	Result	Size	Time	Cycles	GPU	SM Frequency	Process	Attributes
Current	550 - brute_force	(1024, 1, 1)x(1024, 1, 1)	2.07 ms	1,208,822	0 - Tesla T4	584.99 Mhz	[17816] exe	

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

Compare | Tools | View | Export

GPU Speed Of Light Throughput

All

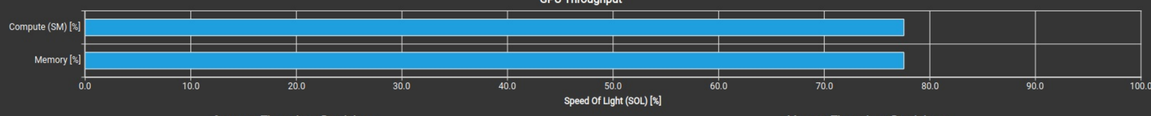
High-level overview of the throughput for compute and memory resources of the GPU. For each unit, the throughput reports the achieved percentage of utilization with respect to the theoretical maximum. Breakdowns show the throughput for each individual sub-metric of Compute and Memory to clearly identify 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 [%]	77.55	Duration [ms]	2.07
Memory Throughput [%]	77.55	Elapsed Cycles [cycle]	1,208,822
L1/TEX Cache Throughput [%]	85.91	SM Active Cycles [cycle]	1,181,678.48
L2 Cache Throughput [%]	15.90	SM Frequency [Mhz]	584.99
DRAM Throughput [%]	3.14	DRAM Frequency [Ghz]	4.99

Balanced Throughput Compute and Memory are well-balanced: To reduce runtime, both computation and memory traffic must be reduced. Check both the » Compute Workload Analysis and » Memory Workload Analysis sections.

Roofline Analysis The ratio of peak float (fp32) to double (fp64) performance on this device is 32:1. The kernel achieved 8% of this device's fp32 peak performance and 0% of its fp64 peak performance. See the Kernel Profiling Guide for more details on roofline analysis.

GPU Throughput



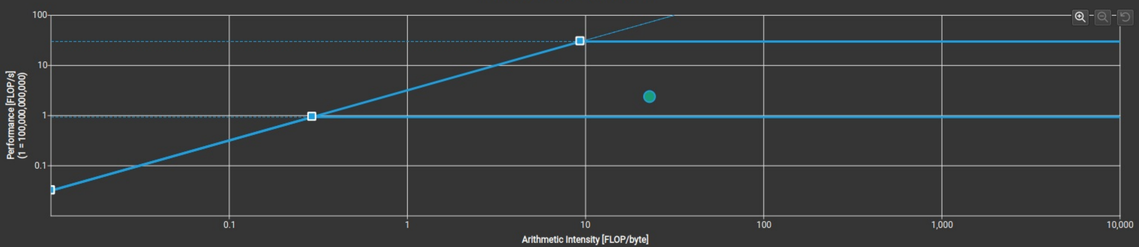
Compute Throughput Breakdown

SM: Inst Executed Pipe Lsu [%]	77.55
SM: Issue Active [%]	54.30
SM: Inst Executed [%]	54.30
SM: Pipe Alu Cycles Active [%]	44.00
SM: Inst Executed Pipe Xu [%]	27.92
SM: Mio Inst Issued [%]	26.84
SM: Pipe Fma Cycles Active [%]	24.58
SM: Mio2rf Writeback Active [%]	16.98
SM: Mio Pq Read Cycles Active [%]	5.54
SM: Mio Pq Write Cycles Active [%]	5.54
SM: Inst Executed Pipe Cbu Pred On Any [%]	5.11
SM: Inst Executed Pipe Adu [%]	2.98
SM: Inst Executed Pipe Uniform [%]	2.12
SM: Memory Throughput Internal Activity [%]	0
SM: Pipe Tensor Cycles Active [%]	0
SM: Pipe Shared Cycles Active [%]	0
SM: Pipe Fp64 Cycles Active [%]	0
IDC: Request Cycles Active [%]	0
SM: Instruction Throughput Internal Activity [%]	0
SM: Inst Executed Pipe Tex [%]	0
SM: Inst Executed Pipe Ipa [%]	0
SM: Inst Executed Pipe Fp16 [%]	0

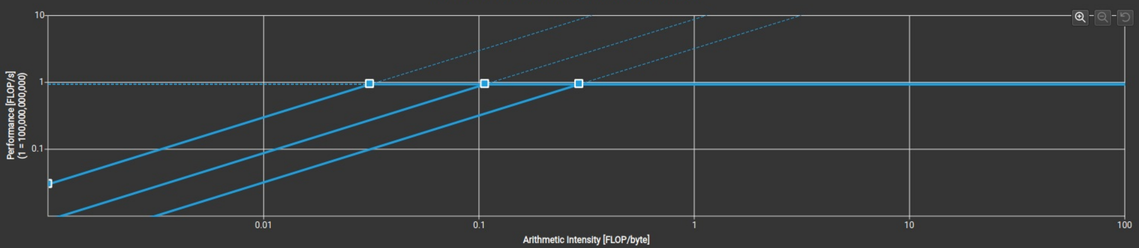
Memory Throughput Breakdown

L1: Lsuin Requests [%]	77.55
L1: Data Pipe Lsu Wavefronts [%]	42.99
L1: Lsu Writeback Active [%]	33.89
L1: M L1tex2xbar Req Cycles Active [%]	18.56
L2: Xbar2lts Cycles Active [%]	15.90
L2: T Sectors [%]	13.41
L1: Data Bank Reads [%]	10.90
L2: D Sectors [%]	3.61
GPU: Compute Memory Access Throughput Internal Activity [%]	3.54
L2: T Tag Requests [%]	3.54
DRAM: Cycles Active [%]	3.14
DRAM: Dram Sectors [%]	2.29
L1: Data Bank Writes [%]	2.07
L1: M Xbar2l1tex Read Sectors [%]	0.82
L2: Lts2xbar Cycles Active [%]	0.70
L2: D Sectors Fill Device [%]	0.04
L1: Texin Sm2tex Req Cycles Active [%]	0.00
L1: F Wavefronts [%]	0.00
L1: Tex Writeback Active [%]	0
L2: D Atomic Input Cycles Active [%]	0
GPU: Compute Memory Request Throughput Internal Activity [%]	0
L2: D Sectors Fill Sysmem [%]	0
L1: Data Pipe Tex Wavefronts [%]	0

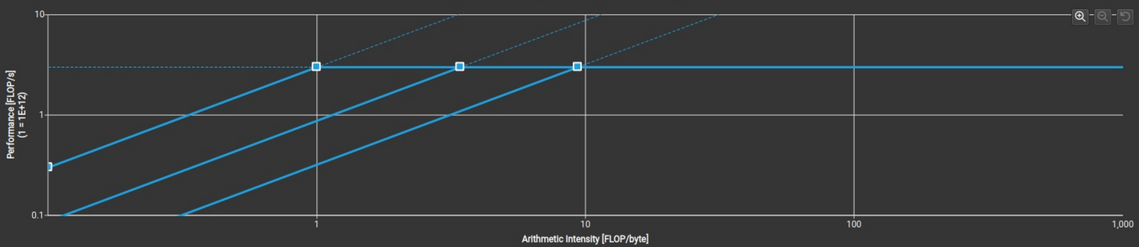
Floating Point Operations Roofline



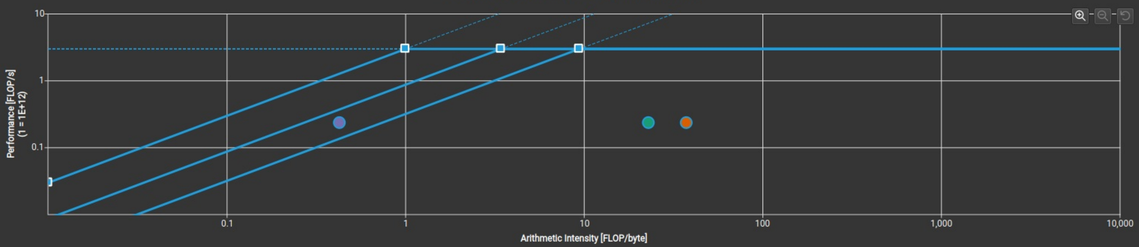
Floating Point Operations Roofline (Double Precision)



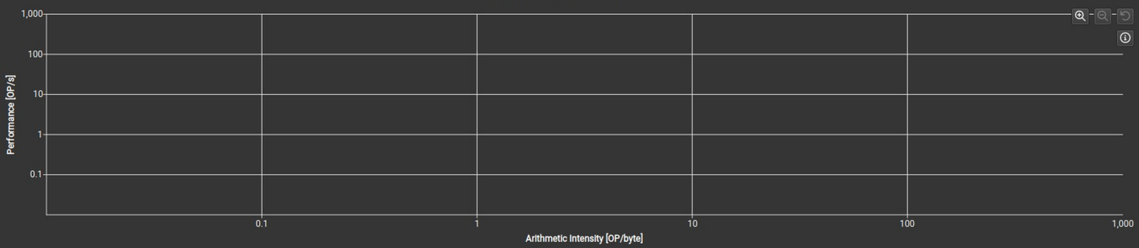
Floating Point Operations Roofline (Half Precision)



Floating Point Operations Roofline (Single Precision)



Tensor Core Operations Roofline



	# Operations	# Operations / Cycle	# Operations / s	Peak % ?eak Operations / Cycle	Peak Operations / s
Src:fp16,bf16,tf32 Dst:fp32	0	0	0	0	40,960	23,991.73
Src:fp16 Dst:fp16	0	0	0	0	40,960	23,991.73
Src:int1	0	0	0	0	81,920	47,983.46
Src:int4	0	0	0	0	81,920	47,983.46
Src:int8	0	0	0	0	81,920	47,983.46

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]	1,181,678.48	Average L1 Active Cycles [cycle]	1,181,678.48
Average L2 Active Cycles [cycle]	1,108,771.25	Average SMSP Active Cycles [cycle]	1,134,258.76
Average DRAM Active Cycles [cycle]	323,801	Total SM Elapsed Cycles [cycle]	48,413,816
Total L1 Elapsed Cycles [cycle]	48,413,816	Total L2 Elapsed Cycles [cycle]	56,534,912
Total SMSP Elapsed Cycles [cycle]	193,655,264	Total DRAM Elapsed Cycles [cycle]	82,554,880

L2 Slices Workload Imbalance Est. Speedup: 6.67% One or more L2 Slices have a much higher number of active cycles than the average number of active cycles. Additionally, other L2 Slices have a much lower number of active cycles than the average number of active cycles. Maximum instance value is 10.62% above the average, while the minimum instance value is 9.34% below the average.

Workload Distribution

	Average	Min	Max	Sum
SM Active Cycles	1,181,678.48	1,158,180	1,203,825	47,267,139
SMSP Active Cycles	1,134,258.76	1,053,541	1,181,601	181,481,401
L1 Active Cycles	1,181,678.48	1,158,180	1,203,825	47,267,139
L2 Active Cycles	1,108,771.25	1,005,243	1,240,520	35,480,680
DRAM Active Cycles	323,801	299,896	340,180	2,590,408