David Chen (sc60)

MP2

Baseline GPU

metrics	ece511-rodinia-3.1	ece511-parboil	ece511-polybench
gpu_tot_sim_cycle	60957	377668	507579
gpu_tot_ipc	6117.812	4392.527	1222.169
L2_BW_total	547.506	263.1033	446.3286
gpu_tot_occupancy	81.5024	41.2161	85.848
L1D_total_cache_miss_rate	0.9884	0.2529	0.245
L2_total_cache_miss_rate	0.4577	0.1519	0.7108
total dram reads	266798	393088	2097152
total dram writes	79612	20868	2022040

Comparison of More SMs

metrics	**rodinia (SM_120)**	**rodinia (SM_160)**	**rodinia (SM_200)**	**parboil (SM_120)**	**parboil (SM_160)**	**parboil (SM_200)**	**polybench (SM_120)**	**polybench (SM_160)**	**polybench (SM_200)**
gpu_tot_sim_cycle	**49944**	**43802**	**42883**	**395641**	**422472**	**404720**	**417514**	**397609**	**399792**
gpu_tot_ipc	**7466.832**	**8513.845**	**8696.3**	**4192.985**	**3926.691**	**4098.925**	**1485.813**	**1560.195**	**1551.676**
L2_BW_total	**670.5238**	**765.2008**	**782.1277**	**312.7711**	**350.6137**	**426.2301**	**544.9571**	**574.4338**	**572.6094**
gpu_tot_occupancy	**78.6879**	**77.7808**	**77.2156**	**27.3806**	**20.5646**	**16.4548**	**86.2606**	**86.2187**	**85.9181**
L1D_total_cache_miss_rate	**0.9921**	**0.993**	**0.9937**	**0.3183**	**0.3838**	**0.4493**	**0.2461**	**0.247**	**0.2476**
L2_total_cache_miss_rate	**0.4562**	**0.4558**	**0.4555**	**0.122**	**0.1019**	**0.0875**	**0.7077**	**0.7052**	**0.7043**
total dram reads	**266928**	**266951**	**266986**	**393088**	**393088**	**393088**	**2097211**	**2098060**	**2102591**
total dram writes	**77689**	**76402**	**75539**	**18476**	**18432**	**18432**	**2021620**	**2021192**	**2020728**

Effect of Memory Coalescing

metrics	**ece511-rodinia-3.1**	**ece511-parboil**	**ece511-polybench**
num_tot_coalesced_accesses	**878800**	**9698304**	**24080908**
num_tot_uncoalesced_accesses	**4768841**	**77586432**	**167608360**
num_coalesced_load	**703696**	**9563136**	**21984780**
num_coalesced_store	**175104**	**135168**	**2096128**
num_uncoalesced_load	**3720265**	**76505088**	**150847524**
num_uncoalesced_store	**1048576**	**1081344**	**16760836**

Questions

a. For Rodinia, IPC increases and total cycles decreases with the increase of number of SMs, so having more SMs will help, but the increase will be marginal.
   For Parboil, IPC and total cycles plateau beyond 120 SMs. So more SMs will not help.
   For Polybench, IPC and total cycles plateau beyond 160 SMs. So more SMs will not help.

b. As the number of SMs increases, occupancy decreases. High occupancy indicates that the kernel is highly parallelizable, but for the same kernel, occupancy is not directly related to the execution time.

c. More SMs result in more L2 bandwidth being used, but only up until the plateau point. More SMs means more “cores” accessing L2 cache which will drive up the L2 bandwidth usage. However, beyond the plateau point, more SMs are not utilized because of the parallelism limitation within the kernel, so L2 bandwidth will not change after that point.
d. Number of SMs does not have an effect on off-chip memory bandwidth, which is different than L2 bandwidth. The reason for this is that since global memory (off-chip memory) access is expensive, A piece of data is often read to the global memory once and cached in L2 for future access. Therefore, global memory bandwidth is usually independent with the amount of SMs. Depending on if the program is memory bounded, increase DRAM bandwidth may or may not help. If the program is not memory bounded, then more bandwidth will not help with performance.
e. Rodinia, Parboil and Polybench has a coalesced memory access rate of 15.6%, 11.1%, 12.6% respectively. Intuitively, Rodinia solves a series of differential equations over a region of space. Therefore, adjacent thread can load adjacent memories, therefore having a better memory coalescing rate.