	Result	Size	Time	Cycles	GPU	SM Frequency	Process	Attributes
Current	559 - brute_force_coarsening	(32768, 1, 1)x(32, 1, 1)	2.22 ms	1,297,840	0 - Tesla T4	584.98 Mhz	[9877] exe	

Summary | Details | Source | Context | Comments | Raw | Session

Compare | Tools | View | Export

GPU Speed Of Light Throughput

All

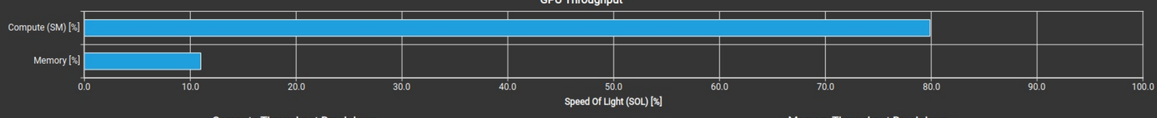
High-level overview of the throughput for compute and memory resources of the GPU. For each unit, the throughput reports the achieved percentage of utilization with respect to the theoretical maximum. Breakdowns show the throughput for each individual sub-metric of Compute and Memory to clearly identify the highest contributor. High-level overview of the utilization for compute and memory resources of the GPU presented as a roofline chart.

Compute (SM) Throughput [%]	79.89	Duration [ms]	2.22
Memory Throughput [%]	11.00	Elapsed Cycles [cycle]	1,297,840
L1/TEX Cache Throughput [%]	11.85	SM Active Cycles [cycle]	1,292,665.05
L2 Cache Throughput [%]	0.65	SM Frequency [Mhz]	584.98
DRAM Throughput [%]	1.95	DRAM Frequency [Ghz]	4.99

High Compute Throughput Compute is more heavily utilized than Memory: Look at the Compute Workload Analysis section to see what the compute pipelines are spending their time doing. Also, consider whether any computation is redundant and could be reduced or moved to look-up tables.

FP64/32 Utilization
Est. Local Speedup: 69.58% The ratio of peak float (fp32) to double (fp64) performance on this device is 32:1. The kernel achieved close to 1% of this device's fp32 peak performance and 60% of its fp64 peak performance. If Compute Workload Analysis determines that this kernel is fp64 bound, consider using 32-bit precision floating point operations to improve its performance. See the Kernel Profiling Guide for more details on roofline analysis.

GPU Throughput



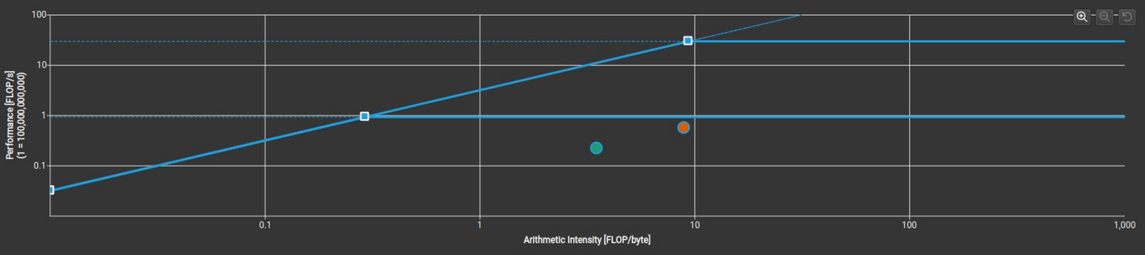
Compute Throughput Breakdown

SM: Pipe Fp64 Cycles Active [%]	79.89
SM: Pipe Alu Cycles Active [%]	48.46
SM: Issue Active [%]	43.28
SM: Inst Executed [%]	43.28
SM: Mio Pq Write Cycles Active [%]	31.16
SM: Inst Executed Pipe Adu [%]	23.43
IDC: Request Cycles Active [%]	15.79
SM: Pipe Fma Cycles Active [%]	15.21
SM: Mio Inst Issued [%]	14.80
SM: Inst Executed Pipe Lsu [%]	11.00
SM: Mio Pq Read Cycles Active [%]	10.68
SM: Inst Executed Pipe Cbu Pred On Any [%]	8.53
SM: Mio2rf Writeback Active [%]	5.29
SM: Inst Executed Pipe Xu [%]	0.13
SM: Inst Executed Pipe Uniform [%]	0.03
SM: Pipe Tensor Cycles Active [%]	0
SM: Inst Executed Pipe Fp16 [%]	0
SM: Pipe Shared Cycles Active [%]	0
SM: Memory Throughput Internal Activity [%]	0
SM: Instruction Throughput Internal Activity [%]	0
SM: Inst Executed Pipe Tex [%]	0
SM: Inst Executed Pipe Ipa [%]	0

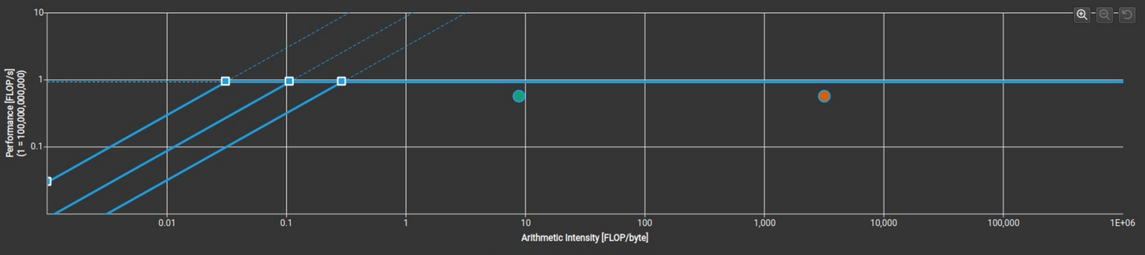
Memory Throughput Breakdown

L1: Lsuin Requests [%]	11.00
L1: Data Pipe Lsu Wavefronts [%]	5.92
L1: Lsu Writeback Active [%]	2.62
DRAM: Cycles Active [%]	1.95
DRAM: Dram Sectors [%]	1.42
L1: M L1tex2xbar Req Cycles Active [%]	0.76
L1: Data Bank Reads [%]	0.69
L2: T Sectors [%]	0.65
L2: Xbar2lts Cycles Active [%]	0.65
L1: Data Bank Writes [%]	0.40
L2: D Sectors [%]	0.31
GPU: Compute Memory Access Throughput Internal Activity [%]	0.16
L2: T Tag Requests [%]	0.16
L2: Lts2xbar Cycles Active [%]	0.00
L1: Texin Sm2tex Req Cycles Active [%]	0.00
L1: F Wavefronts [%]	0.00
L2: D Sectors Fill Device [%]	0.00
L2: D Atomic Input Cycles Active [%]	0
L1: Tex Writeback Active [%]	0
L2: D Sectors Fill Sysmem [%]	0
L1: M Xbar2l1tex Read Sectors [%]	0
GPU: Compute Memory Request Throughput Internal Activity [%]	0
L1: Data Pipe Tex Wavefronts [%]	0

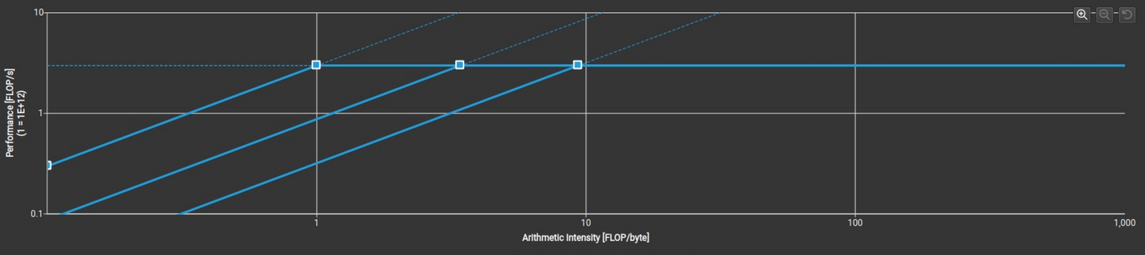
Floating Point Operations Roofline



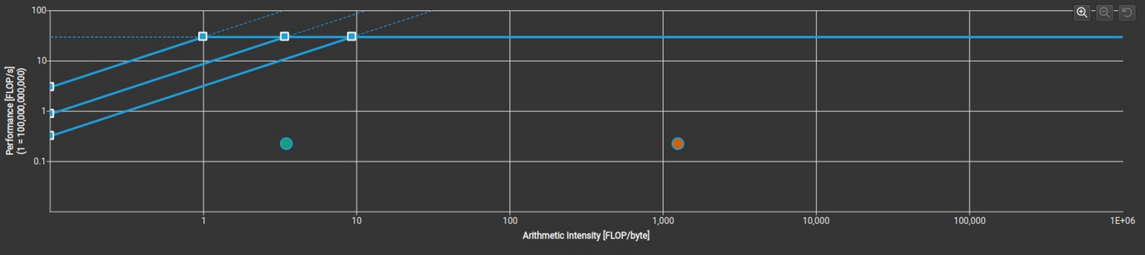
Floating Point Operations Roofline (Double Precision)



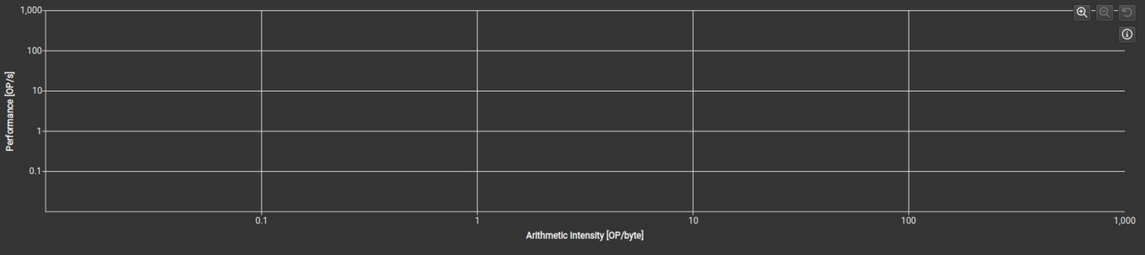
Floating Point Operations Roofline (Half Precision)



Floating Point Operations Roofline (Single Precision)



Tensor Core Operations Roofline



	# Operations	# Operations / Cycle	# Operations / s	Peak % ?eak Operations / Cycle	Peak Operations / Cycle	Peak Operations / s
Src:fp16,bf16,tf32 Dst:fp32	0	0	0	0	40,960	23,929.12
Src:fp16 Dst:fp16	0	0	0	0	40,960	23,929.12
Src:int1	0	0	0	0	81,920	47,858.24
Src:int4	0	0	0	0	81,920	47,858.24
Src:int8	0	0	0	0	81,920	47,858.24

GPU and Memory Workload Distribution

Analysis of workload distribution in active cycles of SM, SMP, SMSP, L1 & L2 caches, and DRAM

Average SM Active Cycles [cycle]	1,292,665.05	Average L1 Active Cycles [cycle]	1,292,665.05
Average L2 Active Cycles [cycle]	115,613.81	Average SMSP Active Cycles [cycle]	1,290,493.82
Average DRAM Active Cycles [cycle]	216,066	Total SM Elapsed Cycles [cycle]	51,844,680
Total L1 Elapsed Cycles [cycle]	51,844,680	Total L2 Elapsed Cycles [cycle]	60,698,688
Total SMSP Elapsed Cycles [cycle]	207,378,720	Total DRAM Elapsed Cycles [cycle]	88,602,624

Workload Distribution

	Average	Min	Max	Sum
SM Active Cycles	1,292,665.05	1,289,842	1,296,384	51,706,602
SMSP Active Cycles	1,290,493.82	1,284,541	1,296,366	206,479,011
L1 Active Cycles	1,292,665.05	1,289,842	1,296,384	51,706,602
L2 Active Cycles	115,613.81	110,260	126,974	3,699,642
DRAM Active Cycles	216,066	215,416	216,684	1,728,528