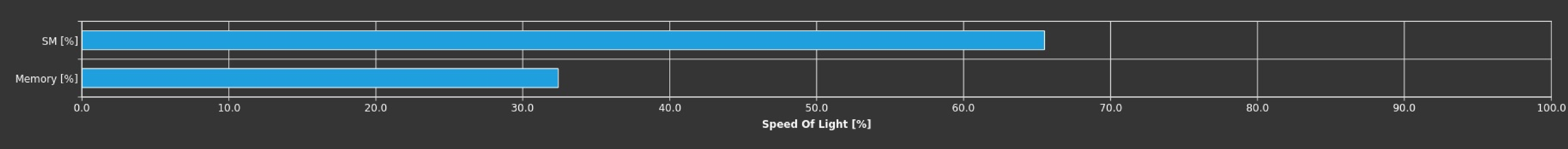
Page: Details | Launch: txcall_cal_flux_residual_kernel_4 | Add Baseline | Apply Rules | Save as PDF

Current 206 - ptxcall_cal_flux_residual_kernel_4 (9600, 1, 1) **Time:** 2.30 msecond **Cycles:** 1,945,424 **Regs:** 168 **GPU:** Quadro M5000 **SM Frequency:** 845.40 Mhz **CC:** 5.2

GPU Speed Of Light

SOL SM [%]	65.51	Duration [msecond]	2.30
SOL Memory [%]	32.42	Elapsed Cycles [cycle]	19,45,424.75
SOL TEX [%]	14.37	SM Frequency [Mhz]	845.40
SOL L2 [%]	32.42	Memory Frequency [Ghz]	6.60
SOL FB [%]	6.08		

GPU Utilization



Recommendations

Bottleneck [Warning] Compute is more heavily utilized than Memory: Look at `Compute Workload Analysis` report section to see what the compute pipelines are spending their time doing. Also, consider whether any computation is redundant and could be reduced.

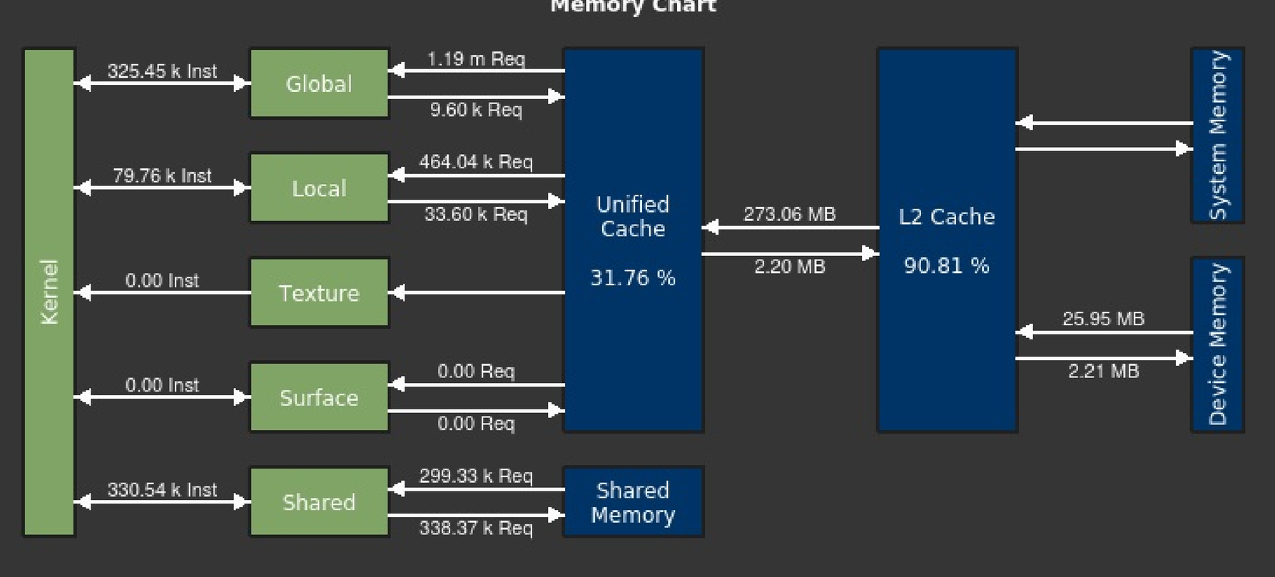
Compute Workload Analysis

Executed Ipc Elapsed [inst/cycle]	0.23	SM Busy [%]	65.51
Executed Ipc Active [inst/cycle]	0.29	Issue Slots Busy [%]	4.87
Issued Ipc Active [inst/cycle]	0.29		

Memory Workload Analysis

Memory Throughput [Gbyte/second]	11.95	Mem Busy [%]	32.42
L1 Hit Rate [%]	31.76	Max Bandwidth [%]	31.69
L2 Hit Rate [%]	90.81	Mem Pipes Busy [%]	1.68

Memory Chart



Shared Memory

	Instructions	Requests	% Peak	Bank Conflicts
Shared Load	1,58,604	2,99,328	1.24	28
Shared Store	1,71,940	3,38,374	1.40	0
Shared Atomic	0	-	-	-
Total	3,30,544	6,37,702	2.64	28

First-Level Cache

	Instructions	SM->TEX Requests	% Peak	Hit Rate	TEX->L2 Requests	% Peak	L2->TEX Returns	% Peak	TEX->SM Returns	% Peak
Global Load Cached	3,24,246	79,672	0.33	38.75	-	-	13,354	0.02	28,10,110	4.51
Global Load Uncached		11,06,304	4.57	-	-	-	84,64,629	13.60		
Local Load Cached	75,564	4,800	0.02	0	-	-	4,800	0.01		
Local Load Uncached		4,59,240	1.90	-	-	-	4,64,796	0.75		
Surface Load	0	0	0	-	-	-	0	0		
Texture Load	0									
Global Store	1,200	9,600	0.04	-	38,400	0.06	-	-	-	-
Local Store	4,200	33,600	0.14	-	33,600	0.05	-	-	-	-
Surface Store	0	0	0	-	0	0	-	-	-	-
Global Reduction	0	0	0	-	0	0	-	-	-	-
Surface Reduction	0	0	0	-	0	0	-	-	-	-
Global Atomic	0	0	0	-	0	0	0	0	0	0
Global Atomic Cas		0	0	-					0	0
Surface Atomic	0	0	0	-	0	0	0	0	0	0
Surface Atomic Cas		0	0	-					0	0
Loads	3,99,810	16,50,016	6.82	0.09	-	-	89,47,579	14.37	28,10,110	4.51
Stores	5,400	43,200	0.18	-	72,000	0.12	-	-	-	-
Total	4,05,210	16,93,216	7.00	0.09	72,000	0.12	89,47,579	14.37	28,10,110	4.51

Second-Level Cache

	TEX->L2 Requests	% Peak	L2->TEX Returns	% Peak	Total Bytes	Total Throughput
Global Load Cached	-	-	13,354	0.05	4,27,328	18,56,99,187.90
Global Load Uncached	-	-	84,64,629	29.98	27,08,68,128	1,17,70,81,57,192.12
Local Load Cached	-	-	4,800	0.02	1,53,600	6,67,48,247.86
Local Load Uncached	-	-	4,64,796	1.65	1,48,73,472	6,46,33,99,710.76
Surface Load	-	-	0	0	0	0
Texture Load						
Global Store	38,400	0.14	-	-	12,28,800	53,39,85,982.87
Local Store	33,600		-	-	10,75,200	46,72,37,735.01
Surface Store	0	0	-	-	0	0
Global Reduction	0	0	-	-	0	0
Surface Reduction	0	0	-	-	0	0
Global Atomic	0	0	0	0	0	0
Global Atomic Cas						
Surface Atomic	0	0	0	0	0	0
Surface Atomic Cas						
Loads	-	-	89,47,579	31.69	28,63,22,528	1,24,42,40,04,338.64
Stores	72,000	0.14	-	-	23,04,000	1,00,12,23,717.88
Total	72,000	0.14	89,47,579	31.69	28,86,26,528	1,25,42,52,28,056.51

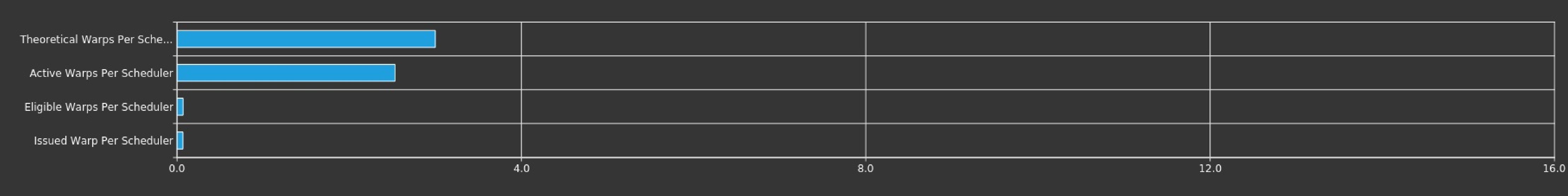
Device Memory

	L2<->FB Sectors	% Peak	Bytes	Throughput
Load	8,50,492	5.60	2,72,15,744	11,82,68,43,920.35
Store	72,465	0.48	23,18,880	1,00,76,89,954.39
Total	9,22,957	6.08	2,95,34,624	12,83,45,33,874.74

Scheduler Statistics

Active Warps Per Scheduler [warp/cycle]	2.53	Instructions Per Active Issue Slot [inst/issue]	1.08
Eligible Warps Per Scheduler [warp/cycle]	0.07	No Eligible [%]	92.21
Issued Warp Per Scheduler [issue/cycle]	0.07	One or More Eligible [%]	6.87

Warps Per Scheduler



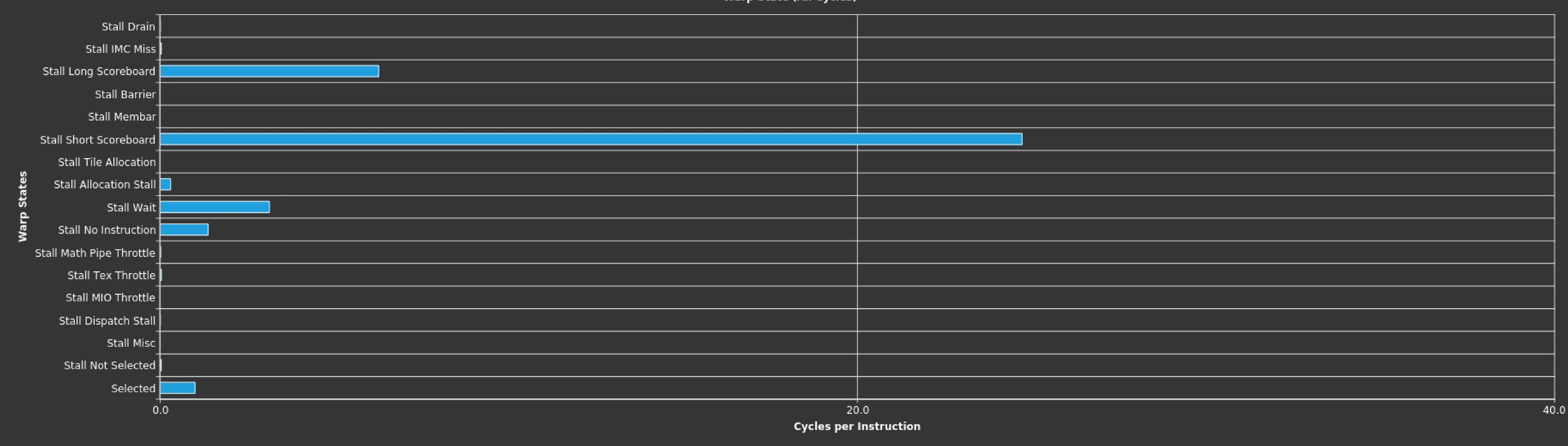
Recommendations

Issue Slot Utilization [Warning] Every scheduler is capable of issuing two instructions per cycle, but for this kernel each scheduler only issues an instruction every 14.6 cycles. This might leave hardware resources underutilized and may lead to less optimal performance. Out of the maximum of 16 warps per scheduler, this kernel allocates an average of 2.53 active warps per scheduler, but only an average of 0.07 warps were eligible per cycle. Eligible warps are the subset of active warps that are ready to issue their next instruction. Every cycle with no eligible warp results in no instruction being issued and the issue slot remains unused. To increase the number of eligible warps either increase the number of active warps or reduce the time the active warps are stalled.

Warp State Statistics

Warp Cycles Per Issued Instruction [cycle/inst]	34.16	Avg. Active Threads Per Warp [thread/inst]	28.14
Warp Cycles Per Issue Active [cycle/issue]	36.83	Avg. Not Predicated Off Threads Per Warp [thread/inst]	27.62
Warp Cycles Per Executed Instruction [cycle/inst]	34.18		

Warp State (All Cycles)



Recommendations

CPI Stall 'Short Scoreboard' [Warning] On average each warp of this kernel spends 24.7 cycles being stalled waiting for a scoreboard dependency on an MIO operation (not to TEX or L1). This represents about 67.1% of the total average of 36.8 cycles between issuing two instructions. The primary reason for a high number of stalls due to short scoreboards is typically memory operations to shared memory, but other contributors include frequent execution of special math instructions (e.g. MUFU) or dynamic branching (e.g. BRX, JMX). Consult the Memory Workload Analysis section to verify if there are shared memory operations and reduce bank conflicts, if reported.

CPI Stall 'Drain' Warp stall analysis for 'Drain' issues — Apply

CPI Stall 'LG Throttle' Warp stall analysis for 'LG Throttle' issues — Apply

CPI Stall 'Long Scoreboard' Warp stall analysis for 'Long Scoreboard' issues — Apply

CPI Stall 'Math Pipe Throttle' Warp stall analysis for 'Math Pipe Throttle' issues — Apply

CPI Stall 'Membar' Warp stall analysis for 'Membar' issues — Apply

CPI Stall 'MIO Throttle' Warp stall analysis for 'MIO Throttle' issues — Apply

CPI Stall 'Misc' Warp stall analysis for 'Misc' issues — Apply

CPI Stall 'No Instructions' Warp stall analysis for 'No Instructions' issues — Apply

CPI Stall 'Not Selected' Warp stall analysis for 'Not Selected' issues — Apply

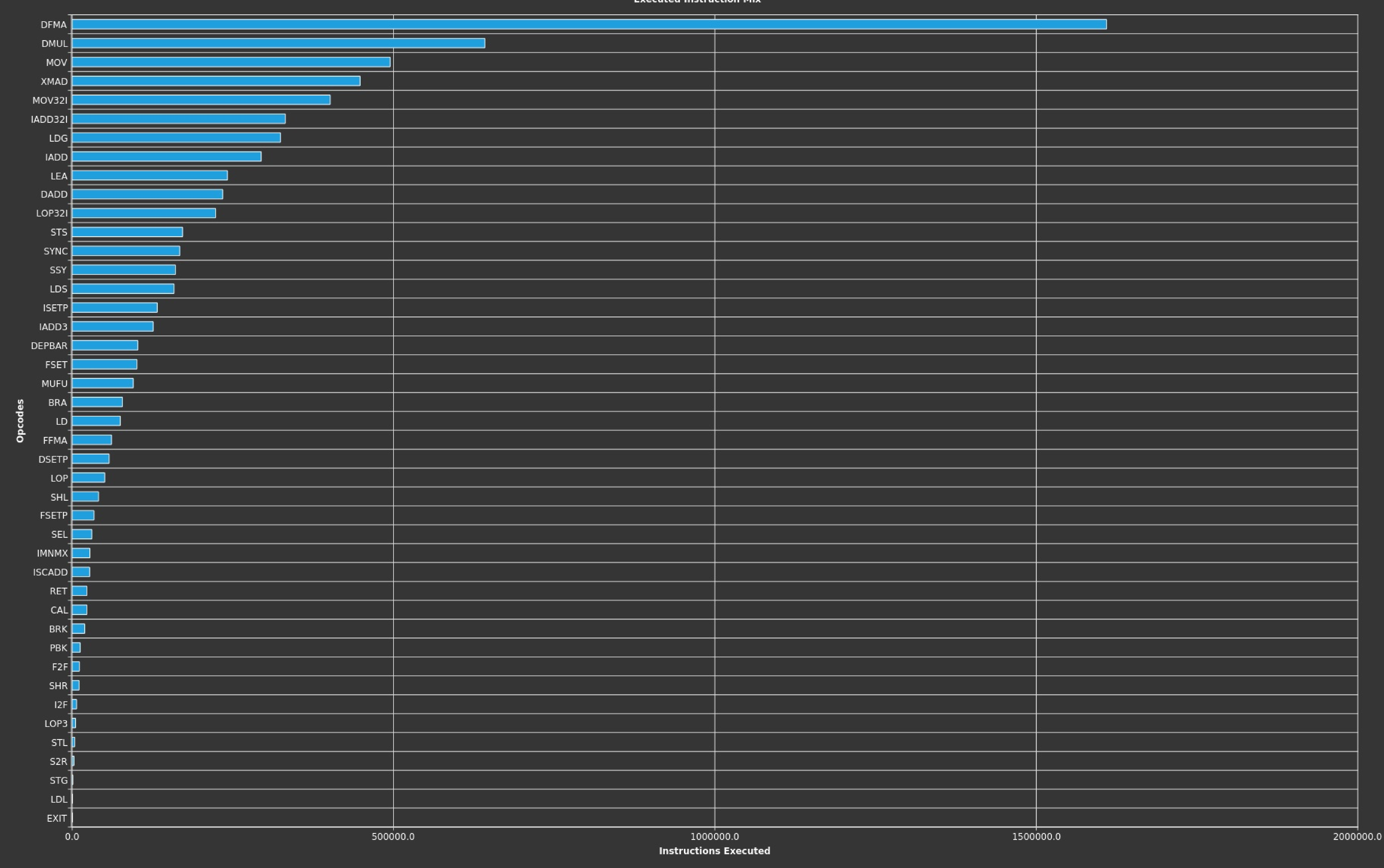
CPI Stall 'Sleeping' Warp stall analysis for 'Sleeping' issues — Apply

CPI Stall 'TEX Throttle' Warp stall analysis for 'TEX Throttle' issues — Apply

Instruction Statistics

Executed Instructions [inst]	70,67,939	Avg. Executed Instructions Per Scheduler [inst]	1,10,436.55
Issued Instructions [inst]	70,71,571	Avg. Issued Instructions Per Scheduler [inst]	1,10,493.30

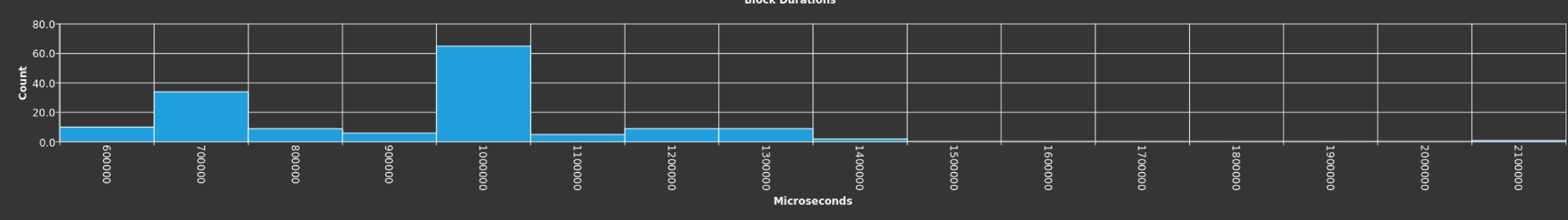
Executed Instruction Mix



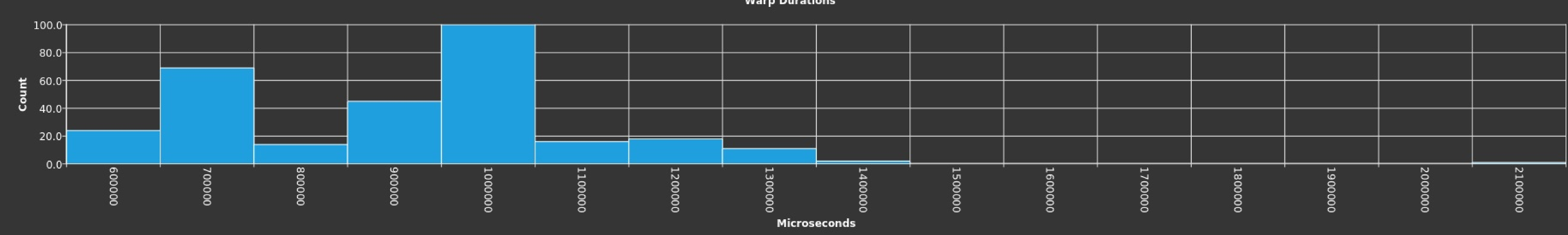
Launch Statistics

Grid Size	150	Registers Per Thread [register/thread]	168
Block Size	64	Static Shared Memory Per Block [Kbyte/block]	12
Threads [thread]	9,600	Dynamic Shared Memory Per Block [byte/block]	0
Waves Per SM	1.56	Shared Memory Configuration Size [Kbyte]	48

Block Durations



Warp Durations



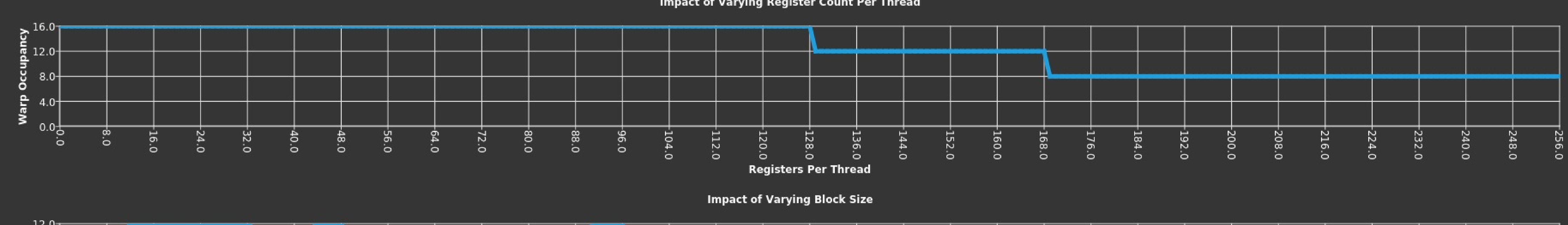
Recommendations

Launch Configuration Kernel launch configuration analysis — Apply

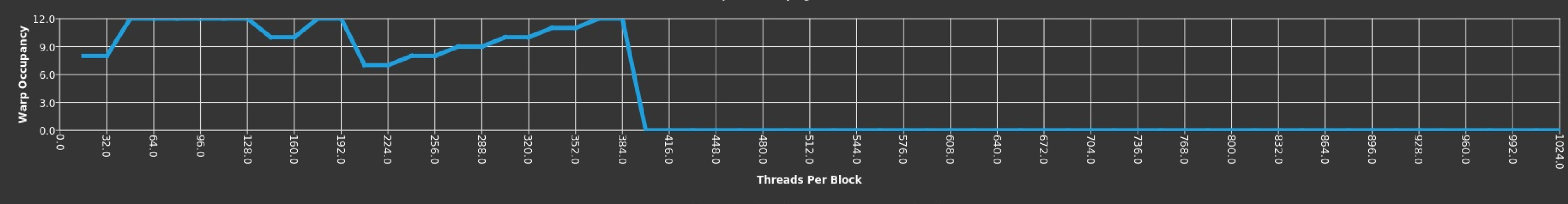
Occupancy

Theoretical Occupancy [%]	18.75	Block Limit Registers [register]	6
Theoretical Active Warps per SM [warp/cycle]	12	Block Limit Local Mem [byte]	8
Achieved Occupancy [%]	15.64	Block Limit Warps [warp]	32
Achieved Active Warps Per SM [warp/cycle]	10.01	Block Limit SM [block]	32

Impact of Varying Register Count Per Thread



Impact of Varying Block Size



Impact of Varying Shared Memory Usage Per Block