Cache Performance Measurement

Problem Description

We are measuring cache performance using PAPI library present in C. We need to measure the miss rate for L1 cache, L2 cache and total number of misprediction and misprediction rate. And need to provide comments about which cache implementation is better it terms of Miss rate execution time branch misprediction and stall cycles produced. The PAPI counters used to measure the aforementioned attributes are as follows:

PAPI_L1_TCM - Level 1 data cache misses

PAPI_L1_TCA - Level 1 data cache accesses

PAPI_L2_TCM - Level 2 data cache misses

PAPI_L2_TCA - Level 2 data cache accesses

PAPI_RES_STL - Cycles stalled on any resource.

PAPI_BR_MSP - Conditional branch instructions mispredicted.

Solution Strategy

To understand the behavior of cache using matrix multiplication. While the program is running we are only interested in the data fetched by the processor as multiplication needs data elements from the array. And therefore, to measure the behavior of cache and better understand the performance we should only considering **data cache** misses and total data cache access. This will help in analyzing the behavior of matrix multiplication for simple and blocked multiplication with different dimensions and block sizes better. But, our analysis is limited on the total L1 and L2 cache misses.

We use only 3 hardware counters, since we have limited hardware resources. PAPI_L1_TCM and PAPI_L1_TCA both are derived events. They are mentioned below:

PAPI_L1_TCM = PAPI_L1_DCM + PAPI_L1_ICM

L1 cache misses = L1 data misses + L1 Instruction misses.

Similarly, for Total cache accesses.

PAPI_L1_TCA = PAPI_L1_DCA + PAPI_L1_ICA

L1 total cache access = L1 data access + L1 Instruction cache access.

This implies that to count the value for each of the above events we require 2 event counters to be calculated. So, if we were to calculate both total cache misses and total cache access simultaneously then we would require 4 hardware counters. Since, we cannot employ 4 counters due to limited hardware counters.

We are running the multiplication program twice in the program. First, we will measure the total cache misses. Second, we reset the resultant matrix to make sure that total cache access for L1 we calculate the total cache accesses for L1. We can easily deduce the miss rate from above information.

Cache Performance Measurement

As Level 3 cache is unavailable in the server cache hierarchy. I have chosen to measure the number of mispredictions and number of stall cycles. These counters are available and not derived from any other counters and thus we can use these 2 counters in one program and measure them accurately using PAPI. We are using the 512 and 1024 as two dimensions of matrix and block size of 16, 32, 64, and 128.

Results

Description of Resources

We are using the whale cluster for conducting this experiment.

The hardware information that was used when the experiment was conducted are as follows:

PAPI Version – 5.6.0.0

Operating System – Linux

Vendor – AMD

Model – QUAD Core AMD Opteron Processor

CPU Max MHz – 2211

Number of Cores – 8

Number of Hardware Counters – 3

Cores per socket – 4

Socket – 2

The above information can be fetched using the "papi_avail" command or we can use the "PAPI_get_hardware_info" method.

Measurements

Level 1 Cache

Level 1 Cache Simple Multiply 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	134625982	3899223909	0.034526353	3931933
2	134624315	3898952686	0.034528327	3849733
3	134628216	3899297103	0.034526278	4486878

Average Miss rate = 0.034526986

Average execution time = 4089514.667 = 4.089 sec

Level 1 Cache Block(16) 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	23973921	4485219397	0.005345094	2134380
2	24009884	4484722606	0.005353705	2279611
3	24034361	4485151888	0.00535865	2066298

Cache Performance Measurement

Average Miss rate = 0.005352483

Average execution time = 2160096.333 = 2.160 sec

Level 1 Cache Block(32) 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	135715951	4319665016	0.031418166	2136327
2	135719897	4319952188	0.031416991	2168186
3	135748462	4320023399	0.031423085	2103557

Average Miss rate = 0.031419414

Average execution time = 2136023.333 = 2.136 sec

Level 1 Cache Block(64) 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	134860898	4235410040	0.031841285	2373631
2	134878254	4233988878	0.031856072	2353184
3	134943406	4234084867	0.031870737	2388753

Average Miss rate = 0.031856031

Average execution time = 2371856 = 2.371 sec

Level 1 Cache Block(128) 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	134728525	4192466258	0.032135864	2832238
2	134750171	4192330939	0.032142064	2857429
3	134755242	4191683868	0.032148236	2956069

Average Miss rate = 0.032142055

Average execution time = 2881912 = 2.881 sec

Level 1 Cache Simple Multiply 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	1075224990	31166618899	0.03449925	126489172
2	1075237017	31166212729	0.034500086	122130600
3	1075199069	31168143443	0.034496731	121919435

Cache Performance Measurement

Average Miss rate = 0.034498689

Average execution time = 123513069 = 123.513 sec

Level 1 Cache Block(16) 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	1094933613	35908398374	0.03049241	17844072
2	1092953905	35897360904	0.030446637	17903978
3	1094346801	35879051771	0.030500996	17693648

Average Miss rate = 0.030480014

Average execution time = 17813899.33 = 17.813 sec

Level 1 Cache Block(32) 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	1083486786	34546395348	0.031363237	16809583
2	1083475319	34551677434	0.031358111	16843989
3	1083473382	34559293474	0.031351144	16609663

Average Miss rate = 0.031357497

Average execution time = 16754411.67 = 16.754 sec

Level 1 Cache Block(64) 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	1077650249	33855327562	0.031831039	21023602
2	1077748279	33832391171	0.031855516	21076587
3	1077592734	33858512775	0.031826346	20590410

Average Miss rate = 0.031837634

Average execution time = 20896866.33 = 20.896 sec

Level 1 Cache Block(128) 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	1076270013	33514177910	0.032113872	27868387
2	1075906730	33512736070	0.032104413	26019707
3	1076055874	33505916690	0.032115399	26520845

Cache Performance Measurement

Average Miss rate = 0.032111228

Average execution time = 26802979.67 = 26.802 sec

Level 2 Cache

Level 2 Cache Simple Multiply 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	136671204	320556539	0.426356001	5848845
2	137438567	459774905	0.298925769	7302365
3	137515437	261495700	0.525880299	5323271

Average Miss rate = 0.417054023

Average execution time = 6158160.333 = 6.158 sec

Level 2 Cache Block(16) 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	1828478	37749634	0.048436973	2408611
2	1855906	43992653	0.042186726	2424184
3	1814793	34320822	0.052877317	2409427

Average Miss rate = 0.047833672

Average execution time = 2414074 = 2.414 sec

Level 2 Cache Block(32) 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	888270	161815054	0.005489415	2241760
2	986089	155719325	0.006332477	2253291
3	878069	149908400	0.00585737	2243149

Average Miss rate = 0.005893087

Average execution time = 2246066.667 = 2.246 sec

Level 2 Cache Block(64) 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	782987	210118463	0.003726407	2413580
2	26627741	239080446	0.111375654	2879400
3	4533148	219178961	0.020682405	2450197

Cache Performance Measurement

Average Miss rate = 0.045261489

Average execution time = 2581059 = 2.581 sec

Level 2 Cache Block(128) 512				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	57154738	226747274	0.252063617	3054822
2	74545297	249667369	0.298578454	3325971
3	58470762	227609122	0.256891119	3055303

Average Miss rate = 0.26917773

Average execution time = 3145365.333 = 3.145 sec

Level 2 Cache Simple Multiply 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	1080806033	2323967113	0.465069418	127136871
2	1083111301	2314003412	0.468068152	126629236
3	1081948599	2373378361	0.45586857	127294296

Average Miss rate = 0.463002047

Average execution time = 127020134.3 = 127.020 sec

Level 2 Cache Block(16) 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	12534220	1294763287	0.009680704	20352150
2	12121112	1270131363	0.009543196	20295404
3	12580386	1322321227	0.009513865	20358485

Average Miss rate = 0.009579255

Average execution time = 20335346.33= 20.335 sec

Level 2 Cache Block(32) 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	6140324	1254208698	0.004895775	18035177
2	6137645	1253979442	0.004894534	17996395
3	6190132	1259040662	0.004916547	17942845

Cache Performance Measurement

Average Miss rate = 0.004902285

Average execution time = 17991472.33 = 17.991 sec

Level 2 Cache Block(64) 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	283667484	1773175174	0.159977135	23197544
2	289483343	1793265167	0.161428075	22971578
3	292126134	1764557994	0.165552016	23193531

Average Miss rate = 0.162319075

Average execution time = 23120884.33 = 23.120 sec

Level 2 Cache Block(128) 1024				
Index	Misses	Total cache access	Miss rate	Execution time in micro sec
1	1003123341	2061344036	0.486635575	28749487
2	1000801079	2072003089	0.483011384	28992373
3	985693941	2032394778	0.484991377	28567518

Average Miss rate = 0.484879445

Average execution time = 28769792.67 = 28.769 sec

Number of Branch Misprediction and Stall Cycles

Misprediction and stall cycle Simple Multiply 512			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	6080059915	265440	3789328
2	9624043485	266426	5393415
3	14305466502	267745	7512423

Average # of Branch Misprediction = 266537

Average # of stall cycles = 10003189967

Average execution time = 5565055.333 = 5.565 sec

Misprediction and stall cycle Block(16) 512			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	2541527921	8949374	2345501
2	2564956757	8949514	2355923
3	2637221464	8948263	2387658

Cache Performance Measurement

Average # of Branch Misprediction = 8949050.333

Average # of stall cycles = 2581235381

Average execution time = 2363027.333 = 2.363 sec

Misprediction and stall cycle Block(32) 512			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	2754437028	4331983	2390517
2	2818874334	4331501	2417249
3	2967126021	4332182	2485063

Average # of Branch Misprediction = 4331888.667

Average # of stall cycles = 2846812461

Average execution time = 2430943 = 2.430 sec

Misprediction and stall cycle Block(64) 512			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	2720728013	2132460	2350620
2	2756845974	2131743	2366218
3	3732736558	2132330	2807954

Average # of Branch Misprediction = 2132177.667

Average # of stall cycles = 3070103515

Average execution time = 2508264 = 2.508 sec

Misprediction and stall cycle Block(128) 512			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	3962481898	1058013	2899232
2	3844783147	1058906	2847934
3	4984909345	1058198	3362346

Average # of Branch Misprediction = 1058372.333

Average # of stall cycles = 4264058130

Average execution time = 3036504 = 3.036 sec

Cache Performance Measurement

Misprediction and stall cycle Simple Multiply 1024			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	256675057495	1114928	124443085
2	258067845749	1111700	125064135
3	257113682696	1111736	124634603

Average # of Branch Misprediction = 1112788

Average # of stall cycles = 257285528647

Average execution time = 124713941 = 124.713 sec

Misprediction and stall cycle Block(16) 1024			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	25778614397	71577755	21180138
2	25841457713	71577446	21222615
3	25875158736	71577602	21253108

Average # of Branch Misprediction = 71577601

Average # of stall cycles = 25831743615

Average execution time = 21218620.33 = 21.218 sec

Misprediction and stall cycle Block(32) 1024			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	21703466076	34646576	18963621
2	21791902849	34646735	19010666
3	21768908794	34646779	18996714

Average # of Branch Misprediction = 34646696.67

Average # of stall cycles = 21754759240

Average execution time = 18990333.67 = 18.990 sec

Misprediction and stall cycle Block(64) 1024			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	29210883978	17049940	22156975
2	27559133484	17049763	21413706
3	28988302443	17050003	22062121

Cache Performance Measurement

Average # of Branch Misprediction = 17049902

Average # of stall cycles = 28586106635

Average execution time = 21877600.67 = 21.877 sec

Misprediction and stall cycle Block(128) 1024			
Index	# of Stall Cycles	# of Misprediction	Execution time in micro sec
1	42052027410	8466214	27885456
2	42671283096	8466754	28165252
3	42880853249	8466722	28260673

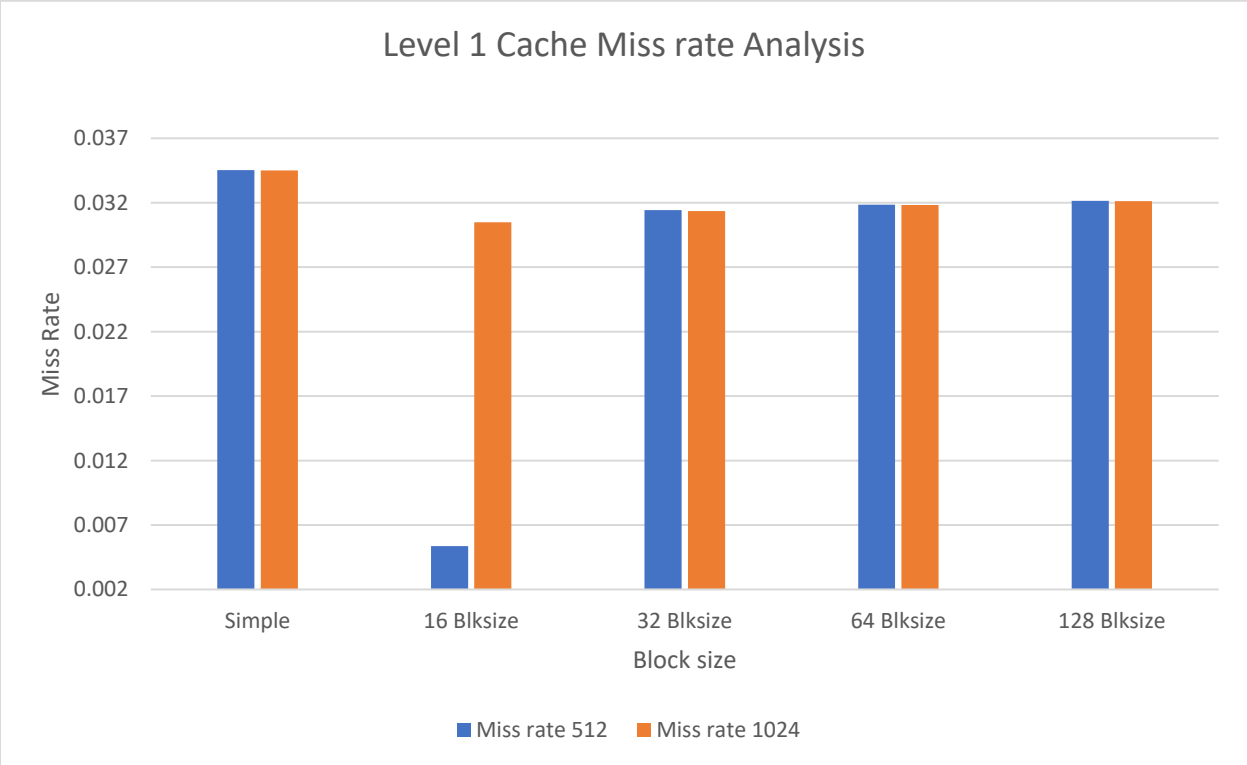
Average # of Branch Misprediction = 8466563.333

Average # of stall cycles = 42534721252

Average execution time = 28103793.67 = 28.103 sec

Comments

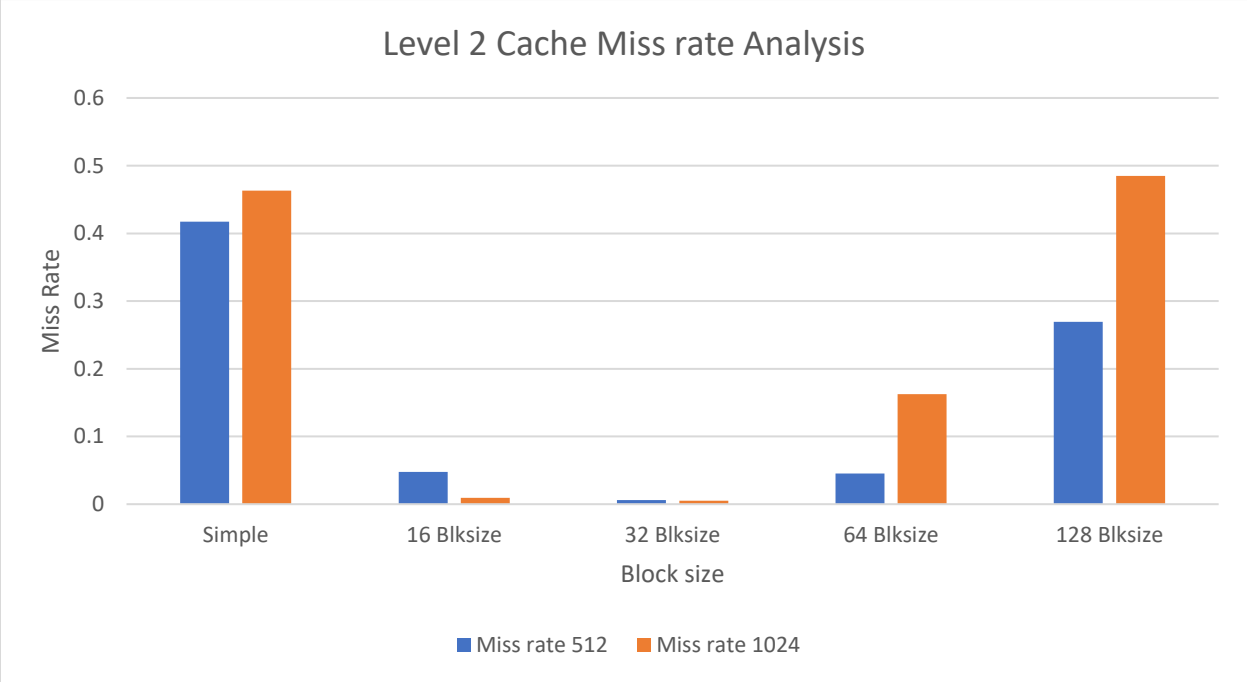
Level 1 Cache



Cache Performance Measurement

1) Simple cache has the highest miss rate for both 512 and 1024-dimension matrix multiplication.
2) Block size of 16 has the lowest miss rate for both 512 and 1024-dimension matrix multiplication.
3) We observe that the miss rate for both the matrix dimension 512 and 1024 are nearby except for block size 16. This may be due to block size of 16 is the smallest. We might have a larger block size for Simple cache. And as we know with greater block size we need greater bandwidth to fetch the data from main memory to cache. This could be one of the reasons that other Block size implementations have higher miss rate.
4) On the other hand, we notice that miss rate for block size 16, matrix multiplication for dimension 1024 is 6 times higher than dimension 512.

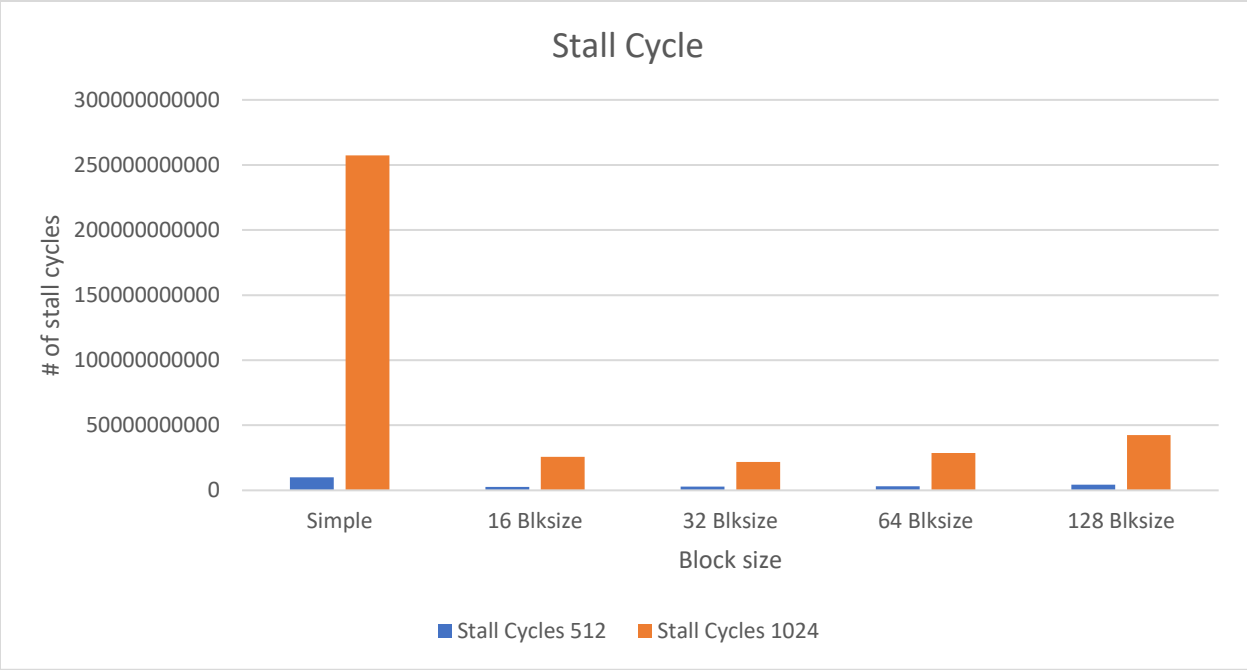
Level 2 Cache



1) We notice that miss rates for block size of 128 has highest value for 1024-dimension matrix multiplication. On the other hand, 512-dimension matrix multiplication the highest miss rate is found for Simple cache.
2) We find that block size of 32 has the lowest miss rate for both 512 and 1024-dimesion matrix multiplication.
3) We observe miss rate for 1024-dimension matrix multiplication is higher than 512-dimension matrix multiplication except for block size of 16.

Cache Performance Measurement

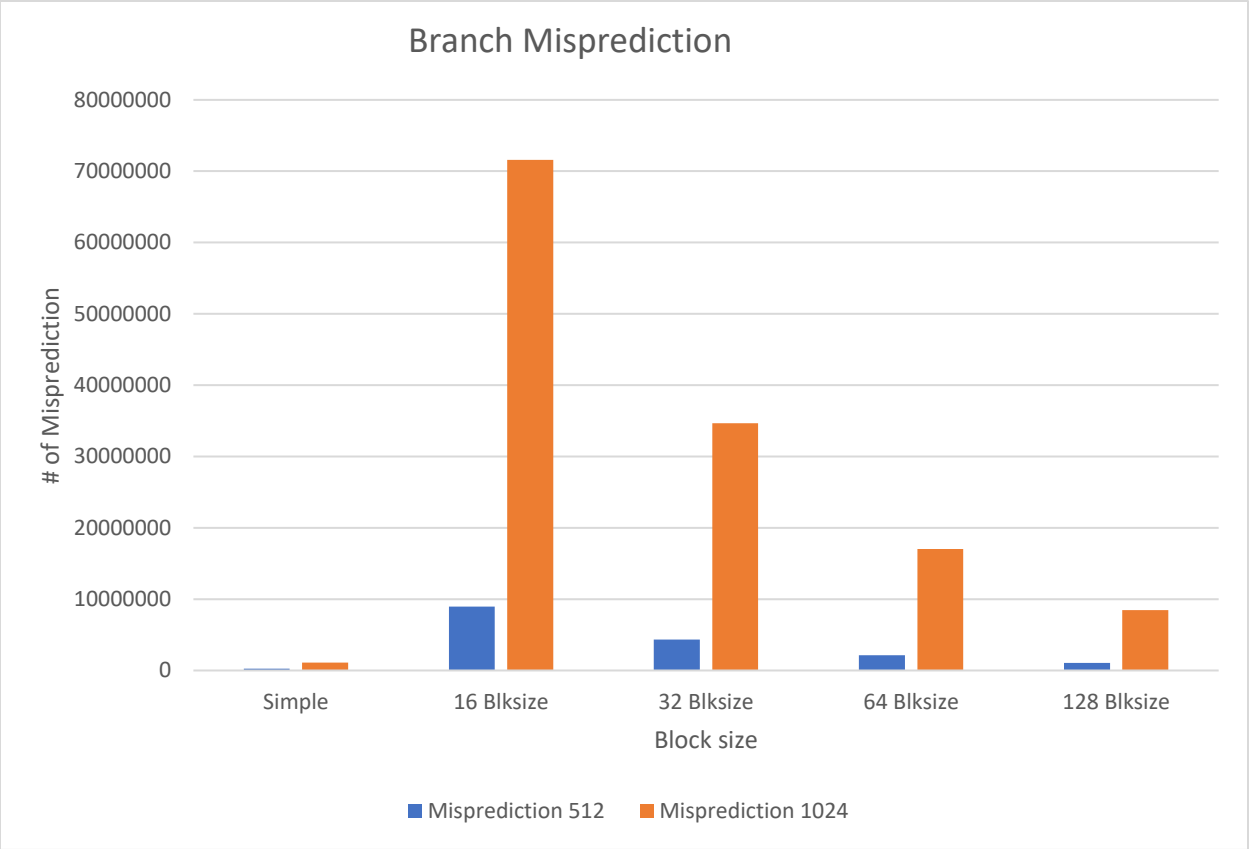
Number of Stall Cycles



1) Highest number of stalls are produced in Simple cache for 512-dimension matrix multiplication.
2) Least number of stalls are produced in block size of 16 for 512-dimension matrix multiplication.
3) Highest number of stalls are produced in Simple cache for 1024-dimension matrix multiplication.
4) Least number of stalls are produced in block size of 32 for 1024-dimension matrix multiplication.

Cache Performance Measurement

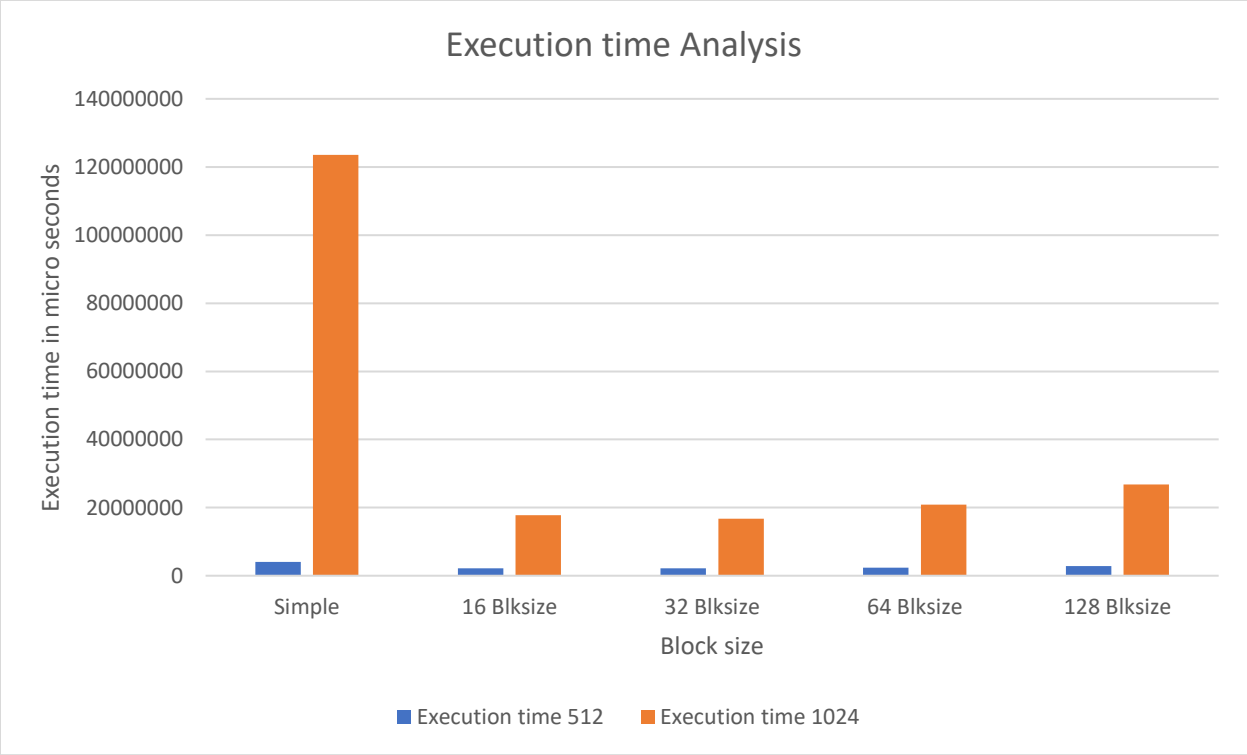
Number of Branch Misprediction



1) Block size of 16 for 1024-dimesion matrix multiplication has the highest misprediction on branches.
2) Simple cache for 1024-dimension matrix multiplication has the lowest misprediction on branches.
3) Block size of 16 for 512-dimension matrix multiplication has the highest misprediction on branches.
4) Simple cache for 512-dimension matrix multiplication has the lowest misprediction on branches.
5) We notice a pattern in Misprediction with increasing block size the number of branch misprediction decreases.

Cache Performance Measurement

Execution Time



1) Simple cache for 1024-dimension matrix multiplication takes the highest execution time.
2) Block size of 32 for 1024-dimension matrix multiplication takes least execution time.
3) Simple cache for 512-dimension matrix multiplication takes highest execution time.
4) Block size of 32 for 512-dimension matrix multiplication takes least execution time.

Final Comments

Looking at all the plots, it seems that measurements for block size 16 has anomalous behavior. It would be more pertinent to say that block size of 16 is for the matrix multiplication program used for measurements is the best cache implementation. Since, multiplication is performed with least miss rate for both L1 and L2 cache, least time for execution and less number of stall cycles are produced.

On the contrary, Simple cache has performed the worst for matrix multiplication program. Since, it is obvious from the plots. It has highest miss rate for L1 and L2 cache, highest number of stall cycles produced and the highest execution time for the program.