

NC State University
Department of Electrical and Computer Engineering
ECE 463/563: Fall 2017
Project #1: Cache Design, Memory Hierarchy Design

by

<< KENNY KADAYIL PAUL >>

NCSU Honor Pledge: "I have neither given nor received unauthorized aid on this test or assignment."

Student's electronic signature: KENNY KADAYIL PAUL

(sign by typing your name)

Course number: 563

(463 or 563 ?)

1. Abstract

For this project, a cache simulator is implemented. This cache simulator can simulate multiple level of caches and configure various cache parameters like block size, total cache size, set associativity, replacement policies like LRU and LFU as well as write policies such as write back and write through. To further enhance the cache performance a two level hierarchy of L1 and L2 cache was used. The L2 parameters also were configurable except for the block size. The impact of victim cache as a part of L1 cache and how it affects the overall hit/miss rate were also studied. After configuring caches, all the associated parameters was displayed and the average access time for the cache was calculated.

2. Experiments

2.1 Effect of block size on performance

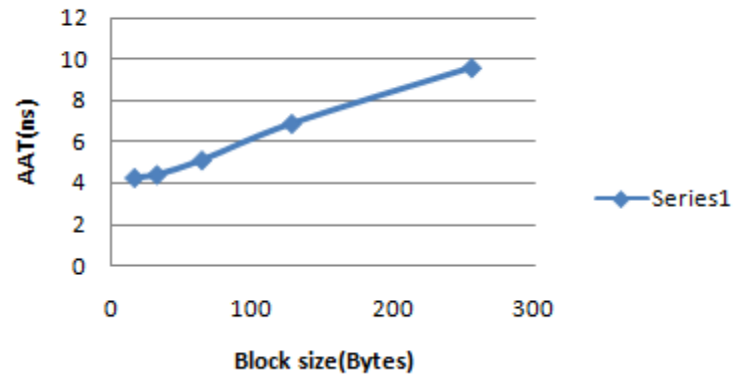
Large block size does not cause much impact on miss rate which more or less remains the same but tends to increase the AAT. This trend is constantly seen for all the trace files. For the vortex trace a block size of 32 tends to decrease the AAT and miss rate which might be due to the fact that larger block size takes advantage of the spatial locality and reduces the compulsory misses.

2.1.1 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
Variable	1024	1	LRU	Write back	gcc_trace

Cache	Block size (bytes)	Miss rate	AAT (ns)
L1	16	19.22%	4.24
L1	32	19.35%	4.39
L1	64	21.49%	5.10
L1	128	26.69%	6.88
L1	256	31.88%	9.60

AAT vs Block size

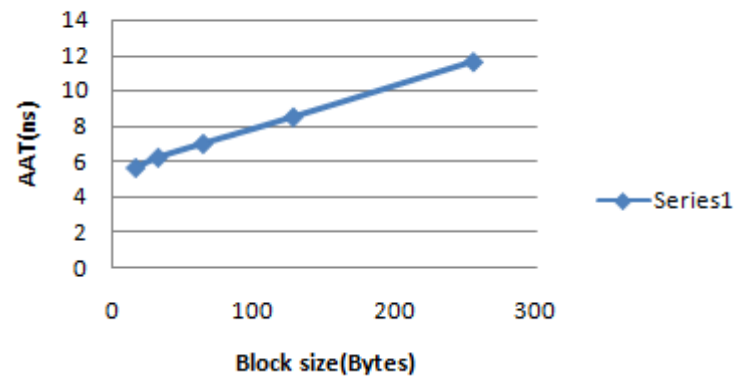


2.1.2 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
Variable	1024	1	LRU	Write back	perl_trace

Cache	Block size (bytes)	Miss rate	AAT (ns)
L1	16	25.99%	5.63
L1	32	28.01%	6.21
L1	64	30.06%	6.99
L1	128	33.45%	8.50
L1	256	39.14%	11.63

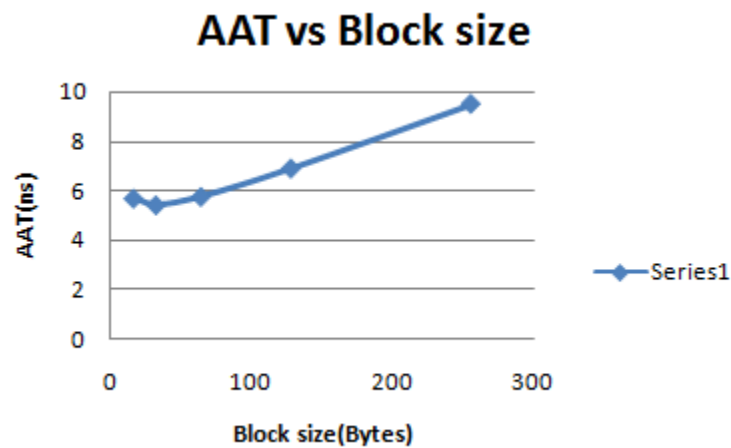
AAT vs Block size



2.1.3 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
Variable	1024	1	LRU	Write back	go_trace

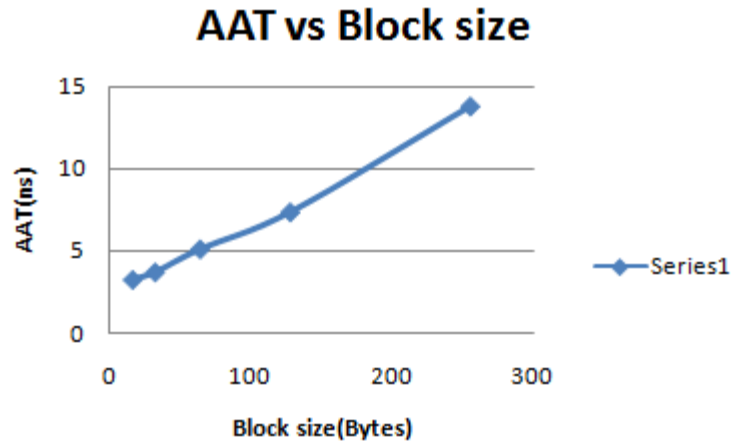
Cache	Block size (bytes)	Miss rate	AAT (ns)
L1	16	26.23%	5.68
L1	32	24.24%	5.41
L1	64	24.43%	5.75
L1	128	26.77%	6.90
L1	256	31.59%	9.52



2.1.4 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
Variable	1024	1	LRU	Write back	vortex_trace

Cache	Block size (bytes)	Miss rate	AAT (ns)
L1	16	14.31%	3.23
L1	32	16.01%	3.69
L1	64	21.47%	5.10
L1	128	28.69%	7.36
L1	256	46.98%	13.83



2.2 Effect of cache size on performance

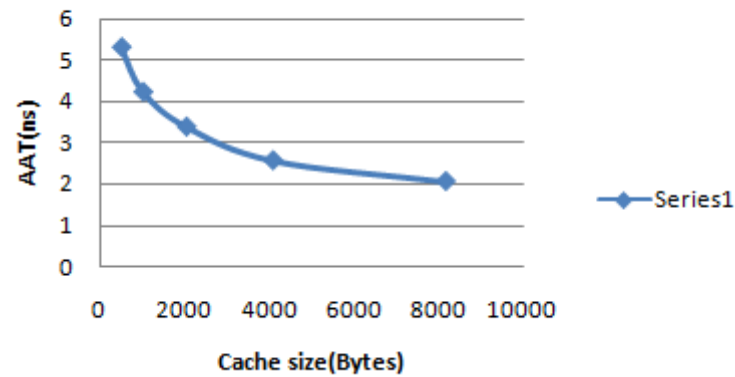
Larger the cache less is the miss rate and AAT. This might be due to the fact that the cache size reduces the number of capacity misses. This trend is constantly seen in all the trace files. But beyond certain limit of 4KB, the AAT and miss rate saturate and does not decrease appreciably.

2.2.1 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
16	Variable	1	LRU	Write back	gcc_trace

Cache	Cache size (bytes)	Miss rate	AAT (ns)
L1	512	24.54%	5.33
L1	1024	19.22%	4.24
L1	2048	15.06%	3.39
L1	4096	11.02%	2.57
L1	8192	8.39%	2.06

AAT vs Cache size

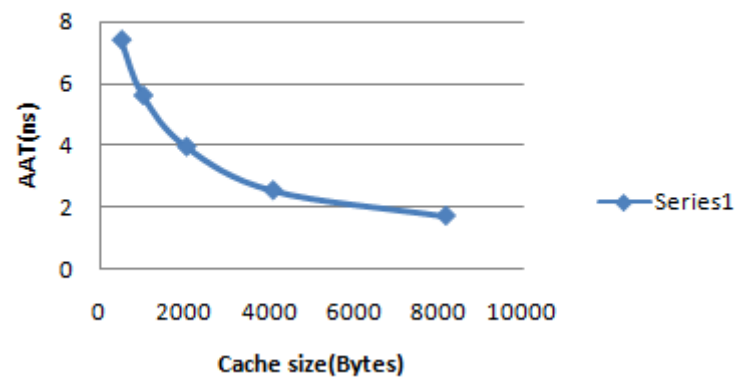


2.2.2 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
16	Variable	1	LRU	Write back	perl_trace

Cache	Cache size (bytes)	Miss rate	AAT (ns)
L1	512	34.84%	7.44
L1	1024	25.99%	5.63
L1	2048	17.84%	3.96
L1	4096	10.85%	2.54
L1	8192	6.72%	1.71

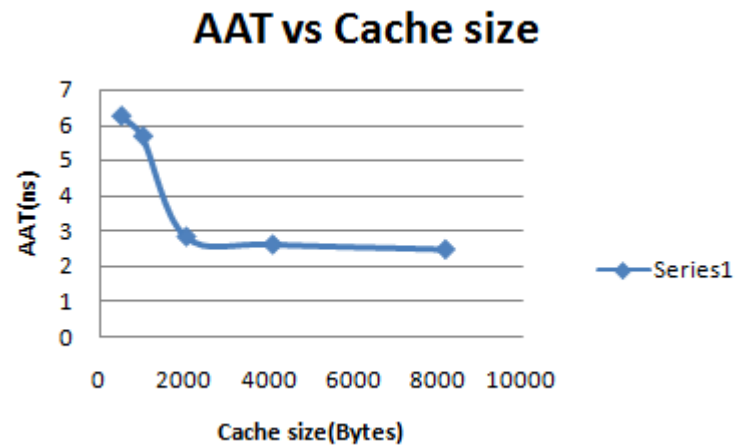
AAT vs Cache size



2.2.3 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
16	Variable	1	LRU	Write back	go_trace

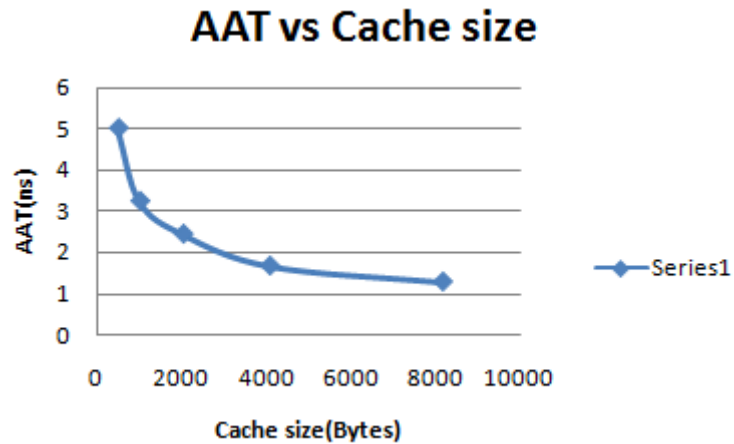
Cache	Cache size (bytes)	Miss rate	AAT (ns)
L1	512	29.04%	6.25
L1	1024	26.23%	5.68
L1	2048	12.42%	2.85
L1	4096	11.28%	2.63
L1	8192	10.50%	2.49



2.2.4 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
16	Variable	1	LRU	Write back	vortex_trace

Cache	Cache size (bytes)	Miss rate	AAT (ns)
L1	512	22.95%	5.00
L1	1024	14.31%	3.23
L1	2048	10.34%	2.43
L1	4096	6.54%	1.66
L1	8192	4.57%	1.27



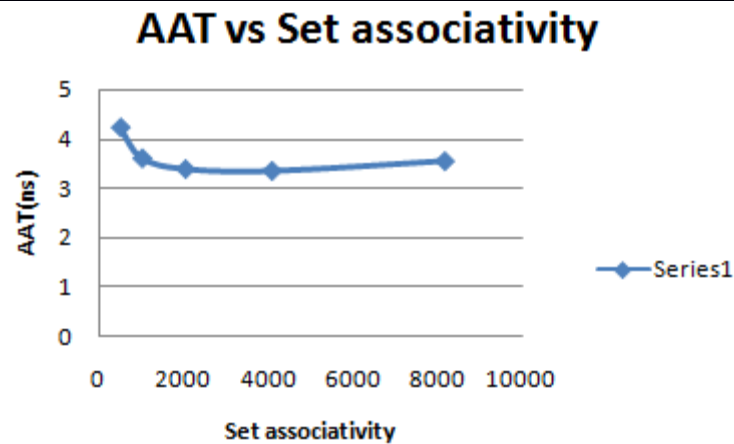
2.3 Effect of set associativity on performance

Higher set associativity tends to decrease the miss rate and AAT. Set associativity takes advantage of temporal locality which points out that the block recently used tends to get used again. Also this reduces the number of conflict misses. From the below data we see that beyond set associativity value of 8, there is an increase in both the miss rate and AAT. Thus the associativity of 8 is optimal.

2.3.1 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
16	1024	Variable	LRU	Write back	gcc_trace

Cache	Cache size	Associativity	Miss rate	AAT (ns)
L1	1024	1	19.22%	4.24
L1	1024	2	16.00%	3.61
L1	1024	4	14.73%	3.39
L1	1024	8	14.01%	3.35
L1	1024	16	14.03%	3.55

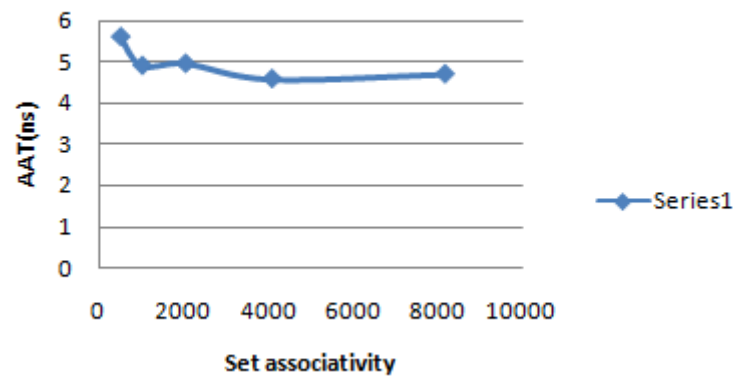


2.3.2 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
16	1024	Variable	LRU	Write back	perl_trace

Cache	Cache size	Associativity	Miss rate	AAT (ns)
L1	1024	1	25.99%	5.63
L1	1024	2	22.36%	4.91
L1	1024	4	22.36%	4.96
L1	1024	8	20.04%	4.58
L1	1024	16	19.65%	4.70

AAT vs Set associativity

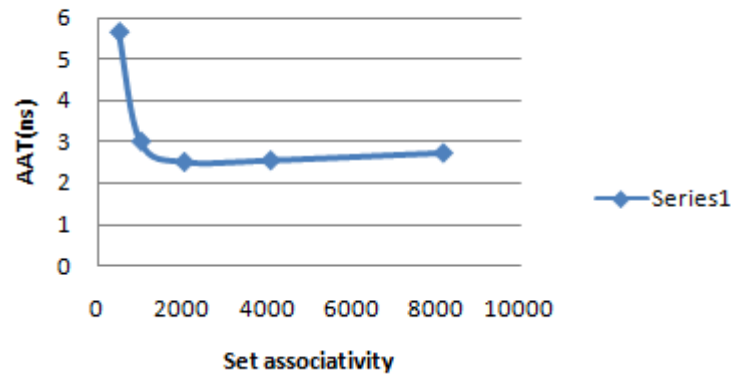


2.3.3 Cache parameters

Block size	Cache size	Associativity	Replacement policy	Write policy	Trace file
16	1024	Variable	LRU	Write back	go_trace

Cache	Cache size	Associativity	Miss rate	AAT (ns)
L1	1024	1	26.23%	5.68
L1	1024	2	13.17%	3.02
L1	1024	4	10.47%	2.52
L1	1024	8	10.16%	2.56
L1	1024	16	10.09%	2.74

AAT vs Set associativity

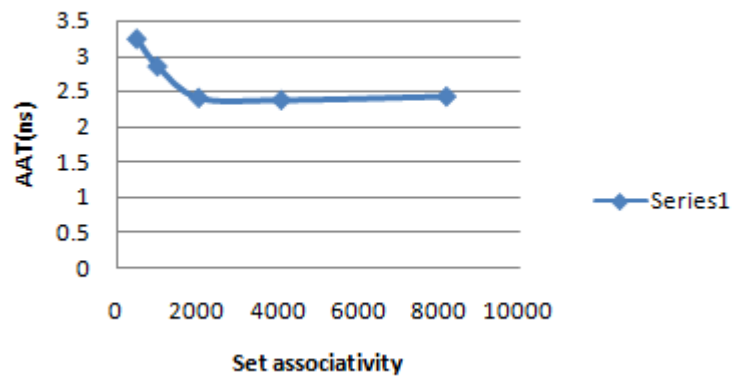


2.3.4 Cache parameters

Block size	No. of sets	Associativity	Replacement policy	Write policy	Trace file
16	64	Variable	LRU	Write back	vortex_trace

Cache	Cache size	Associativity	Miss rate	AAT (ns)
L1	1024	1	14.31%	3.23
L1	1024	2	12.30%	2.85
L1	1024	4	9.92%	2.41
L1	1024	8	9.31%	2.38
L1	1024	16	8.57%	2.43

AAT vs Set associativity



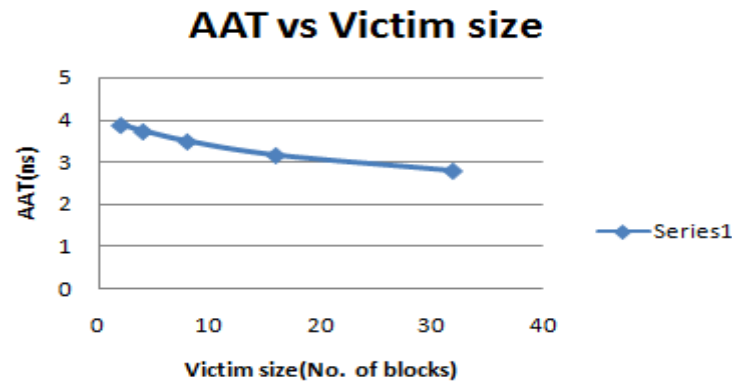
2.4 Effect of victim cache on performance

The trend we see from the below data is that victim cache tends to enhance the hit rate and when more reduces both AAT and miss rate. But beyond a value of 16 blocks in the victim cache, the decrease in the parameters are not appreciable.

2.4.1 Cache parameters

Block size	Cache size	Associativity	Victim size	Replacement	Write policy	Trace file
16	1024	1	Variable	LRU	Write back	gcc_trace

Cache	Block size	No. of victim blocks	Miss rate	AAT (ns)
L1	16	2	17.44%	3.87
L1	16	4	16.67%	3.72
L1	16	8	15.49%	3.48
L1	16	16	13.89%	3.15
L1	16	32	12.08%	2.78

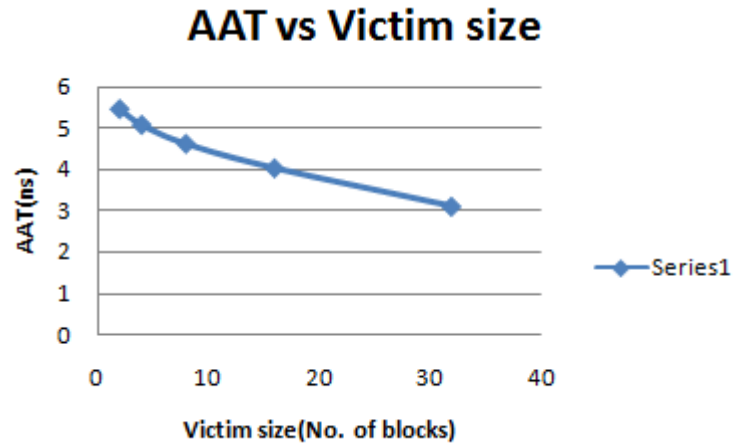


2.4.2 Cache parameters

Block size	Cache size	Associativity	Victim Size	Replacement	Write policy	Trace file
16	1024	1	Variable	LRU	Write back	perl_trace

Cache	Block size	No. of victim blocks	Miss rate	AAT (ns)
L1	16	2	25.10%	5.45
L1	16	4	23.25%	5.07
L1	16	8	21.02%	4.61

L1	16	16	18.15%	4.02
L1	16	32	13.59%	3.09

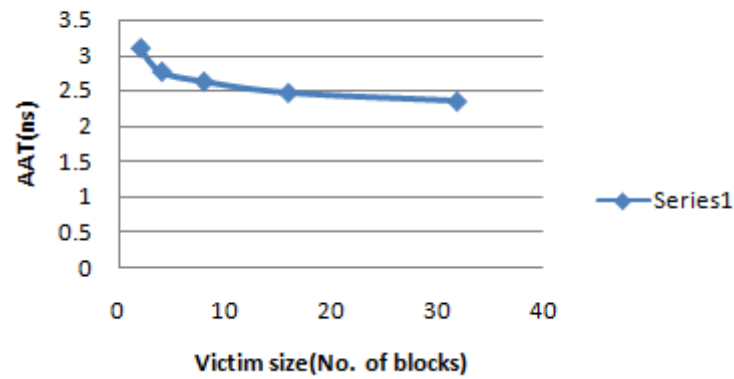


2.4.3 Cache parameters

Block size	Cache size	Associativity	Victim Size	Replacement	Write policy	Trace file
16	1024	1	Variable	LRU	Write back	go_trace

Cache	Block size	No. of victim blocks	Miss rate	AAT (ns)
L1	16	2	13.66%	3.10
L1	16	4	12.06%	2.77
L1	16	8	11.36%	2.63
L1	16	16	10.60%	2.47
L1	16	32	10.00%	2.35

AAT vs Victim size

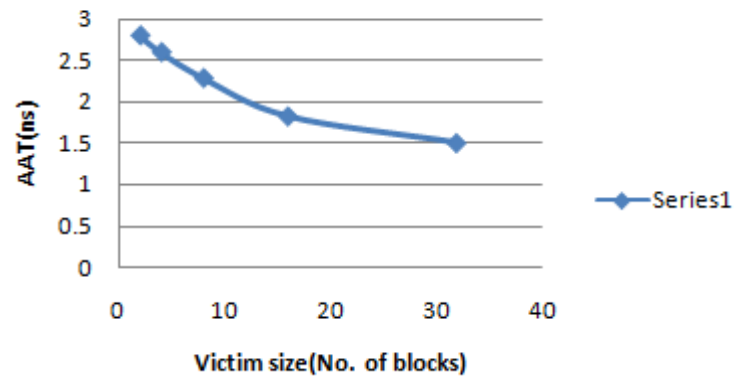


2.4.4 Cache parameters

Block size	Cache size	Associativity	Victim size	Replacement	Write policy	Trace file
16	1024	1	Variable	LRU	Write back	vortex_trace

Cache	Block size	No. of victim blocks	Miss rate	AAT (ns)
L1	16	2	12.21%	2.80
L1	16	4	11.22%	2.60
L1	16	8	9.72%	2.29
L1	16	16	7.46%	1.83
L1	16	32	5.89%	1.51

AAT vs Victim size



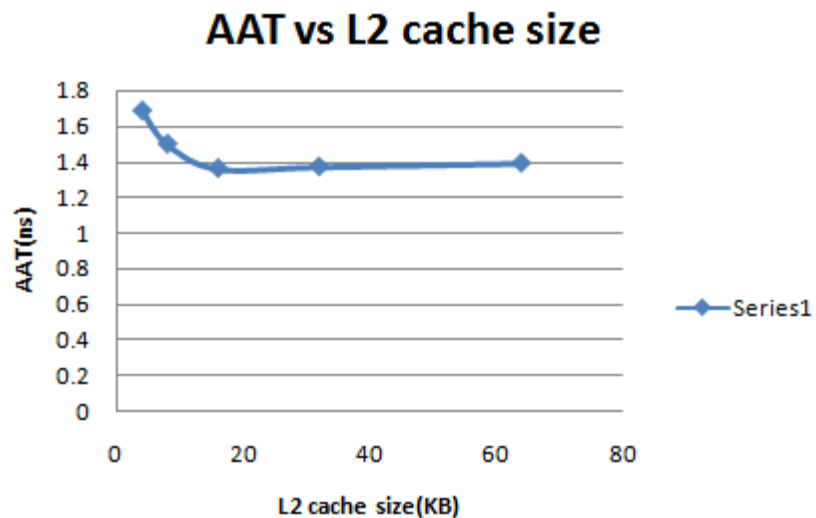
2.5 Effect of L2 cache on performance

Larger block L2 cache size follows the same trend as that for L1 cache size and decrease the overall AAT. But beyond a size of 32 KB, we observe AAT remains almost the same.

2.5.1 Cache parameters

Block size	L1 Cache size	Associativity	Victim Size	L2 Cache size	Replacement	Write policy	Trace file
16	1024	8	128	Variable	LRU	Write back	gcc_trace

Cache	Cache size (Kilobytes)	L1 Miss Rate	AAT (ns)
L2	4	15.49	1.69
L2	8	15.49	1.50
L2	16	15.49	1.36
L2	32	15.49	1.37
L2	64	15.49	1.39
L2	128	15.49	1.44

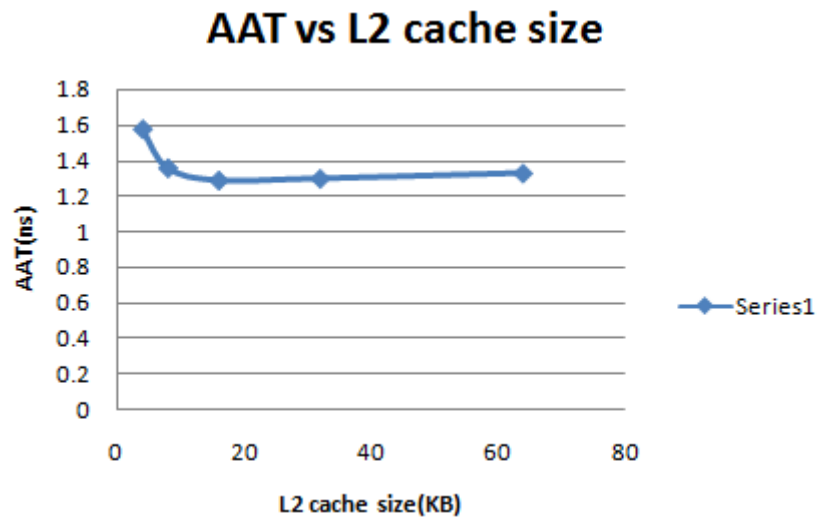


2.5.2 Cache parameters

Block size	L1 Cache size	Associativity	Victim Size	L2 Cache size	Replacement	Write policy	Trace file
16	1024	8	128	Variable	LRU	Write back	go_trace

Cache	Cache size (Kilobytes)	L1 Miss Rate	AAT (ns)
L2	4	21.02	1.58

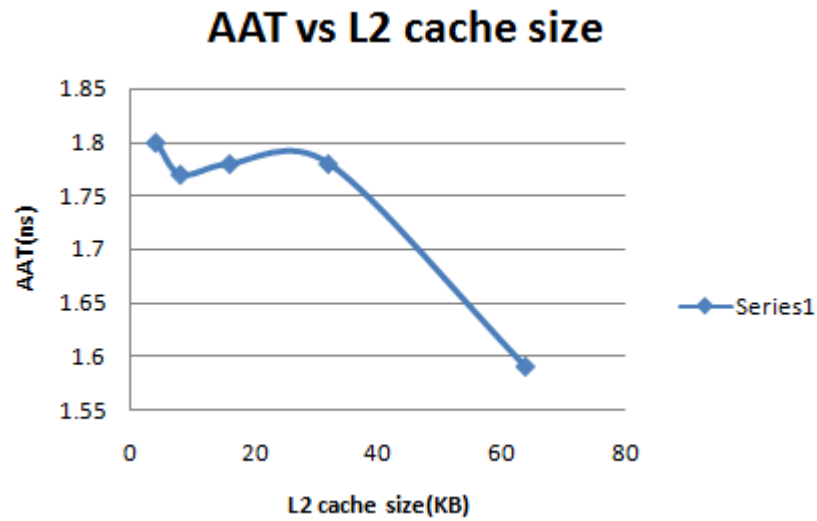
L2	8	21.02	1.36
L2	16	21.02	1.29
L2	32	21.02	1.30
L2	64	21.02	1.33
L2	128	21.02	1.39



2.5.3 Cache parameters

Block size	L1 Cache size	Associativity	Victim Size	L2 Cache size	Replacement	Write policy	Trace file
16	1024	8	128	Variable	LRU	Write back	perl_trace

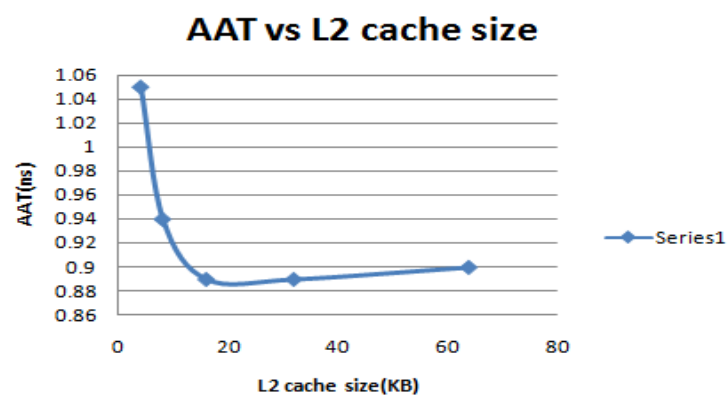
Cache	Cache size (Kilobytes)	L1 Miss Rate	AAT (ns)
L2	4	11.36	1.80
L2	8	11.36	1.77
L2	16	11.36	1.78
L2	32	11.36	1.78
L2	64	11.36	1.59
L2	128	11.36	1.57



2.5.4 Cache parameters

Block size	L1 Cache size	Associativity	Victim Size	L2 Cache size	Replacement	Write policy	Trace file
16	1024	8	128	Variable	LRU	Write back	vortex_trace

Cache	Cache size (Kilobytes)	L1 Miss Rate	AAT (ns)
L2	4	9.72	1.05
L2	8	9.72	0.94
L2	16	9.72	0.89
L2	32	9.72	0.89
L2	64	9.72	0.90
L2	128	9.72	0.94



From the data collected, the following configuration is best suited for the minimum miss rate and AAT

L1 Block size = 32 B

L1 Cache size = 4 KB

L1 Set associativity = 8

L1 Victim size blocks = 16 blocks or 32 blocks

L2 Block size = 32 B

L2 Cache size = 32 KB

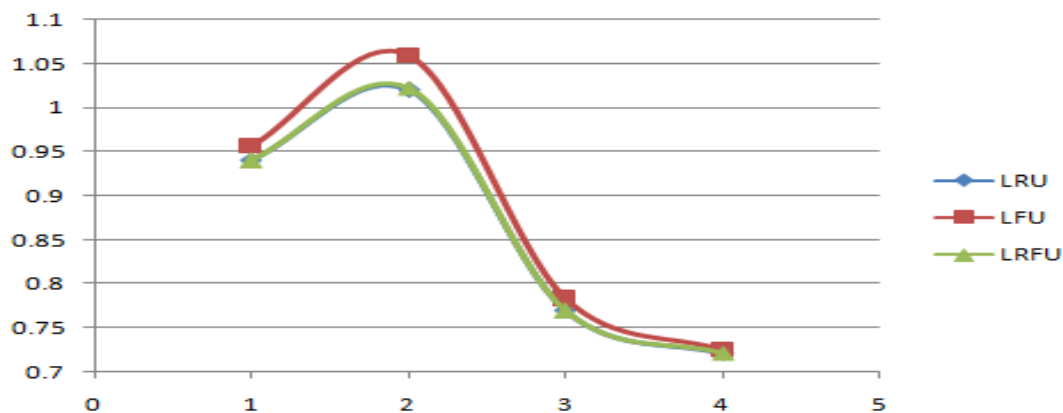
L2 Set associativity = 8

2.5 Effect of LRFU replacement policy on performance

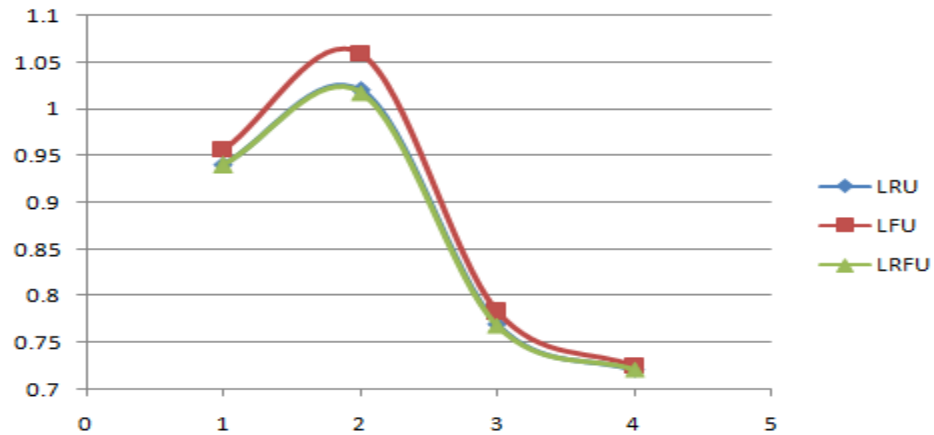
2.5.4 Cache parameters

Block size	L1 Cache size	Associativity	Victim Size	L2 Cache size	Replacement	Write policy	Trace file
64	8192	4	512	16384	LRFU	Write back	variable

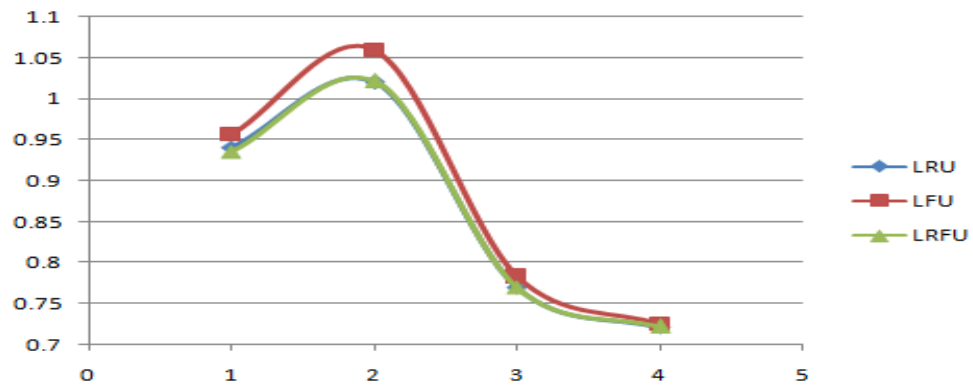
Trace file	Lamda	LRU AAT	LFU AAT	LRFU AAT
Gcc_trace	0.25	0.9411	0.9567	0.9399
Go_trace	0.25	1.0199	1.0596	1.0220
Perl_trace	0.25	0.7701	0.7833	0.7694
Vortex_trace	0.25	0.7211	0.7249	0.7212



Trace file	Lamda	LRU AAT	LFU AAT	LRFU AAT
Gcc_trace	0.3	0.9411	0.9567	0.9397
Go_trace	0.3	1.0199	1.0596	1.0178
Perl_trace	0.3	0.7701	0.7833	0.7689
Vortex_trace	0.3	0.7211	0.7249	0.7212



Trace file	Lamda	LRU AAT	LFU AAT	LRFU AAT
Gcc_trace	0.5	0.9411	0.9567	0.9361
Go_trace	0.5	1.0199	1.0596	1.0215
Perl_trace	0.5	0.7701	0.7833	0.7701
Vortex_trace	0.5	0.7211	0.7249	0.7221



From the values we can see that the LRFU replacement policy has slight edge over both LRU and LFU for varied traces for different values of lamda.