Problem No. 1 (27 points)

(a) (4 points) Why does a “victim cache” always outperform a “miss cache”?

(b) (4 points) Why is the “miss cache” more effective in mitigating data conflict misses as compared to instruction conflict misses?

(c) (5 points) Consider a system with a two-level cache hierarchy. The L2 cache is 128KB, 8-way set-associative with a block size of 128 bytes.

To cut down the number of main memory (DRAM) accesses, the system designers are exploring the following two optimizations to the cache hierarchy:

**Optimization-1:** Add one way to each set in the L2 cache, resulting in a 144KB 9-way set-associative L2 cache.

**Optimization-2:** Add a 128-entry fully-associative victim cache.

Which one of the above two optimizations would be more effective in reducing the number of DRAM accesses? Why?

(d) (14 points) Consider a computer system with a two-level cache hierarchy, consisting of unified L1 and L2 caches. The cache statistics for a program are as follows:

- L1 cache miss rate is 50 misses per 1000 instructions,
- L2 cache hit ratio is 60%.

Assume that we add a prefetcher to the L2 cache. The prefetcher issues 50 prefetches per 1000 instructions, and reduces the L2 cache miss rate to 5 misses per 1000 instructions. Answer the following statistics:

i. What is the “coverage” of the prefetcher?

ii. What is the “accuracy” of the prefetcher?

iii. Does the prefetcher increase or decrease the total DRAM traffic? Compute the percentage increase/decrease in DRAM traffic, after the prefetcher is added.
**Problem No. 2 (15 points)**

(a) (3 points) What is the main difference between vertical waste and horizontal waste in superscalar processors?

(b) (4 points) Identify TWO reasons why adding simultaneous multithreading (SMT) would increase hardware complexity of the register file in a superscalar processor.

(c) (4 points) In a SMT processor, describe ONE advantage and ONE disadvantage of having separate branch predictors for each thread.

(d) (4 points) Identify ONE advantage and ONE disadvantage of SM:Single Issue design as compared to SM:Limited Connection design (both designs from the Tullsen et al. ISCA 1995 SMT paper).

**Problem No. 3 (8 points)**

(a) (4 points) Advanced branch prediction was used in the Pentium M processor as a power-saving feature. Briefly explain why.

(b) (4 points) In the Pentium M processor, assume that a new microarchitecture feature improves performance by 3% while increasing power consumption by 6%. Would that feature be included as a performance optimization? Justify your yes/no answer.