

GPU Speed Of Light Throughput

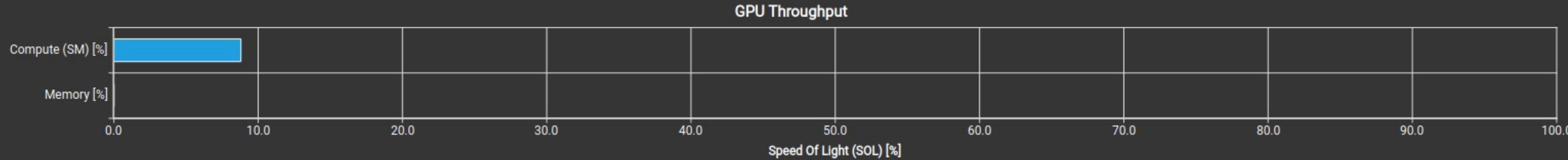
All

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.

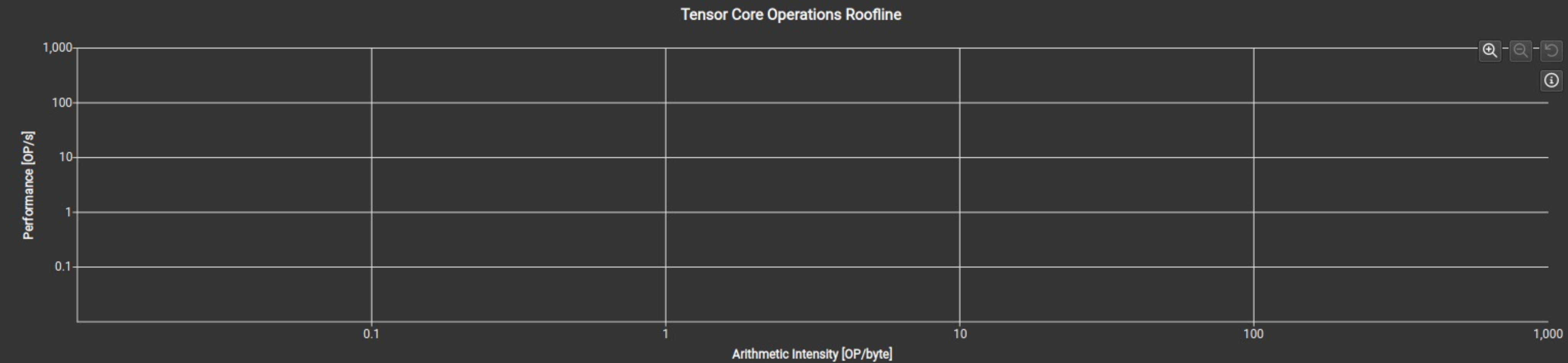
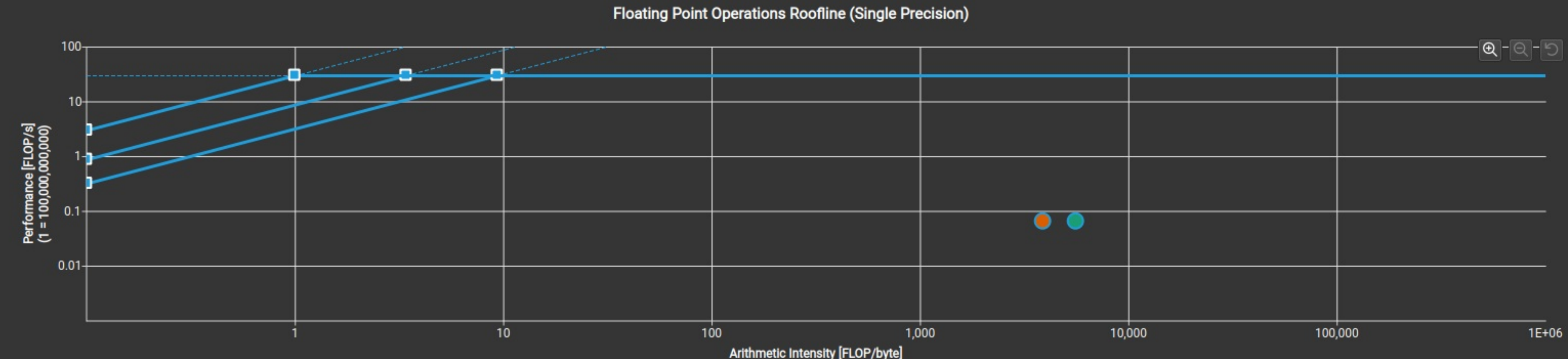
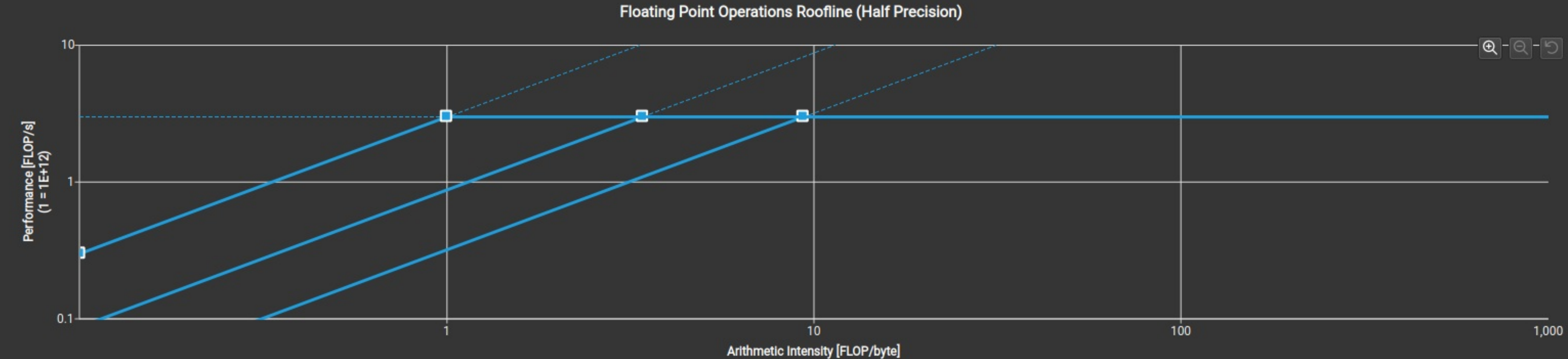
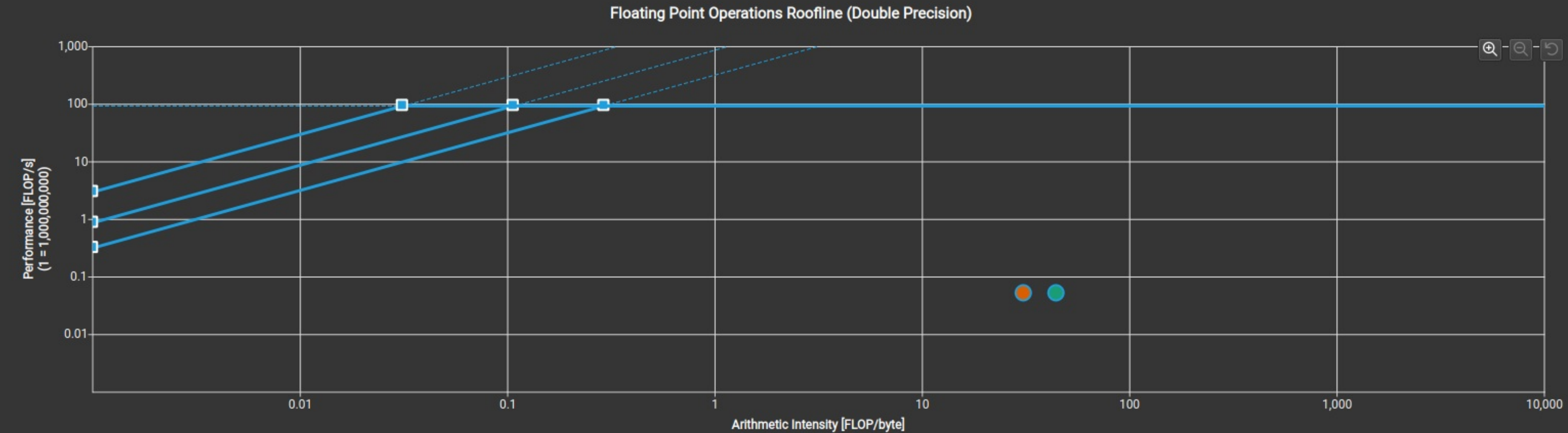
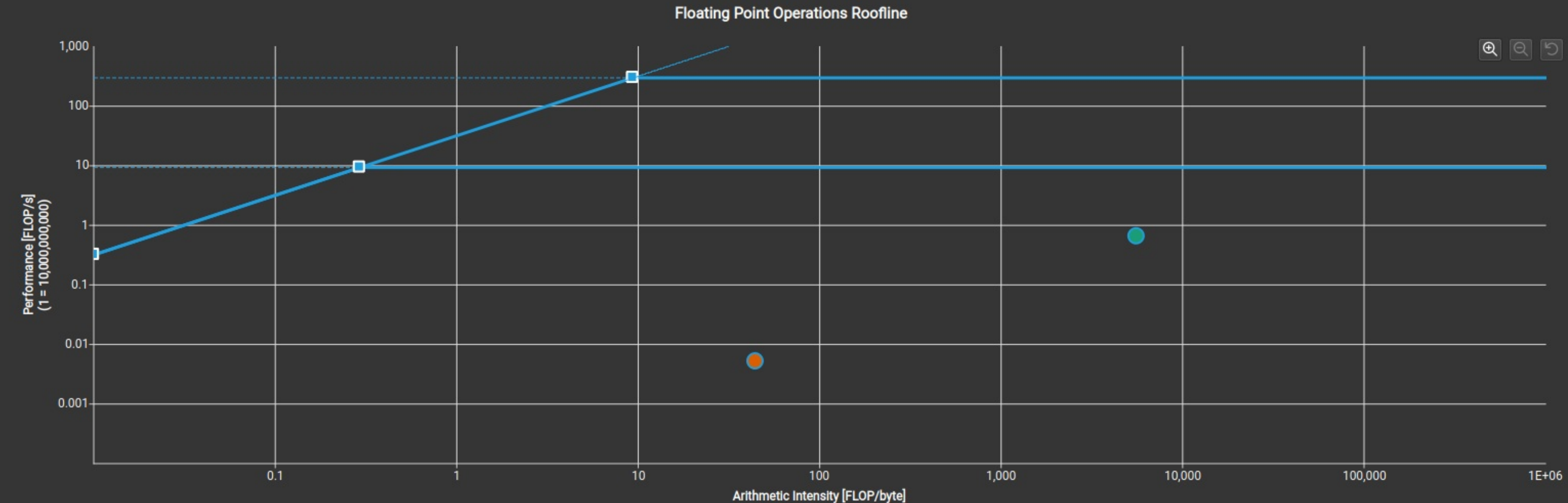
Compute (SM) Throughput [%]	8.81	Duration [ms]	9.39
Memory Throughput [%]	0.00	Elapsed Cycles [cycle]	5,491,400
L1/TEX Cache Throughput [%]	0.01	SM Active Cycles [cycle]	541,596.55
L2 Cache Throughput [%]	0.00	SM Frequency [Mhz]	584.99
DRAM Throughput [%]	0.00	DRAM Frequency [ghz]	5.00

Small Grid This kernel grid is too small to fill the available resources on this device, resulting in only 0.1 full waves across all SMs. Look at [Launch Statistics](#) for more details.

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



Compute Throughput Breakdown		Memory Throughput Breakdown	
SM: Pipe Alu Cycles Active [%]	8.81	L2: T Sectors [%]	0.00
SM: Issue Active [%]	6.59	L1: M L1tex2xbar Req Cycles Active [%]	0.00
SM: Inst Executed [%]	6.59	L2: Xbar2lts Cycles Active [%]	0.00
SM: Inst Executed Pipe Adu [%]	3.88	GPU: Compute Memory Access Throughput Internal Activity [%]	0.00
IDC: Request Cycles Active [%]	2.57	L2: D Sectors Fill Device [%]	0.00
SM: Pipe Fma Cycles Active [%]	1.93	L2: T Tag Requests [%]	0.00
SM: Mio Pq Write Cycles Active [%]	1.31	L2: D Sectors [%]	0.00
SM: Mio Inst Issued [%]	1.30	DRAM: Cycles Active [%]	0.00
SM: Mio Pq Read Cycles Active [%]	1.29	L1: Lsuin Requests [%]	0.00
SM: Inst Executed Pipe Cbu Pred On Any [%]	0.97	DRAM: Dram Sectors [%]	0.00
SM: Mio2rf Writeback Active [%]	0.64	L1: Data Pipe Lsu Wavefronts [%]	0.00
SM: Pipe Fp64 Cycles Active [%]	0.07	L2: Lts2xbar Cycles Active [%]	0.00
SM: Inst Executed Pipe Lsu [%]	0.00	L1: Data Bank Reads [%]	0.00
SM: Inst Executed Pipe Xu [%]	0.00	L1: Data Bank Writes [%]	0.00
SM: Memory Throughput Internal Activity [%]	0	L1: Lsu Writeback Active [%]	0.00
SM: Pipe Tensor Cycles Active [%]	0	L1: Texin Sm2tex Req Cycles Active [%]	0.00
SM: Pipe Shared Cycles Active [%]	0	L1: F Wavefronts [%]	0.00
SM: Instruction Throughput Internal Activity [%]	0	L1: Tex Writeback Active [%]	0
SM: Inst Executed Pipe Uniform [%]	0	L2: D Atomic Input Cycles Active [%]	0
SM: Inst Executed Pipe Tex [%]	0	L1: M Xbar2l1tex Read Sectors [%]	0
SM: Inst Executed Pipe lpa [%]	0	L2: D Sectors Fill System [%]	0
SM: Inst Executed Pipe Fp16 [%]	0	L1: Data Pipe Tex Wavefronts [%]	0
		GPU: Compute Memory Request Throughput Internal Activity [%]	0



	# Operations	# Operations / Cycle	# Operations / s	Peak %	Peak Operations / Cycle	Peak Operations / s
Src:fp16,bf16,tf32 Dst:fp32	0	0	0	0	40,960	23,957.45
Src:fp16 Dst:fp16	0	0	0	0	40,960	23,957.45
Src:int1	0	0	0	0	81,920	47,914.91
Src:int4	0	0	0	0	81,920	47,914.91
Src:int8	0	0	0	0	81,920	47,914.91

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]	541,596.55	Average L1 Active Cycles [cycle]	541,596.55
Average L2 Active Cycles [cycle]	2,837.66	Average SMSP Active Cycles [cycle]	519,678.04
Average DRAM Active Cycles [cycle]	168	Total SM Elapsed Cycles [cycle]	219,620,216
Total L1 Elapsed Cycles [cycle]	219,620,216	Total L2 Elapsed Cycles [cycle]	256,828,544
Total SMSP Elapsed Cycles [cycle]	878,480,864	Total DRAM Elapsed Cycles [cycle]	375,191,552

SMs Workload Imbalance Est. Speedup: 8.89% One or more SMs have a much lower number of active cycles than the average number of active cycles. Maximum instance value is 90.13% above the average, while the minimum instance value is 100.00% below the average.

SMSPs Workload Imbalance Est. Speedup: 8.57% One or more SMSPs have a much lower number of active cycles than the average number of active cycles. Maximum instance value is 90.53% above the average, while the minimum instance value is 100.00% below the average.

L1 Slices Workload Imbalance Est. Speedup: 8.89% One or more L1 Slices have a much lower number of active cycles than the average number of active cycles. Maximum instance value is 90.13% above the average, while the minimum instance value is 100.00% below the average.

Workload Distribution				
	Average	Min	Max	Sum
SM Active Cycles	541,596.55	0	5,488,814	21,663,862
SMSP Active Cycles	519,678.04	0	5,488,790	83,148,486
L1 Active Cycles	541,596.55	0	5,488,814	21,663,862
L2 Active Cycles	2,837.66	1,410	4,334	90,805
DRAM Active Cycles	168	112	264	1,344