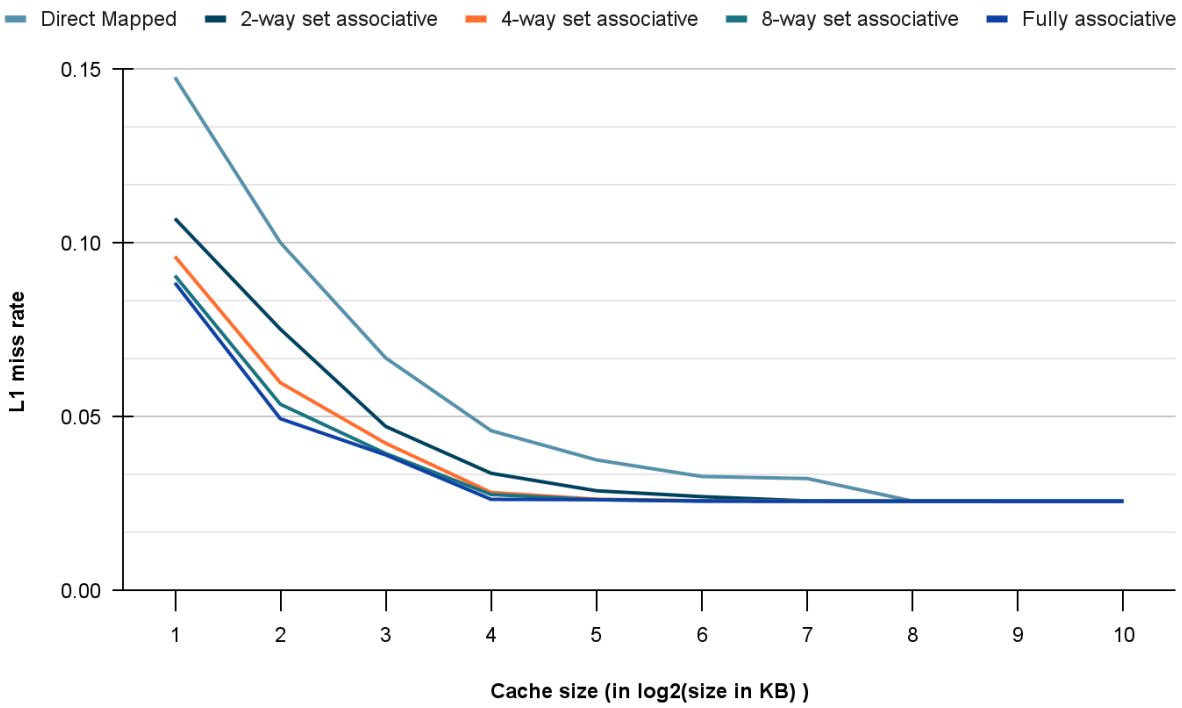
CS6600 : COMPUTER ARCHITECTURE

CACHE SIMULATOR REPORT

ROLL NUMBER: CS24M012

L1 Cache Investigation: SIZE and ASSOC

Plot 1: (Cache Size Vs L1 miss rate)



Trends in the Graph:

As the cache size increases, L1 miss rate decreases for the cache with associativity held constant. Larger caches have more capacity to store blocks and distribute across many sets. So miss rate decreases even though associativity is held constant.

For the same cache size, miss rate decreases with the increase in associativity exhibiting the fully associative cache with lowest miss rate and direct mapped cache with highest miss rate. As the cache is held constant, increasing associativity can reduce conflict misses.

After reaching a certain level of cache size, increasing cache size or associativity shows that rate of decrease of miss rate is almost zero. (approx constant miss rate)

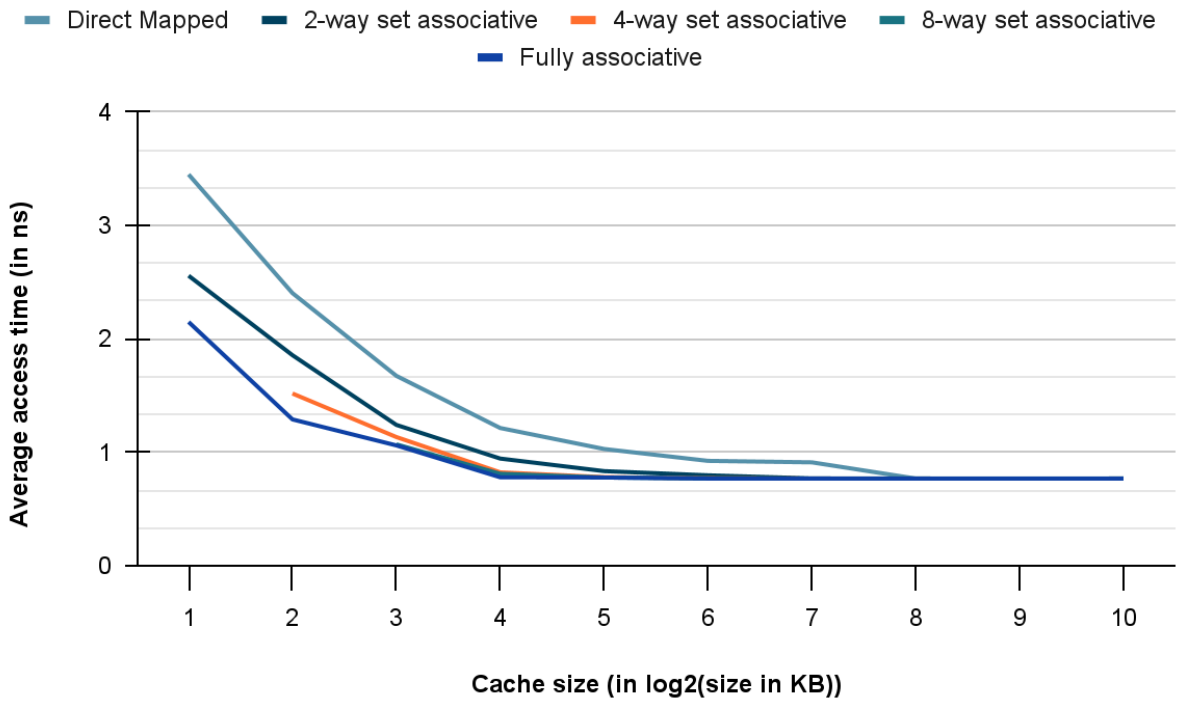
Compulsory Miss rate:

We know that, fully associative cache contains only compulsory misses and capacity misses. As the size of FA cache increases, capacity misses gradually reduces. And from the graph its clear that above 2^8 = 256KB, miss rate is not changing drastically or almost constant which implies capacity of the cache is no longer effecting miss rate. The only misses which are not effected by assoc or capacity of cache are compulsory misses. Hence we can consider compulsory miss rate as the miss rate of 1MB(= 2^{10} KB) fully associative cache without any issues.
Therefore, compulsory miss rate = 0.02582. (2.58% of requests are compulsory misses)

Conflict Miss rate:

We know that, # of misses in FA cache = # of compulsory misses + # of capacity misses
So for each cache size, we can simulate a FA cache of same size to get combined miss rate of compulsory misses and capacity misses. This leaves us the with the fact that for each cache size with associativity k, # of conflict misses = total # of misses - (# of misses of FA cache with same size). This was we can estimate conflict misses for a cache with any cache size and associativity.
For 2KB cache, conflict miss rates are as follows:
Direct mapped : 0.06 , assoc = 2: 0.02, assoc = 4: 0.007, assoc = 8: 0.002 , FA: 0

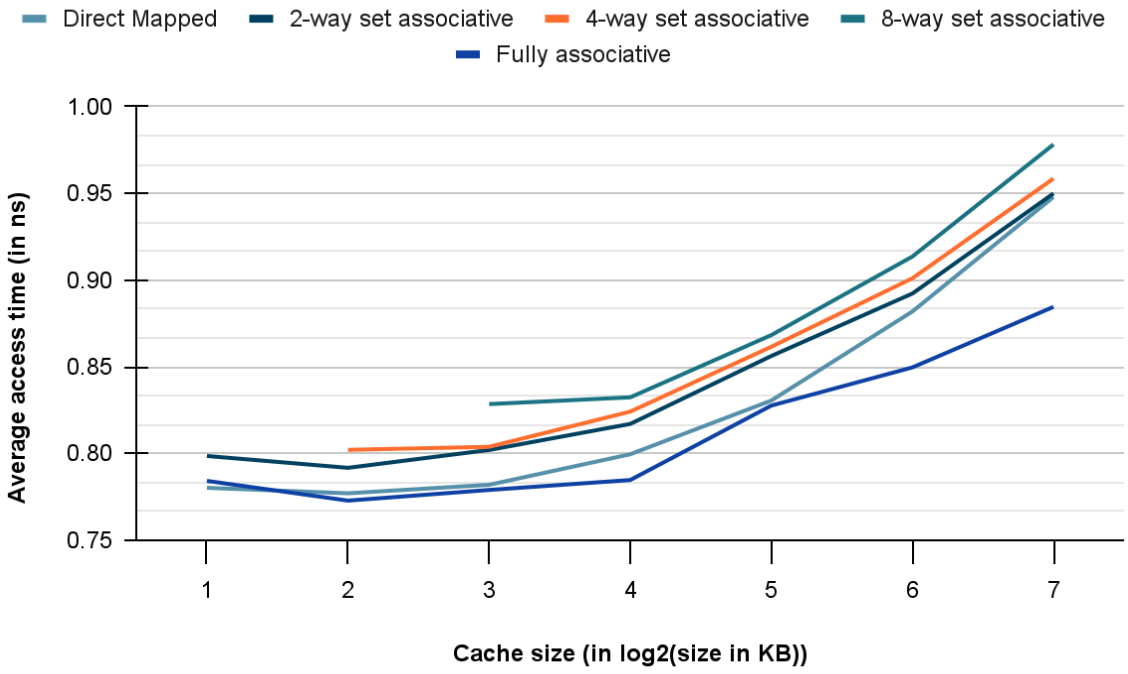
Plot 2: (Cache Size Vs Average access time)



Similar to plot1, just its plotted against average access time instead of L1 miss rate. Since we are plotting for the L1 only cache configuration, average access time is directly proportional to L1 miss rate. So the pattern of plot almost remains the same as plot 1. Few configurations for which cacti failed are not plotted

From the graph, it's clear that for associativity = 4/8, cache sizes from almost from around 32KB-1MB it achieved AAT to be almost constant which is the minimum. For associativity = 2, cache sizes from almost around 64KB it achieved AAT to be almost constant same as previous. Whereas for direct mapped cache to reach same constant value of AAT, it should have a cache size of around 256KB.
We can say the configuration which yields best is 32KB - 256KB cache with associativity as 4.

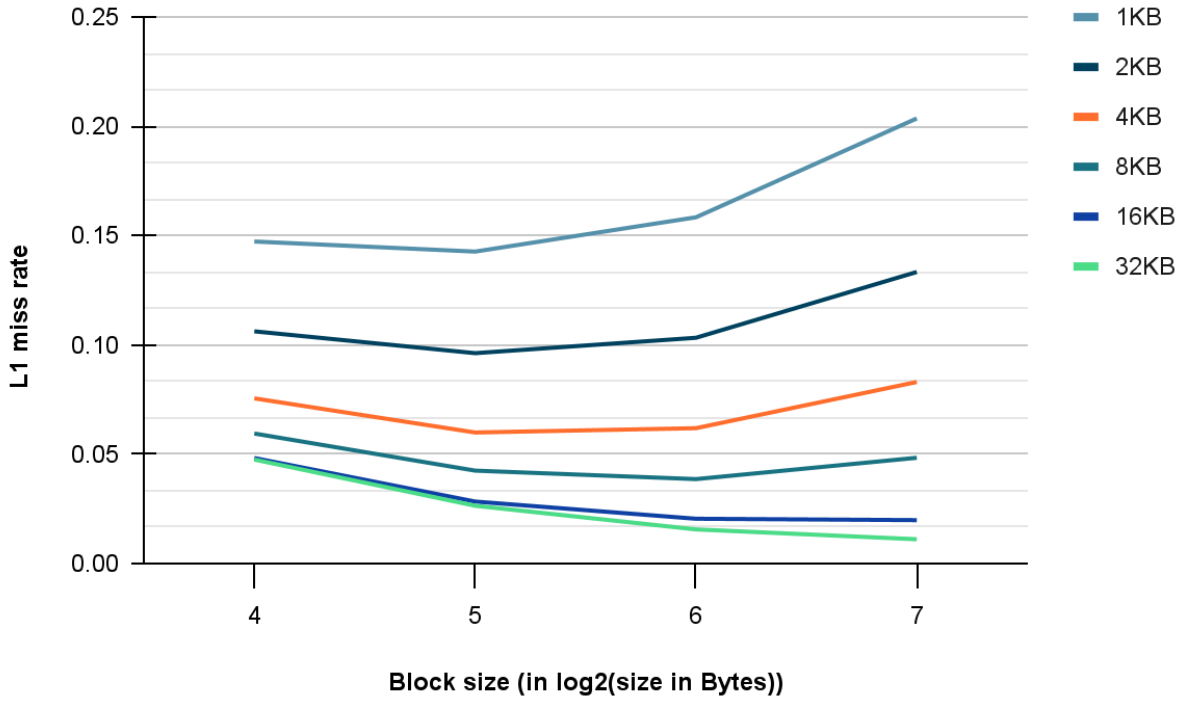
Plot 3: (Cache Size Vs Average access time)



L1 cache configuration that yields best AAT is fully associative with size of 4KB.
Both the plots have almost around 0.77ns as their optimal AAT. There is improvement of around 1.02 %.
While for total area metric, when L2 included in plot3 config has more area than plot2 configuration. (Plot3 area = 1.17844 mm^2, Plot2 area = 0.183217 mm^2)
EDP: Plot3 = 1.478 * 10^{-11} J.s, Plot2 = 9.979 * 10^{-11} J.s

L1 Cache Investigation: SIZE and BLOCKSIZE

Plot 4: (Cache Size Vs L1 miss rate)



Spatial locality vs Cache pollution:

It is evident from the graph the tradeoff between the spatial locality and cache pollution. For smaller cache sizes (1KB - 8KB), when block size is increased generally it benefits spatial locality and miss rate can reduce. But if we observe the graph, initially for these cache sizes, increase in block size reduced the miss rate but after a certain block size, cache pollution was happening and as block size increased, miss rate also increased.

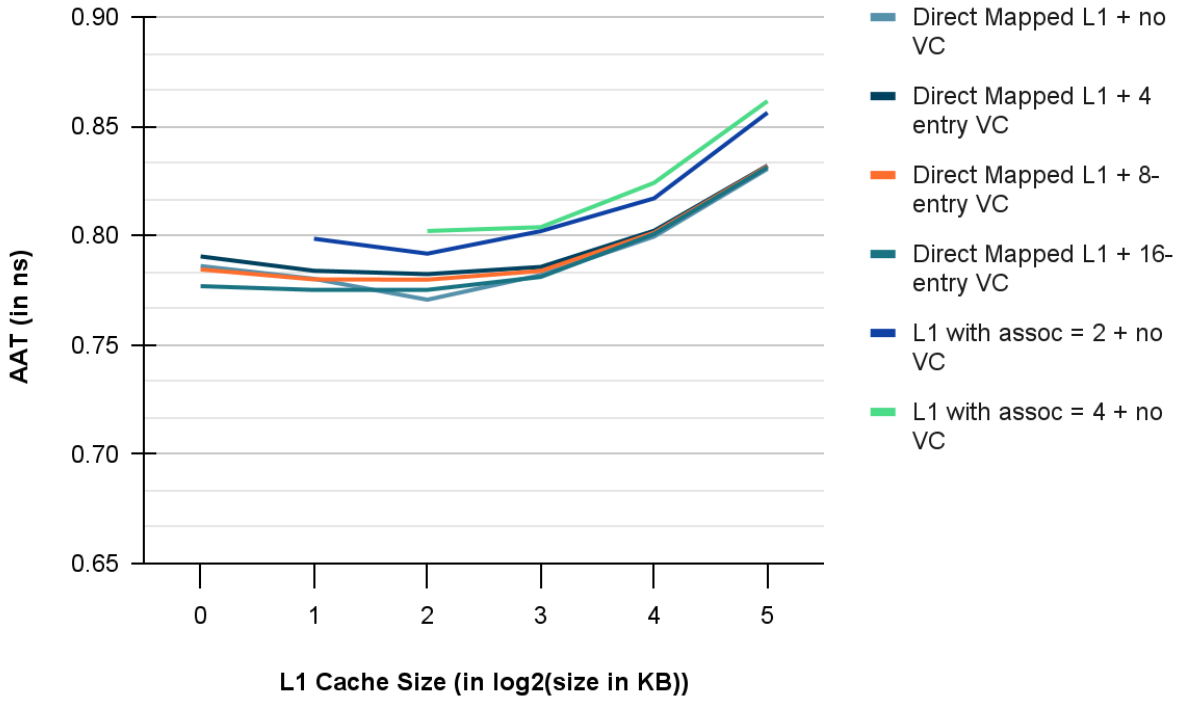
Whereas for cache of sizes of 16KB & 32KB, when block size is increased from 16B to 128B spatial locality became more beneficial and hence miss rate is decreasing.

We can say that larger cache may benefit more from spatial locality and smaller caches, cache pollution may dominate leading to higher miss rates after a certain block size.

→ Optimal blocksize for a 4-way set associative 8KB L1 cache would be 2^6 = 64B
(From the graph...deepest point for 8KB cache)

Victim Cache Investigation

Plot 7: (L1 Size Vs AAT)



→ Compared to the 2-way set-associative L1 cache, direct mapped L1 cache with or without Victim Cache has lesser AAT implying a better performance.

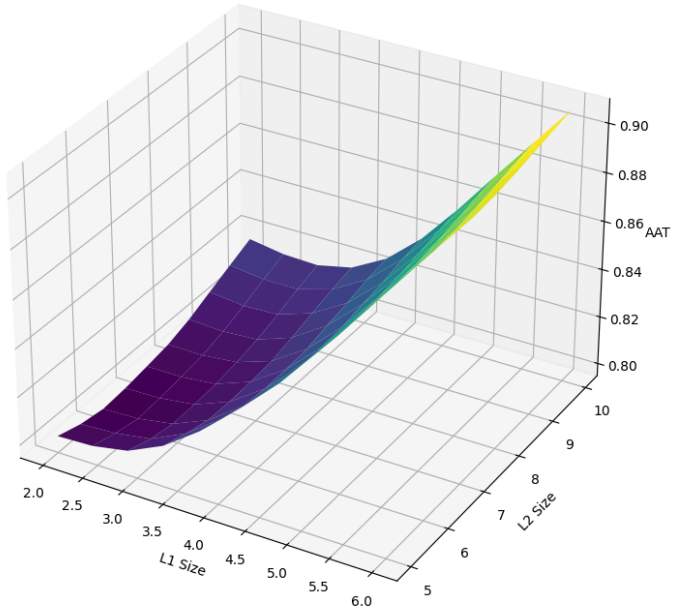
→For most of the cache sizes, direct mapped cache with 16-entry victim cache yields the best AAT. Else if cache sizes above 4KB, then direct mapped cache with no Victim cache yields best AAT.

→2-way set-associative L1 cache with no VC has a smaller total area but yields an AAT that is within 5% of best AAT as said above.

L1 + L2 Co-optimization

Here X, Y axis shows values of cache sizes in log2(size in KB) val.

Plot 5:



Plot 6: