	Result	Size	Time	Cycles	GPU	SM Frequency	Process	Attributes
Current	558 - brute_force	(1024, 1, 1)x(1024, 1, 1)	1.62 ms	947,097	0 - Tesla T4	584.98 Mhz	[15952] exe	

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

Compare | Tools | View | Export

GPU Speed Of Light Throughput

All

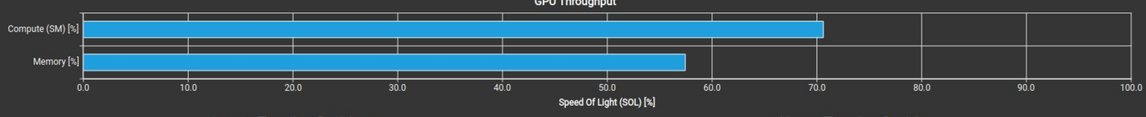
High-level overview of the throughput for compute and memory resources of the GPU. For each unit, the throughput reports the achieved percentage of utilization with respect to the theoretical maximum. Breakdowns show the throughput for each individual sub-metric of Compute and Memory to clearly identify 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 [%]	70.62	Duration [ms]	1.62
Memory Throughput [%]	57.45	Elapsed Cycles [cycle]	947,097
L1/TEX Cache Throughput [%]	70.28	SM Active Cycles [cycle]	921,547.93
L2 Cache Throughput [%]	19.30	SM Frequency [Mhz]	584.98
DRAM Throughput [%]	3.80	DRAM Frequency [Ghz]	4.98

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.

Roofline Analysis The ratio of peak float (fp32) to double (fp64) performance on this device is 32:1. The kernel achieved 10% 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



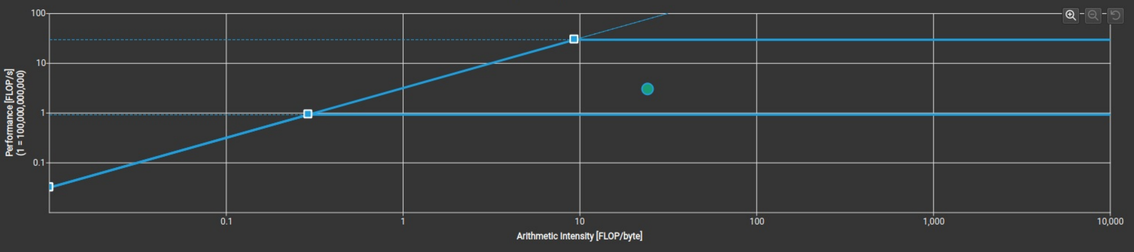
Compute Throughput Breakdown

SM: Issue Active [%]	70.62
SM: Inst Executed [%]	70.62
SM: Inst Executed Pipe Lsu [%]	57.45
SM: Pipe Alu Cycles Active [%]	55.90
SM: Inst Executed Pipe Adu [%]	42.64
IDC: Request Cycles Active [%]	41.65
SM: Pipe Fma Cycles Active [%]	37.84
SM: Inst Executed Pipe Xu [%]	35.88
SM: Mio Inst Issued [%]	33.36
SM: Mio Pq Write Cycles Active [%]	26.00
SM: Mio Pq Read Cycles Active [%]	26.00
SM: Mio2rf Writeback Active [%]	21.89
SM: Inst Executed Pipe Cbu Pred On Any [%]	6.04
SM: Inst Executed Pipe Uniform [%]	0.04
SM: Pipe Tensor Cycles Active [%]	0
SM: Pipe Shared Cycles Active [%]	0
SM: Pipe Fp64 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
SM: Inst Executed Pipe Fp16 [%]	0

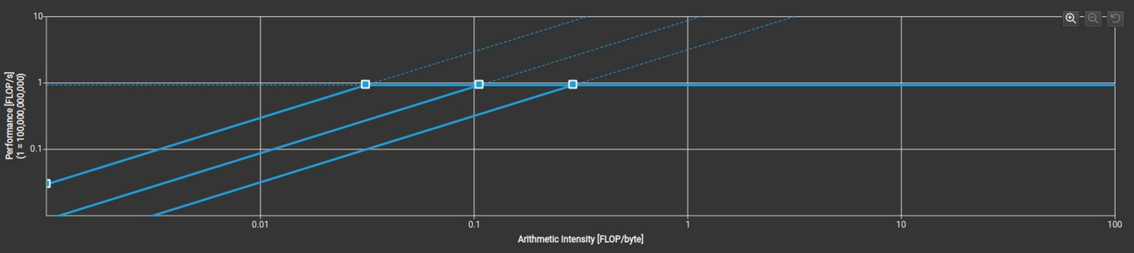
Memory Throughput Breakdown

L1: Lsuin Requests [%]	57.45
L1: Data Pipe Lsu Wavefronts [%]	35.17
L1: Lsu Writeback Active [%]	22.88
L1: M L1tex2xbar Req Cycles Active [%]	22.67
L2: Xbar2lts Cycles Active [%]	19.30
L2: T Sectors [%]	16.33
L1: Data Bank Reads [%]	13.57
L2: D Sectors [%]	4.38
GPU: Compute Memory Access Throughput Internal Activity [%]	4.32
L2: T Tag Requests [%]	4.32
DRAM: Cycles Active [%]	3.80
DRAM: Dram Sectors [%]	2.77
L1: Data Bank Writes [%]	2.65
L1: M Xbar2l1tex Read Sectors [%]	0.99
L2: Lts2xbar Cycles Active [%]	0.88
L2: D Sectors Fill Device [%]	0.06
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

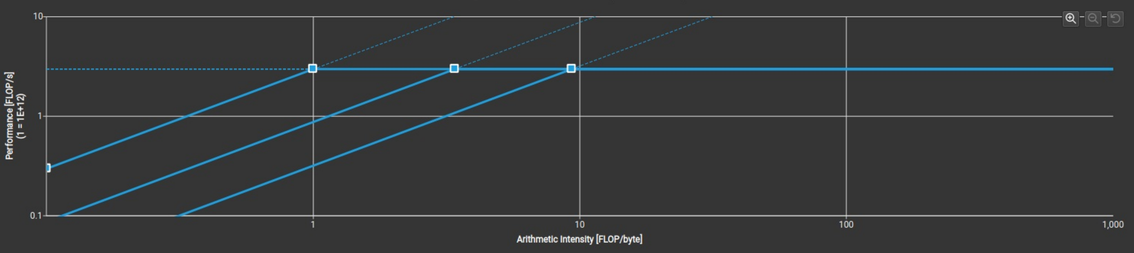
Floating Point Operations Roofline



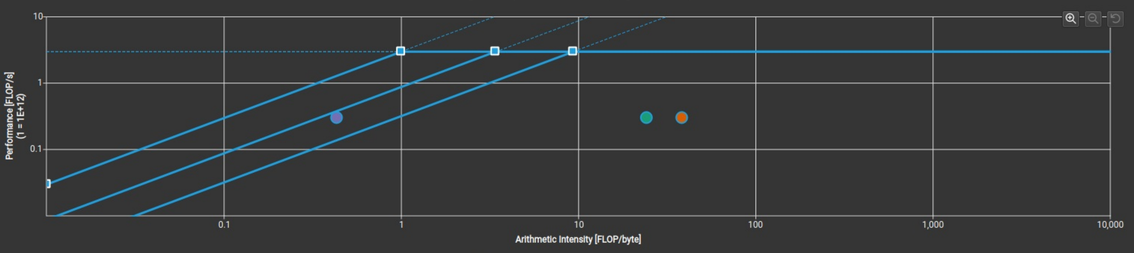
Floating Point Operations Roofline (Double Precision)



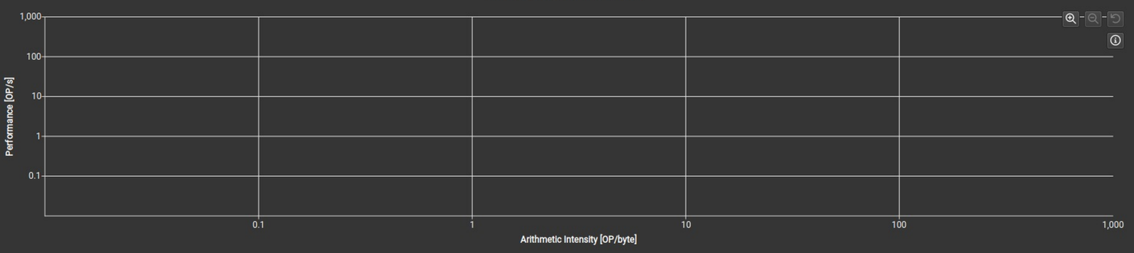
Floating Point Operations Roofline (Half Precision)



Floating Point Operations Roofline (Single Precision)



Tensor Core Operations Roofline



	# Operations	# Operations / Cycle	# Operations / s	Peak % ?eak Operations / Cycle	Peak Operations / Cycle	Peak Operations / s
Src:fp16,bf16,tf32 Dst:fp32	0	0	0	0	40,960	23,828.09
Src:fp16 Dst:fp16	0	0	0	0	40,960	23,828.09
Src:int1	0	0	0	0	81,920	47,656.19
Src:int4	0	0	0	0	81,920	47,656.19
Src:int8	0	0	0	0	81,920	47,656.19

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]	921,547.93	Average L1 Active Cycles [cycle]	921,547.93
Average L2 Active Cycles [cycle]	969,520.03	Average SMSP Active Cycles [cycle]	883,611.76
Average DRAM Active Cycles [cycle]	306,623.50	Total SM Elapsed Cycles [cycle]	37,673,704
Total L1 Elapsed Cycles [cycle]	37,673,704	Total L2 Elapsed Cycles [cycle]	44,294,240
Total SMSP Elapsed Cycles [cycle]	150,694,816	Total DRAM Elapsed Cycles [cycle]	64,507,904

L2 Slices Workload Imbalance Est. Speedup: 5.11% One or more L2 Slices have a much lower number of active cycles than the average number of active cycles. Maximum instance value is 7.29% above the average, while the minimum instance value is 10.42% below the average.

Workload Distribution

	Average	Min	Max	Sum
SM Active Cycles	921,547.93	903,225	940,126	36,861,917
SMSP Active Cycles	883,611.76	844,199	929,447	141,377,882
L1 Active Cycles	921,547.93	903,225	940,126	36,861,917
L2 Active Cycles	969,520.03	868,464	1,045,791	31,024,641
DRAM Active Cycles	306,623.50	278,084	339,472	2,452,988