



TABLE OF CONTENTS

Chapter 1. Release Notes	1
1.1. CUDA 8.0	1
1.2. CUDA 7.5	2
1.3. CUDA 7.0	2
1.4. CUDA 6.5	3
1.5. CUDA 6.0	4
1.6. CUDA 5.5	4
1.7. CUDA 5.0	5
1.8. CUDA 4.2	6
1.9. CUDA 4.1	6
Chapter 2. Getting Started	
2.1. Getting CUDA Samples	7
Windows	
Linux	7
Mac OSX	
2.2. Building Samples	
Windows	
Linux	
Mac	
2.3. CUDA Cross-Platform Samples	
TARGET_ARCH	
TARGET_OS	
TARGET_FS	
Copying Libraries	
2.4. Using CUDA Samples to Create Your Own CUDA Projects	
2.4.1. Creating CUDA Projects for Windows	
2.4.2. Creating CUDA Projects for Linux	
2.4.3. Creating CUDA Projects for Mac OS X	
Chapter 3. Samples Reference	
3.1. Simple Reference	
asyncAPI	
cdpSimplePrint - Simple Print (CUDA Dynamic Parallelism)	
cdpSimpleQuicksort - Simple Quicksort (CUDA Dynamic Parallelism)	
clock - Clock	
clock_nvrtc - Clock libNVRTC	
cppIntegration - C++ Integration	
cppOverload	
cudaOpenMP	
fp16ScalarProduct - FP16 Scalar Product	
inlinePTX - Using Inline PTX	17

	inlinePTX_nvrtc - Using Inline PTX with libNVRTC	. 17
	matrixMul - Matrix Multiplication (CUDA Runtime API Version)	. 17
	matrixMul_nvrtc - Matrix Multiplication with libNVRTC	.18
	matrixMulCUBLAS - Matrix Multiplication (CUBLAS)	18
	matrixMulDrv - Matrix Multiplication (CUDA Driver API Version)	19
	simpleAssert	. 19
	simpleAssert_nvrtc - simpleAssert with libNVRTC	20
	simpleAtomicIntrinsics - Simple Atomic Intrinsics	
	simpleAtomicIntrinsics_nvrtc - Simple Atomic Intrinsics with libNVRTC	.20
	simpleCallback - Simple CUDA Callbacks	.21
	simpleCubemapTexture - Simple Cubemap Texture	21
	simpleIPC	.21
	simpleLayeredTexture - Simple Layered Texture	22
	simpleMPI	22
	simpleMultiCopy - Simple Multi Copy and Compute	
	simpleMultiGPU - Simple Multi-GPU	. 23
	simpleOccupancy	
	simpleP2P - Simple Peer-to-Peer Transfers with Multi-GPU	
	simplePitchLinearTexture - Pitch Linear Texture	.24
	simplePrintf	. 25
	simpleSeparateCompilation - Simple Static GPU Device Library	25
	simpleStreams	. 25
	simpleSurfaceWrite - Simple Surface Write	. 26
	simpleTemplates - Simple Templates	
	simpleTemplates_nvrtc - Simple Templates with libNVRTC	26
	simpleTexture - Simple Texture	.27
	simpleTextureDrv - Simple Texture (Driver Version)	.27
	simpleVoteIntrinsics - Simple Vote Intrinsics	27
	simpleVoteIntrinsics_nvrtc - Simple Vote Intrinsics with libNVRTC	.28
	simpleZeroCopy	. 28
	systemWideAtomics - System wide Atomics	28
	template - Template	
	UnifiedMemoryStreams - Unified Memory Streams	
	vectorAdd - Vector Addition	.30
	vectorAdd_nvrtc - Vector Addition with libNVRTC	.30
	vectorAddDrv - Vector Addition Driver API	30
3.	2. Utilities Reference	31
	bandwidthTest - Bandwidth Test	.31
	deviceQuery - Device Query	.31
	deviceQueryDrv - Device Query Driver API	31
	$\verb p2pB and \verb widthLatencyTest - Peer-to-Peer B and \verb widthLatencyTest with Multi-GPUs$	32
	topologyQuery - Topology Query	.32
3	3. Graphics Reference	33

	bindlessTexture - Bindless Texture	33
	Mandelbrot	.33
	marchingCubes - Marching Cubes Isosurfaces	34
	simpleD3D10 - Simple Direct3D10 (Vertex Array)	. 34
	simpleD3D10RenderTarget - Simple Direct3D10 Render Target	35
	simpleD3D10Texture - Simple D3D10 Texture	35
	simpleD3D11Texture - Simple D3D11 Texture	36
	simpleD3D9 - Simple Direct3D9 (Vertex Arrays)	.36
	simpleD3D9Texture - Simple D3D9 Texture	37
	simpleGL - Simple OpenGL	37
	simpleGLES - Simple OpenGLES	. 38
	simpleGLES_EGLOutput - Simple OpenGLES EGLOutput	38
	simpleGLES_screen - Simple OpenGLES on Screen	39
	simpleTexture3D - Simple Texture 3D	39
	SLID3D10Texture - SLI D3D10 Texture	.40
	volumeFiltering - Volumetric Filtering with 3D Textures and Surface Writes	. 40
	volumeRender - Volume Rendering with 3D Textures	41
3.	4. Imaging Reference	42
	bicubicTexture - Bicubic B-spline Interoplation	
	bilateralFilter - Bilateral Filter	
	boxFilter - Box Filter	. 43
	convolutionFFT2D - FFT-Based 2D Convolution	
	convolutionSeparable - CUDA Separable Convolution	. 44
	convolutionTexture - Texture-based Separable Convolution	
	cudaDecodeD3D9 - CUDA Video Decoder D3D9 API	. 44
	cudaDecodeGL - CUDA Video Decoder GL API	.45
	dct8x8 - DCT8x8	. 46
	dwtHaar1D - 1D Discrete Haar Wavelet Decomposition	.46
	dxtc - DirectX Texture Compressor (DXTC)	46
	CUDA_EGLStreams_Interop - EGLStreams CUDA Interop	
	histogram - CUDA Histogram	47
	HSOpticalFlow - Optical Flow	
	imageDenoising - Image denoising	
	postProcessGL - Post-Process in OpenGL	
	recursiveGaussian - Recursive Gaussian Filter	
	simpleCUDA2GL - CUDA and OpenGL Interop of Images	. 49
	SobelFilter - Sobel Filter	
	stereoDisparity - Stereo Disparity Computation (SAD SIMD Intrinsics)	
3.	5. Finance Reference	
	binomialOptions - Binomial Option Pricing	
	binomialOptions_nvrtc - Binomial Option Pricing with libNVRTC	
	BlackScholes - Black-Scholes Option Pricing	
	BlackScholes_nvrtc - Black-Scholes Option Pricing with libNVRTC	

	MonteCarloMultiGPU - Monte Carlo Option Pricing with Multi-GPU support	. 52
	quasirandomGenerator - Niederreiter Quasirandom Sequence Generator	. 53
	quasirandomGenerator_nvrtc - Niederreiter Quasirandom Sequence Generator with	
	libNVRTC	. 53
	SobolQRNG - Sobol Quasirandom Number Generator	. 53
3.	.6. Simulations Reference	. 54
	fluidsD3D9 - Fluids (Direct3D Version)	54
	fluidsGL - Fluids (OpenGL Version)	54
	fluidsGLES - Fluids (OpenGLES Version)	. 55
	nbody - CUDA N-Body Simulation	55
	nbody_opengles - CUDA N-Body Simulation with GLES	. 56
	nbody_screen - CUDA N-Body Simulation on Screen	. 56
	oceanFFT - CUDA FFT Ocean Simulation	57
	particles - Particles	. 57
	smokeParticles - Smoke Particles	58
	VFlockingD3D10	. 58
3.	.7. Advanced Reference	. 59
	alignedTypes - Aligned Types	59
	c++11_cuda - C++11 CUDA	59
	cdpAdvancedQuicksort - Advanced Quicksort (CUDA Dynamic Parallelism)	60
	cdpBezierTessellation - Bezier Line Tessellation (CUDA Dynamic Parallelism)	60
	cdpLUDecomposition - LU Decomposition (CUDA Dynamic Parallelism)	. 61
	cdpQuadtree - Quad Tree (CUDA Dynamic Parallelism)	61
	concurrentKernels - Concurrent Kernels	61
	eigenvalues - Eigenvalues	62
	fastWalshTransform - Fast Walsh Transform	. 62
	FDTD3d - CUDA C 3D FDTD	62
	FunctionPointers - Function Pointers	. 63
	interval - Interval Computing	. 63
	lineOfSight - Line of Sight	63
	matrixMulDynlinkJIT - Matrix Multiplication (CUDA Driver API version with Dynamic Linking	
	Version)	64
	mergeSort - Merge Sort	. 64
	newdelete - NewDelete	. 64
	ptxjit - PTX Just-in-Time compilation	. 65
	radixSortThrust - CUDA Radix Sort (Thrust Library)	. 65
	reduction - CUDA Parallel Reduction	. 65
	scalarProd - Scalar Product	66
	scan - CUDA Parallel Prefix Sum (Scan)	. 66
	segmentationTreeThrust - CUDA Segmentation Tree Thrust Library	
	shfl_scan - CUDA Parallel Prefix Sum with Shuffle Intrinsics (SHFL_Scan)	
	simpleHyperQ	
	sortingNetworks - CUDA Sorting Networks	

	StreamPriorities - Stream Priorities	. 68
	threadFenceReduction	. 68
	threadMigration - CUDA Context Thread Management	.68
	transpose - Matrix Transpose	. 69
3	.8. Cudalibraries Reference	. 69
	batchCUBLAS	. 69
	BiCGStab	. 69
	boxFilterNPP - Box Filter with NPP	. 70
	cannyEdgeDetectorNPP - Canny Edge Detector NPP	. 70
	conjugateGradient - ConjugateGradient	. 71
	conjugateGradientPrecond - Preconditioned Conjugate Gradient	71
	conjugateGradientUM - ConjugateGradientUM	. 71
	cuHook - CUDA Interception Library	. 72
	cuSolverDn_LinearSolver - cuSolverDn Linear Solver	. 72
	cuSolverRf - cuSolverRf Refactorization	. 72
	cuSolverSp_LinearSolver - cuSolverSp Linear Solver	73
	cuSolverSp_LowlevelCholesky - cuSolverSp LowlevelCholesky Solver	
	cuSolverSp_LowlevelQR - cuSolverSp Lowlevel QR Solver	
	FilterBorderControlNPP - Filter Border Control NPP	. 74
	freeImageInteropNPP - FreeImage and NPP Interopability	
	histEqualizationNPP - Histogram Equalization with NPP	. 75
	jpegNPP - JPEG encode/decode and resize with NPP	
	MC_EstimatePiInlineP - Monte Carlo Estimation of Pi (inline PRNG)	
	MC_EstimatePiInlineQ - Monte Carlo Estimation of Pi (inline QRNG)	
	MC_EstimatePiP - Monte Carlo Estimation of Pi (batch PRNG)	
	MC_EstimatePiQ - Monte Carlo Estimation of Pi (batch QRNG)	
	MC_SingleAsianOptionP - Monte Carlo Single Asian Option	77
	MersenneTwisterGP11213	
	nvgraph_Pagerank - NVGRAPH Page Rank	
	nvgraph_SemiRingSpmv - NVGRAPH Semi-Ring SpMV	
	nvgraph_SSSP - NVGRAPH Single Source Shortest Path	
	randomFog - Random Fog	
	simpleCUBLAS - Simple CUBLAS	
	simpleCUBLASXT - Simple CUBLAS XT	
	simpleCUFFT - Simple CUFFT	
	simpleCUFFT_2d_MGPU - SimpleCUFFT_2d_MGPU	
	simpleCUFFT_callback - Simple CUFFT Callbacks	
	simpleCUFFT_MGPU - Simple CUFFT_MGPU	82
	simpleDevLibCUBLAS - simpleDevLibCUBLAS GPU Device API Library Functions (CUDA	
	Dynamic Parallelism)	
	pter 4. Dependencies	
T	hird-Party Dependencies	
	Franks	0.4

Message Passing Interface	84
Only 64-Bit	84
DirectX	85
OpenGL	85
OpenGL ES	85
OpenMP	85
Screen	85
X11	85
EGL	85
EGLOutput	85
CUDA Features	86
CUFFT Callback Routines	86
CUDA Dynamic Paralellism	86
CUBLAS	86
CUDA Interprocess Communication	86
CUFFT	86
CURAND	86
CUSPARSE	86
CUSOLVER	86
NPP	87
NVGRAPH	87
NVRTC	87
NVCUVID	87
Stream Priorities	87
Unified Virtual Memory	87
16-bit Floating Point	87
C++11 CUDA	87
Chapter 5. Key Concepts and Associated Samples	88
Basic Key Concepts	88
Advanced Key Concepts	94
Chapter 6. CUDA API and Associated Samples	99
CUDA Driver API Samples	99
CUDA Runtime API Samples	103
Chapter 7. Frequently Asked Questions	111

LIST OF TABLES

Table 1	Supported Target Arch/OS Combinations	9
Table 2	Basic Key Concepts and Associated Samples	. 88
Table 3	Advanced Key Concepts and Associated Samples	. 94
Table 4	CUDA Driver API and Associated Samples	99
Table 5	CUDA Runtime API and Associated Samples	104

Chapter 1. RELEASE NOTES

This section describes the release notes for the CUDA Samples only. For the release notes for the whole CUDA Toolkit, please see CUDA Toolkit Release Notes.

1.1. CUDA 8.0

- Added 7_CUDALibraries/FilterBorderControlNPP. Demonstrates how any border version of an NPP filtering function can be used in the most common mode (with border control enabled), can be used to duplicate the results of the equivalent non-border version of the NPP function, and can be used to enable and disable border control on various source image edges depending on what portion of the source image is being used as input.
- Added 7_CUDALibraries/cannyEdgeDetectorNPP. Demonstrates the recommended parameters to use with the nppiFilterCannyBorder_8u_C1R Canny Edge Detection image filter function. This function expects a single channel 8-bit grayscale input image. You can generate a grayscale image from a color image by first calling nppiColorToGray() or nppiRGBToGray(). The Canny Edge Detection function combines and improves on the techniques required to produce an edge detection image using multiple steps.
- Added 7_CUDALibraries/cuSolverSp_LowlevelCholesky. Demonstrates Cholesky factorization using cuSolverSP's low level APIs.
- Added 7_CUDALibraries/cuSolverSp_LowlevelQR. Demonstrates QR factorization using cuSolverSP's low level APIs.
- Added 7_CUDALibraries/BiCGStab. Demonstrates Bi-Conjugate Gradient Stabilized (BiCGStab) iterative method for nonsymmetric and symmetric positive definite linear systems using CUSPARSE and CUBLAS
- ► Added 7_CUDALibraries/nvgraph_Pagerank. Demonstrates Page Rank computation using nvGRAPH Library.
- Added 7_CUDALibraries/nvgraph_SemiRingSpMV. Demonstrates Semi-Ring SpMV using nvGRAPH Library.
- ► Added **7_CUDALibraries/nvgraph_SSSP**. Demonstrates Single Source Shortest Path(SSSP) computation using nvGRAPH Library.

- ► Added 7_CUDALibraries/simpleCUBLASXT. Demonstrates simple example to use CUBLAS-XT library.
- ► Added 6_Advanced/c++11_cuda. Demonstrates C++11 feature support in CUDA.
- Added **1_Utilities/topologyQuery**. Demonstrates how to query the topology of a system with multiple GPU.
- Added **O_Simple/fp16ScalarProduct**. Demonstrates scalar product calculation of two vectors of FP16 numbers.
- ► Added **O_Simple/systemWideAtomics**. Demonstrates system wide atomic instructions on migratable memory.
- Removed 0_Simple/template_runtime. Its purpose is served by 0_Simple/template.

1.2. CUDA 7.5

- Added 7_CUDALibraries/cuSolverDn_LinearSolver. Demonstrates how to use the CUSOLVER library for performing dense matrix factorization using cuSolverDN's LU, QR and Cholesky factorization functions.
- Added **7_CUDALibraries/cuSolverRf**. Demonstrates how to use cuSolverRF, a sparse re-factorization package of the CUSOLVER library.
- Added 7_CUDALibraries/cuSolverSp_LinearSolver. Demonstrates how to use cuSolverSP which provides sparse set of routines for sparse matrix factorization.
- ► The 2_Graphics/simpleD3D9, 2_Graphics/simpleD3D9Texture, 3_Imaging/cudaDecodeD3D9, and 5_Simulations/fluidsD3D9 samples have been modified to use the Direct3D 9Ex API instead of the Direct3D 9 API.
- The 7_CUDALibraries/grabcutNPP and 7_CUDALibraries/ imageSegmentationNPP samples have been removed. These samples used the NPP graphcut APIs, which have been deprecated in CUDA 7.5.

1.3. CUDA 7.0

- Removed support for Windows 32-bit builds.
- ► The Makefile x86_64=1 and ARMv7=1 options have been deprecated. Please use TARGET_ARCH to set the targeted build architecture instead.
- ► The Makefile GCC option has been deprecated. Please use HOST_COMPILER to set the host compiler instead.
- The CUDA Samples are no longer shipped as prebuilt binaries on Windows. Please use VS Solution files provided to build respective executable.
- Added **O_Simple/clock_nvrtc**. Demonstrates how to compile clock function kernel at runtime using libNVRTC to measure the performance of kernel accurately.
- Added **O_Simple/inlinePTX_nvrtc**. Demonstrates compilation of CUDA kernel having PTX embedded at runtime using libNVRTC.
- ► Added **O_Simple/matrixMul_nvrtc**. Demonstrates compilation of matrix multiplication CUDA kernel at runtime using libNVRTC.
- Added **O_Simple/simpleAssert_nvrtc**. Demonstrates compilation of CUDA kernel having assert() at runtime using libNVRTC.

- Added **O_Simple/simpleAtomicIntrinsics_nvrtc**. Demonstrates compilation of CUDA kernel performing atomic operations at runtime using libNVRTC.
- ► Added O_Simple/simpleTemplates_nvrtc. Demonstrates compilation of templatized dynamically allocated shared memory arrays CUDA kernel at runtime using libNVRTC.
- Added **O_Simple/simpleVoteIntrinsics_nvrtc**. Demonstrates compilation of CUDA kernel which uses vote intrinsics at runtime using libNVRTC.
- Added **O_Simple/vectorAdd_nvrtc**. Demonstrates compilation of CUDA kernel performing vector addition at runtime using libNVRTC.
- Added **4_Finance/binomialOptions_nvrtc**. Demonstrates runtime compilation using libNVRTC of CUDA kernel which evaluates fair call price for a given set of European options under binomial model.
- ▶ Added **4_Finance/BlackScholes_nvrtc**. Demonstrates runtime compilation using libNVRTC of CUDA kernel which evaluates fair call and put prices for a given set of European options by Black-Scholes formula.
- Added 4_Finance/quasirandomGenerator_nvrtc. Demonstrates runtime compilation using libNVRTC of CUDA kernel which implements Niederreiter Quasirandom Sequence Generator and Inverse Cumulative Normal Distribution functions for the generation of Standard Normal Distributions.

1.4. CUDA 6.5

- Added **7_CUDALibraries/cuHook**. Demonstrates how to build and use an intercept library with CUDA.
- Added 7_CUDALibraries/simpleCUFFT_callback. Demonstrates how to compute a 1D-convolution of a signal with a filter using a user-supplied CUFFT callback routine, rather than a separate kernel call.
- Added **7_CUDALibraries/simpleCUFFT_MGPU**. Demonstrates how to compute a 1D-convolution of a signal with a filter by transforming both into frequency domain, multiplying them together, and transforming the signal back to time domain on Multiple GPUs.
- Added 7_CUDALibraries/simpleCUFFT_2d_MGPU. Demonstrates how to compute a 2D-convolution of a signal with a filter by transforming both into frequency domain, multiplying them together, and transforming the signal back to time domain on Multiple GPUs.
- ► Removed **3_Imaging/cudaEncode**. Support for the CUDA Video Encoder (NVCUVENC) has been removed.
- ► Removed **4_Finance/ExcelCUDA2007**. The topic will be covered in a blog post at Parallel Forall.
- ▶ Removed **4_Finance/ExcelCUDA2010**. The topic will be covered in a blog post at Parallel Forall.
- ► The **4_Finance/binomialOptions** sample is now restricted to running on GPUs with SM architecture 2.0 or greater.
- The **4_Finance/quasirandomGenerator** sample is now restricted to running on GPUs with SM architecture 2.0 or greater.

- ► The 7_CUDALibraries/boxFilterNPP sample now demonstrates how to use the static NPP libraries on Linux and Mac.
- ► The 7_CUDALibraries/conjugateGradient sample now demonstrates how to use the static CUBLAS and CUSPARSE libraries on Linux and Mac.
- ► The 7_CUDALibraries/MersenneTwisterGP11213 sample now demonstrates how to use the static CURAND library on Linux and Mac.

1.5. CUDA 6.0

- New featured samples that support a new CUDA 6.0 feature called UVM-Lite
- ► Added O_Simple/UnifiedMemoryStreams new CUDA sample that demonstrates the use of OpenMP and CUDA streams with Unified Memory on a single GPU.
- Added 1_Utilities/p2pBandwidthTestLatency new CUDA sample that demonstrates how measure latency between pairs of GPUs with P2P enabled and P2P disabled.
- ▶ Added **6_Advanced/StreamPriorities** This sample demonstrates basic use of the new CUDA 6.0 feature stream priorities.
- ▶ Added 7_CUDALibraries/ConjugateGradientUM This sample implements a conjugate gradient solver on GPU using cuBLAS and cuSPARSE library, using Unified Memory.

1.6. CUDA 5.5

- Linux makefiles have been updated to generate code for the AMRv7 architecture. Only the ARM hard-float floating point ABI is supported. Both native ARMv7 compilation and cross compilation from x86 is supported
- ▶ Performance improvements in CUDA toolkit for Kepler GPUs (SM 3.0 and SM 3.5)
- Makefiles projects have been updated to properly find search default paths for OpenGL, CUDA, MPI, and OpenMP libraries for all OS Platforms (Mac, Linux x86, Linux ARM).
- Linux and Mac project Makefiles now invoke NVCC for building and linking projects.
- ▶ Added **O_Simple/cppOverload** new CUDA sample that demonstrates how to use C++ overloading with CUDA.
- Added 6_Advanced/cdpBezierTessellation new CUDA sample that demonstrates an advanced method of implementing Bezier Line Tessellation using CUDA Dynamic Parallelism. Requires compute capability 3.5 or higher.
- ► Added **7_CUDALibrariess/jpegNPP** new CUDA sample that demonstrates how to use NPP for JPEG compression on the GPU.
- CUDA Samples now have better integration with Nsight Eclipse IDE.
- 6_Advanced/ptxjit sample now includes a new API to demonstrate PTX linking at the driver level.

1.7. CUDA 5.0

- New directory structure for CUDA samples. Samples are classified accordingly to categories: 0_Simple, 1_Utilities, 2_Graphics, 3_Imaging, 4_Finance, 5_Simulations, 6_Advanced, and 7_CUDALibraries
- Added O_Simple/simpleIPC CUDA Runtime API sample is a very basic sample that demonstrates Inter Process Communication with one process per GPU for computation. Requires Compute Capability 2.0 or higher and a Linux Operating System.
- Added O_Simple/simpleSeparateCompilation demonstrates a CUDA 5.0 feature, the ability to create a GPU device static library and use it within another CUDA kernel. This example demonstrates how to pass in a GPU device function (from the GPU device static library) as a function pointer to be called. Requires Compute Capability 2.0 or higher.
- Added 2_Graphics/bindlessTexture demonstrates use of cudaSurfaceObject, cudaTextureObject, and MipMap support in CUDA. Requires Compute Capability 3.0 or higher.
- Added **3_Imaging/stereoDisparity** demonstrates how to compute a stereo disparity map using SIMD SAD (Sum of Absolute Difference) intrinsics. Requires Compute Capability 2.0 or higher.
- Added **O_Simple/cdpSimpleQuicksort** demonstrates a simple quicksort implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.
- ▶ Added **O_Simple/cdpSimplePrint** demonstrates simple printf implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.
- ▶ Added **6_Advanced/cdpLUDecomposition** demonstrates LU Decomposition implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.
- ▶ Added 6_Advanced/cdpAdvancedQuicksort demonstrates an advanced quicksort implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.
- ► Added 6_Advanced/cdpQuadtree demonstrates Quad Trees implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.
- Added 7_CUDALibraries/simpleDevLibCUBLAS implements a simple cuBLAS function calls that call GPU device API library running cuBLAS functions. cuBLAS device code functions take advantage of CUDA Dynamic Parallelism and requires compute capability of 3.5 or higher.

1.8. CUDA 4.2

► Added **segmentationTreeThrust** - demonstrates a method to build image segmentation trees using Thrust. This algorithm is based on Boruvka's MST algorithm.

1.9. CUDA 4.1

- Added MersenneTwisterGP11213 implements Mersenne Twister GP11213, a pseudorandom number generator using the curand library.
- Added HSOpticalFlow When working with image sequences or video it's often useful to have information about objects movement. Optical flow describes apparent motion of objects in image sequence. This sample is a Horn-Schunck method for optical flow written using CUDA.
- ▶ Added **volumeFiltering** demonstrates basic volume rendering and filtering using 3D textures.
- Added **simpleCubeMapTexture** demonstrates how to use **texcubemap** fetch instruction in a CUDA C program.
- ▶ Added **simpleAssert** demonstrates how to use GPU assert in a CUDA C program.
- Added grabcutNPP CUDA implementation of Rother et al. GrabCut approach using the 8 neighborhood NPP Graphcut primitive introduced in CUDA 4.1. (C. Rother, V. Kolmogorov, A. Blake. GrabCut: Interactive Foreground Extraction Using Iterated Graph Cuts. ACM Transactions on Graphics (SIGGRAPH'04), 2004).

Chapter 2. GETTING STARTED

The CUDA Samples are an educational resource provided to teach CUDA programming concepts. The CUDA Samples are not meant to be used for performance measurements.

For system requirements and installation instructions, please refer to the Linux Installation Guide, the Windows Installation Guide, and the Mac Installation Guide.

2.1. Getting CUDA Samples

Windows

On Windows, the CUDA Samples are installed using the CUDA Toolkit Windows Installer. By default, the CUDA Samples are installed in:

C:\ProgramData\NVIDIA Corporation\CUDA Samples\v8.0\

The installation location can be changed at installation time.

Linux

On Linux, to install the CUDA Samples, the CUDA toolkit must first be installed. See the Linux Installation Guide for more information on how to install the CUDA Toolkit.

Then the CUDA Samples can be installed by running the following command, where <arget_path> is the location where to install the samples:

\$ cuda-install-samples-8.0.sh <target_path>

Mac OSX

On Mac OSX, to install the CUDA Samples, the CUDA toolkit must first be installed. See the Mac Installation Guide for more information on how to install the CUDA Toolkit.

Then the CUDA Samples can be installed by running the following command, where <arget_path> is the location where to install the samples:

\$ cuda-install-samples-8.0.sh <target path>

2.2. Building Samples

Windows

The Windows samples are built using the Visual Studio IDE. Solution files (.sln) are provided for each supported version of Visual Studio, using the format:

```
*_vs<version>.sln - for Visual Studio <version>
```

Complete samples solution files exist at:

```
C:\ProgramData\NVIDIA Corporation\CUDA Samples\v8.0\
```

Each individual sample has its own set of solution files at:

```
C:\ProgramData\NVIDIA Corporation\CUDA Samples\v8.0\<sample dir>\
```

To build/examine all the samples at once, the complete solution files should be used. To build/examine a single sample, the individual sample solution files should be used.



Some samples require that the Microsoft DirectX SDK (June 2010 or newer) be installed and that the VC++ directory paths are properly set up (Tools > Options...). Check DirectX Dependencies section for details.

Linux

The Linux samples are built using makefiles. To use the makefiles, change the current directory to the sample directory you wish to build, and run make:

```
$ cd <sample_dir>
$ make
```

The samples makefiles can take advantage of certain options:

► TARGET_ARCH=<arch> - cross-compile targeting a specific architecture. Allowed architectures are x86_64, armv7l, aarch64, and ppc64le.

By default, TARGET_ARCH is set to HOST_ARCH. On a x86_64 machine, not setting TARGET_ARCH is the equivalent of setting TARGET_ARCH=x86_64.

```
$ make TARGET_ARCH=x86_64
$ make TARGET_ARCH=armv71
$ make TARGET_ARCH=aarch64
$ make TARGET_ARCH=ppc641e
```

See here for more details.

dbg=1 - build with debug symbols

```
$ make dbg=1
```

▶ SMS="A B ..." - override the SM architectures for which the sample will be built, where "A B ..." is a space-delimited list of SM architectures. For example, to generate SASS for SM 20 and SM 30, use SMS="20 30".

```
$ make SMS="20 30"
```

► HOST_COMPILER=<host_compiler> - override the default g++ host compiler. See the Linux Installation Guide for a list of supported host compilers.

```
$ make HOST COMPILER=g++
```

Mac

The Mac samples are built using makefiles. To use the makefiles, change directory into the sample directory you wish to build, and run make:

```
$ cd <sample_dir>
$ make
```

The samples makefiles can take advantage of certain options:

▶ **dbg=1** - build with debug symbols

```
$ make dbg=1
```

▶ SMS="A B ..." - override the SM architectures for which the sample will be built, where "A B ..." is a space-delimited list of SM architectures. For example, to generate SASS for SM 20 and SM 30, use SMS="20 30".

```
$ make SMS="A B ..."
```

► HOST_COMPILER=<host_compiler> - override the default clang host compiler. See the Mac Installation Guide for a list of supported host compilers.

```
$ make HOST COMPILER=clang
```

2.3. CUDA Cross-Platform Samples

This section describes the options used to build cross-platform samples.

TARGET_ARCH=<arch> and TARGET_OS=<os> should be chosen based on the supported targets shown below. TARGET_FS=<path> can be used to point nvcc to libraries and headers used by the sample.

Table 1 Supported Target Arch/OS Combinations

		TARGET OS			
		linux	darwin	android	qnx
TARGET ARCH	x86_64	YES	YES	NO	NO
	armv7l	YES	NO	YES	YES
	aarch64	NO	NO	YES	МО
	ppc64le	YES	NO	NO	NO

TARGET ARCH

The target architecture must be specified when cross-compiling applications. If not specified, it defaults to the host architecture. Allowed architectures are:

- ▶ **x86 64** 64-bit x86 CPU architecture
- ▶ armv71 32-bit ARM CPU architecture, like that found on Jetson TK1
- aarch64 64-bit ARM CPU architecture, found on certain Android systems
- ppc641e 64-bit little-endian IBM POWER8 architecture

TARGET_OS

The target OS must be specified when cross-compiling applications. If not specified, it defaults to the host OS. Allowed OSes are:

- ▶ linux for any Linux distributions
- darwin for Mac OS X
- android for any supported device running Android
- qnx for any supported device running QNX

TARGET_FS

The most reliable method to cross-compile the CUDA Samples is to use the TARGET_FS variable. To do so, mount the target's filesystem on the host, say at /mnt/target. This is typically done using exportfs. In cases where exportfs is unavailable, it is sufficient to copy the target's filesystem to /mnt/target. To cross-compile a sample, execute:

```
$ make TARGET_ARCH=<arch> TARGET_OS=<os> TARGET_FS=/mnt/target
```

Copying Libraries

If the TARGET_FS option is not available, the libraries used should be copied from the target system to the host system, say at <code>/opt/target/libs</code>. If the sample uses GL, the GL headers must also be copied, say at <code>/opt/target/include</code>. The linker must then be told where the libraries are with the <code>-rpath-link</code> and/or <code>-L</code> options. To ignore unresolved symbols from some libraries, use the <code>--unresolved-symbols</code> option as shown below. <code>SAMPLE_ENABLED</code> should be used to force the sample to build. For example, to cross-compile a sample which uses such libraries, execute:

2.4. Using CUDA Samples to Create Your Own CUDA Projects

2.4.1. Creating CUDA Projects for Windows

Creating a new CUDA Program using the CUDA Samples infrastructure is easy. We have provided a **template** project that you can copy and modify to suit your needs. Just follow these steps:

(<category> refers to one of the following folders: 0_Simple, 1_Utilities,
2_Graphics, 3_Imaging, 4_Finance, 5_Simulations, 6_Advanced,
7 CUDALibraries.)

1. Copy the content of:

C:\ProgramData\NVIDIA Corporation\CUDA Samples\v8.0\<category>\template to a directory of your own:

C:\ProgramData\NVIDIA Corporation\CUDA Samples\v8.0\<category>\myproject

- **2.** Edit the filenames of the project to suit your needs.
- Edit the *.sln, *.vcproj and source files.
 Just search and replace all occurrences of template with myproject.
- **4.** Build the 32-bit and/or 64-bit, release or debug configurations using:

```
myproject_vs<version>.sln
```

5. Run myproject.exe from the release or debug directories located in:

C:\ProgramData\NVIDIA Corporation\CUDA Samples\v8.0\bin\win[32|64]\[release| debug]

6. Now modify the code to perform the computation you require. See the *CUDA Programming Guide* for details of programming in CUDA.

2.4.2. Creating CUDA Projects for Linux



The default installation folder <SAMPLES_INSTALL_PATH> is NVIDIA_CUDA_8.0_Samples and <category> is one of the following: 0_Simple, 1_Utilities, 2_Graphics, 3_Imaging, 4_Finance, 5_Simulations, 6_Advanced, 7_CUDALibraries.

Creating a new CUDA Program using the NVIDIA CUDA Samples infrastructure is easy. We have provided a **template** project that you can copy and modify to suit your needs. Just follow these steps:

1. Copy the **template** project:

```
cd <SAMPLES_INSTALL_PATH>/<category>
cp -r template <myproject>cd <SAMPLES_INSTALL_PATH>/<category>
```

2. Edit the filenames of the project to suit your needs:

```
mv template.cu myproject.cu
mv template_cpu.cpp myproject_cpu.cpp
```

3. Edit the Makefile and source files.

Just search and replace all occurrences of template with myproject.

4. Build the project as (release):

make

To build the project as (debug), use "make dbg=1":

make dbg=1

5. Run the program:

```
../../bin/x86 64/linux/release/myproject
```

6. Now modify the code to perform the computation you require. See the *CUDA Programming Guide* for details of programming in CUDA.

2.4.3. Creating CUDA Projects for Mac OS X



The default installation folder <samples_install_path> is: /Developer/NVIDIA/CUDA-8.0/samples

Creating a new CUDA Program using the NVIDIA CUDA Samples infrastructure is easy. We have provided a **template** project that you can copy and modify to suit your needs. Just follow these steps:

(<category> is one of the following: 0_Simple, 1_Utilities, 2_Graphics, 3_Imaging,
4_Finance, 5_Simulations, 6_Advanced, 7_CUDALibraries.)

1. Copy the template project:

```
cd <SAMPLES_INSTALL_PATH>/<category>
cp -r template <myproject>
```

2. Edit the filenames of the project to suit your needs:

```
mv template.cu myproject.cu
mv template_cpu.cpp myproject_cpu.cpp
```

3. Edit the **Makefile** and source files.

Just search and replace all occurrences of template with myproject.

4. Build the project as (release):

make

Note: To build the project as (debug), use "make dbg=1"

make dbg=1

5. Run the program:

```
../../bin/x86 64/darwin/release/myproject
```

(It should print PASSED.)

6. Now modify the code to perform the computation you require. See the *CUDA Programming Guide* for details of programming in CUDA.

Chapter 3. SAMPLES REFERENCE

This document contains a complete listing of the code samples that are included with the NVIDIA CUDA Toolkit. It describes each code sample, lists the minimum GPU specification, and provides links to the source code and white papers if available.

The code samples are divided into the following categories:

Simple Reference

Basic CUDA samples for beginners that illustrate key concepts with using CUDA and CUDA runtime APIs.

Utilities Reference

Utility samples that demonstrate how to query device capabilities and measure GPU/CPU bandwidth.

Graphics Reference

Graphical samples that demonstrate interoperability between CUDA and OpenGL or DirectX.

Imaging Reference

Samples that demonstrate image processing, compression, and data analysis.

Finance Reference

Samples that demonstrate parallel algorithms for financial computing.

Simulations Reference

Samples that illustrate a number of simulation algorithms implemented with CUDA.

Advanced Reference

Samples that illustrate advanced algorithms implemented with CUDA.

Cudalibraries Reference

Samples that illustrate how to use CUDA platform libraries (NPP, cuBLAS, cuFFT, cuSPARSE, and cuRAND).

3.1. Simple Reference

asyncAPI

This sample uses CUDA streams and events to overlap execution on CPU and GPU.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaEventCreate, cudaEventRecord, cudaEventQuery, cudaEventDestroy,

cudaEventElapsedTime, cudaMemcpyAsync

Key Concepts Asynchronous Data Transfers, CUDA Streams and Events

Supported OSes Linux, Windows, OS X

cdpSimplePrint - Simple Print (CUDA Dynamic Parallelism)

This sample demonstrates simple printf implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CDP

Supported SM SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUDA Dynamic Parallelism

Supported OSes Linux, Windows, OS X

cdpSimpleQuicksort - Simple Quicksort (CUDA Dynamic Parallelism)

This sample demonstrates simple quicksort implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CDP

Supported SM SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUDA Dynamic Parallelism

clock - Clock

This example shows how to use the clock function to measure the performance of block of threads of a kernel accurately.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaFree, cudaMemcpy

Key Concepts Performance Strategies

Supported OSes Linux, Windows, OS X

clock_nvrtc - Clock libNVRTC

This example shows how to use the clock function using libNVRTC to measure the performance of block of threads of a kernel accurately.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuMemAlloc, cuLaunchKernel, cuMemcpyHtoD, cuMemFree

Key Concepts Performance Strategies, Runtime Compilation

Supported OSes Linux, Windows, OS X

cppIntegration - C++ Integration

This example demonstrates how to integrate CUDA into an existing C++ application, i.e. the CUDA entry point on host side is only a function which is called from C++ code and only the file containing this function is compiled with nvcc. It also demonstrates that vector types can be used from cpp.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaFree, cudaMemcpy

cppOverload

This sample demonstrates how to use C++ function overloading on the GPU.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaFuncSetCacheConfig, cudaFuncGetAttributes

Key Concepts C++ Function Overloading, CUDA Streams and Events

Supported OSes Linux, Windows, OS X

cudaOpenMP

This sample demonstrates how to use OpenMP API to write an application for multiple GPUs.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies
OpenMP
Supported SM
SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1
Architecture
CUDA API
cudaMalloc, cudaFree, cudaMemcpy
Key Concepts
CUDA Systems Integration, OpenMP, Multithreading

fp16ScalarProduct - FP16 Scalar Product

Calculates scalar product of two vectors of FP16 numbers.

Linux, Windows

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies FP16

Supported SM SM 5.3, SM 6.0, SM 6.1

Architecture

Supported OSes

CUDA API cudaMalloc, cudaMallocHost, cudaMemcpy, cudaFree, cudaFreeHost

Key Concepts CUDA Runtime API

Supported OSes Linux, Windows, OS X

inlinePTX - Using Inline PTX

A simple test application that demonstrates a new CUDA 4.0 ability to embed PTX in a CUDA kernel.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaMallocHost, cudaFree, cudaFreeHost, cudaMemcpy

Key Concepts Performance Strategies, PTX Assembly, CUDA Driver API

Supported OSes Linux, Windows, OS X

inlinePTX_nvrtc - Using Inline PTX with libNVRTC

A simple test application that demonstrates a new CUDA 4.0 ability to embed PTX in a CUDA kernel.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuMemAlloc, cuLaunchKernel, cuMemcpyDtoH

Key Concepts Performance Strategies, PTX Assembly, CUDA Driver API, Runtime Compilation

Supported OSes Linux, Windows, OS X

matrixMul - Matrix Multiplication (CUDA Runtime API Version)

This sample implements matrix multiplication and is exactly the same as Chapter 6 of the programming guide. It has been written for clarity of exposition to illustrate various CUDA programming principles, not with the goal of providing the most performant generic kernel for matrix multiplication. To illustrate GPU performance for matrix

multiply, this sample also shows how to use the new CUDA 4.0 interface for CUBLAS to demonstrate high-performance performance for matrix multiplication.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaEventCreate, cudaEventRecord, cudaEventQuery, cudaEventDestroy,

cudaEventElapsedTime, cudaEventSynchronize, cudaMalloc, cudaFree,

cudaMemcpy

Key Concepts CUDA Runtime API, Linear Algebra

Supported OSes Linux, Windows, OS X

matrixMul_nvrtc - Matrix Multiplication with libNVRTC

This sample implements matrix multiplication and is exactly the same as Chapter 6 of the programming guide. It has been written for clarity of exposition to illustrate various CUDA programming principles, not with the goal of providing the most performant generic kernel for matrix multiplication. To illustrate GPU performance for matrix multiply, this sample also shows how to use the new CUDA 4.0 interface for CUBLAS to demonstrate high-performance performance for matrix multiplication.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuModuleLoad, cuModuleLoadDataEx, cuModuleGetFunction, cuMemAlloc,

cuMemFree, cuMemcpyHtoD, cuMemcpyDtoH, cuLaunchKernel

Key Concepts CUDA Runtime API, Linear Algebra, Runtime Compilation

Supported OSes Linux, Windows, OS X

matrixMulCUBLAS - Matrix Multiplication (CUBLAS)

This sample implements matrix multiplication from Chapter 3 of the programming guide. To illustrate GPU performance for matrix multiply, this sample also shows how to use the new CUDA 4.0 interface for CUBLAS to demonstrate high-performance performance for matrix multiplication.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUBLAS

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaEventCreate, cudaEventRecord, cudaEventQuery, cudaEventDestroy,

cudaEventElapsedTime, cudaMalloc, cudaFree, cudaMemcpy, cublasCreate,

cublasSgemm

Key Concepts CUDA Runtime API, Performance Strategies, Linear Algebra, CUBLAS

Supported OSes Linux, Windows, OS X

matrixMulDrv - Matrix Multiplication (CUDA Driver API Version)

This sample implements matrix multiplication and uses the new CUDA 4.0 kernel launch Driver API. It has been written for clarity of exposition to illustrate various CUDA programming principles, not with the goal of providing the most performant generic kernel for matrix multiplication. CUBLAS provides high-performance matrix multiplication.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuModuleLoad, cuModuleLoadDataEx, cuModuleGetFunction, cuMemAlloc,

cuMemFree, cuMemcpyHtoD, cuMemcpyDtoH, cuLaunchKernel

Key Concepts CUDA Driver API, Matrix Multiply

Supported OSes Linux, Windows, OS X

simpleAssert

This CUDA Runtime API sample is a very basic sample that implements how to use the assert function in the device code. Requires Compute Capability 2.0.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaMallocHost, cudaFree, cudaFreeHost, cudaMemcpy

Key Concepts Assert

simpleAssert_nvrtc - simpleAssert with libNVRTC

This CUDA Runtime API sample is a very basic sample that implements how to use the assert function in the device code. Requires Compute Capability 2.0.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuLaunchKernel

Key Concepts Assert, Runtime Compilation

Supported OSes Linux, Windows

simpleAtomicIntrinsics - Simple Atomic Intrinsics

A simple demonstration of global memory atomic instructions. Requires Compute Capability 2.0 or higher.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaFree, cudaMemcpy, cudaFreeHost

Key Concepts Atomic Intrinsics

Supported OSes Linux, Windows, OS X

simpleAtomicIntrinsics_nvrtc - Simple Atomic Intrinsics with libNVRTC

A simple demonstration of global memory atomic instructions. This sample makes use of NVRTC for Runtime Compilation.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuMemAlloc, cuMemFree, cuMemcpyHtoD, cuLaunchKernel

Key Concepts Atomic Intrinsics, Runtime Compilation

Supported OSes Linux, Windows, OS X

simpleCallback - Simple CUDA Callbacks

This sample implements multi-threaded heterogeneous computing workloads with the new CPU callbacks for CUDA streams and events introduced with CUDA 5.0.

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

CUDA API cudaStreamCreate, cudaMemcpyAsync, cudaStreamAddCallback,

cudaStreamDestroy

Key Concepts CUDA Streams, Callback Functions, Multithreading

Supported OSes Linux, Windows, OS X

simpleCubemapTexture - Simple Cubemap Texture

Simple example that demonstrates how to use a new CUDA 4.1 feature to support cubemap Textures in CUDA C.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaMalloc3DArray, cudaMemcpy3D, cudaCreateChannelDesc,

cudaBindTextureToArray, cudaMalloc, cudaFree, cudaFreeArray, cudaMemcpy

Key Concepts Texture, Volume Processing

Supported OSes Linux, Windows, OS X

simpleIPC

This CUDA Runtime API sample is a very basic sample that demonstrates Inter Process Communication with one process per GPU for computation. Requires Compute Capability 2.0 or higher and a Linux Operating System

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies IPC

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudalpcGetEventHandlet, cudalpcOpenMemHandle, cudalpcCloseMemHandle,

cudaFreeHost, cudaMemcpy

Key Concepts CUDA Systems Integration, Peer to Peer, InterProcess Communication

Supported OSes Linux

simpleLayeredTexture - Simple Layered Texture

Simple example that demonstrates how to use a new CUDA 4.0 feature to support layered Textures in CUDA C.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaMalloc3DArray, cudaMemcpy3D, cudaCreateChannelDesc,

cudaBindTextureToArray, cudaMalloc, cudaFree, cudaFreeArray, cudaMemcpy

Key Concepts Texture, Volume Processing

Supported OSes Linux, Windows, OS X

simpleMPI

Simple example demonstrating how to use MPI in combination with CUDA.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies MP

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMallco, cudaFree, cudaMemcpy

Key Concepts CUDA Systems Integration, MPI, Multithreading

simpleMultiCopy - Simple Multi Copy and Compute

Supported in GPUs with Compute Capability 1.1, overlapping compute with one memcopy is possible from the host system. For Quadro and Tesla GPUs with Compute Capability 2.0, a second overlapped copy operation in either direction at full speed is possible (PCI-e is symmetric). This sample illustrates the usage of CUDA streams to achieve overlapping of kernel execution with data copies to and from the device.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaEventCreate, cudaEventRecord, cudaEventQuery, cudaEventDestroy,

cuda Event Elapsed Time, cuda Memcpy Async

Key Concepts CUDA Streams and Events, Asynchronous Data Transfers, Overlap Compute

and Copy, GPU Performance

Supported OSes Linux, Windows, OS X

simpleMultiGPU - Simple Multi-GPU

This application demonstrates how to use the new CUDA 4.0 API for CUDA context management and multi-threaded access to run CUDA kernels on multiple-GPUs.

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

CUDA API cudaEventCreate, cudaEventRecord, cudaEventQuery, cudaEventDestroy,

cudaEventElapsedTime, cudaMemcpyAsync

Key Concepts Asynchronous Data Transfers, CUDA Streams and Events, Multithreading,

Multi-GPU

Supported OSes Linux, Windows, OS X

simpleOccupancy

This sample demonstrates the basic usage of the CUDA occupancy calculator and occupancy-based launch configurator APIs by launching a kernel with the launch configurator, and measures the utilization difference against a manually configured launch.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Occupancy Calculator

simpleP2P - Simple Peer-to-Peer Transfers with Multi-GPU

This application demonstrates CUDA APIs that support Peer-To-Peer (P2P) copies, Peer-To-Peer (P2P) addressing, and Unified Virtual Memory Addressing (UVA) between multiple GPUs. In general, P2P is supported between two same GPUs with some exceptions, such as some Tesla and Quadro GPUs.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies only-64-bit

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

CUDA API cudaDeviceCanAccessPeer, cudaDeviceEnablePeerAccess,

 $cuda Device Disable Peer Access, \ cuda Event Create With Flags,$

cudaEventElapsedTime, cudaMemcpy

Key Concepts Performance Strategies, Asynchronous Data Transfers, Unified Virtual Address

Space, Peer to Peer Data Transfers, Multi-GPU

Supported OSes Linux, Windows

simplePitchLinearTexture - Pitch Linear Texture

Use of Pitch Linear Textures

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMallocPitch, cudaMallocArray, cudaMemcpy2D, cudaMemcpyToArray,

cudaBindTexture2D, cudaBindTextureToArray, cudaCreateChannelDesc, cudaMalloc, cudaFree, cudaFreeArray, cudaUnbindTexture, cudaMemset2D,

cudaMemcpy2D

Key Concepts Texture, Image Processing

simplePrintf

This CUDA Runtime API sample is a very basic sample that implements how to use the printf function in the device code. Specifically, for devices with compute capability less than 2.0, the function cuPrintf is called; otherwise, printf can be used directly.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaPrintfDisplay, cudaPrintfEnd

Key Concepts Debugging

Supported OSes Linux, Windows, OS X

simpleSeparateCompilation - Simple Static GPU Device Library

This sample demonstrates a CUDA 5.0 feature, the ability to create a GPU device static library and use it within another CUDA kernel. This example demonstrates how to pass in a GPU device function (from the GPU device static library) as a function pointer to be called. This sample requires devices with compute capability 2.0 or higher.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Separate Compilation

Supported OSes Linux, Windows, OS X

simpleStreams

This sample uses CUDA streams to overlap kernel executions with memory copies between the host and a GPU device. This sample uses a new CUDA 4.0 feature that supports pinning of generic host memory. Requires Compute Capability 2.0 or higher.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaEventCreate, cudaEventRecord, cudaEventQuery, cudaEventDestroy,

cuda Event Elapsed Time, cuda Memcpy Async

Key Concepts Asynchronous Data Transfers, CUDA Streams and Events

simpleSurfaceWrite - Simple Surface Write

Simple example that demonstrates the use of 2D surface references (Write-to-Texture)

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaMallocArray, cudaBindSurfaceToArray,

cudaBindTextureToArray, cudaCreateChannelDesc, cudaMalloc, cudaFree,

cudaFreeArray, cudaMemcpy

Key Concepts Texture, Surface Writes, Image Processing

Supported OSes Linux, Windows, OS X

simpleTemplates - Simple Templates

This sample is a templatized version of the template project. It also shows how to correctly templatize dynamically allocated shared memory arrays.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts C++ Templates

Supported OSes Linux, Windows, OS X

simpleTemplates_nvrtc - Simple Templates with libNVRTC

This sample is a templatized version of the template project. It also shows how to correctly templatize dynamically allocated shared memory arrays.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

Key Concepts C++ Templates, Runtime Compilation

simpleTexture - Simple Texture

Simple example that demonstrates use of Textures in CUDA.

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

CUDA API cudaMalloc, cudaMallocArray, cudaMemcpyToArray, cudaCreateChannelDesc,

cudaBindTextureToArray, cudaMalloc, cudaFree, cudaFreeArray, cudaMemcpy

Key Concepts CUDA Runtime API, Texture, Image Processing

Supported OSes Linux, Windows, OS X

simpleTextureDrv - Simple Texture (Driver Version)

Simple example that demonstrates use of Textures in CUDA. This sample uses the new CUDA 4.0 kernel launch Driver API.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuModuleLoad, cuModuleLoadDataEx, cuModuleGetFunction, cuLaunchKernel,

cuCtxSynchronize, cuMemcpyDtoH, cuMemAlloc, cuMemFree, cuArrayCreate,

cuArrayDestroy, cuCtxDetach, cuMemcpy2D, cuModuleGetTexRef, cuTexRefSetArray, cuTexRefSetAddressMode, cuTexRefSetFilterMode,

cuTexRefSetFlags, cuTexRefSetFormat, cuParamSetTexRef

Key Concepts CUDA Driver API, Texture, Image Processing

Supported OSes Linux, Windows, OS X

simpleVoteIntrinsics - Simple Vote Intrinsics

Simple program which demonstrates how to use the Vote (any, all) intrinsic instruction in a CUDA kernel. Requires Compute Capability 2.0 or higher.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMallco, cudaFree, cudaMemcpy, cudaFreeHost

Key Concepts Vote Intrinsics

simpleVoteIntrinsics_nvrtc - Simple Vote Intrinsics with libNVRTC

Simple program which demonstrates how to use the Vote (any, all) intrinsic instruction in a CUDA kernel with runtime compilation using NVRTC APIs. Requires Compute Capability 2.0 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuMemAlloc, cuMemFree, cuMemcpyHtoD, cuMemFree

Key Concepts Vote Intrinsics, CUDA Driver API, Runtime Compilation

Supported OSes Linux, Windows, OS X

simpleZeroCopy

This sample illustrates how to use Zero MemCopy, kernels can read and write directly to pinned system memory.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaEventCreate, cudaEventRecord, cudaEventQuery, cudaEventDestroy,

cudaEventElapsedTime, cudaHostAlloc, cudaHostGetDevicePointer,

cudaHostRegister, cudaHostUnregister, cudaFreeHost

Key Concepts Performance Strategies, Pinned System Paged Memory, Vector Addition

Supported OSes Linux, Windows, OS X

Whitepaper CUDA2.2PinnedMemoryAPIs.pdf

systemWideAtomics - System wide Atomics

A simple demonstration of system wide atomic instructions.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies UVM

Supported SM SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaFree, cudaMemcpy, cudaFreeHost

Key Concepts Atomic Intrinsics, Unified Memory

Supported OSes Linux

template - Template

A trivial template project that can be used as a starting point to create new CUDA projects.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMalloc, cudaFree, cudaDeviceSynchronize, cudaMemcpy

Key Concepts Device Memory Allocation

Supported OSes Linux, Windows, OS X

UnifiedMemoryStreams - Unified Memory Streams

This sample demonstrates the use of OpenMP and streams with Unified Memory on a single GPU.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies UVM, CUBLAS

Supported SM SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaMallocManaged, cudaStreamAttachManagedMem

Key Concepts CUDA Systems Integration, OpenMP, CUBLAS, Multithreading, Unified Memory,

CUDA Streams and Events

vectorAdd - Vector Addition

This CUDA Runtime API sample is a very basic sample that implements element by element vector addition. It is the same as the sample illustrating Chapter 3 of the programming guide with some additions like error checking.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaEventCreate, cudaEventRecord, cudaEventQuery, cudaEventDestroy,

cudaEventElapsedTime, cudaEventSynchronize, cudaMalloc, cudaFree,

cudaMemcpy

Key Concepts CUDA Runtime API, Vector Addition

Supported OSes Linux, Windows, OS X

vectorAdd_nvrtc - Vector Addition with libNVRTC

This CUDA Driver API sample uses NVRTC for runtime compilation of vector addition kernel. Vector addition kernel demonstrated is the same as the sample illustrating Chapter 3 of the programming guide.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuMemAlloc, cuMemFree, cuMemcpyHtoD, cuMemcpyDtoH

Key Concepts CUDA Driver API, Vector Addition, Runtime Compilation

Supported OSes Linux, Windows, OS X

vectorAddDrv - Vector Addition Driver API

This Vector Addition sample is a basic sample that is implemented element by element. It is the same as the sample illustrating Chapter 3 of the programming guide with some additions like error checking. This sample also uses the new CUDA 4.0 kernel launch Driver API.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuModuleLoad, cuModuleLoadDataEx, cuModuleGetFunction, cuMemAlloc,

cuMemFree, cuMemcpyHtoD, cuMemcpyDtoH, cuLaunchKernel

Key Concepts CUDA Driver API, Vector Addition

Supported OSes Linux, Windows, OS X

3.2. Utilities Reference

bandwidthTest - Bandwidth Test

This is a simple test program to measure the memcopy bandwidth of the GPU and memcpy bandwidth across PCI-e. This test application is capable of measuring device to device copy bandwidth, host to device copy bandwidth for pageable and page-locked memory, and device to host copy bandwidth for pageable and page-locked memory.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaSetDevice, cudaHostAlloc, cudaFree, cudaMallocHost, cudaFreeHost,

cudaMemcpy, cudaMemcpyAsync, cudaEventCreate, cudaEventRecord, cudaEventDestroy, cudaDeviceSynchronize, cudaEventElapsedTime

Key Concepts CUDA Streams and Events, Performance Strategies

Supported OSes Linux, Windows, OS X

deviceQuery - Device Query

This sample enumerates the properties of the CUDA devices present in the system.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaSetDevice, cudaGetDeviceCount, cudaGetDeviceProperties,

 $cuda Driver Get Version, \ cuda Runtime Get Version$

Key Concepts CUDA Runtime API, Device Query

Supported OSes Linux, Windows, OS X

deviceQueryDrv - Device Query Driver API

This sample enumerates the properties of the CUDA devices present using CUDA Driver API calls

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API culnit, cuDeviceGetCount, cuDeviceComputeCapability, cuDriverGetVersion,

 $cuDevice Total Mem, \ cuDevice Get Attribute$

Key Concepts CUDA Driver API, Device Query

Supported OSes Linux, Windows, OS X

p2pBandwidthLatencyTest - Peer-to-Peer Bandwidth Latency Test with Multi-GPUs

This application demonstrates the CUDA Peer-To-Peer (P2P) data transfers between pairs of GPUs and computes latency and bandwidth. Tests on GPU pairs using P2P and without P2P are tested.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaDeviceCanAccessPeer, cudaDeviceEnablePeerAccess,

 $cuda Device Disable Peer Access, \ cuda Event Create With Flags,$

cudaEventElapsedTime, cudaMemcpy

Key Concepts Performance Strategies, Asynchronous Data Transfers, Unified Virtual Address

Space, Peer to Peer Data Transfers, Multi-GPU

Supported OSes Linux, Windows, OS X

topologyQuery - Topology Query

A simple exemple on how to query the topology of a system with multiple GPU

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaDeviceGetP2PAttribute, cudaGetDeviceAttribute, cudaGetDeviceCount

Key Concepts Performance Strategies, Multi-GPU

3.3. Graphics Reference

bindlessTexture - Bindless Texture

This example demonstrates use of cudaSurfaceObject, cudaTextureObject, and MipMap support in CUDA. A GPU with Compute Capability SM 3.0 is required to run the sample.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Supported SM SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Texture

Supported OSes Linux, Windows, OS X

Mandelbrot

This sample uses CUDA to compute and display the Mandelbrot or Julia sets interactively. It also illustrates the use of "double single" arithmetic to improve precision when zooming a long way into the pattern. This sample uses double precision. Thanks to Mark Granger of NewTek who submitted this code sample.!

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

 $cuda Graphics Register Resource, \ cuda Graphics GLR egister Buffer,$

cuda Graphics Unregister Resource

Key Concepts Graphics Interop, Data Parallel Algorithms

Supported OSes Linux, Windows, OS X

marchingCubes - Marching Cubes Isosurfaces

This sample extracts a geometric isosurface from a volume dataset using the marching cubes algorithm. It uses the scan (prefix sum) function from the Thrust library to perform stream compaction.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cuda Graphics Unregister Resource

Key Concepts OpenGL Graphics Interop, Vertex Buffers, 3D Graphics, Physically Based

Simulation

Supported OSes Linux, Windows, OS X

simpleD3D10 - Simple Direct3D10 (Vertex Array)

Simple program which demonstrates interoperability between CUDA and Direct3D10. The program generates a vertex array with CUDA and uses Direct3D10 to render the geometry. A Direct3D Capable device is required.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaD3D10GetDevice, cudaD3D10SetDirect3DDevice,

cudaGraphicsD3D10RegisterResource, cudaGraphicsResourceSetMapFlags, cudaGraphicsSubResourceGetMappedArray, cudaMemcpy2DToArray,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, 3D Graphics

Supported OSes Windows

simpleD3D10RenderTarget - Simple Direct3D10 Render Target

Simple program which demonstrates interop of rendertargets between Direct3D10 and CUDA. The program uses RenderTarget positions with CUDA and generates a histogram with visualization. A Direct3D10 Capable device is required.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

CUDA API cudaD3D10GetDevice, cudaD3D10SetDirect3DDevice,

cudaGraphicsD3D10RegisterResource, cudaGraphicsResourceSetMapFlags, cudaGraphicsSubResourceGetMappedArray, cudaMemcpy2DToArray,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Texture

Supported OSes Windows

simpleD3D10Texture - Simple D3D10 Texture

Simple program which demonstrates how to interoperate CUDA with Direct3D10 Texture. The program creates a number of D3D10 Textures (2D, 3D, and CubeMap) which are generated from CUDA kernels. Direct3D then renders the results on the screen. A Direct3D10 Capable device is required.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaD3D10GetDevice, cudaD3D10SetDirect3DDevice,

cudaGraphicsD3D10RegisterResource, cudaGraphicsResourceSetMapFlags, cudaGraphicsSubResourceGetMappedArray, cudaMemcpy2DToArray,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Texture

Supported OSes Windows

simpleD3D11Texture - Simple D3D11 Texture

Simple program which demonstrates Direct3D11 Texture interoperability with CUDA. The program creates a number of D3D11 Textures (2D, 3D, and CubeMap) which are written to from CUDA kernels. Direct3D then renders the results on the screen. A Direct3D Capable device is required.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaD3D11GetDevice, cudaD3D11SetDirect3DDevice,

cudaGraphicsD3D11RegisterResource, cudaGraphicsResourceSetMapFlags, cudaGraphicsSubResourceGetMappedArray, cudaMemcpy2DToArray,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Image Processing

Supported OSes Windows

simpleD3D9 - Simple Direct3D9 (Vertex Arrays)

Simple program which demonstrates interoperability between CUDA and Direct3D9. The program generates a vertex array with CUDA and uses Direct3D9 to render the geometry. A Direct3D capable device is required.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaD3D9GetDevice, cudaD3D9SetDirect3DDevice,

 $cuda Graphics D3D9 Register Resource, \ cuda Graphics Unregister Resource$

Key Concepts Graphics Interop

Supported OSes Windows

simpleD3D9Texture - Simple D3D9 Texture

Simple program which demonstrates Direct3D9 Texture interoperability with CUDA. The program creates a number of D3D9 Textures (2D, 3D, and CubeMap) which are written to from CUDA kernels. Direct3D then renders the results on the screen. A Direct3D capable device is required.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaD3D9GetDevice, cudaD3D9SetDirect3DDevice,

cudaGraphicsD3D9RegisterResource, cudaGraphicsResourceSetMapFlags, cudaGraphicsSubResourceGetMappedArray, cudaMemcpy2DToArray,

cudaMemcpy3D, cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Texture

Supported OSes Windows

simpleGL - Simple OpenGL

Simple program which demonstrates interoperability between CUDA and OpenGL. The program modifies vertex positions with CUDA and uses OpenGL to render the geometry.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Vertex Buffers, 3D Graphics

Supported OSes Linux, Windows, OS X

simpleGLES - Simple OpenGLES

Demonstrates data exchange between CUDA and OpenGL ES (aka Graphics interop). The program modifies vertex positions with CUDA and uses OpenGL ES to render the geometry.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GLES

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGraphicsMapResources, cudaGraphicsUnmapResources,

cudaGraphicsResourceGetMappedPointer, cudaGraphicsRegisterResource,

 $cuda Graphics GLR egister Buffer, \ cuda Graphics Unregister Resource$

Key Concepts Graphics Interop, Vertex Buffers, 3D Graphics

Supported OSes Linux

simpleGLES_EGLOutput - Simple OpenGLES EGLOutput

Demonstrates data exchange between CUDA and OpenGL ES (aka Graphics interop). The program modifies vertex positions with CUDA and uses OpenGL ES to render the geometry, and shows how to render directly to the display using the EGLOutput mechanism and the DRM library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies EGLOutput, GLES

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGraphicsMapResources, cudaGraphicsUnmapResources,

cudaGraphicsResourceGetMappedPointer, cudaGraphicsRegisterResource,

 $cuda Graphics GLR egister Buffer,\ cuda Graphics Unregister Resource$

Key Concepts Graphics Interop, Vertex Buffers, 3D Graphics

Supported OSes Linux

simpleGLES_screen - Simple OpenGLES on Screen

Demonstrates data exchange between CUDA and OpenGL ES (aka Graphics interop). The program modifies vertex positions with CUDA and uses OpenGL ES to render the geometry.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies screen, GLES

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGraphicsMapResources, cudaGraphicsUnmapResources,

cudaGraphicsResourceGetMappedPointer, cudaGraphicsRegisterResource,

 $cuda Graphics GLR egister Buffer,\ cuda Graphics Unregister Resource$

Key Concepts Graphics Interop, Vertex Buffers, 3D Graphics

Supported OSes Linux

simpleTexture3D - Simple Texture 3D

Simple example that demonstrates use of 3D Textures in CUDA.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Image Processing, 3D Textures, Surface Writes

Supported OSes Linux, Windows, OS X

SLID3D10Texture - SLI D3D10 Texture

Simple program which demonstrates SLI with Direct3D10 Texture interoperability with CUDA. The program creates a D3D10 Texture which is written to from a CUDA kernel. Direct3D then renders the results on the screen. A Direct3D Capable device is required.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaD3D10GetDevice, cudaD3D10SetDirect3DDevice,

cudaGraphicsD3D10RegisterResource, cudaGraphicsResourceSetMapFlags, cudaGraphicsSubResourceGetMappedArray, cudaMemcpy2DToArray,

cuda Graphics Unregister Resource

Key Concepts Performance Strategies, Graphics Interop, Image Processing, 2D Textures

Supported OSes Windows

volumeFiltering - Volumetric Filtering with 3D Textures and Surface Writes

This sample demonstrates 3D Volumetric Filtering using 3D Textures and 3D Surface Writes.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

 $cuda Graphics Register Resource, \ cuda Graphics GLR egister Buffer,$

cuda Graphics Unregister Resource

Key Concepts Graphics Interop, Image Processing, 3D Textures, Surface Writes

Supported OSes Linux, Windows, OS X

volumeRender - Volume Rendering with 3D Textures

This sample demonstrates basic volume rendering using 3D Textures.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Image Processing, 3D Textures

3.4. Imaging Reference

bicubicTexture - Bicubic B-spline Interoplation

This sample demonstrates how to efficiently implement a Bicubic B-spline interpolation filter with CUDA texture.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

 $cuda Graphics Unmap Resources, \ cuda Graphics Resource Get Mapped Pointer,$

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cuda Graphics Unregister Resource

Key Concepts Graphics Interop, Image Processing

Supported OSes Linux, Windows, OS X

bilateralFilter - Bilateral Filter

Bilateral filter is an edge-preserving non-linear smoothing filter that is implemented with CUDA with OpenGL rendering. It can be used in image recovery and denoising. Each pixel is weight by considering both the spatial distance and color distance between its neighbors. Reference: "C. Tomasi, R. Manduchi, Bilateral Filtering for Gray and Color Images, proceeding of the ICCV, 1998, http://users.soe.ucsc.edu/~manduchi/Papers/ICCV98.pdf"

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Image Processing

Supported OSes Linux, Windows, OS X

boxFilter - Box Filter

Fast image box filter using CUDA with OpenGL rendering.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

 $cuda Graphics Unmap Resources, \ cuda Graphics Resource Get Mapped Pointer,$

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Image Processing

Supported OSes Linux, Windows, OS X

convolutionFFT2D - FFT-Based 2D Convolution

This sample demonstrates how 2D convolutions with very large kernel sizes can be efficiently implemented using FFT transformations.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUFFT

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cufftPlan2d, cufftExecR2C, cufftExecC2R, cufftDestroy

Key Concepts Image Processing, CUFFT Library

convolutionSeparable - CUDA Separable Convolution

This sample implements a separable convolution filter of a 2D signal with a gaussian kernel.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, Data Parallel Algorithms

Supported OSes Linux, Windows, OS X

Whitepaper convolutionSeparable.pdf

convolutionTexture - Texture-based Separable Convolution

Texture-based implementation of a separable 2D convolution with a gaussian kernel. Used for performance comparison against convolutionSeparable.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, Texture, Data Parallel Algorithms

Supported OSes Linux, Windows, OS X

cudaDecodeD3D9 - CUDA Video Decoder D3D9 API

This sample demonstrates how to efficiently use the CUDA Video Decoder API to decode MPEG-2, VC-1, or H.264 sources. YUV to RGB conversion of video is accomplished with CUDA kernel. The output result is rendered to a D3D9 surface. The decoded video is not displayed on the screen, but with -displayvideo at the command line parameter, the video output can be seen. Requires a Direct3D capable device and Compute Capability 2.0 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuDeviceGet, cuDeviceGetAttribute, cuDeviceComputeCapability,

cuDeviceGetCount, cuDeviceGetName, cuDeviceTotalMem, cuD3D9CtxCreate, cuD3D9GetDevice, cuModuleLoad, cuModuleUnload, cuModuleGetFunction,

cuModuleGetGlobal, cuModuleLoadDataEx, cuModuleGetTexRef,

cuD3D9MapResources, cuD3D9UnmapResources, cuD3D9RegisterResource,

cuD3D9UnregisterResource, cuD3D9ResourceSetMapFlags,

cuD3D9ResourceGetMappedPointer, cuD3D9ResourceGetMappedPitch, cuParamSetv, cuParamSeti, cuParamSetSize, cuLaunchGridAsync,

cuCtxCreate, cuMemAlloc, cuMemFree, cuMemAllocHost, cuMemFreeHost, cuMemcpyDtoHAsync, cuMemsetD8, cuStreamCreate, cuCtxPushCurrent,

 $\verb"cuCtxPopCurrent", \verb"cuvidCreateDecoder", \verb"cuvidDecodePicture", \\$

 $cuvid Map Video Frame, \ cuvid Unmap Video Frame, \ cuvid Destroy Decoder,$

cuvidCtxLockCreate, cuvidCtxLockDestroy, cuCtxDestroy

Key Concepts Graphics Interop, Image Processing, Video Compression

Supported OSes Windows

Whitepaper nvcuvid.pdf

cudaDecodeGL - CUDA Video Decoder GL API

This sample demonstrates how to efficiently use the CUDA Video Decoder API to decode video sources based on MPEG-2, VC-1, and H.264. YUV to RGB conversion of video is accomplished with CUDA kernel. The output result is rendered to a OpenGL surface. The decoded video is black, but can be enabled with -displayvideo added to the command line. Requires Compute Capability 2.0 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL, cuvid

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuDeviceGet, cuDeviceGetAttribute, cuDeviceComputeCapability,

cuDeviceGetCount, cuDeviceGetName, cuDeviceTotalMem,

cuGLCtxCreate, cuGLGetDevice, cuModuleLoad, cuModuleUnload, cuModuleGetFunction, cuModuleGetGlobal, cuModuleLoadDataEx, cuModuleGetTexRef, cuGLMapResources, cuGLUnmapResources,

cuGLRegisterResource, cuGLUnregisterResource, cuGLResourceSetMapFlags,

cuGLResourceGetMappedPointer, cuGLResourceGetMappedPitch, cuParamSetv, cuParamSeti, cuParamSetSize, cuLaunchGridAsync,

cuCtxCreate, cuMemAlloc, cuMemFree, cuMemAllocHost, cuMemFreeHost,

cuMemcpyDtoHAsync, cuMemsetD8, cuStreamCreate, cuCtxPushCurrent,

cuCtxPopCurrent, cuvidCreateDecoder, cuvidDecodePicture,

cuvidMapVideoFrame, cuvidUnmapVideoFrame, cuvidDestroyDecoder,

cuvidCtxLockCreate, cuvidCtxLockDestroy, cuCtxDestroy

Key Concepts Graphics Interop, Image Processing, Video Compression

Supported OSes Linux, Windows

Whitepaper nvcuvid.pdf

dct8x8 - DCT8x8

This sample demonstrates how Discrete Cosine Transform (DCT) for blocks of 8 by 8 pixels can be performed using CUDA: a naive implementation by definition and a more traditional approach used in many libraries. As opposed to implementing DCT in a fragment shader, CUDA allows for an easier and more efficient implementation.

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

Key Concepts Image Processing, Video Compression

Supported OSes Linux, Windows, OS X

Whitepaper dct8x8.pdf

dwtHaar1D - 1D Discrete Haar Wavelet Decomposition

Discrete Haar wavelet decomposition for 1D signals with a length which is a power of 2.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, Video Compression

Supported OSes Linux, Windows, OS X

dxtc - DirectX Texture Compressor (DXTC)

High Quality DXT Compression using CUDA. This example shows how to implement an existing computationally-intensive CPU compression algorithm in parallel on the GPU, and obtain an order of magnitude performance improvement.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, Image Compression

Supported OSes Linux, Windows, OS X

Whitepaper cuda_dxtc.pdf

CUDA_EGLStreams_Interop - EGLStreams CUDA Interop

Demonstrates data exchange between CUDA and EGL Streams.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies EGL

Sw 2.0, Sm 3.0, Sm 3.2, Sm 3.5, Sm 3.7, Sm 5.0, Sm 5.2, Sm 5.3, Sm 6.0, Sm 6.1

Architecture

CUDA API cuDeviceGet, cuDeviceGetAttribute,

 $cuDevice Compute Capability, \ cuDevice Get Count,\\$

cuDeviceGetName, cuGraphicsResourceGetMappedEglFrame,

 ${\it cuEGLStreamConsumerAcquireFrame, cuEGLStreamConsumerReleaseFrame, cuEGLStreamProducerPresentFrame, cuCtxCreate, cuMemAlloc, cuMemFree, cuCtxCreate, cuMemAlloc, cuMemAlloc, cuMemFree, cuCtxCreate, cuMemAlloc, cuMemAllo$

cuMemcpy3D, cuStreamCreate, cuCtxPushCurrent, cuCtxPopCurrent,

cuCtxDestroy

Key Concepts EGLStreams Interop

Supported OSes Linux

histogram - CUDA Histogram

This sample demonstrates efficient implementation of 64-bin and 256-bin histogram.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, Data Parallel Algorithms

Supported OSes Linux, Windows, OS X

Whitepaper histogram.pdf

HSOpticalFlow - Optical Flow

Variational optical flow estimation example. Uses textures for image operations. Shows how simple PDE solver can be accelerated with CUDA.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, Data Parallel Algorithms

Supported OSes Linux, Windows, OS X

Whitepaper OpticalFlow.pdf

imageDenoising - Image denoising

This sample demonstrates two adaptive image denoising techniques: KNN and NLM, based on computation of both geometric and color distance between texels. While both techniques are implemented in the DirectX SDK using shaders, massively speeded up variation of the latter technique, taking advantage of shared memory, is implemented in addition to DirectX counterparts.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing

Supported OSes Linux, Windows, OS X

Whitepaper imageDenoising.pdf

postProcessGL - Post-Process in OpenGL

This sample shows how to post-process an image rendered in OpenGL using CUDA.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cuda Graphics Unregister Resource

Key Concepts Graphics Interop, Image Processing

Supported OSes Linux, Windows, OS X

recursiveGaussian - Recursive Gaussian Filter

This sample implements a Gaussian blur using Deriche's recursive method. The advantage of this method is that the execution time is independent of the filter width.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Image Processing

Supported OSes Linux, Windows, OS X

simpleCUDA2GL - CUDA and OpenGL Interop of Images

This sample shows how to copy CUDA image back to OpenGL using the most efficient methods.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

 $cuda Graphics Register Resource, \ cuda Graphics GLR egister Buffer,$

cuda Graphics Unregister Resource

Key Concepts Graphics Interop, Image Processing, Performance Strategies

Supported OSes Linux, Windows, OS X

SobelFilter - Sobel Filter

This sample implements the Sobel edge detection filter for 8-bit monochrome images.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Image Processing

Supported OSes Linux, Windows, OS X

stereoDisparity - Stereo Disparity Computation (SAD SIMD Intrinsics)

A CUDA program that demonstrates how to compute a stereo disparity map using SIMD SAD (Sum of Absolute Difference) intrinsics. Requires Compute Capability 2.0 or higher.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, Video Intrinsics

3.5. Finance Reference

binomial Options - Binomial Option Pricing

This sample evaluates fair call price for a given set of European options under binomial model.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts
Computational Finance
Supported OSes
Linux, Windows, OS X
Whitepaper
binomialOptions.pdf

binomialOptions_nvrtc - Binomial Option Pricing with libNVRTC

This sample evaluates fair call price for a given set of European options under binomial model. This sample makes use of NVRTC for Runtime Compilation.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Computational Finance, Runtime Compilation

Supported OSes Linux, Windows, OS X

BlackScholes - Black-Scholes Option Pricing

This sample evaluates fair call and put prices for a given set of European options by Black-Scholes formula.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Computational Finance

Supported OSes Linux, Windows, OS X

Whitepaper BlackScholes.pdf

BlackScholes_nvrtc - Black-Scholes Option Pricing with libNVRTC

This sample evaluates fair call and put prices for a given set of European options by Black-Scholes formula, compiling the CUDA kernels involved at runtime using NVRTC.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Computational Finance, Runtime Compilation

Supported OSes Linux, Windows, OS X

MonteCarloMultiGPU - Monte Carlo Option Pricing with Multi-GPU support

This sample evaluates fair call price for a given set of European options using the Monte Carlo approach, taking advantage of all CUDA-capable GPUs installed in the system. This sample use double precision hardware if a GTX 200 class GPU is present. The sample also takes advantage of CUDA 4.0 capability to supporting using a single CPU thread to control multiple GPUs

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CURAND

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Supported OSes Linux, Windows, OS X

Whitepaper MonteCarlo.pdf

quasirandomGenerator - Niederreiter Quasirandom Sequence Generator

This sample implements Niederreiter Quasirandom Sequence Generator and Inverse Cumulative Normal Distribution functions for the generation of Standard Normal Distributions.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Computational Finance

Supported OSes Linux, Windows, OS X

quasirandomGenerator_nvrtc - Niederreiter Quasirandom Sequence Generator with libNVRTC

This sample implements Niederreiter Quasirandom Sequence Generator and Inverse Cumulative Normal Distribution functions for the generation of Standard Normal Distributions, compiling the CUDA kernels involved at runtime using NVRTC.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVRTC

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Computational Finance, Runtime Compilation

Supported OSes Linux, Windows, OS X

SobolQRNG - Sobol Quasirandom Number Generator

This sample implements Sobol Quasirandom Sequence Generator.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Computational Finance
Supported OSes Linux, Windows, OS X

3.6. Simulations Reference

fluidsD3D9 - Fluids (Direct3D Version)

An example of fluid simulation using CUDA and CUFFT, with Direct3D 9 rendering. A Direct3D Capable device is required.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies DirectX

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaD3D9SetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cuda Graphics Unregister Resource

Key Concepts Graphics Interop, CUFFT Library, Physically-Based Simulation

Supported OSes Windows

fluidsGL - Fluids (OpenGL Version)

An example of fluid simulation using CUDA and CUFFT, with OpenGL rendering.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL, CUFFT

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

 $cuda Graphics Unmap Resources, \ cuda Graphics Resource Get Mapped Pointer,$

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, CUFFT Library, Physically-Based Simulation

Whitepaper fluidsGL.pdf

fluidsGLES - Fluids (OpenGLES Version)

An example of fluid simulation using CUDA and CUFFT, with OpenGLES rendering.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GLES, CUFFT

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGraphicsMapResources, cudaGraphicsUnmapResources,

 $cuda Graphics Resource Get Mapped Pointer, \ cuda Graphics Register Resource,$

 $cuda Graphics GLR egister Buffer,\ cuda Graphics Unregister Resource$

Key Concepts Graphics Interop, CUFFT Library, Physically-Based Simulation

Supported OSes Linux

nbody - CUDA N-Body Simulation

This sample demonstrates efficient all-pairs simulation of a gravitational n-body simulation in CUDA. This sample accompanies the GPU Gems 3 chapter "Fast N-Body Simulation with CUDA". With CUDA 5.5, performance on Tesla K20c has increased to over 1.8TFLOP/s single precision. Double Performance has also improved on all Kepler and Fermi GPU architectures as well. Starting in CUDA 4.0, the nBody sample has been updated to take advantage of new features to easily scale the n-body simulation across multiple GPUs in a single PC. Adding "-numbodies=
bodies>" to the command line will allow users to set # of bodies for simulation. Adding "-numdevices=<N>" to the command line option will cause the sample to use N devices (if available) for simulation. In this mode, the position and velocity data for all bodies are read from system memory using "zero copy" rather than from device memory. For a small number of devices (4 or fewer) and a large enough number of bodies, bandwidth is not a bottleneck so we can achieve strong scaling across these devices.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

 $cuda Graphics Unmap Resources, \ cuda Graphics Resource Get Mapped Pointer,$

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Data Parallel Algorithms, Physically-Based Simulation

Supported OSes Linux, Windows, OS X

Whitepaper nbody_gems3_ch31.pdf

nbody_opengles - CUDA N-Body Simulation with GLES

This sample demonstrates efficient all-pairs simulation of a gravitational n-body simulation in CUDA. Unlike the OpenGL nbody sample, there is no user interaction.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GLES

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cuda Graphics Unregister Resource

Key Concepts Graphics Interop, Data Parallel Algorithms, Physically-Based Simulation

Supported OSes Linux

nbody_screen - CUDA N-Body Simulation on Screen

This sample demonstrates efficient all-pairs simulation of a gravitational n-body simulation in CUDA. Unlike the OpenGL nbody sample, there is no user interaction.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies screen, GLES

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

 $cuda Graphics Unmap Resources, \ cuda Graphics Resource Get Mapped Pointer,$

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Data Parallel Algorithms, Physically-Based Simulation

Supported OSes Linux

oceanFFT - CUDA FFT Ocean Simulation

This sample simulates an Ocean height field using CUFFT Library and renders the result using OpenGL.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL, CUFFT

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

 $cuda Graphics Unmap Resources, \ cuda Graphics Resource Get Mapped Pointer,$

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource, cufftPlan2d, cufftExecR2C, cufftExecC2R,

cufftDestroy

Key Concepts Graphics Interop, Image Processing, CUFFT Library

Supported OSes Linux, Windows, OS X

particles - Particles

This sample uses CUDA to simulate and visualize a large set of particles and their physical interaction. Adding "-particles=<N>" to the command line will allow users to set # of particles for simulation. This example implements a uniform grid data structure using either atomic operations or a fast radix sort from the Thrust library

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cuda Graphics Unregister Resource

Key Concepts Graphics Interop, Data Parallel Algorithms, Physically-Based Simulation,

Performance Strategies

Supported OSes Linux, Windows, OS X

Whitepaper particles.pdf

smokeParticles - Smoke Particles

Smoke simulation with volumetric shadows using half-angle slicing technique. Uses CUDA for procedural simulation, Thrust Library for sorting algorithms, and OpenGL for graphics rendering.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cudaGLSetGLDevice, cudaGraphicsMapResources,

cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer,

cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer,

cudaGraphicsUnregisterResource

Key Concepts Graphics Interop, Data Parallel Algorithms, Physically-Based Simulation

Supported OSes Linux, Windows, OS X

Whitepaper smokeParticles.pdf

VFlockingD3D10

The sample models formation of V-shaped flocks by big birds, such as geese and cranes. The algorithms of such flocking are borrowed from the paper "V-like formations in flocks of artificial birds" from Artificial Life, Vol. 14, No. 2, 2008. The sample has CPU-

and GPU-based implementations. Press 'g' to toggle between them. The GPU-based simulation works many times faster than the CPU-based one. The printout in the console window reports the simulation time per step. Press 'r' to reset the initial distribution of birds.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies	DirectX
Supported SM Architecture	SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1
CUDA API	cudaD3D10SetGLDevice, cudaGraphicsMapResources, cudaGraphicsUnmapResources, cudaGraphicsResourceGetMappedPointer, cudaGraphicsRegisterResource, cudaGraphicsGLRegisterBuffer, cudaGraphicsUnregisterResource
Key Concepts	Graphics Interop, Data Parallel Algorithms, Physically-Based Simulation, Performance Strategies
Supported OSes	Windows

3.7. Advanced Reference

alignedTypes - Aligned Types

A simple test, showing huge access speed gap between aligned and misaligned structures.

Supported SM Architecture	SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1
Key Concepts	Performance Strategies
Supported OSes	Linux, Windows, OS X

c++11_cuda - C++11 CUDA

This sample demonstrates C++11 feature support in CUDA. It scans a input text file and prints no. of occurrences of x, y, z, w characters.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CPP11

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CPP11 CUDA

Supported OSes Linux, OS X

cdpAdvancedQuicksort - Advanced Quicksort (CUDA Dynamic Parallelism)

This sample demonstrates an advanced quicksort implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CDP

Supported SM SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUDA Dynamic Parallelism

Supported OSes Linux, Windows, OS X

cdpBezierTessellation - Bezier Line Tessellation (CUDA Dynamic Parallelism)

This sample demonstrates bezier tessellation of lines implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CDP

Supported SM SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUDA Dynamic Parallelism

Supported OSes Linux, Windows, OS X

cdpLUDecomposition - LU Decomposition (CUDA Dynamic Parallelism)

This sample demonstrates LU Decomposition implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CDP, CUBLAS

Supported SM SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUDA Dynamic Parallelism

Supported OSes Linux, Windows, OS X

cdpQuadtree - Quad Tree (CUDA Dynamic Parallelism)

This sample demonstrates Quad Trees implemented using CUDA Dynamic Parallelism. This sample requires devices with compute capability 3.5 or higher.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CDP

Supported SM SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUDA Dynamic Parallelism

Supported OSes Linux, Windows, OS X

concurrent Kernels - Concurrent Kernels

This sample demonstrates the use of CUDA streams for concurrent execution of several kernels on devices of compute capability 2.0 or higher. Devices of compute capability

1.x will run the kernels sequentially. It also illustrates how to introduce dependencies between CUDA streams with the new cudaStreamWaitEvent function introduced in CUDA 3.2

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

Key Concepts Performance Strategies

Supported OSes Linux, Windows, OS X

eigenvalues - Eigenvalues

The computation of all or a subset of all eigenvalues is an important problem in Linear Algebra, statistics, physics, and many other fields. This sample demonstrates a parallel implementation of a bisection algorithm for the computation of all eigenvalues of a tridiagonal symmetric matrix of arbitrary size with CUDA.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra

Supported OSes Linux, Windows, OS X

Whitepaper eigenvalues.pdf

fastWalshTransform - Fast Walsh Transform

Naturally(Hadamard)-ordered Fast Walsh Transform for batching vectors of arbitrary eligible lengths that are power of two in size.

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

Key Concepts Linear Algebra, Data-Parallel Algorithms, Video Compression

Supported OSes Linux, Windows, OS X

FDTD3d - CUDA C 3D FDTD

This sample applies a finite differences time domain progression stencil on a 3D surface.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Performance Strategies

FunctionPointers - Function Pointers

This sample illustrates how to use function pointers and implements the Sobel Edge Detection filter for 8-bit monochrome images.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Graphics Interop, Image Processing

Supported OSes Linux, Windows, OS X

interval - Interval Computing

Interval arithmetic operators example. Uses various C++ features (templates and recursion). The recursive mode requires Compute SM 2.0 capabilities.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Recursion, Templates

Supported OSes Linux, Windows, OS X

lineOfSight - Line of Sight

This sample is an implementation of a simple line-of-sight algorithm: Given a height map and a ray originating at some observation point, it computes all the points along the ray that are visible from the observation point. The implementation is based on the Thrust library (http://code.google.com/p/thrust/).

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

matrixMulDynlinkJIT - Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version)

This sample revisits matrix multiplication using the CUDA driver API. It demonstrates how to link to CUDA driver at runtime and how to use JIT (just-in-time) compilation from PTX code. It has been written for clarity of exposition to illustrate various CUDA programming principles, not with the goal of providing the most performant generic kernel for matrix multiplication. CUBLAS provides high-performance matrix multiplication.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuModuleLoad, cuModuleLoadDataEx, cuModuleGetFunction, cuMemAlloc,

cuMemFree, cuMemcpyHtoD, cuMemcpyDtoH, cuLaunchKernel

Key Concepts CUDA Driver API, CUDA Dynamically Linked Library

Supported OSes Linux, Windows, OS X

mergeSort - Merge Sort

This sample implements a merge sort (also known as Batcher's sort), algorithms belonging to the class of sorting networks. While generally subefficient on large sequences compared to algorithms with better asymptotic algorithmic complexity (i.e. merge sort or radix sort), may be the algorithms of choice for sorting batches of short-to mid-sized (key, value) array pairs. Refer to the excellent tutorial by H. W. Lang http://www.iti.fh-flensburg.de/lang/algorithmen/sortieren/networks/indexen.htm

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Data-Parallel Algorithms

Supported OSes Linux, Windows, OS X

newdelete - NewDelete

This sample demonstrates dynamic global memory allocation through device C++ new and delete operators and virtual function declarations available with CUDA 4.0.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

ptxjit - PTX Just-in-Time compilation

This sample uses the Driver API to just-in-time compile (JIT) a Kernel from PTX code. Additionally, this sample demonstrates the seamless interoperability capability of the CUDA Runtime and CUDA Driver API calls. For CUDA 5.5, this sample shows how to use cuLink* functions to link PTX assembly using the CUDA driver at runtime.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUDA Driver API

Supported OSes Linux, Windows, OS X

radixSortThrust - CUDA Radix Sort (Thrust Library)

This sample demonstrates a very fast and efficient parallel radix sort uses Thrust library (http://code.google.com/p/thrust/). The included RadixSort class can sort either key-value pairs (with float or unsigned integer keys) or keys only. The optimized code in this sample (and also in reduction and scan) uses a technique known as warp-synchronous programming, which relies on the fact that within a warp of threads running on a CUDA GPU, all threads execute instructions synchronously. The code uses this to avoid __syncthreads() when threads within a warp are sharing data via __shared__ memory. It is important to note that for this to work correctly without race conditions on all GPUs, the shared memory used in these warp-synchronous expressions must be declared volatile. If it is not declared volatile, then in the absence of __syncthreads(), the compiler is free to delay stores to __shared__ memory and keep the data in registers (an optimization technique), which will result in incorrect execution. So please heed the use of volatile in these samples and use it in the same way in any code you derive from them.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Data-Parallel Algorithms, Performance Strategies

Supported OSes Linux, Windows, OS X

Whitepaper readme.txt

reduction - CUDA Parallel Reduction

A parallel sum reduction that computes the sum of a large arrays of values. This sample demonstrates several important optimization strategies for 1:Data-Parallel Algorithms like reduction.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Data-Parallel Algorithms, Performance Strategies

Supported OSes Linux, Windows, OS X

scalarProd - Scalar Product

This sample calculates scalar products of a given set of input vector pairs.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra

Supported OSes Linux, Windows, OS X

scan - CUDA Parallel Prefix Sum (Scan)

This example demonstrates an efficient CUDA implementation of parallel prefix sum, also known as "scan". Given an array of numbers, scan computes a new array in which each element is the sum of all the elements before it in the input array.

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Data-Parallel Algorithms, Performance Strategies

Supported OSes Linux, Windows, OS X

segmentationTreeThrust - CUDA Segmentation Tree Thrust Library

This sample demonstrates an approach to the image segmentation trees construction. This method is based on Boruvka's MST algorithm.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Data-Parallel Algorithms, Performance Strategies

Supported OSes Linux, Windows, OS X

shfl_scan - CUDA Parallel Prefix Sum with Shuffle Intrinsics (SHFL_Scan)

This example demonstrates how to use the shuffle intrinsic __shfl_up to perform a scan operation across a thread block. A GPU with Compute Capability SM 3.0. is required to run the sample

Sw 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Data-Parallel Algorithms, Performance Strategies

Supported OSes Linux, Windows, OS X

simpleHyperQ

This sample demonstrates the use of CUDA streams for concurrent execution of several kernels on devices which provide HyperQ (SM 3.5). Devices without HyperQ (SM 2.0 and SM 3.0) will run a maximum of two kernels concurrently.

Sw 2.0, Sw 3.0, Sw 3.2, Sw 3.5, Sw 3.7, Sw 5.0, Sw 5.2, Sw 5.3, Sw 6.0, Sw 6.1

Architecture

Key Concepts CUDA Systems Integration, Performance Strategies

Supported OSes Linux, Windows, OS X

Whitepaper HyperQ.pdf

sortingNetworks - CUDA Sorting Networks

This sample implements bitonic sort and odd-even merge sort (also known as Batcher's sort), algorithms belonging to the class of sorting networks. While generally subefficient, for large sequences compared to algorithms with better asymptotic algorithmic complexity (i.e. merge sort or radix sort), this may be the preferred algorithms of choice for sorting batches of short-sized to mid-sized (key, value) array pairs. Refer to an excellent tutorial by H. W. Lang http://www.iti.fh-flensburg.de/lang/algorithmen/sortieren/networks/indexen.htm

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Data-Parallel Algorithms

Supported OSes Linux, Windows, OS X

StreamPriorities - Stream Priorities

This sample demonstrates basic use of stream priorities.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies Stream-Priorities

Supported SM SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUDA Streams and Events

Supported OSes Linux, OS X

threadFenceReduction

This sample shows how to perform a reduction operation on an array of values using the thread Fence intrinsic to produce a single value in a single kernel (as opposed to two or more kernel calls as shown in the "reduction" CUDA Sample). Single-pass reduction requires global atomic instructions (Compute Capability 2.0 or later) and the _threadfence() intrinsic (CUDA 2.2 or later).

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Data-Parallel Algorithms, Performance Strategies

Supported OSes Linux, Windows, OS X

threadMigration - CUDA Context Thread Management

Simple program illustrating how to the CUDA Context Management API and uses the new CUDA 4.0 parameter passing and CUDA launch API. CUDA contexts can be created separately and attached independently to different threads.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cuCtxCreate, cuCtxDestroy, cuModuleLoad, cuModuleLoadDataEx,

cuModuleGetFunction, cuLaunchKernel, cuMemcpyDtoH, cuCtxPushCurrent,

cuCtxPopCurrent

Key Concepts CUDA Driver API

Supported OSes Linux, Windows, OS X

transpose - Matrix Transpose

This sample demonstrates Matrix Transpose. Different performance are shown to achieve high performance.

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Performance Strategies, Linear Algebra

Supported OSes Linux, Windows, OS X

Whitepaper MatrixTranspose.pdf

3.8. Cudalibraries Reference

batchCUBLAS

A CUDA Sample that demonstrates how using batched CUBLAS API calls to improve overall performance.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUBLAS

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUBLAS Library

Supported OSes Linux, Windows, OS X

BiCGStab

A CUDA Sample that demonstrates Bi-Conjugate Gradient Stabilized (BiCGStab) iterative method for nonsymmetric and symmetric positive definite (s.p.d.) linear systems using CUSPARSE and CUBLAS.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUSPARSE, CUBLAS

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUBLAS Library, CUSPARSE Library

Supported OSes Linux, Windows, OS X

boxFilterNPP - Box Filter with NPP

A NPP CUDA Sample that demonstrates how to use NPP FilterBox function to perform a Box Filter.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies FreeImage, NPP

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Performance Strategies, Image Processing, NPP Library

Supported OSes Linux, Windows, OS X

cannyEdgeDetectorNPP - Canny Edge Detector NPP

An NPP CUDA Sample that demonstrates the recommended parameters to use with the nppiFilterCannyBorder_8u_C1R Canny Edge Detection image filter function. This function expects a single channel 8-bit grayscale input image. You can generate a grayscale image from a color image by first calling nppiColorToGray() or nppiRGBToGray(). The Canny Edge Detection function combines and improves on the techniques required to produce an edge detection image using multiple steps.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies Freelmage, NPP

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Performance Strategies, Image Processing, NPP Library

Supported OSes Linux, Windows, OS X

conjugateGradient - ConjugateGradient

This sample implements a conjugate gradient solver on GPU using CUBLAS and CUSPARSE library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUBLAS, CUSPARSE

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUBLAS Library, CUSPARSE Library

Supported OSes Linux, Windows, OS X

conjugateGradientPrecond - Preconditioned ConjugateGradient

This sample implements a preconditioned conjugate gradient solver on GPU using CUBLAS and CUSPARSE library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUBLAS, CUSPARSE

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUBLAS Library, CUSPARSE Library

Supported OSes Linux, Windows, OS X

conjugateGradientUM - ConjugateGradientUM

This sample implements a conjugate gradient solver on GPU using CUBLAS and CUSPARSE library, using Unified Memory

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies UVM, CUBLAS, CUSPARSE

Supported SM SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Unified Memory, Linear Algebra, CUBLAS Library, CUSPARSE Library

Supported OSes Linux, Windows, OS X

cuHook - CUDA Interception Library

This sample demonstrates how to build and use an intercept library with CUDA. The library has to be loaded via LD_PRELOAD, e.g. LD_PRELOAD=<full_path>/ libcuhook.so.1 ./cuHook

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Supported OSes Linux

cuSolverDn LinearSolver - cuSolverDn Linear Solver

A CUDA Sample that demonstrates cuSolverDN's LU, QR and Cholesky factorization.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUSOLVER, CUBLAS, CUSPARSE

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUSOLVER Library

Supported OSes Linux, Windows, OS X

cuSolverRf - cuSolverRf Refactorization

A CUDA Sample that demonstrates cuSolver's refactorization library - CUSOLVERRF.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUSOLVER, CUBLAS, CUSPARSE

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUSOLVER Library

Supported OSes Linux, Windows, OS X

cuSolverSp_LinearSolver - cuSolverSp Linear Solver

A CUDA Sample that demonstrates cuSolverSP's LU, QR and Cholesky factorization.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUSOLVER, CUSPARSE

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUSOLVER Library

Supported OSes Linux, Windows, OS X

cuSolverSp_LowlevelCholesky - cuSolverSp LowlevelCholesky Solver

A CUDA Sample that demonstrates Cholesky factorization using cuSolverSP's low level APIs.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUSOLVER, CUSPARSE

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUSOLVER Library

Supported OSes Linux, Windows, OS X

cuSolverSp_LowlevelQR - cuSolverSp Lowlevel QR Solver

A CUDA Sample that demonstrates QR factorization using cuSolverSP's low level APIs.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUSOLVER, CUSPARSE

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Linear Algebra, CUSOLVER Library

Supported OSes Linux, Windows, OS X

FilterBorderControlNPP - Filter Border Control NPP

This NPP CUDA Sample demonstrates how any border version of an NPP filtering function can be used in the most common mode (with border control enabled), can be used to duplicate the results of the equivalent non-border version of the NPP function, and can be used to enable and disable border control on various source image edges depending on what portion of the source image is being used as input.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies FreeImage, NPP

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Performance Strategies, Image Processing, NPP Library

Supported OSes Linux, Windows, OS X

freeImageInteropNPP - FreeImage and NPP Interopability

A simple CUDA Sample demonstrate how to use FreeImage library with NPP.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies FreeImage, NPP

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Performance Strategies, Image Processing, NPP Library

Supported OSes Linux, Windows, OS X

histEqualizationNPP - Histogram Equalization with NPP

This CUDA Sample demonstrates how to use NPP for histogram equalization for image data.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies Freelmage, NPP

Sw 2.0, Sm 3.0, Sm 3.2, Sm 3.5, Sm 3.7, Sm 5.0, Sm 5.2, Sm 5.3, Sm 6.0, Sm 6.1

Architecture

Key Concepts Image Processing, Performance Strategies, NPP Library

Supported OSes Linux, Windows, OS X

jpegNPP - JPEG encode/decode and resize with NPP

This sample demonstrates a simple image processing pipline. First, a JPEG file is huffman decoded and inverse DCT transformed and dequantized. Then the different plances are resized. Finally, the resized image is quantized, forward DCT transformed and huffman encoded.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies Freelmage, NPP

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API nppGetGpuComputeCapability, nppiDCTInitAlloc,

nppiDecodeHuffmanScanHost_JPEG_8u16s_P3R,

nppiDCTQuantInv8x8LS_JPEG_16s8u_C1R_NEW, nppiResizeSqrPixel_8u_C1R,

nppiEncodeHuffmanGetSize, nppiDCTFree

Supported OSes Linux, Windows, OS X

MC_EstimatePiInlineP - Monte Carlo Estimation of Pi (inline PRNG)

This sample uses Monte Carlo simulation for Estimation of Pi (using inline PRNG). This sample also uses the NVIDIA CURAND library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

CURAND

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Random Number Generator, Computational Finance, CURAND Library

Supported OSes Linux, Windows, OS X

MC_EstimatePiInlineQ - Monte Carlo Estimation of Pi (inline QRNG)

This sample uses Monte Carlo simulation for Estimation of Pi (using inline QRNG). This sample also uses the NVIDIA CURAND library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies	CURAND
Supported SM Architecture	SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1
Key Concepts	Random Number Generator, Computational Finance, CURAND Library
Supported OSes	Linux, Windows, OS X

MC_EstimatePiP - Monte Carlo Estimation of Pi (batch PRNG)

This sample uses Monte Carlo simulation for Estimation of Pi (using batch PRNG). This sample also uses the NVIDIA CURAND library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies	CURAND
Supported SM Architecture	SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1
Key Concepts	Random Number Generator, Computational Finance, CURAND Library
Supported OSes	Linux, Windows, OS X

MC_EstimatePiQ - Monte Carlo Estimation of Pi (batch QRNG)

This sample uses Monte Carlo simulation for Estimation of Pi (using batch QRNG). This sample also uses the NVIDIA CURAND library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies	CURAND
Supported SM Architecture	SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1
Key Concepts	Random Number Generator, Computational Finance, CURAND Library
Supported OSes	Linux, Windows, OS X

MC_SingleAsianOptionP - Monte Carlo Single Asian Option

This sample uses Monte Carlo to simulate Single Asian Options using the NVIDIA CURAND library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CURAND

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Random Number Generator, Computational Finance, CURAND Library

Supported OSes Linux, Windows, OS X

MersenneTwisterGP11213

This sample demonstrates the Mersenne Twister random number generator GP11213 in cuRAND.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CURAND

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Computational Finance, CURAND Library

Supported OSes Linux, Windows, OS X

nvgraph_Pagerank - NVGRAPH Page Rank

A CUDA Sample that demonstrates Page Rank computation using NVGRAPH Library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVGRAPH

Sw 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Graph Analytics, NVGRAPH Library

Supported OSes Linux, Windows, OS X

nvgraph_SemiRingSpmv - NVGRAPH Semi-Ring SpMV

A CUDA Sample that demonstrates Semi-Ring SpMV using NVGRAPH Library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVGRAPH

Sw 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Graph Analytics, NVGRAPH Library

Supported OSes Linux, Windows, OS X

nvgraph_SSSP - NVGRAPH Single Source Shortest Path

A CUDA Sample that demonstrates Single Source Shortest Path(SSSP) computation using NVGRAPH Library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies NVGRAPH

Supported SM SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Graph Analytics, NVGRAPH Library

Supported OSes Linux, Windows, OS X

randomFog - Random Fog

This sample illustrates pseudo- and quasi- random numbers produced by CURAND.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies X11, GL, CURAND

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts 3D Graphics, CURAND Library

Supported OSes Linux, Windows, OS X

simpleCUBLAS - Simple CUBLAS

Example of using CUBLAS using the new CUBLAS API interface available in CUDA 4.0.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUBLAS

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, CUBLAS Library

Supported OSes Linux, Windows, OS X

simpleCUBLASXT - Simple CUBLAS XT

Example of using CUBLAS-XT library.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUBLAS

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts CUBLAS-XT Library

Supported OSes Linux, Windows, OS X

simpleCUFFT - Simple CUFFT

Example of using CUFFT. In this example, CUFFT is used to compute the 1D-convolution of some signal with some filter by transforming both into frequency

domain, multiplying them together, and transforming the signal back to time domain. cuFFT plans are created using simple and advanced API functions.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUFFT

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, CUFFT Library

Supported OSes Linux, Windows, OS X

simpleCUFFT_2d_MGPU - SimpleCUFFT_2d_MGPU

Example of using CUFFT. In this example, CUFFT is used to compute the 2D-convolution of some signal with some filter by transforming both into frequency domain, multiplying them together, and transforming the signal back to time domain on Multiple GPU.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUFFT

Sw 2.0, SM 3.0, SM 3.2, SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, CUFFT Library

Supported OSes Linux, Windows, OS X

simpleCUFFT_callback - Simple CUFFT Callbacks

Example of using CUFFT. In this example, CUFFT is used to compute the 1D-convolution of some signal with some filter by transforming both into frequency domain, multiplying them together, and transforming the signal back to time domain. The difference between this example and the Simple CUFFT example is that the multiplication step is done by the CUFFT kernel with a user-supplied CUFFT callback routine, rather than by a separate kernel call.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample

will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies callback, CUFFT

Supported SM SM 2.0, SM 3.0, SM 3.2, SM 3.5, SM 5.0, SM 5.3, SM 6.0, SM 6.1

Architecture

Key Concepts Image Processing, CUFFT Library

Supported OSes Linux

simpleCUFFT_MGPU - Simple CUFFT_MGPU

Example of using CUFFT. In this example, CUFFT is used to compute the 1D-convolution of some signal with some filter by transforming both into frequency domain, multiplying them together, and transforming the signal back to time domain on Multiple GPU.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CUFFT

Sw 2.0, Sm 3.0, Sm 3.2, Sm 3.5, Sm 3.7, Sm 5.0, Sm 5.2, Sm 5.3, Sm 6.0, Sm 6.1

Architecture

Key Concepts Image Processing, CUFFT Library

Supported OSes Linux, Windows, OS X

simpleDevLibCUBLAS - simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism)

This sample implements a simple CUBLAS function calls that call GPU device API library running CUBLAS functions. This sample requires a SM 3.5 capable device.

This sample depends on other applications or libraries to be present on the system to either build or run. If these dependencies are not available on the system, the sample will not be installed. If these dependencies are available, but not installed, the sample will waive itself at build time.

Dependencies CDP, CUBLAS

Supported SM SM 3.5, SM 3.7, SM 5.0, SM 5.2, SM 5.3, SM 6.0, SM 6.1

Architecture

CUDA API cublasCreate, cublasSetVector, cublasSgemm, cudaMalloc, cudaFree,

cudaMemcpy

Key Concepts CUDA Dynamic Parallelism, Linear Algebra

Supported OSes Linux, Windows, OS X

Chapter 4. DEPENDENCIES

Some CUDA Samples rely on third-party applications and/or libraries, or features provided by the CUDA Toolkit and Driver, to either build or execute. These dependencies are listed below.

If a sample has a dependency that is not available on the system, the sample will not be installed. If a sample has a third-party dependency that is available on the system, but is not installed, the sample will waive itself at build time.

Each sample's dependencies are listed in the Samples Reference section.

Third-Party Dependencies

These third-party dependencies are required by some CUDA samples. If available, these dependencies are either installed on your system automatically, or are installable via your system's package manager (Linux) or a third-party website.

Freelmage

FreeImage is an open source imaging library. FreeImage can usually be installed on Linux using your distribution's package manager system. FreeImage can also be downloaded from the FreeImage website. FreeImage is also redistributed with the CUDA Samples.

Message Passing Interface

MPI (Message Passing Interface) is an API for communicating data between distributed processes. A MPI compiler can be installed using your Linux distribution's package manager system. It is also available on some online resources, such as Open MPI.

Only 64-Bit

Some samples can only be run on a 64-bit operating system.

DirectX

DirectX is a collection of APIs designed to allow development of multimedia applications on Microsoft platforms. For Microsoft platforms, NVIDIA's CUDA Driver supports DirectX. Several CUDA Samples for Windows demonstrates CUDA-DirectX Interoperability, for building such samples one needs to install Direct X SDK (June 2010 or newer), this is required to be installed on Windows 7, Windows 10 and Windows Server 2008, Other Windows OSes do not need to explicitly install the DirectX SDK.

OpenGL

OpenGL is a graphics library used for 2D and 3D rendering. On systems which support OpenGL, NVIDIA's OpenGL implementation is provided with the CUDA Driver.

OpenGL ES

OpenGL ES is an embedded systems graphics library used for 2D and 3D rendering. On systems which support OpenGL ES, NVIDIA's OpenGL ES implementation is provided with the CUDA Driver.

OpenMP

OpenMP is an API for multiprocessing programming. OpenMP can be installed using your Linux distribution's package manager system. It usually comes preinstalled with GCC. It can also be found at the OpenMP website.

Screen

Screen is a windowing system found on the QNX operating system. Screen is usually found as part of the root filesystem.

X11

X11 is a windowing system commonly found on *-nix style operating systems. X11 can be installed using your Linux distribution's package manager, and comes preinstalled on Mac OS X systems.

EGL

EGL is an interface between Khronos rendering APIs (such as OpenGL, OpenGL ES or OpenVG) and the underlying native platform windowing system.

EGLOutput

EGLOutput is a set of EGL extensions which allow EGL to render directly to the display.

CUDA Features

These CUDA features are needed by some CUDA samples. They are provided by either the CUDA Toolkit or CUDA Driver. Some features may not be available on your system.

CUFFT Callback Routines

CUFFT Callback Routines are user-supplied kernel routines that CUFFT will call when loading or storing data. These callback routines are only available on Linux x86_64 and ppc64le systems.

CUDA Dynamic Paralellism

CDP (CUDA Dynamic Paralellism) allows kernels to be launched from threads running on the GPU. CDP is only available on GPUs with SM architecture of 3.5 or above.

CUBLAS

CUBLAS (CUDA Basic Linear Algebra Subroutines) is a GPU-accelerated version of the BLAS library.

CUDA Interprocess Communication

IPC (Interprocess Communication) allows processes to share device pointers. IPC is only available on Linux x86_64 and ppc64le systems.

CUFFT

CUFFT (CUDA Fast Fourier Transform) is a GPU-accelerated FFT library.

CURAND

CURAND (CUDA Random Number Generation) is a GPU-accelerated RNG library.

CUSPARSE

CUSPARSE (CUDA Sparse Matrix) provides linear algebra subroutines used for sparse matrix calculations.

CUSOLVER

CUSOLVER library is a high-level package based on the CUBLAS and CUSPARSE libraries. It combines three separate libraries under a single umbrella, each of which can be used independently or in concert with other toolkit libraries. The intent of CUSOLVER is to provide useful LAPACK-like features, such as common matrix factorization and triangular solve routines for dense matrices, a sparse least-squares solver and an

eigenvalue solver. In addition cuSolver provides a new refactorization library useful for solving sequences of matrices with a shared sparsity pattern.

NPP

NPP (NVIDIA Performance Primitives) provides GPU-accelerated image, video, and signal processing functions.

NVGRAPH

NVGRAPH is a GPU-accelerated graph analytics library..

NVRTC

NVRTC (CUDA RunTime Compilation) is a runtime compilation library for CUDA C++.

NVCUVID

NVCUVID (NVIDIA CUDA Video Decoder) provides GPU-accelerated video decoding capabilities.

Stream Priorities

Stream Priorities allows the creation of streams with specified priorities. Stream Priorities is only available on GPUs with SM architecture of 3.5 or above.

Unified Virtual Memory

UVM (Unified Virtual Memory) enables memory that can be accessed by both the CPU and GPU without explicit copying between the two. UVM is only available on Linux and Windows systems.

16-bit Floating Point

FP16 is a 16-bit floating-point format. One bit is used for the sign, five bits for the exponent, and ten bits for the mantissa. FP16 is only available on specific mobile platforms.

C++11 CUDA

NVCC Support of C++11 features.

Chapter 5. KEY CONCEPTS AND ASSOCIATED SAMPLES

The tables below describe the key concepts of the CUDA Toolkit and lists the samples that illustrate how that concept is used.

Basic Key Concepts

Basic Concepts demonstrates how to make use of CUDA features.

Table 2 Basic Key Concepts and Associated Samples

Basic Key Concept	Description	Samples
3D Graphics	3D Rendering	Random Fog, Simple Direct3D10 (Vertex Array), Simple OpenGL, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen
3D Textures	Volume Textures	Simple Texture 3D
Assert	GPU Assert	simpleAssert, simpleAssert with libNVRTC
Asynchronous Data Transfers	Overlapping I/O and Compute	Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Multi Copy and Compute, Simple Multi-GPU, Simple Peer- to-Peer Transfers with Multi-GPU, asyncAPI, simpleStreams
Atomic Intrinsics	Using atomics with GPU kernels	Simple Atomic Intrinsics, Simple Atomic Intrinsics with libNVRTC, System wide Atomics
C++ Function Overloading	Use C++ overloading with GPU kernels	cppOverload

Basic Key		
Concept	Description	Samples
C++ Templates	Using Templates with GPU kernels	Simple Templates, Simple Templates with libNVRTC
CUBLAS	CUDA BLAS samples	Matrix Multiplication (CUBLAS), Unified Memory Streams
CUBLAS Library	CUDA BLAS samples	BiCGStab, Simple CUBLAS, batchCUBLAS
CUBLAS-XT Library		Simple CUBLAS XT
CUDA Driver API	Samples that show the CUDA Driver API	Device Query Driver API, Matrix Multiplication (CUDA Driver API Version), Simple Texture (Driver Version), Simple Vote Intrinsics with libNVRTC, Using Inline PTX, Using Inline PTX with libNVRTC, Vector Addition Driver API, Vector Addition with libNVRTC
CUDA Dynamic Parallelism	Dynamic Parallelism with GPU Kernels (SM 3.5)	Simple Print (CUDA Dynamic Parallelism), simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism)
CUDA Runtime API	Samples that use the Runtime API	Device Query, FP16 Scalar Product, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Matrix Multiplication with libNVRTC, Simple Texture, Vector Addition
CUDA Streams	Stream API definies a sequence of operations that can be overlapped with I/O	Simple CUDA Callbacks
CUDA Streams and Events	Synchronizing Kernels with Event Timers and Streams	Bandwidth Test, Simple Multi Copy and Compute, Simple Multi-GPU, Unified Memory Streams, asyncAPI, cppOverload, simpleStreams
CUDA Systems Integration	Samples that integrate with Multi Process (OpenMP, IPC, and MPI)	Unified Memory Streams, cudaOpenMP, simpleIPC, simpleMPI

Basic Key Concept	Description	Samples
CUFFT Library	Samples that use the CUDA FFT accelerated library	Simple CUFFT, Simple CUFFT Callbacks, Simple CUFFT_MGPU, SimpleCUFFT_2d_MGPU
CURAND Library	Samples that use the CUDA random number generator	MersenneTwisterGP11213, Random Fog
CUSOLVER Library	Samples that use the cuSOLVER accelerated library	cuSolverDn Linear Solver , cuSolverRf Refactorization, cuSolverSp Linear Solver , cuSolverSp Lowlevel QR Solver, cuSolverSp LowlevelCholesky Solver
CUSPARSE Library	Samples that use the cuSPARSE (Sparse Vector Matrix Multiply) functions	BiCGStab
Callback Functions	Creating Callback functions with GPU kernels	Simple CUDA Callbacks
Computationa Finance	Finance Algorithms	Black-Scholes Option Pricing, Black- Scholes Option Pricing with libNVRTC, MersenneTwisterGP11213
Data Parallel Algorithms	Samples that show good usage of Data Parallel Algorithms	CUDA Separable Convolution, Texture- based Separable Convolution
Debugging	Samples useful for debugging	simplePrintf
Device Memory Allocation	Samples that show GPU Device side memory allocation	Template
Device Query	Sample showing simple device query of information	Device Query, Device Query Driver API
EGLStreams Interop	Samples demonstrating how to use EGL Streams and CUDA Interop.	EGLStreams CUDA Interop
GPU Performance	Samples demonstrating high performance and data I/O	Simple Multi Copy and Compute
Graph Analytics	Samples demonstrating how to use graph analytics with CUDA	NVGRAPH Page Rank, NVGRAPH Semi-Ring SpMV , NVGRAPH Single Source Shortest Path
Graphics Interop	Samples that demonstrate interop between graphics APIs and CUDA	Bicubic B-spline Interoplation, Bilateral Filter, Box Filter, CUDA and OpenGL

Basic Key		
Concept	Description	Samples
		Interop of Images, Simple D3D10 Texture, Simple D3D11 Texture, Simple D3D9 Texture, Simple Direct3D10 (Vertex Array), Simple Direct3D10 Render Target, Simple Direct3D9 (Vertex Arrays), Simple OpenGL, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen, Simple Texture 3D
Image Processing	Samples that demonstrate image processing algorithms in CUDA	Bicubic B-spline Interoplation, Bilateral Filter, Box Filter, Box Filter with NPP, CUDA Separable Convolution, CUDA and OpenGL Interop of Images, Canny Edge Detector NPP, Filter Border Control NPP, Freelmage and NPP Interopability, Histogram Equalization with NPP, Pitch Linear Texture, Simple CUBLAS, Simple CUFFT, Simple CUFFT Callbacks, Simple CUFFT_MGPU, Simple D3D11 Texture, Simple Surface Write, Simple Texture, Simple Texture (Driver Version), Simple Texture 3D, SimpleCUFFT_2d_MGPU, Texture-based Separable Convolution
InterProcess Communication	Samples that demonstrate Inter Process nCommunication between processes	simpleIPC
Linear Algebra	Samples demonstrating linear algebra with CUDA	BiCGStab, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Matrix Multiplication with libNVRTC, batchCUBLAS, cuSolverDn Linear Solver, cuSolverRf Refactorization, cuSolverSp Linear Solver, cuSolverSp Lowlevel QR Solver, cuSolverSp LowlevelCholesky Solver, simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism)
MPI	Samples demonstrating how to use CUDA with MPI programs	simpleMPI

Basic Key Concept	Description	Samples
Matrix Multiply	Samples demonstrating matrix multiply CUDA	Matrix Multiplication (CUDA Driver API Version)
Multi-GPU	Samples demonstrating how to take advantage of multiple GPUs and CUDA	Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Multi-GPU, Simple Peer- to-Peer Transfers with Multi-GPU, Topology Query
Multithreading	Samples demonstrating how to use multithreading with CUDA	Simple CUDA Callbacks, Simple Multi-GPU, Unified Memory Streams, cudaOpenMP, simpleMPI
NPP Library	Samples demonstrating how to use NPP (NVIDIA Performance Primitives) for image processing	Box Filter with NPP, Canny Edge Detector NPP, Filter Border Control NPP, FreeImage and NPP Interopability, Histogram Equalization with NPP
NVGRAPH Library	nvGRAPH library	NVGRAPH Page Rank, NVGRAPH Semi-Ring SpMV , NVGRAPH Single Source Shortest Path
Occupancy Calculator	Samples demonstrating how to use the CUDA Occupancy Calculator	simpleOccupancy
OpenMP	Samples demonstrating how to use OpenMP	Unified Memory Streams, cudaOpenMP
Overlap Compute and Copy	Samples demonstrating how to overlap Compute and Data I/O	Simple Multi Copy and Compute
PTX Assembly	Samples demonstrating how to use PTX code with CUDA	Using Inline PTX, Using Inline PTX with libNVRTC
Peer to Peer	Samples demonstrating how to handle P2P data transfers between multiple GPUs	simpleIPC
Peer to Peer Data Transfers	Samples demonstrating how to handle P2P data transfers between multiple GPUs	Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Peer-to-Peer Transfers with Multi-GPU
Performance Strategies	Samples demonstrating high performance with CUDA	Bandwidth Test, Box Filter with NPP, CUDA and OpenGL Interop of Images, Canny Edge Detector NPP, Clock, Clock libNVRTC, Filter Border Control NPP, FreeImage and NPP Interopability, Histogram Equalization

Pasis Voy		
Basic Key Concept	Description	Samples
		with NPP, Matrix Multiplication (CUBLAS), Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Peer-to-Peer Transfers with Multi-GPU, Topology Query, Using Inline PTX, Using Inline PTX with libNVRTC, simpleZeroCopy
Pinned System Paged Memory	Samples demonstrating how to properly handle data I/O efficiently between the CPU host and GPU video memory	simpleZeroCopy
Separate Compilation	Samples demonstrating how to use CUDA library linking	Simple Static GPU Device Library
Surface Writes	Samples demonstrating how to use Surface Writes with GPU kernels	Simple Surface Write, Simple Texture 3D
Texture	Samples demonstrating how to use textures GPU kernels	Pitch Linear Texture, Simple Cubemap Texture, Simple D3D10 Texture, Simple D3D9 Texture, Simple Direct3D10 Render Target, Simple Layered Texture, Simple Surface Write, Simple Texture, Simple Texture (Driver Version), Texture-based Separable Convolution
Unified Memory	Samples demonstrating how to use Unified Memory	ConjugateGradientUM, System wide Atomics, Unified Memory Streams
Unified Virtual Address Space	Samples demonstrating how to use UVA with CUDA programs	Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Peer-to-Peer Transfers with Multi-GPU
Vector Addition	Samples demonstrating how to use Vector Addition with CUDA programs	Vector Addition, Vector Addition Driver API, Vector Addition with libNVRTC, simpleZeroCopy
Vertex Buffers	Samples demonstrating how to use Vertex Buffers with CUDA kernels	Simple OpenGL, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen
Volume Processing	Samples demonstrating how to use 3D Textures for volume rendering	Simple Cubemap Texture, Simple Layered Texture

Basic Key Concept	Description	Samples
Vote Intrinsics	Samples demonstrating how to use vote intrinsics with CUDA	Simple Vote Intrinsics, Simple Vote Intrinsics with libNVRTC

Advanced Key Concepts

Advanced Concepts demonstrate advanced techniques and algorithms implemented with CUDA.

Table 3 Advanced Key Concepts and Associated Samples

Advanced Key Concept	Description	Samples
2D Textures	Texture Mapping	SLI D3D10 Texture
3D Graphics	3D Rendering	Marching Cubes Isosurfaces
3D Textures	Volume Textures	Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
CPP11 CUDA	Samples demonstrating how to use C++11 feature support in CUDA.	C++11 CUDA
CUBLAS Library	CUDA BLAS samples	ConjugateGradient, ConjugateGradientUM, Preconditioned Conjugate Gradient
CUDA Driver API	Samples that show the CUDA Driver API	CUDA Context Thread Management, Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version), PTX Just-in- Time compilation
CUDA Dynamic Parallelism	Dynamic Parallelism with GPU Kernels (SM 3.5)	Advanced Quicksort (CUDA Dynamic Parallelism), Bezier Line Tessellation (CUDA Dynamic Parallelism), LU Decomposition (CUDA Dynamic Parallelism), Quad Tree (CUDA Dynamic Parallelism), Simple Quicksort (CUDA Dynamic Parallelism)
CUDA Dynamically Linked Library	Dynamic loading of the CUDA DLL using CUDA Driver API	Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version)

Advaraad		
Advanced Key Concept	Description	Samples
CUDA Streams and Events	Synchronizing Kernels with Event Timers and Streams	Stream Priorities
CUDA Systems Integration	Samples that integrate with Multi Process (OpenMP, IPC, and MPI)	simpleHyperQ
CUFFT Library	Samples that use the CUDA FFT accelerated library	CUDA FFT Ocean Simulation, FFT-Based 2D Convolution, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version)
CURAND Library	Samples that use the CUDA random number generator	Monte Carlo Estimation of Pi (batch PRNG), Monte Carlo Estimation of Pi (batch QRNG), Monte Carlo Estimation of Pi (inline PRNG), Monte Carlo Estimation of Pi (inline QRNG), Monte Carlo Single Asian Option
CUSPARSE Library	Samples that use the cuSPARSE (Sparse Vector Matrix Multiply) functions	ConjugateGradient, ConjugateGradientUM, Preconditioned Conjugate Gradient
Computationa Finance	. Finance Algorithms	Binomial Option Pricing, Binomial Option Pricing with libNVRTC, Monte Carlo Estimation of Pi (batch PRNG), Monte Carlo Estimation of Pi (batch QRNG), Monte Carlo Estimation of Pi (inline PRNG), Monte Carlo Estimation of Pi (inline QRNG), Monte Carlo Single Asian Option, Niederreiter Quasirandom Sequence Generator, Niederreiter Quasirandom Sequence Generator with libNVRTC, Sobol Quasirandom Number Generator
Data Parallel Algorithms	Samples that show good usage of Data Parallel Algorithms	CUDA Histogram, CUDA N-Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, Mandelbrot, Optical Flow, Particles, Smoke Particles, VFlockingD3D10
Data-Parallel Algorithms	Samples that show good usage of Data Parallel Algorithms	CUDA Parallel Prefix Sum (Scan), CUDA Parallel Prefix Sum with Shuffle Intrinsics (SHFL_Scan), CUDA Parallel Reduction,

Advanced		
Key		
Concept	Description	Samples
		CUDA Radix Sort (Thrust Library), CUDA Segmentation Tree Thrust Library, CUDA Sorting Networks, Fast Walsh Transform, Merge Sort, threadFenceReduction
Graphics Interop	Samples that demonstrate interop between graphics APIs and CUDA	Bindless Texture, CUDA FFT Ocean Simulation, CUDA N-Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version), Function Pointers, Mandelbrot, Particles, Post-Process in OpenGL, Recursive Gaussian Filter, SLI D3D10 Texture, Smoke Particles, Sobel Filter, VFlockingD3D10, Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
Image Compression	Samples that demonstrate image and video compression	DirectX Texture Compressor (DXTC)
Image Processing	Samples that demonstrate image processing algorithms in CUDA	1D Discrete Haar Wavelet Decomposition, CUDA FFT Ocean Simulation, CUDA Histogram, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, DCT8x8, DirectX Texture Compressor (DXTC), FFT- Based 2D Convolution, Function Pointers, Image denoising, Optical Flow, Post- Process in OpenGL, Recursive Gaussian Filter, SLI D3D10 Texture, Sobel Filter, Stereo Disparity Computation (SAD SIMD Intrinsics), Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
Linear Algebra	Samples demonstrating linear algebra with CUDA	ConjugateGradient, ConjugateGradientUM, Eigenvalues, Fast Walsh Transform, Matrix Transpose, Preconditioned Conjugate Gradient, Scalar Product

Advanced		
Key	Description	Camples
Concept	Description	Samples
OpenGL Graphics Interop	Samples demonstrating how to use interoperability CUDA with OpenGL Marching Cubes Isosurfaces	
Performance Strategies	Samples demonstrating high performance with CUDA	Aligned Types, CUDA C 3D FDTD, CUDA Parallel Prefix Sum (Scan), CUDA Parallel Prefix Sum with Shuffle Intrinsics (SHFL_Scan), CUDA Parallel Reduction, CUDA Radix Sort (Thrust Library), CUDA Segmentation Tree Thrust Library, Concurrent Kernels, Matrix Transpose, Particles, SLI D3D10 Texture, VFlockingD3D10, simpleHyperQ, threadFenceReduction
Physically Based Simulation	Samples demonstrating high performance collisions and/or physocal interactions	Marching Cubes Isosurfaces
Physically- Based Simulation	Samples demonstrating high performance collisions and/or physocal interactions	CUDA N-Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version), Particles, Smoke Particles, VFlockingD3D10
Random Number Generator	Samples demonstrating how to use random number generation with CUDA	Monte Carlo Estimation of Pi (batch PRNG), Monte Carlo Estimation of Pi (batch QRNG), Monte Carlo Estimation of Pi (inline PRNG), Monte Carlo Estimation of Pi (inline QRNG), Monte Carlo Single Asian Option
Recursion	Samples demonstrating recursion on CUDA	Interval Computing
Runtime Compilation	Samples demonstrating how to use NVRTC APIs for runtime compilation of CUDA Kernels	Binomial Option Pricing with libNVRTC, Black-Scholes Option Pricing with libNVRTC, Clock libNVRTC, Matrix Multiplication with libNVRTC, Niederreiter Quasirandom Sequence Generator with libNVRTC, Simple Atomic Intrinsics with libNVRTC, Simple Templates with libNVRTC, Simple Vote Intrinsics with libNVRTC, Using Inline

Advanced Key Concept	Description	Samples
		PTX with libNVRTC, Vector Addition with libNVRTC, simpleAssert with libNVRTC
Surface Writes	Samples demonstrating how to use Surface Writes with GPU kernels	Volumetric Filtering with 3D Textures and Surface Writes
Templates	Samples demonstrating how to use templates GPU kernels	Interval Computing
Texture	Samples demonstrating how to use textures GPU kernels	Bindless Texture
Vertex Buffers	Samples demonstrating how to use Vertex Buffers with CUDA kernels	Marching Cubes Isosurfaces
Video Compression	Samples demonstrating how to use video compression with CUDA	1D Discrete Haar Wavelet Decomposition, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, DCT8x8, Fast Walsh Transform
Video Intrinsics	Samples demonstrating how to use video intrinsics with CUDA	Stereo Disparity Computation (SAD SIMD Intrinsics)

Chapter 6. CUDA API AND ASSOCIATED SAMPLES

The tables below list the samples associated with each CUDA API.

CUDA Driver API Samples

The table below lists the samples associated with each CUDA Driver API.

Table 4 CUDA Driver API and Associated Samples

CUDA Driver API	Samples
cuArrayCreate	Simple Texture (Driver Version)
cuArrayDestroy	Simple Texture (Driver Version)
cuCtxCreate	CUDA Context Thread Management, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, EGLStreams CUDA Interop
cuCtxDestroy	CUDA Context Thread Management, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, EGLStreams CUDA Interop
cuCtxDetach	Simple Texture (Driver Version)
cuCtxPopCurrent	CUDA Context Thread Management, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, EGLStreams CUDA Interop
cuCtxPushCurrent	CUDA Context Thread Management, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, EGLStreams CUDA Interop
cuCtxSynchronize	Simple Texture (Driver Version)

CUDA Driver API	Samples
cuD3D9CtxCreate	CUDA Video Decoder D3D9 API
cuD3D9GetDevice	CUDA Video Decoder D3D9 API
cuD3D9MapResources	CUDA Video Decoder D3D9 API
cuD3D9RegisterResource	CUDA Video Decoder D3D9 API
cuD3D9ResourceGetMappedPitch	CUDA Video Decoder D3D9 API
cuD3D9ResourceGetMappedPointer	CUDA Video Decoder D3D9 API
cuD3D9ResourceSetMapFlags	CUDA Video Decoder D3D9 API
cuD3D9UnmapResources	CUDA Video Decoder D3D9 API
cuD3D9UnregisterResource	CUDA Video Decoder D3D9 API
cuDeviceComputeCapability	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Device Query Driver API, EGLStreams CUDA Interop
cuDeviceGet	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, EGLStreams CUDA Interop
cuDeviceGetAttribute	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Device Query Driver API, EGLStreams CUDA Interop
cuDeviceGetCount	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Device Query Driver API, EGLStreams CUDA Interop
cuDeviceGetName	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, EGLStreams CUDA Interop
cuDeviceTotalMem	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Device Query Driver API
cuDriverGetVersion	Device Query Driver API
cuEGLStreamConsumerAcquireFrame	EGLStreams CUDA Interop
cuEGLStreamConsumerReleaseFrame	EGLStreams CUDA Interop
cuEGLStreamProducerPresentFrame	EGLStreams CUDA Interop
cuGLCtxCreate	CUDA Video Decoder GL API
cuGLGetDevice	CUDA Video Decoder GL API
cuGLMapResources	CUDA Video Decoder GL API

CUDA Driver API	Samples
cuGLRegisterResource	CUDA Video Decoder GL API
cuGLResourceGetMappedPitch	CUDA Video Decoder GL API
cuGLResourceGetMappedPointer	CUDA Video Decoder GL API
cuGLResourceSetMapFlags	CUDA Video Decoder GL API
cuGLUnmapResources	CUDA Video Decoder GL API
cuGLUnregisterResource	CUDA Video Decoder GL API
cuGraphicsResourceGetMappedEglFrame	EGLStreams CUDA Interop
culnit	Device Query Driver API
cuLaunchGridAsync	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuLaunchKernel	CUDA Context Thread Management, Clock libNVRTC, Matrix Multiplication (CUDA Driver API Version), Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version), Matrix Multiplication with libNVRTC, Simple Atomic Intrinsics with libNVRTC, Simple Texture (Driver Version), Using Inline PTX with libNVRTC, Vector Addition Driver API, simpleAssert with libNVRTC
cuMemAlloc	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Clock libNVRTC, EGLStreams CUDA Interop, Matrix Multiplication (CUDA Driver API Version), Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version), Matrix Multiplication with libNVRTC, Simple Atomic Intrinsics with libNVRTC, Simple Texture (Driver Version), Simple Vote Intrinsics with libNVRTC, Using Inline PTX with libNVRTC, Vector Addition Driver API, Vector Addition with libNVRTC
cuMemAllocHost	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuMemFree	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Clock libNVRTC, EGLStreams CUDA Interop, Matrix Multiplication (CUDA Driver API Version), Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version), Matrix Multiplication with libNVRTC, Simple Atomic Intrinsics with libNVRTC, Simple Texture (Driver Version), Simple Vote Intrinsics with libNVRTC, Vector Addition Driver API, Vector Addition with libNVRTC

CUDA Driver API	Samples
cuMemFreeHost	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuMemcpy2D	Simple Texture (Driver Version)
cuMemcpy3D	EGLStreams CUDA Interop
cuMemcpyDtoH	CUDA Context Thread Management, Matrix Multiplication (CUDA Driver API Version), Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version), Matrix Multiplication with libNVRTC, Simple Texture (Driver Version), Using Inline PTX with libNVRTC, Vector Addition Driver API, Vector Addition with libNVRTC
cuMemcpyDtoHAsync	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuMemcpyHtoD	Clock libNVRTC, Matrix Multiplication (CUDA Driver API Version), Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version), Matrix Multiplication with libNVRTC, Simple Atomic Intrinsics with libNVRTC, Simple Vote Intrinsics with libNVRTC, Vector Addition Driver API, Vector Addition with libNVRTC
cuMemsetD8	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuModuleGetFunction	CUDA Context Thread Management, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Matrix Multiplication (CUDA Driver API Version), Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version), Matrix Multiplication with libNVRTC, Simple Texture (Driver Version), Vector Addition Driver API
cuModuleGetGlobal	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuModuleGetTexRef	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Simple Texture (Driver Version)
cuModuleLoad	CUDA Context Thread Management, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Matrix Multiplication (CUDA Driver API Version), Matrix Multiplication (CUDA Driver API version with Dynamic Linking Version), Matrix Multiplication with libNVRTC, Simple Texture (Driver Version), Vector Addition Driver API
cuModuleLoadDataEx	CUDA Context Thread Management, CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, Matrix Multiplication (CUDA Driver API Version), Matrix Multiplication (CUDA

CUDA Driver API	Samples
	Driver API version with Dynamic Linking Version), Matrix Multiplication with libNVRTC, Simple Texture (Driver Version), Vector Addition Driver API
cuModuleUnload	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuParamSetSize	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuParamSetTexRef	Simple Texture (Driver Version)
cuParamSeti	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuParamSetv	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuStreamCreate	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API, EGLStreams CUDA Interop
cuTexRefSetAddressMode	Simple Texture (Driver Version)
cuTexRefSetArray	Simple Texture (Driver Version)
cuTexRefSetFilterMode	Simple Texture (Driver Version)
cuTexRefSetFlags	Simple Texture (Driver Version)
cuTexRefSetFormat	Simple Texture (Driver Version)
cuvidCreateDecoder	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuvidCtxLockCreate	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuvidCtxLockDestroy	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuvidDecodePicture	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuvidDestroyDecoder	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuvidMapVideoFrame	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API
cuvidUnmapVideoFrame	CUDA Video Decoder D3D9 API, CUDA Video Decoder GL API

CUDA Runtime API Samples

The table below lists the samples associated with each CUDA Runtime API.

Table 5 CUDA Runtime API and Associated Samples

CUDA Runtime API	Samples
cublasCreate	Matrix Multiplication (CUBLAS), simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism)
cublasSetVector	simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism)
cublasSgemm	Matrix Multiplication (CUBLAS), simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism)
cudaBindSurfaceToArray	Simple Surface Write
cudaBindTexture2D	Pitch Linear Texture
cudaBindTextureToArray	Pitch Linear Texture, Simple Cubemap Texture, Simple Layered Texture, Simple Surface Write, Simple Texture
cudaCreateChannelDesc	Pitch Linear Texture, Simple Cubemap Texture, Simple Layered Texture, Simple Surface Write, Simple Texture
cudaD3D10GetDevice	SLI D3D10 Texture, Simple D3D10 Texture, Simple Direct3D10 (Vertex Array), Simple Direct3D10 Render Target
cudaD3D10SetDirect3DDevice	SLI D3D10 Texture, Simple D3D10 Texture, Simple Direct3D10 (Vertex Array), Simple Direct3D10 Render Target
cudaD3D10SetGLDevice	VFlockingD3D10
cudaD3D11GetDevice	Simple D3D11 Texture
cudaD3D11SetDirect3DDevice	Simple D3D11 Texture
cudaD3D9GetDevice	Simple D3D9 Texture, Simple Direct3D9 (Vertex Arrays)
cudaD3D9SetDirect3DDevice	Simple D3D9 Texture, Simple Direct3D9 (Vertex Arrays)
cudaD3D9SetGLDevice	Fluids (Direct3D Version)
cudaDeviceCanAccessPeer	Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Peer-to-Peer Transfers with Multi-GPU
cudaDeviceDisablePeerAccess	Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Peer-to-Peer Transfers with Multi-GPU
cudaDeviceEnablePeerAccess	Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Peer-to-Peer Transfers with Multi-GPU
cudaDeviceGetP2PAttribute	Topology Query

CUDA Runtime API	Samples
cudaDeviceSynchronize	Bandwidth Test, Template
cudaDriverGetVersion	Device Query
cudaEventCreate	Bandwidth Test, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Simple Multi Copy and Compute, Simple Multi-GPU, Vector Addition, asyncAPI, simpleStreams, simpleZeroCopy
cudaEventCreateWithFlags	Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Peer-to-Peer Transfers with Multi-GPU
cudaEventDestroy	Bandwidth Test, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Simple Multi Copy and Compute, Simple Multi-GPU, Vector Addition, asyncAPI, simpleStreams, simpleZeroCopy
cudaEventElapsedTime	Bandwidth Test, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Multi Copy and Compute, Simple Multi-GPU, Simple Peer-to- Peer Transfers with Multi-GPU, Vector Addition, asyncAPI, simpleStreams, simpleZeroCopy
cudaEventQuery	Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Simple Multi Copy and Compute, Simple Multi-GPU, Vector Addition, asyncAPI, simpleStreams, simpleZeroCopy
cudaEventRecord	Bandwidth Test, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Simple Multi Copy and Compute, Simple Multi-GPU, Vector Addition, asyncAPI, simpleStreams, simpleZeroCopy
cudaEventSynchronize	Matrix Multiplication (CUDA Runtime API Version), Vector Addition
cudaFree	Bandwidth Test, C++ Integration, Clock, FP16 Scalar Product, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Pitch Linear Texture, Simple Atomic Intrinsics, Simple Cubemap Texture, Simple Layered Texture, Simple Surface Write, Simple Texture, Simple Vote Intrinsics, System wide Atomics, Template, Using Inline PTX, Vector Addition, cudaOpenMP, simpleAssert, simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism), simpleMPI

CUDA Runtime API	Samples
cudaFreeArray	Pitch Linear Texture, Simple Cubemap Texture, Simple Layered Texture, Simple Surface Write, Simple Texture
cudaFreeHost	Bandwidth Test, FP16 Scalar Product, Simple Atomic Intrinsics, Simple Vote Intrinsics, System wide Atomics, Using Inline PTX, simpleAssert, simpleIPC, simpleZeroCopy
cudaFuncGetAttributes	cppOverload
cudaFuncSetCacheConfig	cppOverload
cudaGLSetGLDevice	Bicubic B-spline Interoplation, Bilateral Filter, Bindless Texture, Box Filter, CUDA FFT Ocean Simulation, CUDA N- Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, CUDA and OpenGL Interop of Images, Fluids (OpenGL Version), Mandelbrot, Marching Cubes Isosurfaces, Particles, Post-Process in OpenGL, Recursive Gaussian Filter, Simple OpenGL, Simple Texture 3D, Smoke Particles, Sobel Filter, Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
cudaGetDeviceAttribute	Topology Query
cudaGetDeviceCount	Device Query, Topology Query
cudaGetDeviceProperties	Device Query
cudaGraphicsD3D10RegisterResource	SLI D3D10 Texture, Simple D3D10 Texture, Simple Direct3D10 (Vertex Array), Simple Direct3D10 Render Target
cudaGraphicsD3D11RegisterResource	Simple D3D11 Texture
cudaGraphicsD3D9RegisterResource	Simple D3D9 Texture, Simple Direct3D9 (Vertex Arrays)
cudaGraphicsGLRegisterBuffer	Bicubic B-spline Interoplation, Bilateral Filter, Bindless Texture, Box Filter, CUDA FFT Ocean Simulation, CUDA N- Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, CUDA and OpenGL Interop of Images, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version), Mandelbrot, Marching Cubes Isosurfaces, Particles, Post-Process in OpenGL, Recursive Gaussian Filter, Simple OpenGL, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen, Simple Texture 3D, Smoke Particles, Sobel Filter, VFlockingD3D10, Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes

CUDA Runtime API	Samples
cudaGraphicsMapResources	Bicubic B-spline Interoplation, Bilateral Filter, Bindless Texture, Box Filter, CUDA FFT Ocean Simulation, CUDA N-Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, CUDA and OpenGL Interop of Images, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version), Mandelbrot, Marching Cubes Isosurfaces, Particles, Post-Process in OpenGL, Recursive Gaussian Filter, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen, Simple Texture 3D, Smoke Particles, Sobel Filter, VFlockingD3D10, Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
cudaGraphicsRegisterResource	Bicubic B-spline Interoplation, Bilateral Filter, Bindless Texture, Box Filter, CUDA FFT Ocean Simulation, CUDA N-Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, CUDA and OpenGL Interop of Images, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version), Mandelbrot, Marching Cubes Isosurfaces, Particles, Post-Process in OpenGL, Recursive Gaussian Filter, Simple OpenGL, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen, Simple Texture 3D, Smoke Particles, Sobel Filter, VFlockingD3D10, Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
cudaGraphicsResourceGetMappedPointer	Bicubic B-spline Interoplation, Bilateral Filter, Bindless Texture, Box Filter, CUDA FFT Ocean Simulation, CUDA N-Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, CUDA and OpenGL Interop of Images, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version), Mandelbrot, Marching Cubes Isosurfaces, Particles, Post-Process in OpenGL, Recursive Gaussian Filter, Simple OpenGLE, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen, Simple Texture 3D, Smoke Particles, Sobel Filter, VFlockingD3D10, Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
cudaGraphicsResourceSetMapFlags	SLI D3D10 Texture, Simple D3D10 Texture, Simple D3D11 Texture, Simple D3D9 Texture, Simple Direct3D10 (Vertex Array), Simple Direct3D10 Render Target

CUDA Runtime API	Samples
cuda Graphics Sub Resource Get Mapped Arra	/ SLI D3D10 Texture, Simple D3D10 Texture, Simple D3D11 Texture, Simple D3D9 Texture, Simple Direct3D10 (Vertex Array), Simple Direct3D10 Render Target
cudaGraphicsUnmapResources	Bicubic B-spline Interoplation, Bilateral Filter, Bindless Texture, Box Filter, CUDA FFT Ocean Simulation, CUDA N-Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, CUDA and OpenGL Interop of Images, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version), Mandelbrot, Marching Cubes Isosurfaces, Particles, Post-Process in OpenGL, Recursive Gaussian Filter, Simple OpenGL, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen, Simple Texture 3D, Smoke Particles, Sobel Filter, VFlockingD3D10, Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
cudaGraphicsUnregisterResource	Bicubic B-spline Interoplation, Bilateral Filter, Bindless Texture, Box Filter, CUDA FFT Ocean Simulation, CUDA N-Body Simulation, CUDA N-Body Simulation on Screen, CUDA N-Body Simulation with GLES, CUDA and OpenGL Interop of Images, Fluids (Direct3D Version), Fluids (OpenGL Version), Fluids (OpenGLES Version), Mandelbrot, Marching Cubes Isosurfaces, Particles, Post-Process in OpenGL, Recursive Gaussian Filter, SLI D3D10 Texture, Simple D3D10 Texture, Simple D3D11 Texture, Simple D3D9 Texture, Simple Direct3D10 (Vertex Array), Simple Direct3D10 Render Target, Simple Direct3D9 (Vertex Arrays), Simple OpenGL, Simple OpenGLES, Simple OpenGLES EGLOutput, Simple OpenGLES on Screen, Simple Texture 3D, Smoke Particles, Sobel Filter, VFlockingD3D10, Volume Rendering with 3D Textures, Volumetric Filtering with 3D Textures and Surface Writes
cudaHostAlloc	Bandwidth Test, simpleZeroCopy
cudaHostGetDevicePointer	simpleZeroCopy
cudaHostRegister	simpleZeroCopy
cudaHostUnregister	simpleZeroCopy
cudalpcCloseMemHandle	simpleIPC
cudalpcGetEventHandlet	simpleIPC

CUDA Runtime API	Samples
cudalpcOpenMemHandle	simpleIPC
cudaMallco	Simple Vote Intrinsics, simpleMPI
cudaMalloc	C++ Integration, Clock, FP16 Scalar Product, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Pitch Linear Texture, Simple Atomic Intrinsics, Simple Cubemap Texture, Simple Layered Texture, Simple Surface Write, Simple Texture, System wide Atomics, Template, Using Inline PTX, Vector Addition, cudaOpenMP, simpleAssert, simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism)
cudaMalloc3DArray	Simple Cubemap Texture, Simple Layered Texture
cudaMallocArray	Pitch Linear Texture, Simple Surface Write, Simple Texture
cudaMallocHost	Bandwidth Test, FP16 Scalar Product, Using Inline PTX, simpleAssert
cudaMallocManaged	Unified Memory Streams
cudaMallocPitch	Pitch Linear Texture
cudaMemcpy	Bandwidth Test, C++ Integration, Clock, FP16 Scalar Product, Matrix Multiplication (CUBLAS), Matrix Multiplication (CUDA Runtime API Version), Peer-to-Peer Bandwidth Latency Test with Multi-GPUs, Simple Atomic Intrinsics, Simple Cubemap Texture, Simple Layered Texture, Simple Peer-to-Peer Transfers with Multi-GPU, Simple Surface Write, Simple Texture, Simple Vote Intrinsics, System wide Atomics, Template, Using Inline PTX, Vector Addition, cudaOpenMP, simpleAssert, simpleDevLibCUBLAS GPU Device API Library Functions (CUDA Dynamic Parallelism), simpleIPC, simpleMPI
cudaMemcpy2D	Pitch Linear Texture
cudaMemcpy2DToArray	SLI D3D10 Texture, Simple D3D10 Texture, Simple D3D11 Texture, Simple D3D9 Texture, Simple Direct3D10 (Vertex Array), Simple Direct3D10 Render Target
cudaMemcpy3D	Simple Cubemap Texture, Simple D3D9 Texture, Simple Layered Texture
cudaMemcpyAsync	Bandwidth Test, Simple CUDA Callbacks, Simple Multi Copy and Compute, Simple Multi-GPU, asyncAPI, simpleStreams

CUDA Runtime API	Samples
cudaMemcpyToArray	Pitch Linear Texture, Simple Texture
cudaMemset2D	Pitch Linear Texture
cudaPrintfDisplay	simplePrintf
cudaPrintfEnd	simplePrintf
cudaRuntimeGetVersion	Device Query
cudaSetDevice	Bandwidth Test, Device Query
cudaStreamAddCallback	Simple CUDA Callbacks
cudaStreamAttachManagedMem	Unified Memory Streams
cudaStreamCreate	Simple CUDA Callbacks
cudaStreamDestroy	Simple CUDA Callbacks
cudaUnbindTexture	Pitch Linear Texture
cufftDestroy	CUDA FFT Ocean Simulation, FFT-Based 2D Convolution
cufftExecC2R	CUDA FFT Ocean Simulation, FFT-Based 2D Convolution
cufftExecR2C	CUDA FFT Ocean Simulation, FFT-Based 2D Convolution
cufftPlan2d	CUDA FFT Ocean Simulation, FFT-Based 2D Convolution
nppGetGpuComputeCapability	JPEG encode/decode and resize with NPP
nppiDCTFree	JPEG encode/decode and resize with NPP
nppiDCTInitAlloc	JPEG encode/decode and resize with NPP
nppiDCTQuantInv8x8LS_JPEG_16s8u_C1R_	NETG encode/decode and resize with NPP
nppiDecodeHuffmanScanHost_JPEG_8u16	s 即孫 encode/decode and resize with NPP
nppiEncodeHuffmanGetSize	JPEG encode/decode and resize with NPP
nppiResizeSqrPixel_8u_C1R	JPEG encode/decode and resize with NPP

Chapter 7. FREQUENTLY ASKED QUESTIONS

Answers to frequently asked questions about CUDA can be found at http://developer.nvidia.com/cuda-faq and in the CUDA Toolkit Release Notes.

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 and the NVIDIA logo are trademarks 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

© 2007-2016 NVIDIA Corporation. All rights reserved.

