDA_MEMCPY3D_st, 544
- reserved1
 - CUDA_MEMCPY3D_st, 544
- resType
 - CUDA_RESOURCE_DESC_st, 548
 - cudaResourceDesc, 572
- rint
 - CUDA_MATH_DOUBLE, 487
- rintf
 - CUDA_MATH_SINGLE, 461
- round
 - CUDA_MATH_DOUBLE, 487
- roundf
 - CUDA_MATH_SINGLE, 461
- rsqrt
 - CUDA_MATH_DOUBLE, 488
- rsqrtf
 - CUDA_MATH_SINGLE, 461
- scalbln
 - CUDA_MATH_DOUBLE, 488
- scalblnf
 - CUDA_MATH_SINGLE, 461
- scalbn
 - CUDA_MATH_DOUBLE, 488
- scalbnf
 - CUDA_MATH_SINGLE, 462
- sharedMemPerBlock
 - cudaDeviceProp, 558
 - CUdevprop_st, 578
- sharedSizeBytes
 - cudaFuncAttributes, 561
- signbit
 - CUDA_MATH_DOUBLE, 488
 - CUDA_MATH_SINGLE, 462
- SIMDWidth
 - CUdevprop_st, 578
- sin
 - CUDA_MATH_DOUBLE, 489
- sincos
 - CUDA_MATH_DOUBLE, 489
- sincosf
 - CUDA_MATH_SINGLE, 462
- sincospi
 - CUDA_MATH_DOUBLE, 489
- sincospif
 - CUDA_MATH_SINGLE, 462
- sinf
 - CUDA_MATH_SINGLE, 463
- Single Precision Ininsics, 493
- Single Precision Mathematical Functions, 440
- sinh
 - CUDA_MATH_DOUBLE, 489
- sinhf
 - CUDA_MATH_SINGLE, 463
- sinpi
 - CUDA_MATH_DOUBLE, 490
- sinpif
 - CUDA_MATH_SINGLE, 463
- sizeInBytes
 - CUDA_RESOURCE_DESC_st, 548
 - cudaResourceDesc, 572
- sqrt
 - CUDA_MATH_DOUBLE, 490
- sqrtf
 - CUDA_MATH_SINGLE, 463
- srcArray
 - CUDA_MEMCPY2D_st, 539
 - CUDA_MEMCPY3D_PEER_st, 541
 - CUDA_MEMCPY3D_st, 544
 - cudaMemcpy3DParms, 565
 - cudaMemcpy3DPeerParms, 567
- srcContext
 - CUDA_MEMCPY3D_PEER_st, 541
- srcDevice
 - CUDA_MEMCPY2D_st, 539
 - CUDA_MEMCPY3D_PEER_st, 541
 - CUDA_MEMCPY3D_st, 544
 - cudaMemcpy3DPeerParms, 567
- srcHeight
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545
- srcHost
 - CUDA_MEMCPY2D_st, 539
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545
- srcLOD
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545
- srcMemoryType
 - CUDA_MEMCPY2D_st, 539
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545

- srcPitch
 - CUDA_MEMCPY2D_st, 539
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545
- srcPos
 - cudaMemcpy3DParms, 565
 - cudaMemcpy3DPeerParms, 567
- srcPtr
 - cudaMemcpy3DParms, 565
 - cudaMemcpy3DPeerParms, 568
- srcXInBytes
 - CUDA_MEMCPY2D_st, 539
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545
- srcY
 - CUDA_MEMCPY2D_st, 539
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545
- srcZ
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545
- sRGB
 - cudaTextureDesc, 577
 - textureReference, 584
- Stream Management, 42, 333
- Surface Object Management, 167, 376
- Surface Reference Management, 159, 368
- surfaceAlignment
 - cudaDeviceProp, 559
- surfaceReference, 582
 - channelDesc, 582
- tan
 - CUDA_MATH_DOUBLE, 490
- tanf
 - CUDA_MATH_SINGLE, 464
- tanh
 - CUDA_MATH_DOUBLE, 490
- tanhf
 - CUDA_MATH_SINGLE, 464
- tccDriver
 - cudaDeviceProp, 559
- Texture Object Management, 161, 370
- Texture Reference Management, 153, 353
- textureAlign
 - CUdevprop_st, 579
- textureAlignment
 - cudaDeviceProp, 559
- texturePitchAlignment
 - cudaDeviceProp, 559
- textureReference, 583
 - addressMode, 583
 - channelDesc, 583
 - filterMode, 583
 - maxAnisotropy, 583
 - maxMipmapLevelClamp, 583
 - minMipmapLevelClamp, 583
 - mipmapFilterMode, 583
 - mipmapLevelBias, 583
 - normalized, 584
 - sRGB, 584
- tgamma
 - CUDA_MATH_DOUBLE, 491
- tgammaf
 - CUDA_MATH_SINGLE, 464
- totalConstantMemory
 - CUdevprop_st, 579
- totalConstMem
 - cudaDeviceProp, 559
- totalGlobalMem
 - cudaDeviceProp, 559
- trunc
 - CUDA_MATH_DOUBLE, 491
- truncf
 - CUDA_MATH_SINGLE, 465
- Type Casting Ininsics, 518
- Unified Addressing, 96, 329
- unifiedAddressing
 - cudaDeviceProp, 559
- VDPAA Interoperability, 145, 435
- Version Management, 169, 243
- w
 - cudaChannelFormatDesc, 553
- warpSize
 - cudaDeviceProp, 559
- Width
 - CUDA_ARRAY3D_DESCRIPTOR_st, 536
 - CUDA_ARRAY_DESCRIPTOR_st, 537
- width
 - CUDA_RESOURCE_DESC_st, 548
 - CUDA_RESOURCE_VIEW_DESC_st, 549
 - cudaExtent, 560
 - cudaResourceDesc, 573
 - cudaResourceViewDesc, 574
- WidthInBytes
 - CUDA_MEMCPY2D_st, 539
 - CUDA_MEMCPY3D_PEER_st, 542
 - CUDA_MEMCPY3D_st, 545
- x
 - cudaChannelFormatDesc, 553
 - cudaPos, 571
- xsize
 - cudaPitchedPtr, 569
- y

[cudaChannelFormatDesc](#), 553
 [cudaPos](#), 571

y0
 [CUDA_MATH_DOUBLE](#), 491

y0f
 [CUDA_MATH_SINGLE](#), 465

y1
 [CUDA_MATH_DOUBLE](#), 491

y1f
 [CUDA_MATH_SINGLE](#), 465

yn
 [CUDA_MATH_DOUBLE](#), 492

ynf
 [CUDA_MATH_SINGLE](#), 465

ysize
 [cudaPitchedPtr](#), 569

z
 [cudaChannelFormatDesc](#), 553
 [cudaPos](#), 571

Notice

ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE.

Information furnished is believed to be accurate and reliable. However, NVIDIA Corporation assumes no responsibility for the consequences of use of such information or for any infringement of patents or other rights of third parties that may result from its use. No license is granted by implication of otherwise under any patent rights of NVIDIA Corporation. Specifications mentioned in this publication are subject to change without notice. This publication supersedes and replaces all other information previously supplied. NVIDIA Corporation products are not authorized as critical components in life support devices or systems without express written approval of NVIDIA Corporation.

Trademarks

NVIDIA, the NVIDIA logo, GeForce, Tesla, and Quadro are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

© 2007-2012 NVIDIA Corporation. All rights reserved.