COMPUTER ORGANIZATION AND ARCHITECTURE
COA LAB ASSIGNMENT – 4

Name: Y.Krishna vamshi
Roll No: 21CS01019

Plot of IPC vs Configurations:

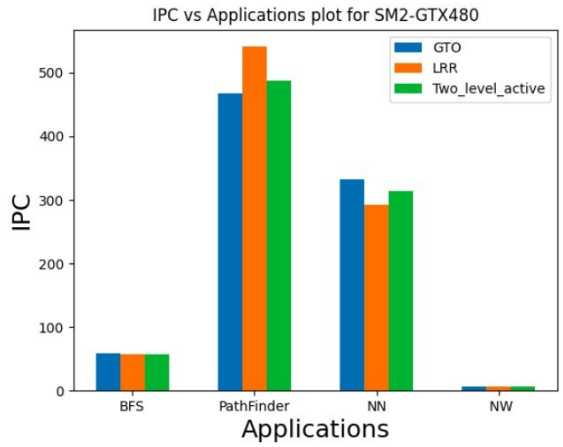
For each application,I have calculated the values of IPC for the following configurations:

- SM2_GTX480
- SM3_KEPLER_TITAN
- SM6_TitanX
- SM7_QV100
- SM7_TITANV
- SM75_RTX2060

and for each configuration, I have run for each warp schedulers (GTO - Greedy-then-oldest ,LRR - Loose Round Robin, and TL - Two Level).

PLOT:

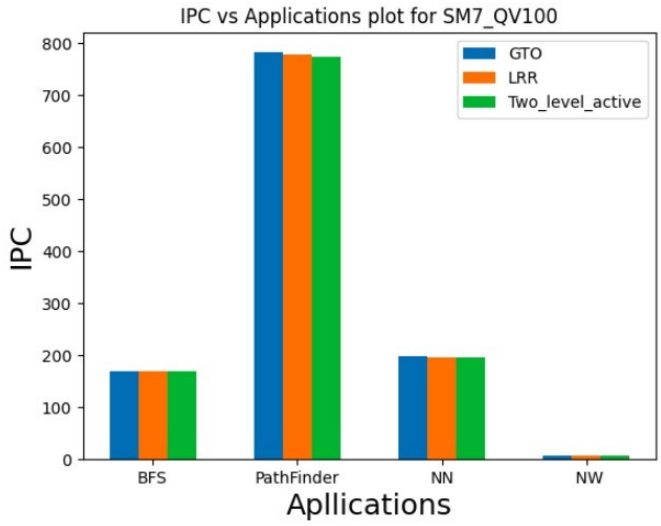
SM2-GTX480:



Supporting Data:

			SM2_GTX480				
			IPS	CPS	IPC	RUN_TIME	L1D MR
BFS	GTO		140474	2397	58.60408844	219	0.8225
	LRR		212164	3712	57.15625	145	0.823
	TWO_LEVEL_ACTIVE		146494	2549	57.47116516	210	0.8209
PathFinder	GTO		1031370	2205	467.7414966	61	0.5671
	LRR		1398080	2584	541.0526316	45	0.567
	TWO_LEVEL_ACTIVE		911791	1870	487.5887701	69	0.567
NN	GTO		600526	1811	331.5991165	2	0.6009
	LRR		600526	2059	291.6590578	2	0.6
	TWO_LEVEL_ACTIVE		600526	1911	314.2469911	2	
NW	GTO		58284	9428	6.182011031	107	0.7509
	LRR		62364	10083	6.185063969	100	0.7509
	TWO_LEVEL_ACTIVE		43009	6966	6.174131496	145	0.7509

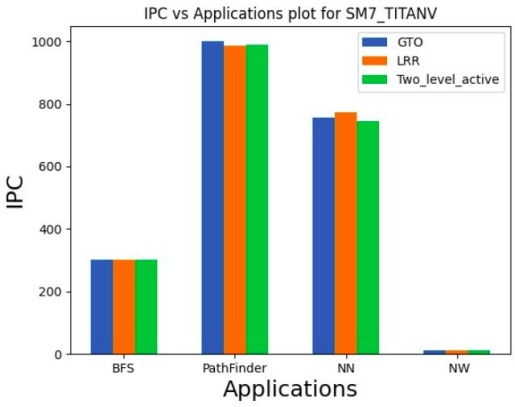
SM7-QV100:



Supporting Data:

SM7_QV100					
IPS	CPS	IPC	RUN_TIME	L1 MR	L2 MR
136123	803	169.5180573	226	0.4894	0
175793	1039	169.1944177	175	0.4881	0
112688	666	169.2012012	273	0.4895	0
547074	699	782.6523605	115	1	0.0123
796374	1024	777.7089844	79	1	0.0123
452615	585	773.7008547	139	1	0.0123
300263	1513	198.4553866	4	0.6	0.3334
300263	1524	197.0229659	4	0.6	0.3334
200175	1017	196.8289086	6	0.6	0.3334
18024	2502	7.20383693	346	0.8661	0
19013	2639	7.204622963	328	0.8661	0
13074	1817	7.195376995	477	0.8661	0

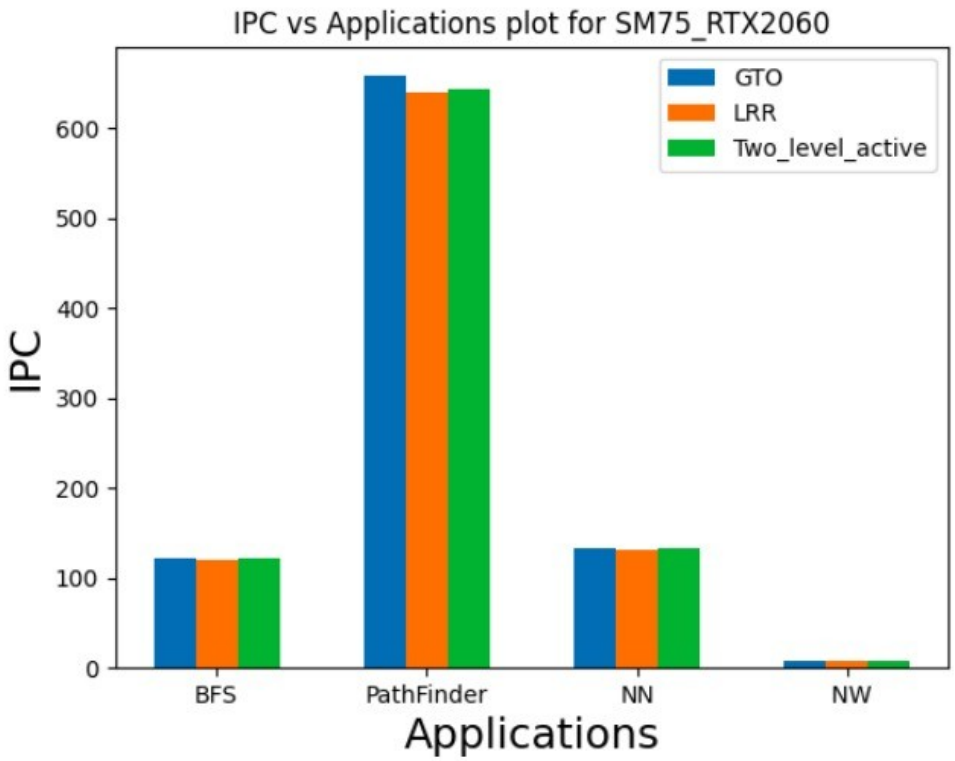
SM7-TITANV:



Supporting Data:

SM7_TITANV					
IPS	CPS	IPC	RUN_TIME	L1 MR	L2 MR
133755	444	301.25	230	0.4931	0.0317
221322	735	301.1183673	139	0.4936	0.0301
116972	388	301.4742268	263	0.493	0.031
367915	368	999.7690217	171	1	0.282
911791	924	986.7867965	69	1	0.282
487702	493	989.2535497	129	1	0.282
400350	529	756.805293	3	0.6	0.3334
400350	518	772.8764479	3	0.6	0.3334
300263	403	745.0694789	4	0.6	0.3334
30721	2720	11.29448529	203	0.8661	0
32146	2846	11.29515109	194	0.8661	0
21140	1876	11.26865672	295	0.8661	0

SM75-RTX2060:



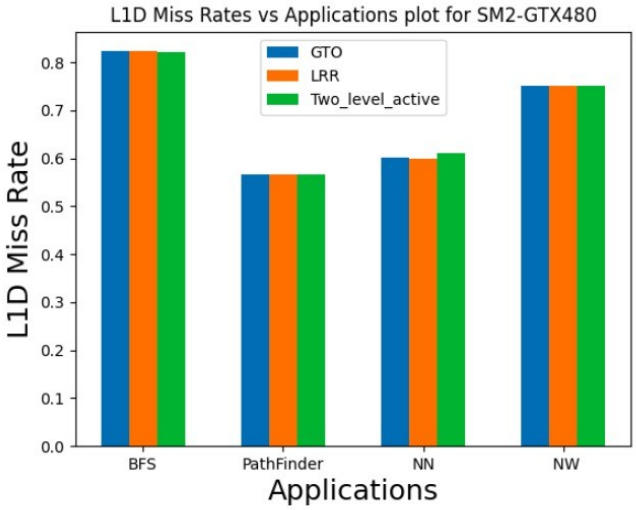
Supporting Data

SM75_RTX2060					
IPS	CPS	IPC	RUN_TIME	L1 MR	L2 MR
201070	1650	121.8606061	153	0.5869	0
287512	2388	120.39866	107	0.5865	0
184214	1506	122.3200531	167	0.5875	0
786420	1196	657.541806	80	1	0.824
1143883	1791	638.6839754	55	1	0.824
706894	1100	642.6309091	89	1	0.824
400350	3003	133.3166833	3	0.6	0.3334
400350	3074	130.2374756	3	0.6	0.3334
300263	2252	133.3317052	4	0.6	0.3334
43308	6089	7.112497947	144	0.8661	0
44866	6308	7.112555485	139	0.8661	0
31338	4393	7.133621671	199	0.8661	0

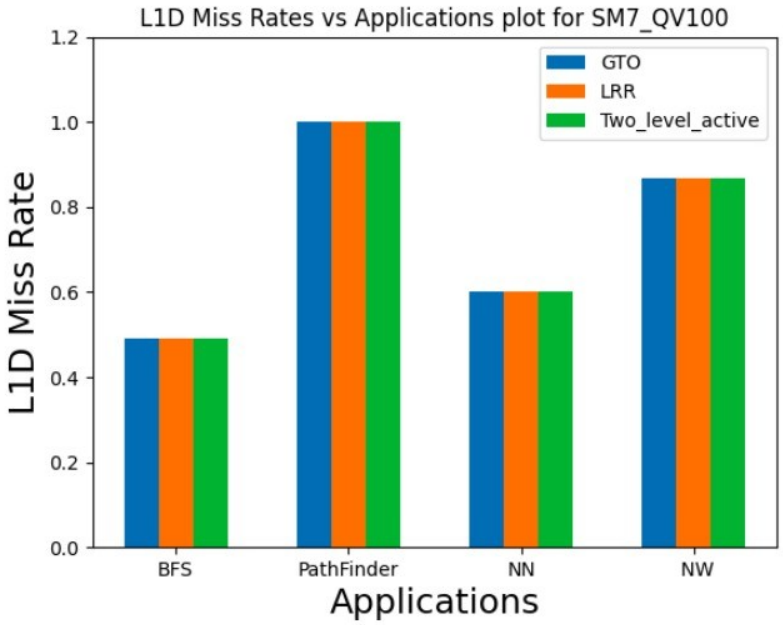
Q2: Plot of L1D Miss Rates:

Supporting data for L1D and L2 Miss rates are also provided in above screenshots.

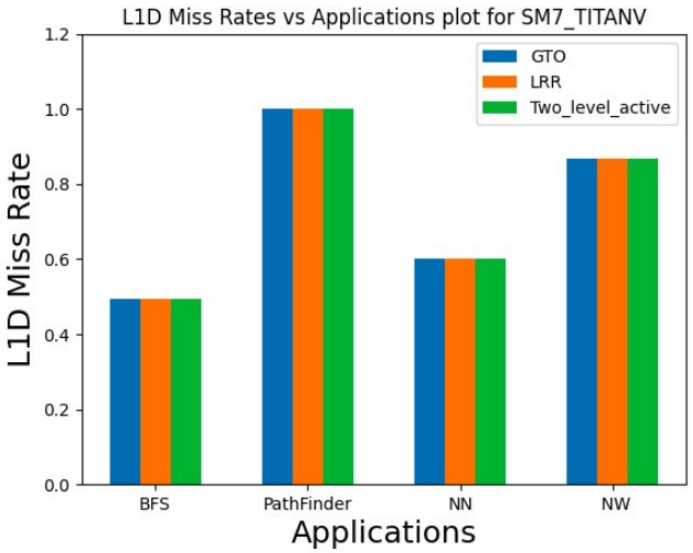
SM2-GTX480:



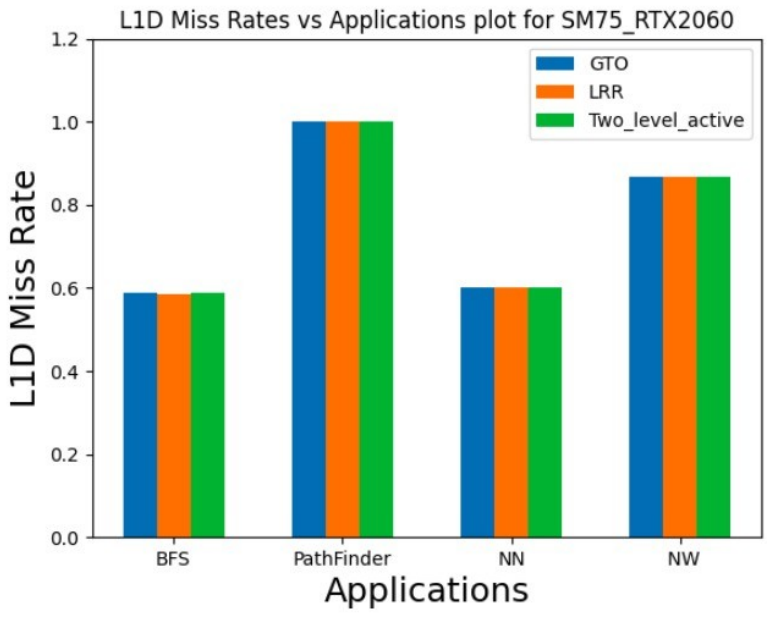
SM7-QV100:



SM7-TITANV:

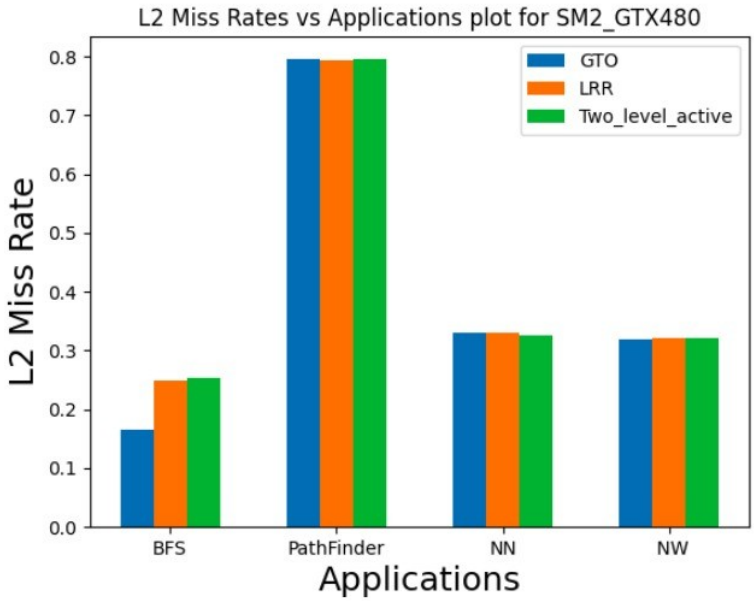


SM75-RTX2060:

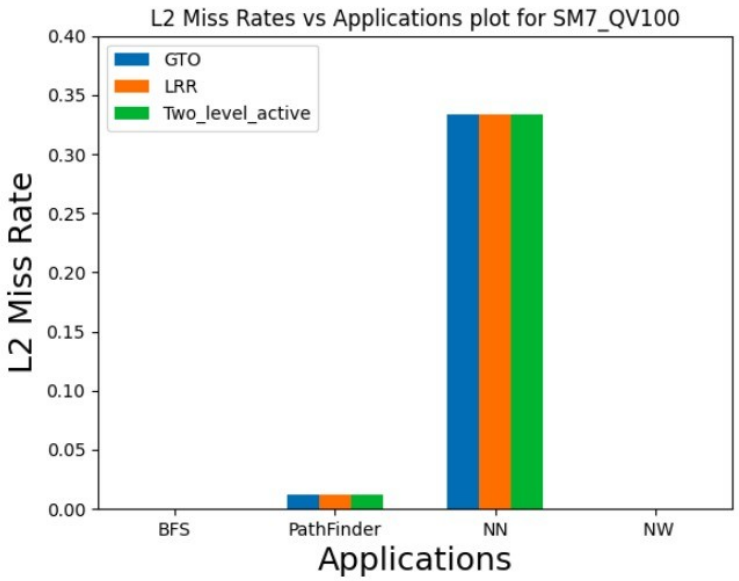


Q2 – b: L2 Miss Rates PLOT:

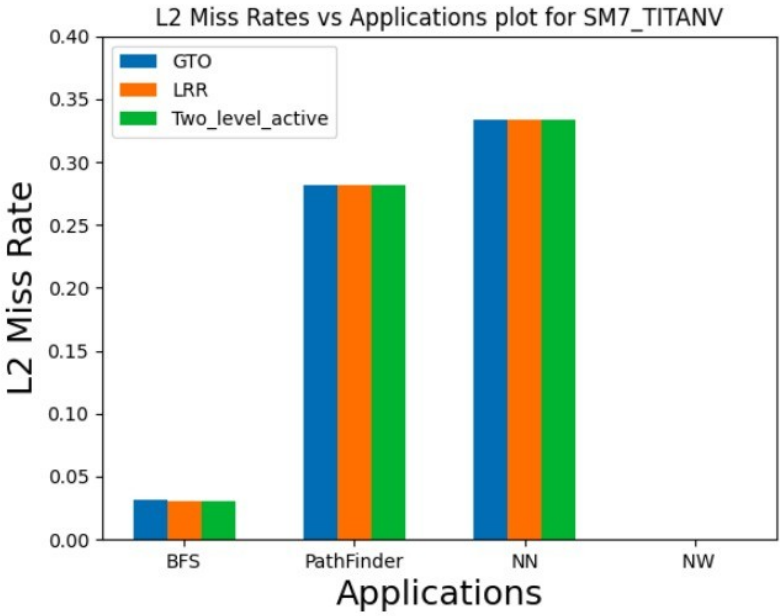
SM2-GTX480:



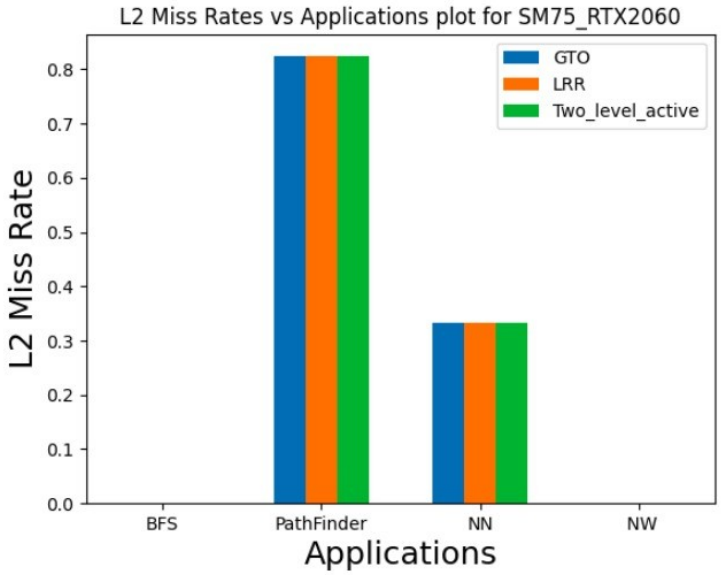
SM7-QV100:



SM7-TITANV:



SM75-RTX2060:



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).

Data for 32KB:

		SM2_GTX480	
		L1D MR	L2 MR
BFS	GTO	0.8225	0.1651
NN	GTO	0.6009	0.3304
NW	GTO	0.7509	0.3192

Data for 8MB:

		SM2_GTX480	
		L1D MR	L2 MR
BFS	GTO	0.3257	0.4795
NN	GTO	0.6	0.3309
NW	GTO	0.7481	0.3214

- On increasing L1D cache size, We can store more data in L1D cache and more number of accesses will be hit.
- As number of access are hit, miss rate decreases and access rate increases.
- We can see above observation from provided supporting data.
- As number of L1D access are hit,access for L2 will be decreased.
- Though L2 access decreases,L2D miss rate may increase or decrease, because both number of access and hit access will change.

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?

Data for L1D = 32KB:

		Power	Total Average Power
BFS	RFP	7.15	62.6126
	DRAMP	0.0948	
	L1P	3.8962	
	L2P	0.923	
NN	RFP	9.312	79.2563
	DRAMP	0.0579366	
	L1P	5.1959	
	L2P	10.0307	

Data for L1D = 8 MB:

		Power	Total Average Power
BFS	RFP	7.15057	94.2833
	DRAMP	0.094	
	L1P	3.895	
	L2P	0.923	
NN	RFP	9.31267	79.2728
	DRAMP	0.0579366	
	L1P	5.1959	
	L2P	10.0307	
NW	RFP	0.0302	21.7539
	DRAMP	0.069	
	L1P	0.008432	
	L2P	0.0679054	

- Here we can observe that as L1D cache size increases, percentage of power consumed by L1d,L2,Dram decreases.
- As L1D cache size increases, L1D hit rate increases.
- This is because fetching data from on-chip caches is generally more power-efficient compared to fetching from off-chip memory.