

COA LAB ASSIGNMENT 4

Team Members:

21CS01024

21CS01037

21CS01043

21CS01050

21CS01015

Configuration: SM2_GTX480

Scheduler: gto

L1D_total_cache_miss_rate = 0.8225

L2_total_cache_miss_rate = 0.1651

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 52 sec (172 sec)

gpgpu_simulation_rate = 178859 (inst/sec)

gpgpu_simulation_rate = 3052 (cycle/sec)

gpgpu_silicon_slowdown = 229357x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler: lrr

L1D_total_cache_miss_rate = 0.8230

L2_total_cache_miss_rate = 0.2493

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 41 sec (161 sec)

gpgpu_simulation_rate = 191079 (inst/sec)

gpgpu_simulation_rate = 3343 (cycle/sec)

gpgpu_silicon_slowdown = 209392x

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler: two level

L1D_total_cache_miss_rate = 0.8209

L2_total_cache_miss_rate = 0.2537

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 35 sec (155 sec)

gpgpu_simulation_rate = 198476 (inst/sec)

gpgpu_simulation_rate = 3453 (cycle/sec)

gpgpu_silicon_slowdown = 202722x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Configuration: SM3_KEPLER_TITAN

Scheduler:gto

L1D_total_cache_accesses = 0

L1D_total_cache_misses = 0

L2_total_cache_miss_rate = 0.0906

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 18 sec (138 sec)

gpgpu_simulation_rate = 222926 (inst/sec)

gpgpu_simulation_rate = 1933 (cycle/sec)

gpgpu_silicon_slowdown = 433005x

Kernel Executed 10 times

Result stored in result.txt

Scheduler:lrr

L1D_total_cache_accesses = 0

L1D_total_cache_misses = 0

L2_total_cache_miss_rate = 0.0913

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 15 sec (135 sec)

gpgpu_simulation_rate = 227880 (inst/sec)

gpgpu_simulation_rate = 2037 (cycle/sec)

gpgpu_silicon_slowdown = 410898x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:two level

L1D_total_cache_accesses = 0

L1D_total_cache_misses = 0

L2_total_cache_miss_rate = 0.0909

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 17 sec (137 sec)

gpgpu_simulation_rate = 224553 (inst/sec)

gpgpu_simulation_rate = 1990 (cycle/sec)

gpgpu_silicon_slowdown = 420603x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Configuration:SM6_TITANX

Scheduler:gto

L1D_total_cache_accesses = 0

L1D_total_cache_misses = 0

L2_total_cache_miss_rate = 0.0218

gpgpu_simulation_time = 0 days, 0 hrs, 3 min, 14 sec (194 sec)

gpgpu_simulation_rate = 158576 (inst/sec)

gpgpu_simulation_rate = 1564 (cycle/sec)

gpgpu_silicon_slowdown = 906010x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:lr

L1D_total_cache_accesses = 0

L1D_total_cache_misses = 0

L2_total_cache_miss_rate = 0.0258

gpgpu_simulation_time = 0 days, 0 hrs, 3 min, 9 sec (189 sec)

gpgpu_simulation_rate = 162771 (inst/sec)

gpgpu_simulation_rate = 1663 (cycle/sec)

gpgpu_silicon_slowdown = 852074x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:two level

L1D_total_cache_accesses = 0

L1D_total_cache_misses = 0

L2_total_cache_miss_rate = 0.0255

gpgpu_simulation_time = 0 days, 0 hrs, 3 min, 24 sec (204 sec)

gpgpu_simulation_rate = 150803 (inst/sec)

gpgpu_simulation_rate = 1537 (cycle/sec)

gpgpu_silicon_slowdown = 921925x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Configuration:SM7_QV100

Scheduler:gto

L1D_total_cache_miss_rate = 0.4894

L2_total_cache_miss_rate = 0.0000

gpgpu_simulation_time = 0 days, 0 hrs, 4 min, 14 sec (254 sec)

gpgpu_simulation_rate = 121117 (inst/sec)

gpgpu_simulation_rate = 714 (cycle/sec)

gpgpu_silicon_slowdown = 1585434x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:lrr

L1D_total_cache_miss_rate = 0.4895

L2_total_cache_miss_rate = 0.0000

gpgpu_simulation_time = 0 days, 0 hrs, 4 min, 6 sec (246 sec)

gpgpu_simulation_rate = 125056 (inst/sec)

gpgpu_simulation_rate = 739 (cycle/sec)

gpgpu_silicon_slowdown = 1531799x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:two level

L1D_total_cache_miss_rate = 0.4895

L2_total_cache_miss_rate = 0.0000

gpgpu_simulation_time = 0 days, 0 hrs, 4 min, 16 sec (256 sec)

gpgpu_simulation_rate = 120171 (inst/sec)

gpgpu_simulation_rate = 710 (cycle/sec)

gpgpu_silicon_slowdown = 1594366x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Configuration:SM6_TITANX

Scheduler:gto

L1D_total_cache_miss_rate = 0.4936

L2_total_cache_miss_rate = 0.0317

gpgpu_simulation_time = 0 days, 0 hrs, 4 min, 21 sec (261 sec)

gpgpu_simulation_rate = 117869 (inst/sec)

gpgpu_simulation_rate = 391 (cycle/sec)

gpgpu_silicon_slowdown = 3069053x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:lr

L1D_total_cache_miss_rate = 0.4936

L2_total_cache_miss_rate = 0.0301

gpgpu_simulation_time = 0 days, 0 hrs, 4 min, 20 sec (260 sec)

gpgpu_simulation_rate = 118322 (inst/sec)

gpgpu_simulation_rate = 393 (cycle/sec)

gpgpu_silicon_slowdown = 3053435x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:two level

L1D_total_cache_miss_rate = 0.4937

L2_total_cache_miss_rate = 0.0310

gpgpu_simulation_time = 0 days, 0 hrs, 4 min, 30 sec (270 sec)

gpgpu_simulation_rate = 113940 (inst/sec)

gpgpu_simulation_rate = 378 (cycle/sec)

gpgpu_silicon_slowdown = 3174603x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Configuration:SM75_RTX2070

Scheduler:gto

L1D_total_cache_miss_rate = 0.5869

L2_total_cache_miss_rate = 0.0000

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 35 sec (155 sec)

gpgpu_simulation_rate = 198476 (inst/sec)

gpgpu_simulation_rate = 1628 (cycle/sec)

gpgpu_silicon_slowdown = 838452x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:lrr

L1D_total_cache_miss_rate = 0.5865

L2_total_cache_miss_rate = 0.0000

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 38 sec (158 sec)

gpgpu_simulation_rate = 194707 (inst/sec)

gpgpu_simulation_rate = 1617 (cycle/sec)

gpgpu_silicon_slowdown = 844155x

Kernel Executed 10 times

Result stored in result.txt

GPGPU-Sim: *** exit detected ***

Scheduler:two level

L1D_total_cache_miss_rate = 0.5875

L2_total_cache_miss_rate = 0.0000

gpgpu_simulation_time = 0 days, 0 hrs, 2 min, 52 sec (172 sec)

gpgpu_simulation_rate = 178859 (inst/sec)

gpgpu_simulation_rate = 1462 (cycle/sec)

gpgpu_silicon_slowdown = 933652x

Kernel Executed 10 times

Result stored in result.txt

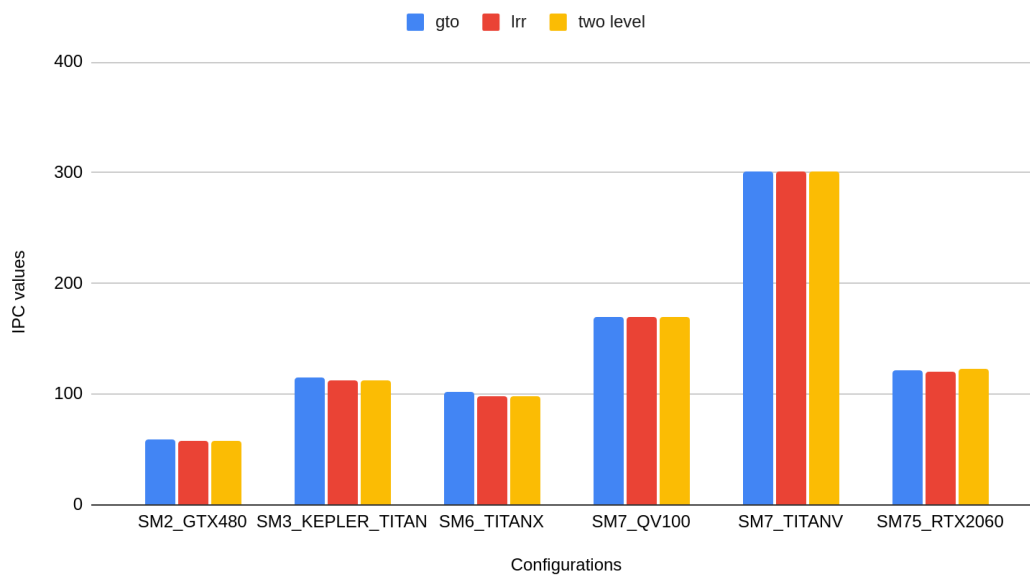
GPGPU-Sim: *** exit detected ***

Question-1:

| | | | |
|------------|-------------|-------------|-------------|
| BFS | | | |
| IPC values | | | |
| | gto | lrr | two level |
| SM2_GTX480 | 58.60386632 | 57.15794197 | 57.47929337 |

| | | | |
|------------------|-------------|-------------|-------------|
| SM3_KEPLER_TITAN | 115.3264356 | 111.8703976 | 112.8407035 |
| SM6_TITANX | 101.3913043 | 97.87793145 | 98.1151594 |
| SM7_QV100 | 169.6316527 | 169.2232747 | 169.2549296 |
| SM7_TITANV | 301.455243 | 301.0737913 | 301.4285714 |
| SM75_RTX2060 | 121.9140049 | 120.4124923 | 122.3385773 |

Diff warp Schedulers(for BFS)

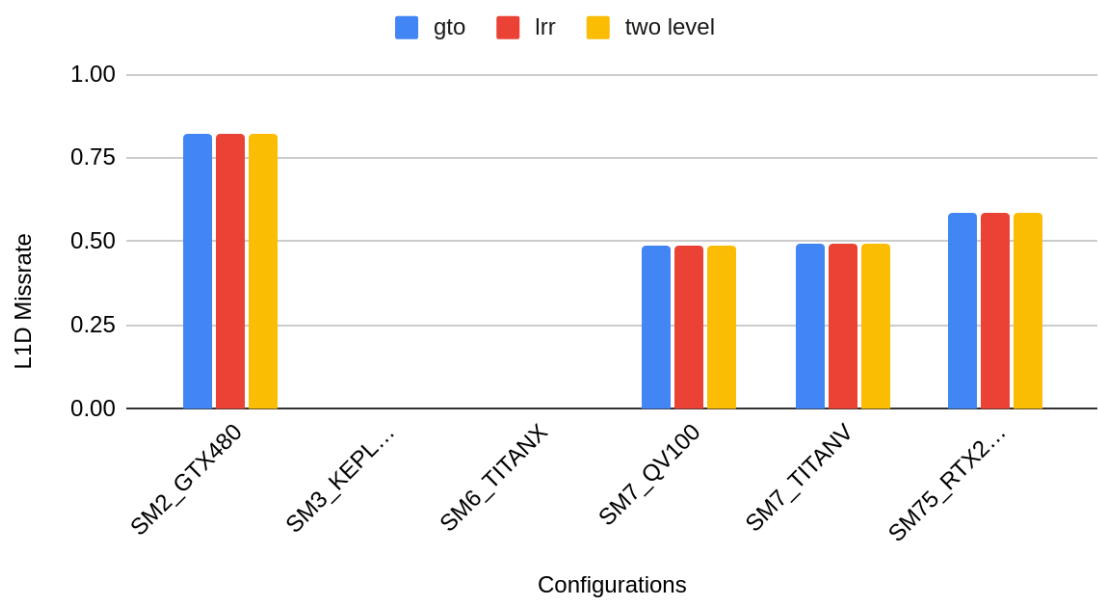


Question-2:

(i)L1D Missrate

| | | | |
|------------------|--------|--------|-----------|
| BFS | | | |
| L1D Missrate | | | |
| | gto | lrr | two level |
| SM2_GTX480 | 0.8225 | 0.823 | 0.8209 |
| SM3_KEPLER_TITAN | | | |
| SM6_TITANX | | | |
| SM7_QV100 | 0.4894 | 0.4895 | 0.4895 |
| SM7_TITANV | 0.4936 | 0.4936 | 0.4937 |
| SM75_RTX2060 | 0.5869 | 0.5865 | 0.5875 |

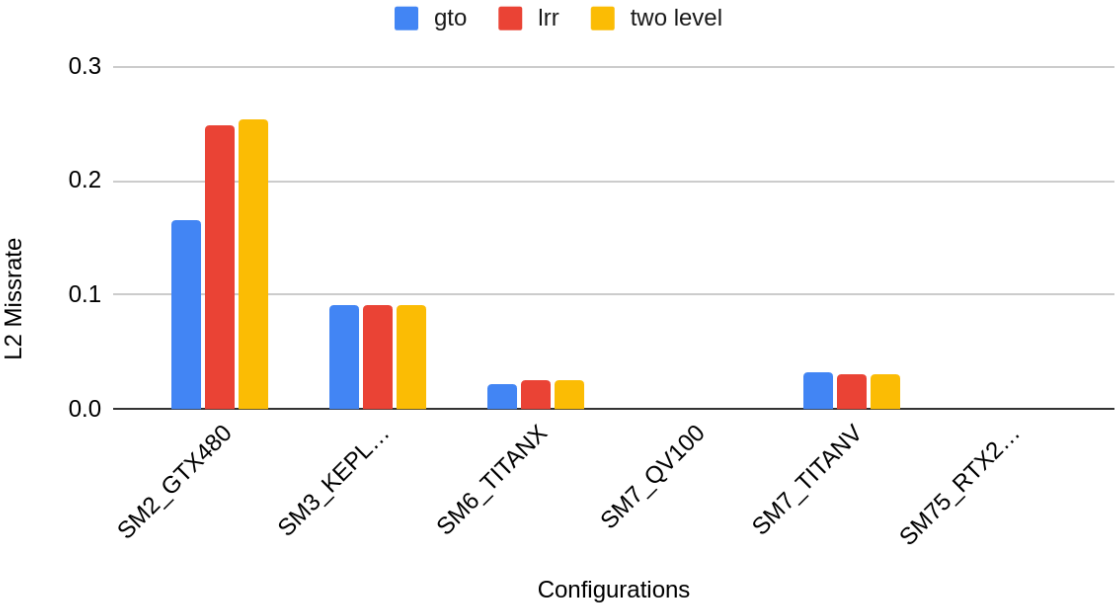
Diff WarpSchedulers(For BFS)



(ii)L2 Missrate

| | | | |
|------------------|--------|--------|-----------|
| BFS | | | |
| L2 Missrate | | | |
| | gto | lrr | two level |
| SM2_GTX480 | 0.1651 | 0.2493 | 0.2537 |
| SM3_KEPLER_TITAN | 0.0906 | 0.0913 | 0.0909 |
| SM6_TITANX | 0.0218 | 0.0258 | 0.0255 |
| SM7_QV100 | 0 | 0 | 0 |
| SM7_TITANV | 0.0317 | 0.0301 | 0.031 |
| SM75_RTX2060 | 0 | 0 | 0 |

Diff WarpSchedulers(For BFS)



3. Categorize the applications w.r.t the L1D and L2 Cache hit rates. What changes do you observe w.r.t L1D and L2 cache hit rates when the L1D cache size is increased from 32KB to 8MB? (use warp scheduler GTO)

- ``<nsets>:<bsize>:<assoc>``: These parameters define the basic properties of the cache.
- ``<nsets>`` is set to 32, which means there are 32 sets in the cache.
- ``<bsize>`` is set to 128, which implies that each cache block (line) is 128 bytes in size.
- ``<assoc>`` is set to 2048, indicating a high level of associativity. There are 2048 cache lines per set.

| Configuration | L1D | | L2 | |
|------------------|--------|--------|--------|--------|
| | 32KB | 8MB | 32KB | 8MB |
| SM2_GTX480 | 0.1775 | 0.6743 | 0.8349 | 0.5205 |
| SM3_KEPLER_TITAN | 1 | 1 | 0.9094 | 0.9094 |
| SM6_TITANX | 1 | 1 | 0.9782 | 0.9782 |
| SM7_QV100 | 0.5106 | 0.5106 | 1 | 1 |
| SM7_TITANV | 0.5064 | 0.5064 | 0.9683 | 0.9683 |
| SM75_RTX2060 | 0.4131 | 0.5768 | 1 | 1 |

4. What percentage of power is consumed by Execution units, DRAM, Register Files in each application run? Do you notice any correlation between the L1D cache hit rates observed in Question 3 and the Power consumption between different applications?

| Configuration | Exe. Time | DRAM | Reg. Files | Tot. power | %EU | %DRAM | %RF |
|------------------|-----------|------|------------|------------|-------|-------|------|
| SM2_GTX480 | 28.38 | 0 | 66.28 | 94.21 | 29.1 | 0.01 | 70 |
| SM3_KEPLER_TITAN | 72.65 | 0 | 6.17 | 78.89 | 92.1 | 0.08 | 7.8 |
| SM6_TITANX | 35.98 | 0 | 36.21 | 72.1 | 49.81 | 0.07 | 50.2 |
| SM7_QV100 | 48.29 | 0 | 73.21 | 120.7 | 40.29 | 0.13 | 59.9 |
| SM7_TITANV | 111.32 | 0 | 20.78 | 132.1 | 84.6 | 0.73 | 15.7 |
| SM75_RTX2060 | 40.41 | 0 | 20.81 | 61.32 | 65.81 | 0.05 | 34.0 |

- **Execution Unit Power Consumption:** In general, cache size and execution unit power consumption have a positive correlation. This discovery is illustrated in the context of particular GPU models, where an increase in cache capacity is matched by an increase in the percentage of power allotted to execution units.
- **DRAM Power Consumption:** Because managing a larger cache requires more power, the DRAM's power consumption is positively correlated with cache size. The observable increase in power consumption of dynamic random-access memory (DRAM) and the accompanying rise in the fraction of power given to DRAM can be seen in some graphics processing unit (GPU) models with an 8 MB cache, which supports the aforementioned observation.
- **Aggregate Power:** In some circumstances, a larger cache results in a higher overall power usage, mostly as a result of the execution units' and DRAM modules' higher power requirements.