
Index

- `__device__` 24
- `__global__` 25
- `__host__` 24
- `<<<>>>` 27, 30
- `<random>` 179
- `<thread>` 225
- 3D rank formula 33
- `a_less` device function 136
- AABB 276
- active-warps 37
- affine transformation 162
- affine3D kernel 163
- AI 435
 - Deep Learning 436
 - optimised frameworks 435
- `argc` 439
- `argv` 439
- astrophotography 290
- `asyncDiskIO` example results 232
- `asyncDiskIO` program support functions 227
- asynchronous memory transfer 185
- `atan2` 254
- `atexit` function 216
- Atomic operations 24, 382
- `atomicAdd` 382
- `atomicCAS` 384
- `atomicMax` 117
- attenuation length 271
- `auto` 75
- AVX 393
- Axes Aligned Bounding Block 276

- backward projection 241
- `backward_project` kernel 265
- base LOR 247
- Batcher network 137
- `batcher9` kernel 139
- Beowulf PC clusters 314
- BF16 359
- BGO *See* bismuth germanate
- bilinear device function 148
- bilinear interpolation 143
- binary IO in C 415
- `binary_partition` 103
- bismuth germanate 271

- bit 402
- BL *See* base LOR
- BLAS 65
- `blockDim` 20
- `blockIdx` 20
- Boolean values 404
- BP *See* backward projection
- B-spline interpolation 143
- buffers ping-pong operation 188
- byte 402

- cache 12
- cache lines 13
- cache-coherency 382
- caching scheme 13
- cascade function 119
- cast C++ 444
- cast intrinsic function 386
- CC *See* compute capability
- `cgwarp` 79
- `checkCudaErrors` 349
- chessboard update pattern 199
- `chrono.h` 4
- coalesced groups 92
- coalesced memory access 374
- `coalesced_group` 96
- `coalesced_threads` 97
- code repository xxi
- coincidence event 240
- compute capability 15, 45, 373
- const C++ 446
- const qualifier 57
- constant memory *See* GPU constant memory
- `constexpr` 259
- container class 441
- `coop3D` kernel 77
- cost function 167
- `cost_functor` implementation 169
- `cost_functor` optimisation code 171
- `costfun_sumsq` kernel 167
- CPU
 - ALU 11
 - AVX 13
 - cache lines 13
 - execute 11

- fetch 11
- Intel Haswell caching 12
- master clock 10
- memory 10
- register file 11
- SSE 13
- cpusum program 2
- cr_Ptr 60
- cstring 445
- CUB 409
- cuBLAS 66
- CUDA
 - API reference 407
 - C++_Programming_Guide 379
 - computational libraries 407
 - compute-sanitizer 356
 - cudaMalloc 190
 - cudaMallocHost 190
 - documentation 406
 - event 218
 - event status 218
 - event timing 220
 - examples 374
 - graphs 233
 - graphs example 234
 - kernel functions 2, 8
 - libc++ project 409
 - libraries 406
 - occupancy_calculator 379
 - programming support 406
 - samples 380
 - SDK 374
 - SDK 11.0 72
 - streams and events 209
 - third party extensions 408
 - toolkit 377
 - toolkit evolution 378
 - unified memory 295
 - unified virtual addressing 298
 - waves 19
 - zero-copy memory 301
- CUDA_DEVICE_MAX_CONNECTIONS 211, 214
- cudaChooseDevice 297
- cudaDeviceGetCacheConfig 297
- cudaDeviceGetSharedMemConfig 297
- cudaDeviceReset 215, 297
- cudaDeviceSetCacheConfig 297
- cudaDeviceSetSharedMemConfig 297
- cudaDeviceSynchronize 47, 221
- cudaErrorNotReady 218
- cudaEventCreate 210, 218
- cudaEventDestroy 210
- cudaEventElapsedTime 210, 219, 225
- cudaEventQuery 210, 218
- cudaEventRecord 210, 218, 223
- cudaEventSynchronize 210
- cudaFreeHost 301
- cuda-gdb 356
- cudaGetDevice 297
- cudaGetDeviceCount 297
- cudaGetDeviceFlags 297
- cudaGetDeviceProperties 297
- cudaGraphInstantiate 238
- cudaGraphLaunch 238
- cudaHostAlloc 301
- cudaHostAllocDefault 301
- cudaHostAllocMapped 301
- cudaHostAllocPortable 301
- cudaHostAllocWriteCombined 301
- cudaHostGetDevicePointer 301
- cudaMallocHost 301
- cuda-memchk 354
- cudaMemcpy 299
- cudaMemcpyAsync 214
- cudaMemcpyDefault 299
- cudaMemcpyDeviceToDevice 299
- cudaMemcpyDeviceToHost 299
- cudaMemcpyHostToDevice 299
- cudaMemcpyHostToHost 299
- cudaPeekAtLastError 349
- cudaProfilerStart 332
- cudaProfilerStop 332
- cudaSelectDevice 295
- cudaSetDevice 297
- cudaSetDeviceFlags 297
- cudaStream object 210
- cudaStreamBeginCapture 238
- cudaStreamCreate 210
- cudaStreamDestroy 210
- cudaStreamEndCapture 238
- cudaStreamQuery 210
- cudaStreamSynchronize 210, 223
- cudaStreamWaitEvent 210, 223, 225
- cudaSuccess 218
- cuDNN 435
- cuRand 178
- cuRand Device API 191
- cuRand Host API 185
- cuRand library 185
- cuRand MRG32k3a 191
- cuRand MTGP32 192
- cuRand Philox4x32_10 192
- cuRand XORWOW 191
- CX header files 410
 - cx.h 411
 - cx.h listing 410
 - cxbinio.h 411
 - cxbinio.h listing 415, 418
 - cxconfun.h 411
 - cxconfun.h listing 433
 - cxtextures.h 411
 - cxtextures.h listing 425
 - cxtimers.h 411
 - cxtimers.h listing 422

450

cx::ok 354
 cxbino 158

 D2H *See* device to host transfer
 DALI 436
 deadlock 92, 95
 deadlock kernel 94
 deadlock_coalesced kernel 97
 debugging 325
 default stream 210
 depth of interaction 270
 Derenzo phantom 271
 results 269
 device driver 103
 device memory *See* GPU main memory
 DGX Systems 437
 Dicom 158
 DIGITS 436
 dim3 28
 disk IO asynchronous 227
 disk IO overhead 225
 distributed memory model 315
 DOI *See* depth of interaction
 double buffering 110
 DX11 337
 DX12 337
 dynamic shared memory 51

 event2 program 221
 event2 program timelines 226

 filter9PT kernel 128
 filter9PT_2 kernel 130
 filter9PT_3 kernel 131
 find_spot kernel 257
 float_as_uint 386
 floating point denormalised 405
 flourine 18 239
 flush to zero 329
 Flynn's Taxonomy 22
 forward projection 241
 forward_project kernel 262
 FP *See* forward projection
 FP16 358
 FP32 358
 frame buffer *See* GPU main memory
 FTZ *See* flush to zero
 full block simulation 275
 fullsim program 244
 fullsim program structs 248
 functor 170

 gaming 14
 GeForce GTX 14
 GeForce RTX 14
 GPGPU 373
 GPU
 Architecture 14–16
 clz intrinsic function 49

Index

constant memory 16, 129
 division by zero 266
 generations 375
 graphics processing unit 1
 local memory 17
 main memory 16
 NVIDIA models 14
 register file 17
 shared memory 17
 texture hardware 144
 texture interpolation modes 144
 texture linear filtering 144
 texture linear interpolation 145
 texture memory 17
 texture nearest point sampling
 144
 texture table lookup 145
 gpu_log kernel 326
 gpmult0 kernel 58
 gpmult2 kernel 61
 gpsum program 7
 gputiled kernel 62
 gputiled1 kernel 65
 grid wide synchronisation
 103
 grid.sync 77
 grid3D kernel 31
 grid3D_linear kernel 34
 gridDim 20
 GTX1080 15

 halo 111
 hexadecimal 402
 high pass filter 128
 hostmult0 function 54
 hostmult1 function 56
 HPC 293

 ICC *see* Intel compiler
 Intel compiler 394
 image display with OpenCV 154
 image processing filters 127
 image registration 167
 image registration results 175
 image rotation 146
 ImageJ 157
 inactive threads 91
 indexing with lamdba function 60
 InfiniBand 293
 Intel compiler 393–394
 Intel intrinsics library 397
 Intel System Studio 394
 interp3D device function 166
 interp3D kernel 166
 intrinsic functions warp level 90
 inverse transform method 196, 251
 Ising model 198
 phase transition 199
 results 207

- Jacobi iteration 107
- kernel 25
- key_to_lor device function 260
- labeled_partition 103
- lambda function lam_check 279
- lane 28
- Laplacian 106
- latency 12
- latency hiding 37–39, 232
- launch configuration 19
- lerp function 147
- line of response 239
- Linux debugging 356
- log(1+x) series 325
- loop unrolling 48, 65
- LOR *See* line of response
- LOR array 242
- lor_to_key device function 260
- low pass filter 128
- make_parms utility function 174
- makefile 387
- mashdata kernel 214
- matmult kernel 361
- matmultS kernel 363
- median filter 135
- median finding network 138
- median9 device function 137
- memchk 352
- memcpy_async 103
- message passing interface 295, 313
- MIMD 23
- MISD 23
- MLEM 286
 - definition 240
 - formula 241
- MLEM iteration times 267
- MMX 393
- modern C++ 374
- Moore's Law 2
- Morton ordering 153, 431
- MPI 1, 313–323
 - MPI_Allreduce 316
 - MPI_Alltoall 320
 - MPI_Barrier 320
 - MPI_Bcast 316
 - MPI_Comm_rank 316
 - MPI_Comm_size 316
 - MPI_Comm_split 320
 - MPI_Finalize 295
 - MPI_Gather 316
 - MPI_Init 316
 - MPI_Recv 320
 - MPI_Reduce 316
 - MPI_Scan 320
 - MPI_Scatter 316
 - MPI_Send 320
 - tutorial 315
 - version history 314
- mpialltoall example 321
- mpialltoall matrix transpose 323
- mpic++ 319
- mpireduce program 316
- mpirun 319
- MRI
 - axial slice 164
 - coronal slice 164
 - head scan 164
 - sagittal slice 164
 - scan 162
 - transverse slice 164
- multiGPU program 299
- multi-grid group 103
- MUTEX 386
- naïve sorting algorithm 137
- NCCL 436
- nearest device function 148
- NeMo 437
- node 293, 314
- non-linear filter 135
- normalised cumulative distribution 196
- Nsight Compute 338
 - compute workload analysis 341
 - GPU speed of light 340
 - instruction statistics 344
 - launch statistics 344
 - memory workload analysis 341
 - occupancy 346
 - scheduler statistics 342
 - source counters 346
 - warp state statistics 343
- Nsight Eclipse Edition 356
- Nsight Systems 325, 336
- nsys 337
- NVCC
 - command line 387
 - compiler 65, 387
 - nvcc.exe 379
 - restrict 57
- NVIDIA Cg 373
- NVIDIA GPU Cloud 437
- NVIDIA Tools Extension Library 333
- NVIDIA Visual Profiler 211
- NVLINK 293
- nvprof example 332
- nvprof use of 330
- nvprof.exe 379
- NVTX *See* NVIDIA Tools Extension Library
- NVVP 211
- NVVP timelines 216
- occupancy 20
- ok function 349
- ompsum 4

452

OpenCV
 GpuMat 207
 imread 160
 imshow 161
 imwrite 161
 Mat object 160
 waitkey 161
 OpenGL 337
 OpenMP 1, 399
 ompsum 3
 Visual Studio C++ 6
 OptiX 408
 ordered subsets expectation maximisation 243
 OSEM
 full iteration 268
 method 268
 ordered subsets EM 243
 subset definitions 268
 overflow error 403

 P2P *See* peer to peer
 p2ptest kernel 299
 parallel primitive 6
 parallel programming 9
 parallel projection 268
 parallel reduction 40–51
 paramset struct 169
 Pascal Architecture 15
 peer to peer 299
 PET
 block detectors 274
 cylindrical detector 239
 map array 246
 phantom data 244
 photomultiplier 239
 scanner 239
 scintillating crystals 239
 touching cylinder 274
 pi, calculation of 179
 piG example 193
 piH example
 basic version 180
 cuRand Host API 186
 faster host RNG 182
 OMP version 183
 using cudaMemcpyAsync 188
 using pinned memory 188
 pinned memory 214
 pipeline example 212
 pixel addressing 142
 pixel aliasing 147
 pixel interpolation 142
 pointer aliasing 57
 Poisson's equation 107
 pol2cart program 244
 poluse program 244
 positron annihilation 239
 power consumption 375

Index

pragma 5
 printf, debugging with 347
 pseudorandom number generator 178
 PTX code 279, 389
 PTX documentation 406
 PyCUDA 437
 Python 435
 Python Toolkit 437

 Quadro 14
 quasi-random number generator 179

 r_ptr 60
 radiation transport calculations 283
 RAI1 439
 rand 178
 random number generator 178
 random numbers true 179
 rank of thread 28
 rank of warp 28
 ray 275
 ray equation 275
 ray tracing 275–276
 ray_to_block device function 276
 ray_to_block2 device function 279
 ray_to_cyl device function 252
 ray_to_cyl_doi device function 272, 276
 raytracing 408
 read after write error 74
 read-after-write 382
 readspot program 244
 reco program 244
 recosem program 244
 reduce_classic function 305
 reduce_classic_pinned function 307
 reduce_coal_any_v1 kernel 100
 reduce_half_v1 kernel 369
 reduce_managed function 310
 reduce_maxdiff kernel 115
 reduce_thrust_pinned function 308
 reduce_zero copy function 309
 reduce0 kernel 41
 reduce1 kernel 44
 reduce2 kernel 46
 reduce3 kernel 48
 reduce4 kernel 49
 reduce5 kernel 73
 reduce6 kernel 81
 reduce7 kernel 83
 reduce7_v1 kernel 86
 reduce7_v1_coal kernel function 98
 reduce8 kernel 85
 reduceT kernel 367
 region of interest 249
 register file *See* GPU register file
 resident warps 37
 restrict keyword 56
 RGB 14, 126

- RGB planar format 126
- RGBA 126
- Richardson–Lucy image deblurring 286
- Richardson–Lucy iteration 286
- RL iteration *See* Richardson–Lucy iteration
- rl_backward device function 286
- rl_deconv host function 288
- rl_forward device function 286
- RNG multiple threads 182
- RNG uncorrelated sequences 191
- ROI *See* region of interest
- rotate1 kernel 149
- rotate2 kernel 151
- RT *See* ray tracing
- RT unit 408
- RTX 2070 18, 377
- RTX 3090 14

- saxpy 66, 444
- SDK volumeRender example 207
- serial bottlenecks 9
- SFU *See* special function units
- sgemm 66
- shared memory 45, 51–53, 382
- shared_example kernel 52
- sharpening filter 128
- shfl_xor 84
- SIMD 23, 393
- simpleCudaGraphs example 237
- SIMT 23
- simulation 70
- simultaneous update 24
- single pixel noise 136
- SISD 23
- SM *See* symmetric multiprocessor, *See* system matrix
- SM element 241
- smoothing filter 136
- Sobel filter 128, 134
- sobel9PT kernel 135
- special function units 15–16
- speculative execution 12
- stalled threads 37
- star trail 290
- stencil 2D convergence 114
- stencil 9-point 113
- stencil cascade iterations 119
- stencil code performance 114
- stencil2D kernel 107
- stencil2D_sm kernel 112
- stencil3D kernels 124
- stencil9PT kernel 113
- stencils 2D 106
- stencils 3D 123
- struct Roi 251
- swap function 136
- symmetric multiprocessor 15
- syncthreads 47, 72
- syncthreads weak form 93
- synwarp 72
- system matrix 240
- system matrix building 259

- templates C++ 443
- Tensor Cores
 - reduction 366
 - sparsity 371
- TensorRT 436
- ternary operator 439
- Tesla 14
- texture mapping unit 151
- TF32 359
- this_grid 75
- this_thread_block 75
- thread class C++11 1
- thread divergence 89
- thread block 9, 19
- thread block 3D addressing 32–34
- threadIdx 20
- thread-linear addressing 29
- thrust 8, 408
- thrustHvecPin 187
- tiled matrix multiplication 62
- tiled partition multiwarp 76
- tiled partition 75
- tiles overlapping 111
- TMU *See* texture mapping unit
- TOP500 list 375
- track_ray device function 281
- trapezoidal rule 2
- Triton 436
- trivially parallel 22
- twos complement 403

- UM *See* unified memory
- underflow error 403
- unified memory 302
- use_fast_math 388
- vector loading 86
- video ram *See* GPU main memory
- Visual Studio
 - break points 350
 - CUDA debugging 352
 - debugging 349
 - locals window 351
 - Nsight addon 352
 - project properties 350
 - set command line arguments 350
 - start debugging 350
 - step into 351
 - step over 351

- voxel 162
- voxel polar 240
- voxgen kernel 249

454

voxgen_block kernel 283
voxgen_doi device function 276
Vulcan 337

warp 15, 72
warp engine 15, 16
warp matrix functions 360
 fill_fragment 360
 fragment 360
 load_matrix_sync 360

Index

mma_sync 360
 store_matrix_sync 360
warp shuffle functions 91
warp-only kernel coding 83
warpSize 20
WDDM drive 337
WE *See* warp engine
workstation multi-GPU 293

zoomfrom kernel 119