Simulation of Cache

Project Report

Instructor	*: Prof. Yan Solihin*
Student	*: Pratheek Bramhasamudra Mallikarjuna*
ID	*: 200065594*
E-mail	*: pbramha@ncsu.edu*

I, Pratheek Bramhasamudra Mallikarjuna, hereby declare that, I have complied with all the requirements necessary for maintaining academic integrity to the best of my knowledge while doing this project.

Pratheek Bramhasamudra Mallikarjuna

Table of Contents

Introduction

Year after year Processors are becoming faster and faster. The rate at which processor speed increased was large compared to rate at which memory access speed increased. Because of which accessing DRAM became the bottleneck of computer systems. To solve this problem memory hierarchy was introduced. Here processor talks to small faster memories called cache, which in turn talks to physical memory. Also, cache serves as temporary storage of recently used data, which works of well because of the locality nature of the memory access of the programs.

This project deals with simulation of the memory hierarchy with L1 and L2 caches. It also deals with implementing cache inclusion protocols like Inclusive Policy, Non-Inclusive policy and Exclusive policy. Along with these, it also deals with implementation of Replacement policies like LRU, FIFO and Pseudo-LRU. Various tests are conducted on this implementation to see the effects of Block size, associativity, Cache Size, Cache Inclusion policies and Replacement policies etc.

Approach

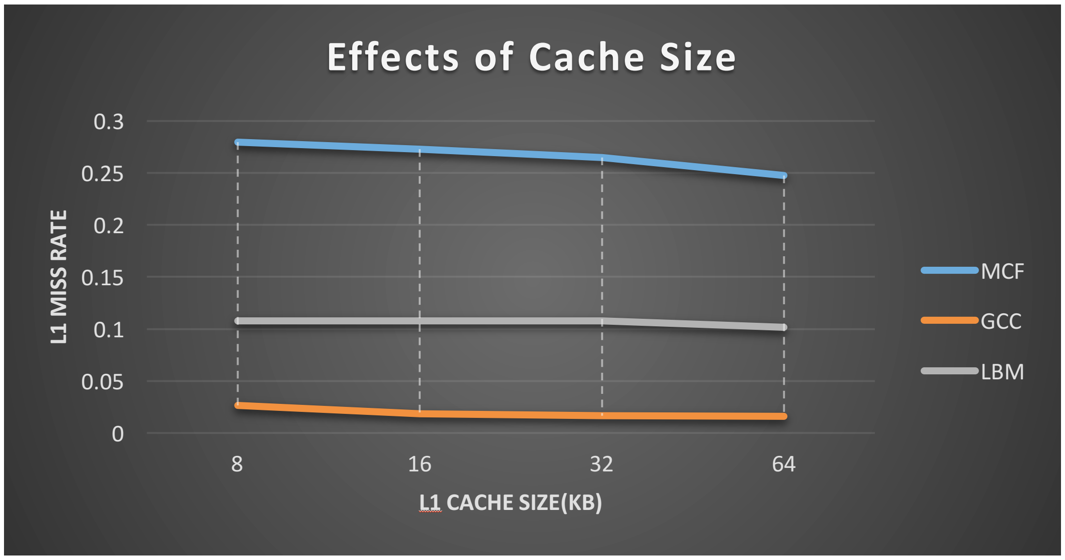
The memory requests from the processors are read from a text file and fed to cache block. A single *Cache Class* is designed to handle L1, L2 and L3 etc. Cache class constructor allocates required memory while destructor de-allocates which object is destroyed. Two main member functions are *read* and *write*. They call various other member functions to achieve the required specifications. Separate methods have been implemented to update LRU, get LRU block, evicting a block etc., which makes the project modular. It can be easily extended to more cache levels making in highly scalable.

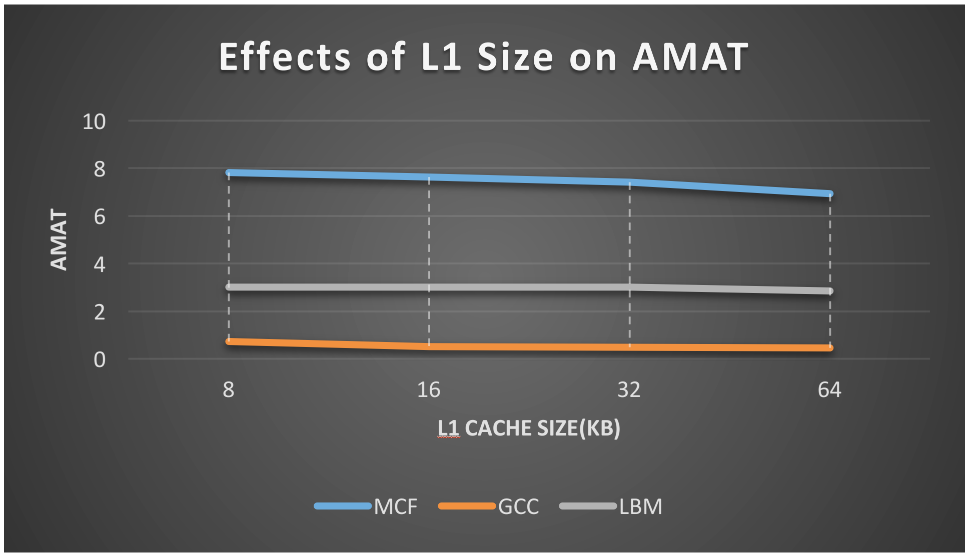
Results

1. Cache Explorations

1.1 Effects of L1 Cache Size

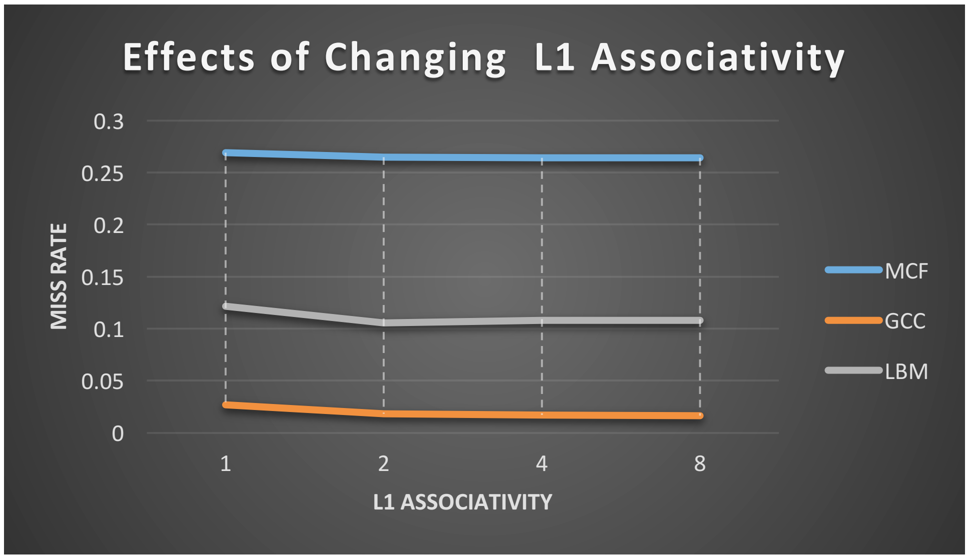
L1 block size is 64 and associativity at 4. As we can see in Figure 1, miss rates decrease with increase in cache size. Also, AMAT decreases with increase in cache size as shown in the following figures.

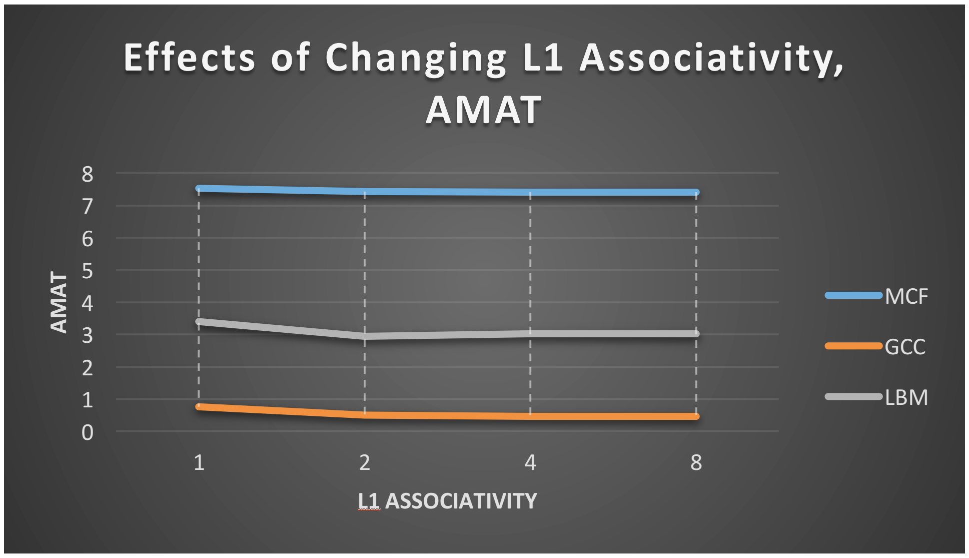




1.2 Effects of L1 Cache Associativity

We can see that, as associativity increases, miss rate decreases initially and then it gradually gets saturated. The trend continues for AMAT as well because of the diminishing returns.



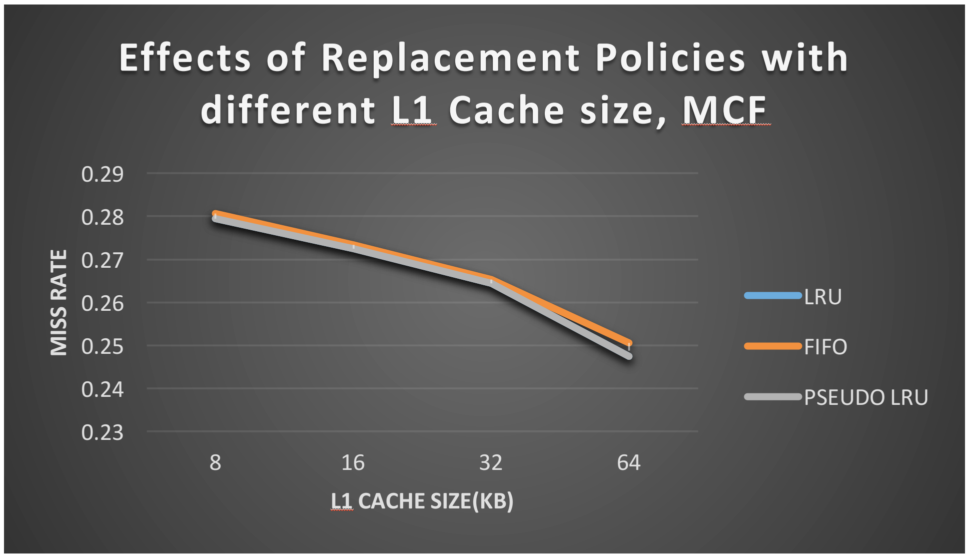


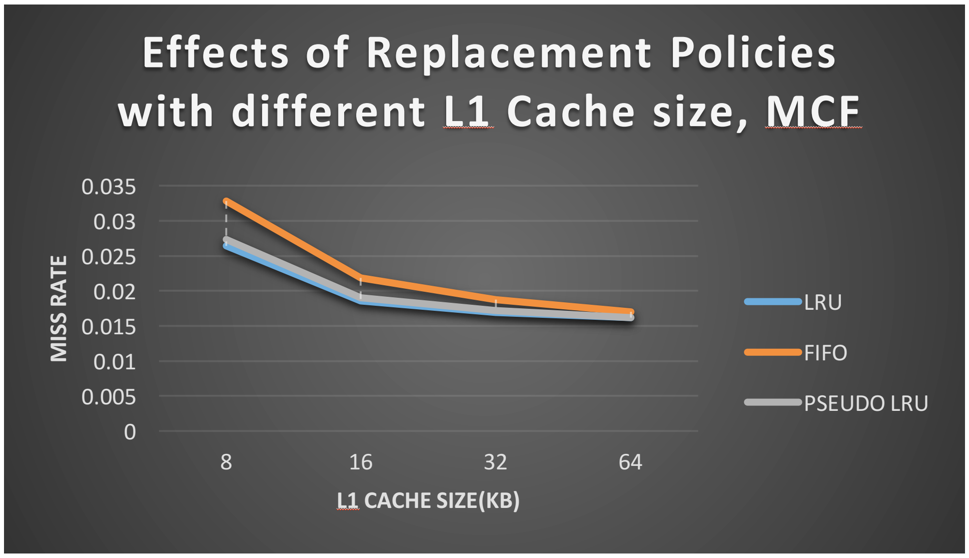
1.3 Discussions

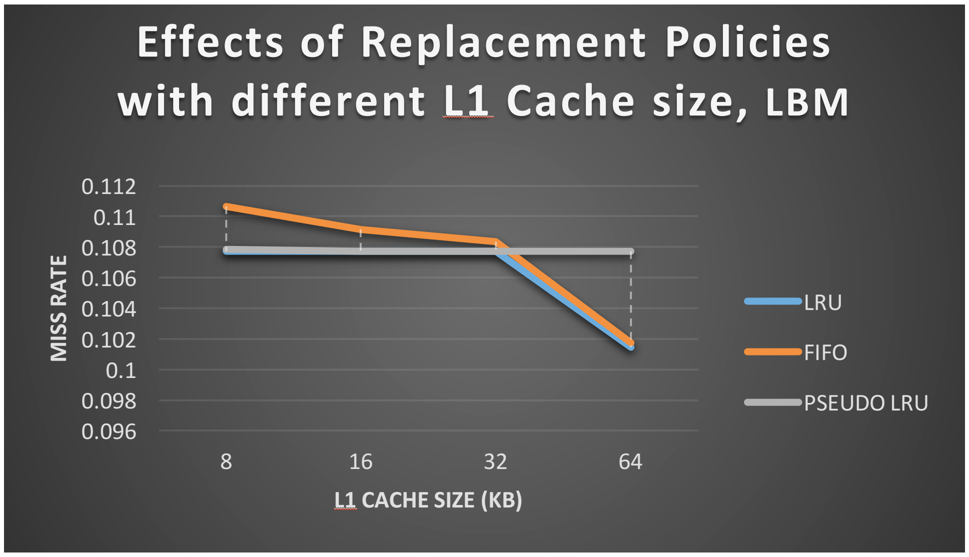
- AMAT tends to decrease with increase in cache size or associativity, but with diminishing returns.
- Since there is no L2 cache, low miss L1 miss rate gives low AMAT.
- According to the AMAT graphs, Cache size of 16/32 gives high performance for moderately less cache size. Since, it gives good performance for associativity of 2/4 as per both the figures, we can go with associativity of 2/4.

2. Cache Replacement Policy

The performance of different Replacement Policies with different L1 cache size is depicted in the following figures. Performance of the policies seems to be on par with each other, except that FIFO policy seems to have higher miss rate for low cache size.



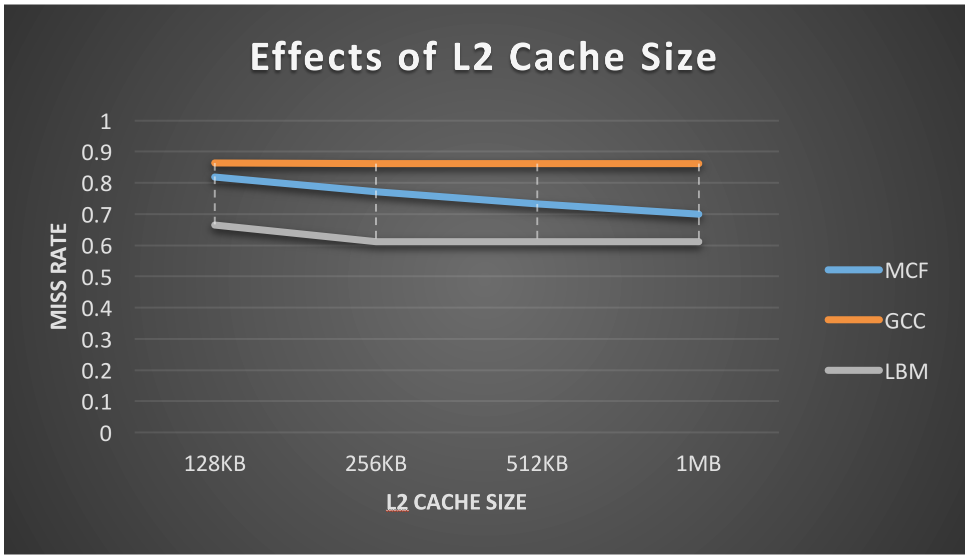




3. L2 Cache Exploration

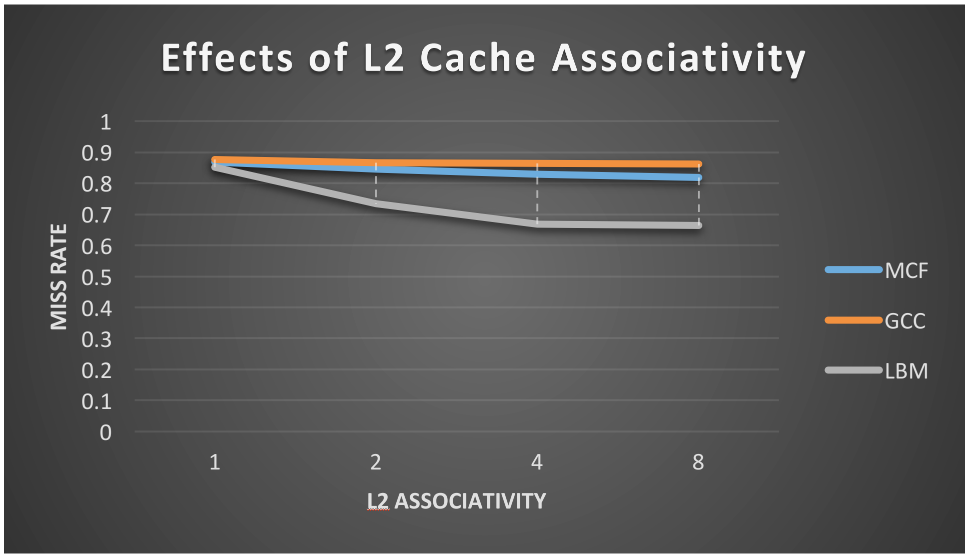
3. 1 Effects of L2 Cache size

As we can see, increasing L2 cache size has less effect on miss rates, however, even though miss rates decrease in the beginning, they tend to saturate quickly. 256 – 512 KB of cache size seem to give god enough performance considering their low area.



3. 2 Effects of L2 Cache Associativity

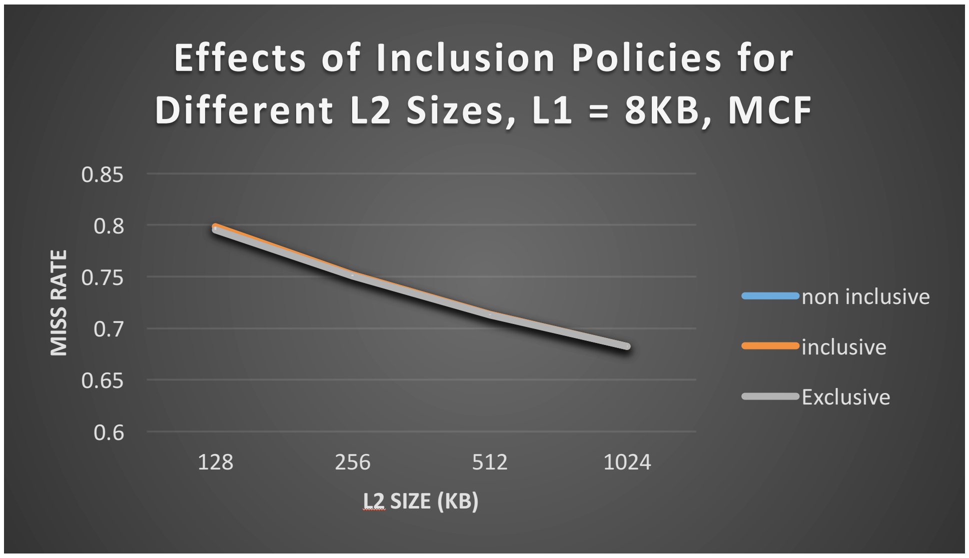
Even here, even though miss rates decrease in the beginning, they tend to saturate for higher associativity values. An associativity of 4 seems to be a good pick.

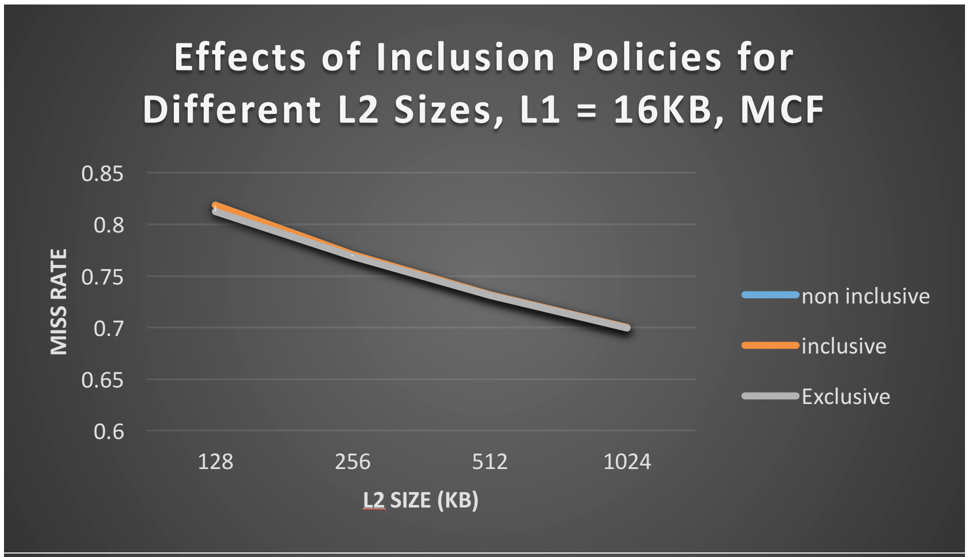


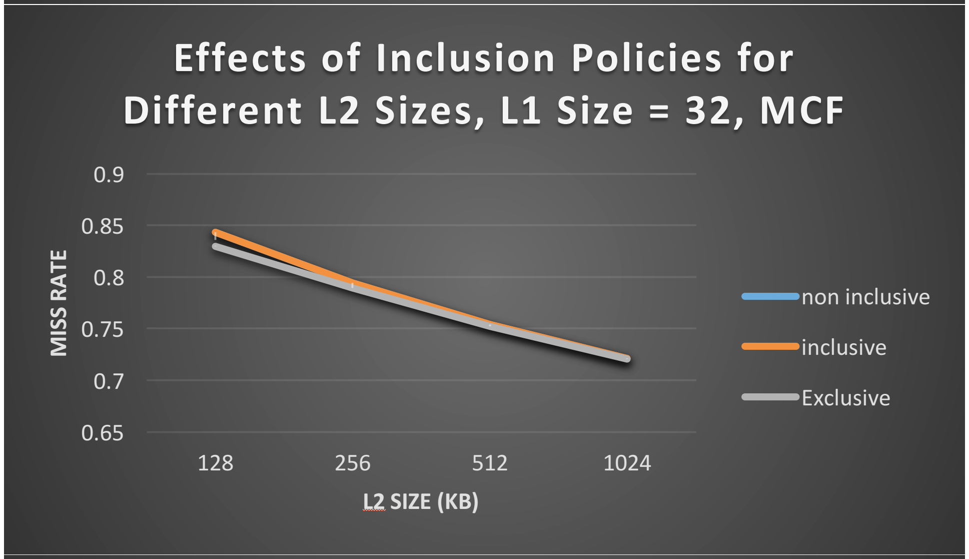
4. Inclusion Policy

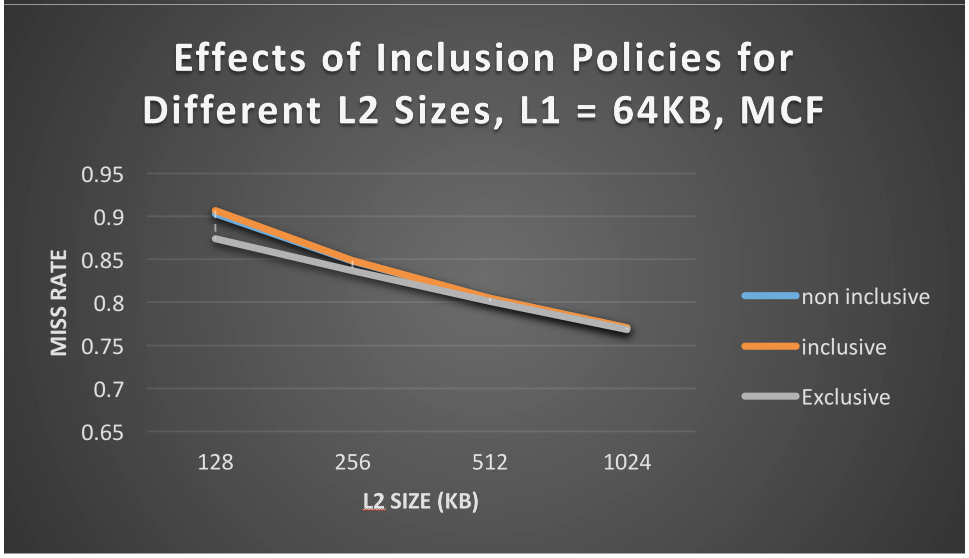
Performance is analyzed for various L1 sizes and L2 sizes for different Inclusion policies. Experiment is repeated for different traces. Following are the plots.

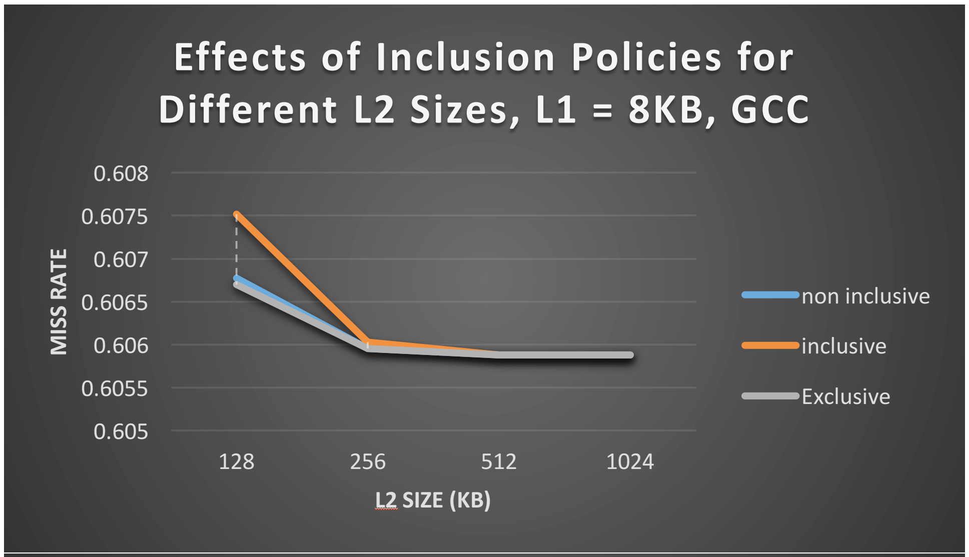
4.1 Miss Rates

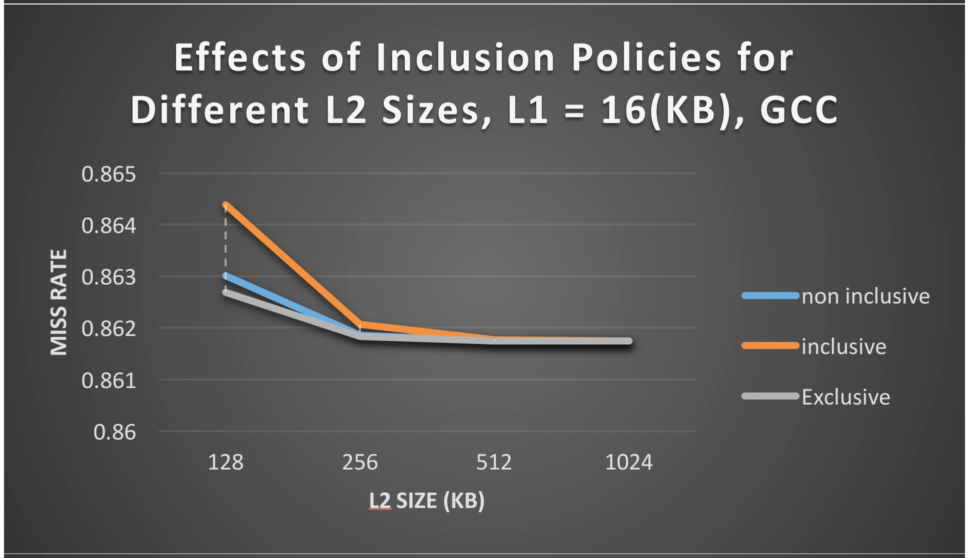


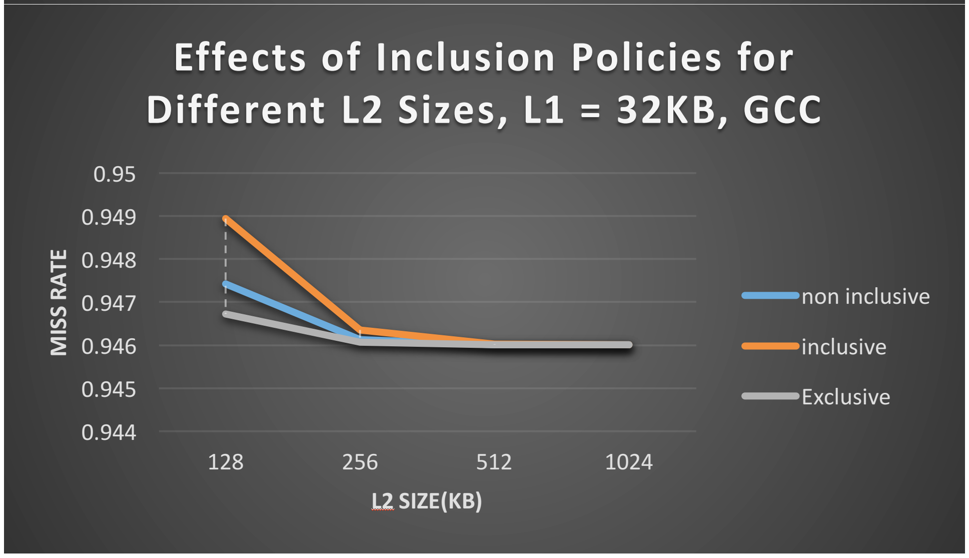

Effects of Inclusion Policies for Different L2 Sizes, L1 = 16KB, MCF
MISS RATE
0.85
0.8
0.75
0.7
0.65
0.6
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

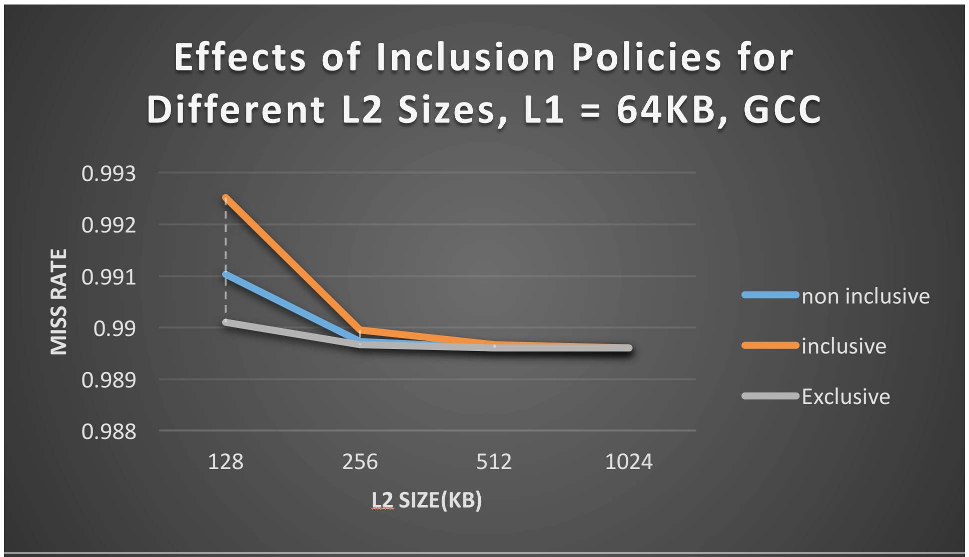

Effects of Inclusion Policies for Different L2 Sizes, L1 Size = 32, MCF
MISS RATE
0.9
0.85
0.8
0.75
0.7
0.65
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

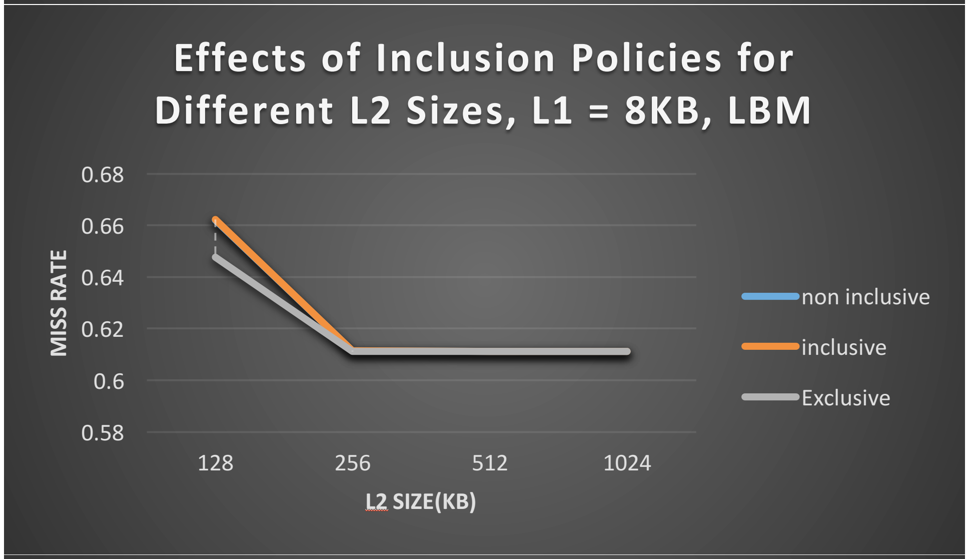

Effects of Inclusion Policies for Different L2 Sizes, L1 = 64KB, MCF
MISS RATE
0.95
0.9
0.85
0.8
0.75
0.7
0.65
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

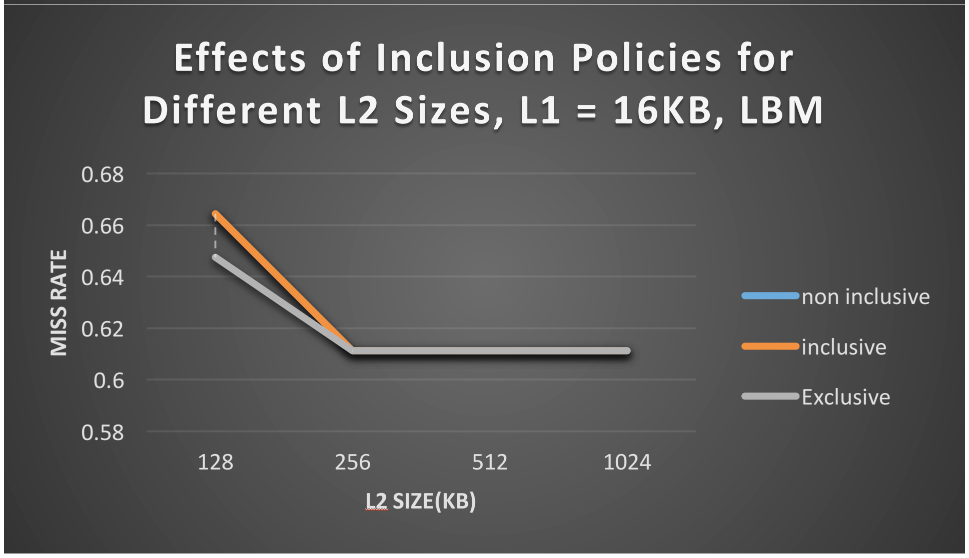

Effects of Inclusion Policies for Different L2 Sizes, L1 = 8KB, GCC
MISS RATE
0.608
0.6075
0.607
0.6065
0.606
0.6055
0.605
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

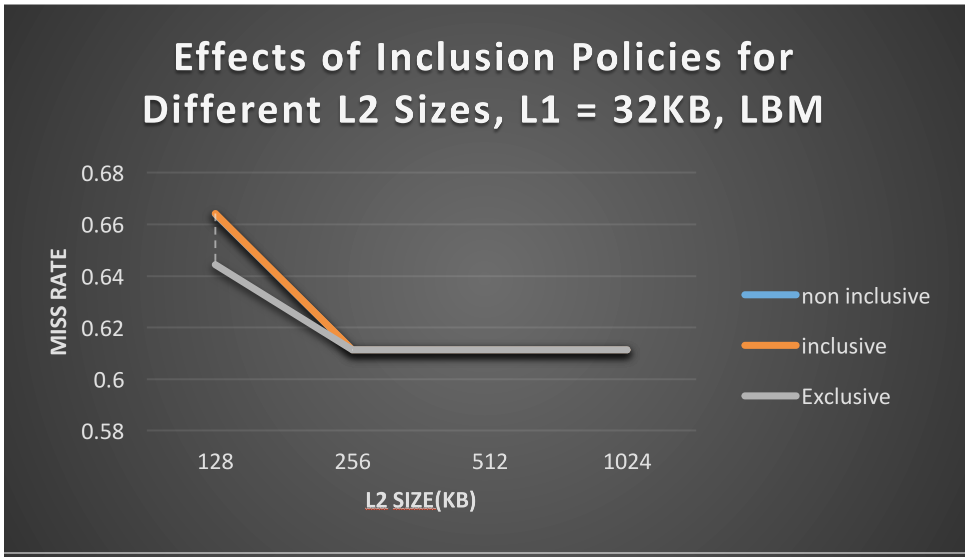

Effects of Inclusion Policies for Different L2 Sizes, L1 = 16(KB), GCC
MISS RATE
0.865
0.864
0.863
0.862
0.861
0.86
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

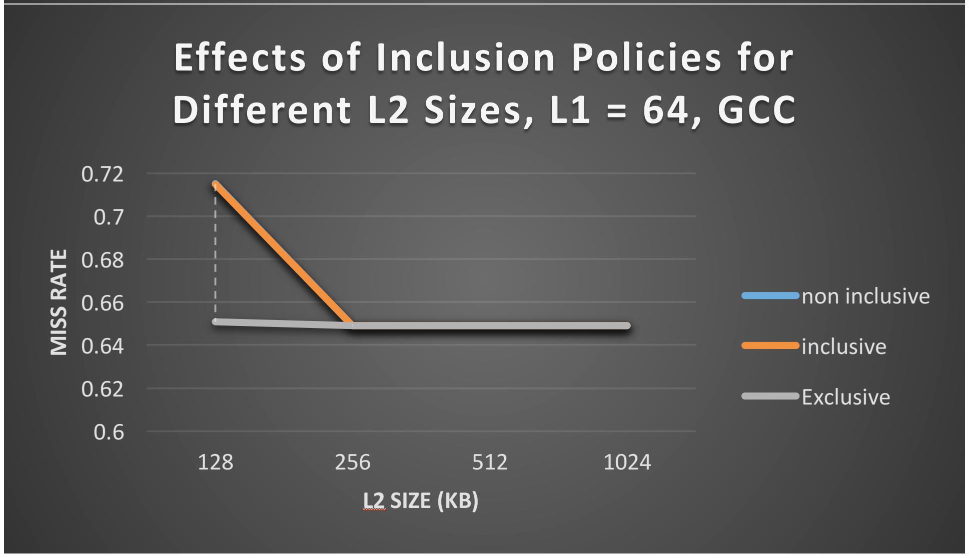

Effects of Inclusion Policies for Different L2 Sizes, L1 = 32KB, GCC
MISS RATE
0.95
0.949
0.948
0.947
0.946
0.945
0.944
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive


Effects of Inclusion Policies for Different L2 Sizes, L1 = 64KB, GCC
MISS RATE
0.993
0.992
0.991
0.99
0.989
0.988
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive

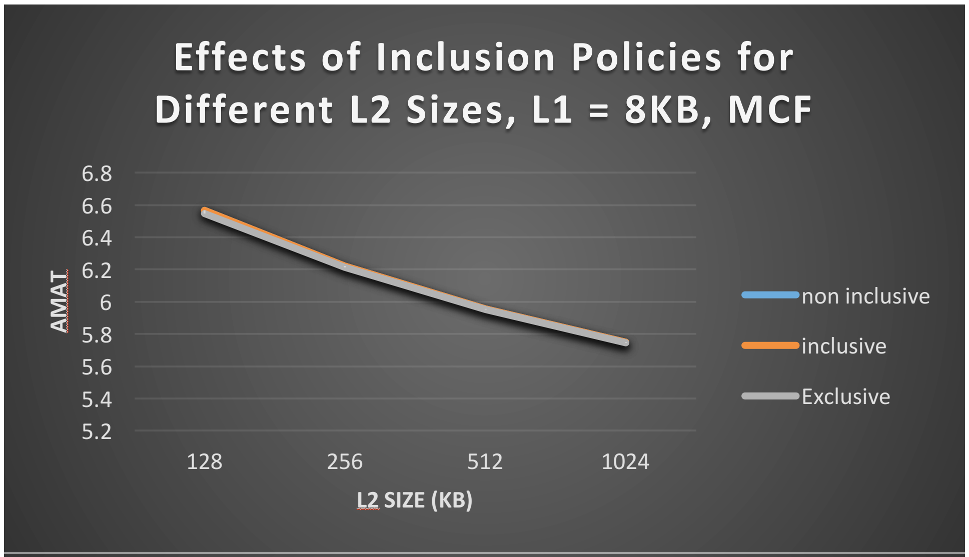

Effects of Inclusion Policies for Different L2 Sizes, L1 = 8KB, LBM
MISS RATE
0.68
0.66
0.64
0.62
0.6
0.58
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive

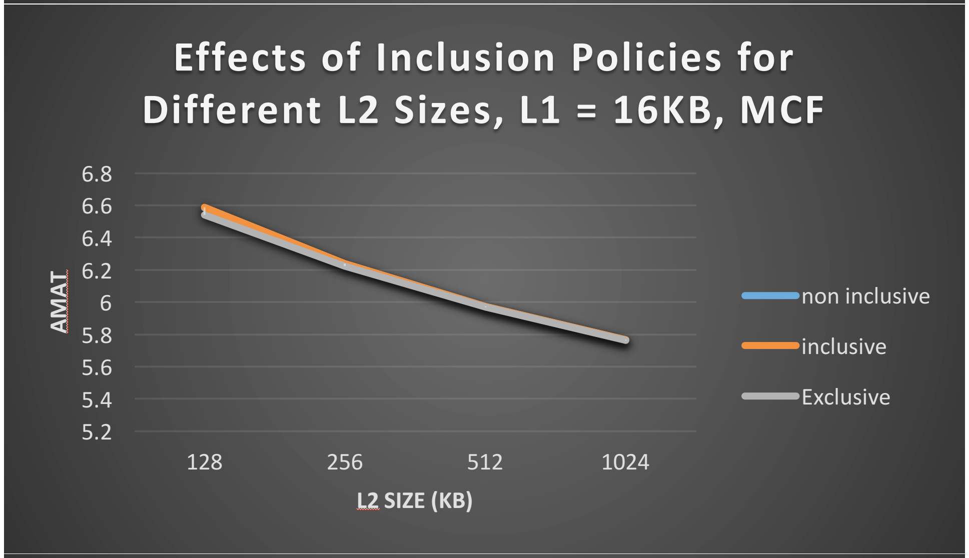

Effects of Inclusion Policies for Different L2 Sizes, L1 = 16KB, LBM
MISS RATE
0.68
0.66
0.64
0.62
0.6
0.58
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive

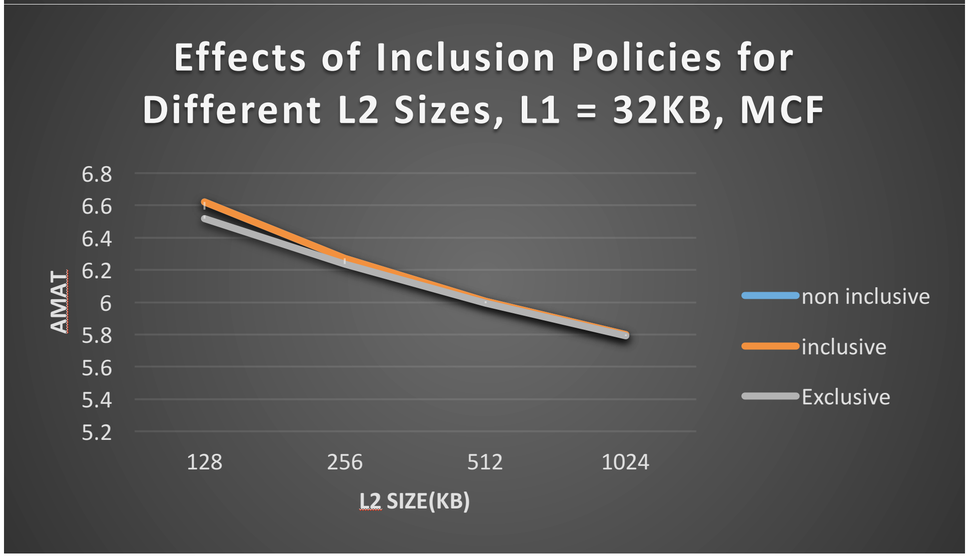

Effects of Inclusion Policies for Different L2 Sizes, L1 = 32KB, LBM
MISS RATE
0.68
0.66
0.64
0.62
0.6
0.58
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive

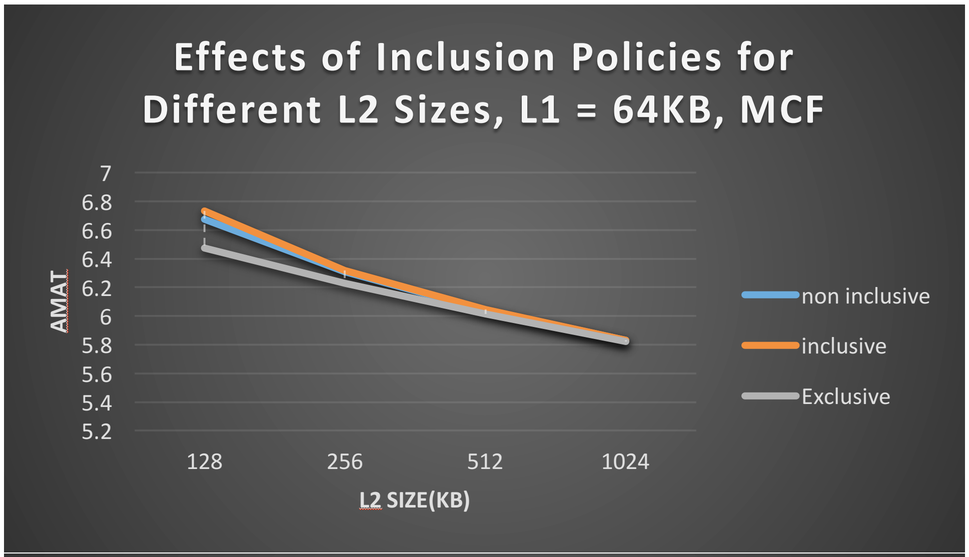

Effects of Inclusion Policies for Different L2 Sizes, L1 = 64, GCC
MISS RATE
0.72
0.7
0.68
0.66
0.64
0.62
0.6
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

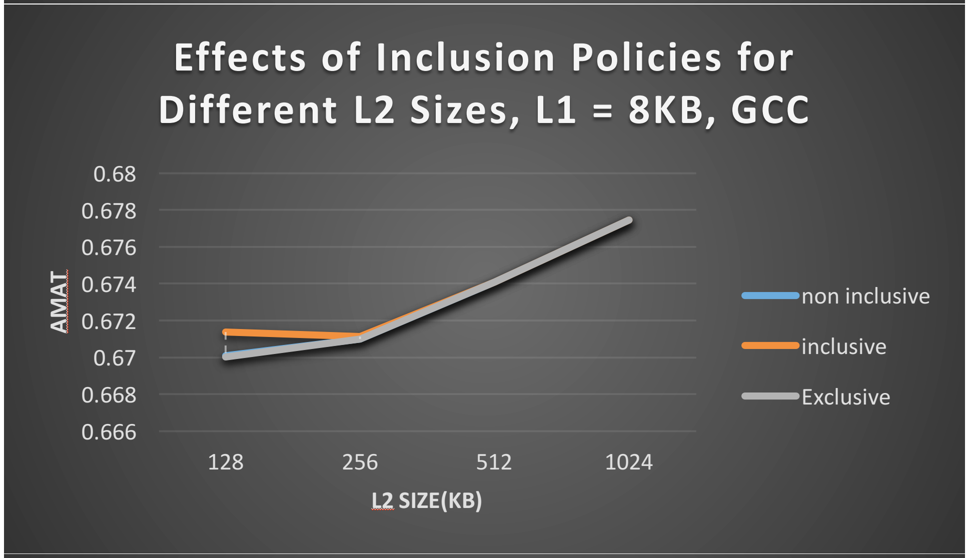
4.2 AMAT

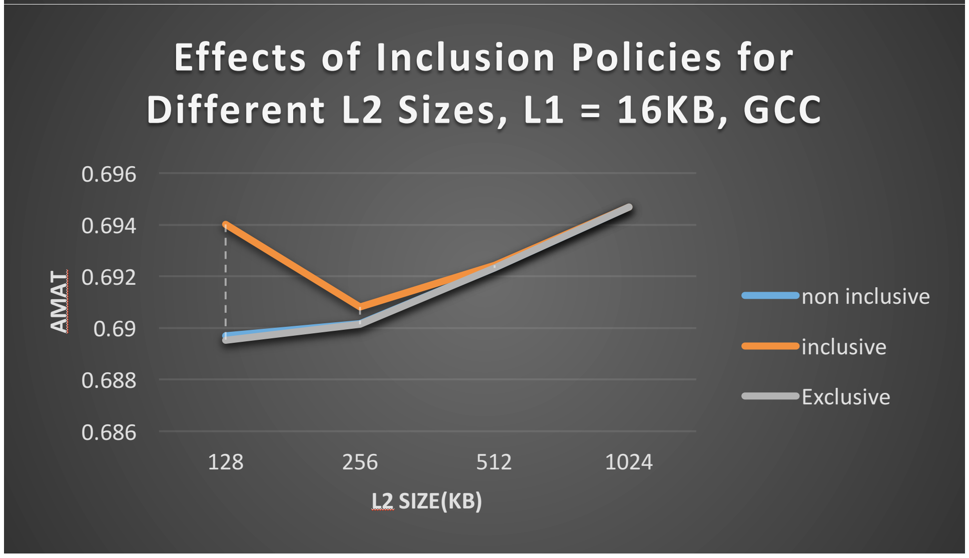

Effects of Inclusion Policies for Different L2 Sizes, L1 = 8KB, MCF
AMAT
6.8
6.6
6.4
6.2
6
5.8
5.6
5.4
5.2
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

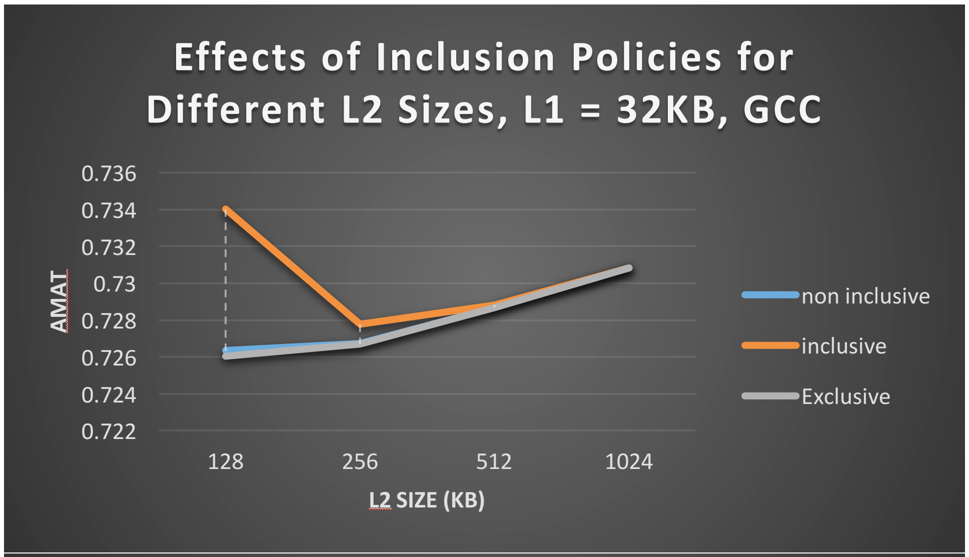

Effects of Inclusion Policies for Different L2 Sizes, L1 = 16KB, MCF
AMAT
6.8
6.6
6.4
6.2
6
5.8
5.6
5.4
5.2
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

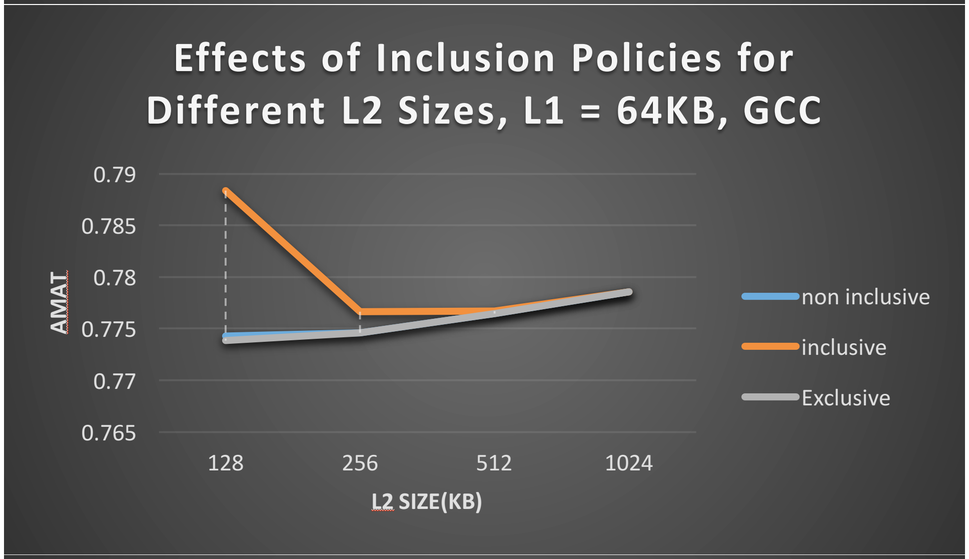

Effects of Inclusion Policies for Different L2 Sizes, L1 = 32KB, MCF
AMAT
6.8
6.6
6.4
6.2
6
5.8
5.6
5.4
5.2
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive

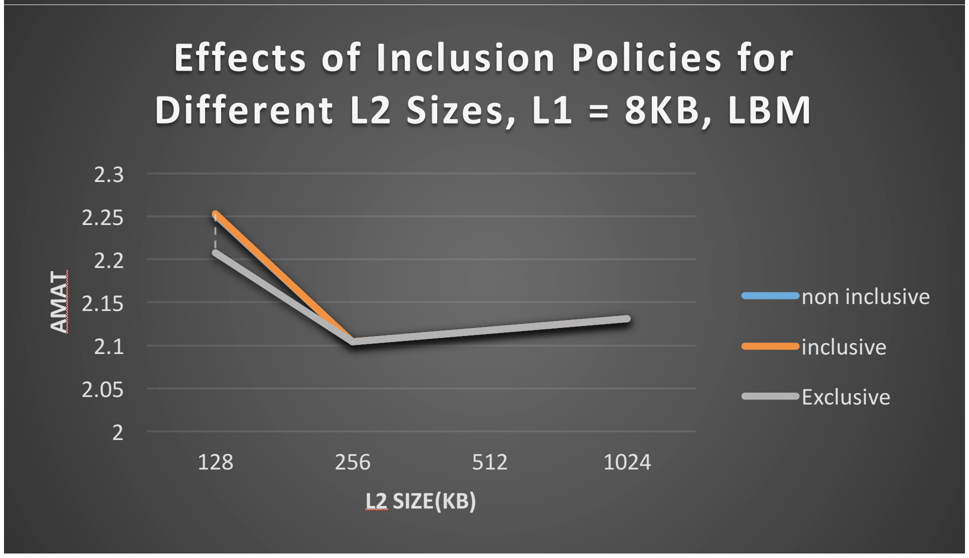

Effects of Inclusion Policies for Different L2 Sizes, L1 = 64KB, MCF
AMAT
7
6.8
6.6
6.4
6.2
6
5.8
5.6
5.4
5.2
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive

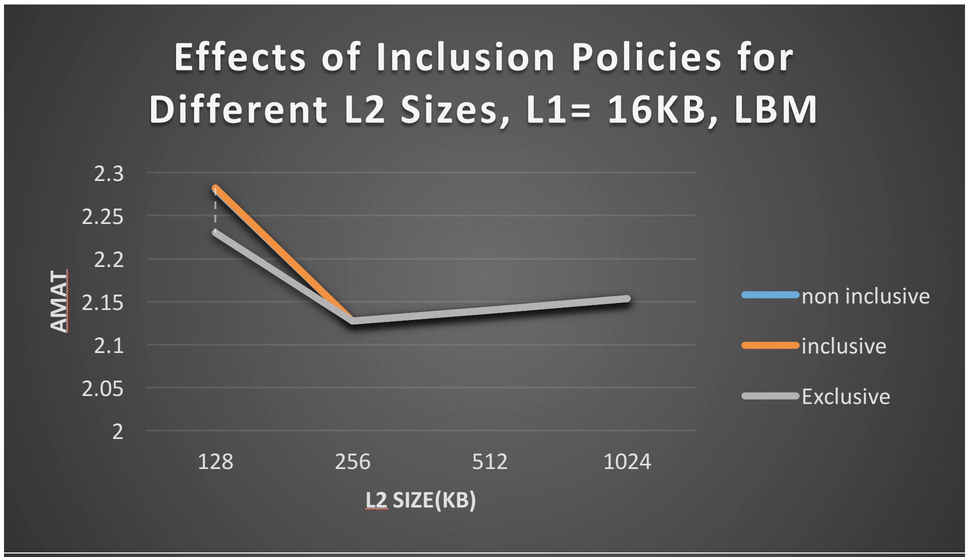

Effects of Inclusion Policies for Different L2 Sizes, L1 = 8KB, GCC
AMAT
0.68
0.678
0.676
0.674
0.672
0.67
0.668
0.666
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive

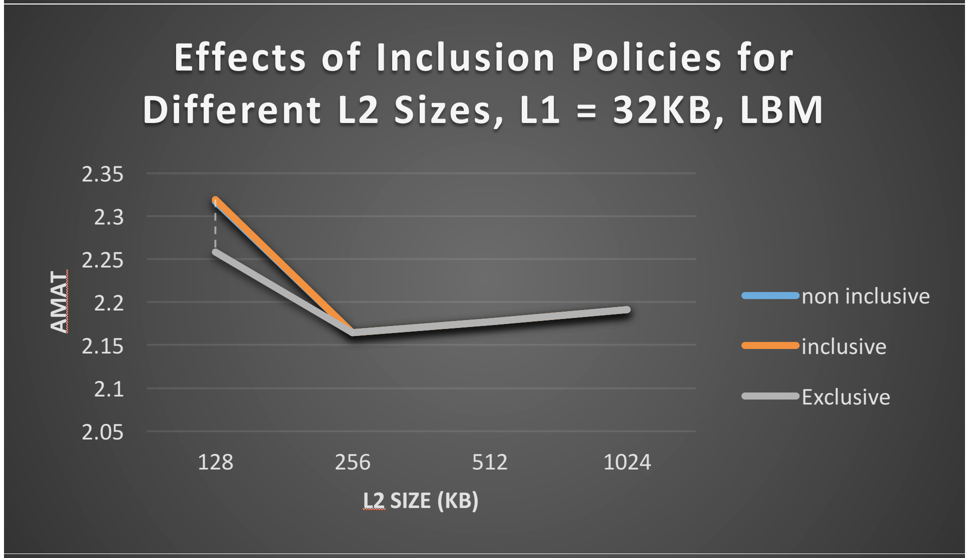

Effects of Inclusion Policies for Different L2 Sizes, L1 = 16KB, GCC
AMAT
0.696
0.694
0.692
0.69
0.688
0.686
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive

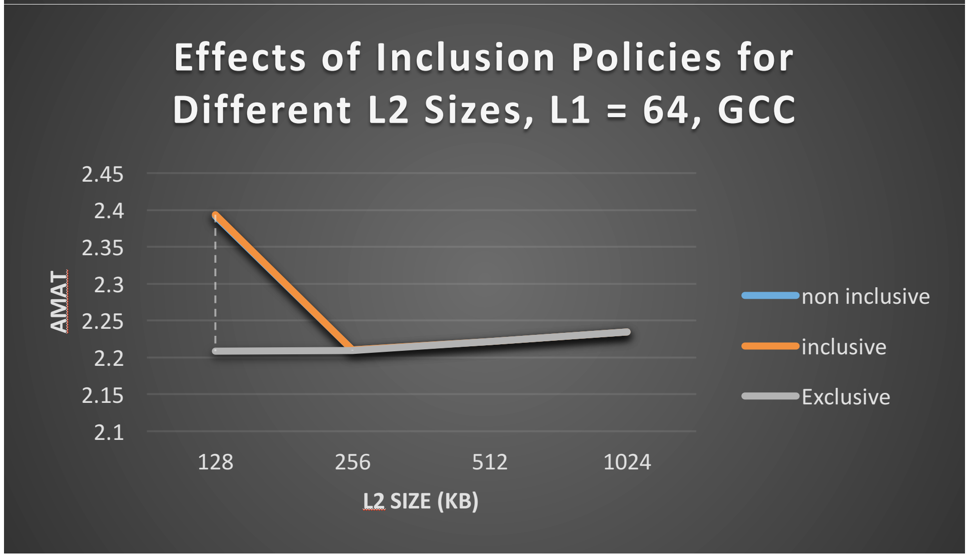

Effects of Inclusion Policies for Different L2 Sizes, L1 = 32KB, GCC
AMAT
0.736
0.734
0.732
0.73
0.728
0.726
0.724
0.722
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive


Effects of Inclusion Policies for Different L2 Sizes, L1 = 64KB, GCC
AMAT
0.79
0.785
0.78
0.775
0.77
0.765
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive


Effects of Inclusion Policies for Different L2 Sizes, L1 = 8KB, LBM
AMAT
2.3
2.25
2.2
2.15
2.1
2.05
2
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive


Effects of Inclusion Policies for Different L2 Sizes, L1= 16KB, LBM
AMAT
2.3
2.25
2.2
2.15
2.1
2.05
2
128
256
512
1024
L2 SIZE(KB)
non inclusive
inclusive
Exclusive


Effects of Inclusion Policies for Different L2 Sizes, L1 = 32KB, LBM
AMAT
2.35
2.3
2.25
2.2
2.15
2.1
2.05
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive


Effects of Inclusion Policies for Different L2 Sizes, L1 = 64, GCC
AMAT
2.45
2.4
2.35
2.3
2.25
2.2
2.15
2.1
128
256
512
1024
L2 SIZE (KB)
non inclusive
inclusive
Exclusive

4.1 Discussions

- From the plots, we can see that, exclusive policy gives better performance with large L1 cache size. This is because of fewer write-backs from L1 to L2. However, with higher L2 size, exclusive almost have same performance as non-inclusive.

- Also, from the plots, inclusive caches tend to perform lower with bigger L1 cache, because of the wasted space, especially when L2 cache is small.

- Another trend we can see is, even though AMAT decreases for increase in L2 cache size, it increases if the L2 cache size is too high. This is because of the additional hit latency of L2 due to increases size.