Noah Perryman

Patrick Lyons

COE 1541 - Project 2

For experiment 1 the goal was to examine the effect of the block size by simulating a system with equal data and instruction cache sizes. Three different sizes for each of the caches (1KB, 16KB, and 128KB) were simulated with 4-way set associativity and four possible block sizes (4B, 16B, 64B, and 256B). For the first long trace file and second long trace file we found that the instruction cache the miss rates were relatively small (less than 4 %) for all of the possible combinations of block and cache sizes. To effectively make a zero percentage miss rate the cache size needs to be at least 16KB and the block size needs to be at least 16B. For the data cache the miss rate was found to decrease as the cache size increased for both long trace files. For each cache size though, it appeared as though there was an optimal block size where the miss rate would be the lowest. If the block size were increased or decreased this miss rate would increase. This is most likely because of replacement blocks kicking out values that need to be accessed, which is dependent in part on block size. It also appears that as the miss rate decreases the execution time in cycles also decreases.

For experiment 2 the goal was to try and design an optimum cache system based on the data retrieved for experiment 1. The lowest miss rates for both trace files occurred for a cache size of 128KB and a block size of 256B. However, this is costly as it takes up a lot of space and may lead to a lot of empty space and unused data in a retrieved block. To combat this a drop in percentage was taken in order to decrease the size of the blocks and caches. An instruction cache size of 16KB and a block size of 4B was chosen. A data cache size of 16KB and a block size of 256B was chosen. This was because the drop in percentage points for decreasing the block size was too big for the second large trace file for us to justify the change. The instruction cache miss rate for both large trace files was found to be approximately 0 (0.0018 %). The data cache miss rate for the first large trace file was found to be 17.006 % and the miss rate for the second large trace file was found to be 24.239 %. The execution time in cycles for the first large trace file was found to be 233,357,279 and the execution time in cycles for the second large trace file was found to be 333,327,836.

For experiment 3 the goal was to examine the effect of the associativity when considering a 32KB instruction cache and a 32KB data cache. The associativity size was either 1, 4, or 8. As expected it was found that the miss rate of both caches decreased as the associativity increased. This makes sense because as the associativity increases, the likeliness of a block being replaced decreases as there are possible block options when replacement occurs. Given this fact it is clear that the likeliness of a block being hit increases as on the subsequent accesses after it is placed in the cache it is more likely to still be in the cache. However, this decrease in miss rate may not always be helpful. The miss rates were shown to decrease only a small amount the higher the associativity became. While this does still increase the hit rate, this increase may be so small that it might not be justifiable or cost efficient to make the associativity larger.

In conclusion, increasing the cache size of either the instruction cache or data cache decreased the miss rates. However, increasing the block size varied for each cache size, as each cache size seemed to have an optimal block size (with a general trend of higher hit rates the higher the block size after the maximum miss rate). On top of that it was noted that as the associativity increases the miss rate decreases. However, as noted, this increase in associativity may not always be beneficial. Overall it is best to compare the functional capabilities of the cache with the size and cost of the cache. A larger cache with larger blocks and a higher associativity may have a higher performance but it may not be possible to achieve given the cost and/or feasibility.