

8. You are involved in designing a computing system using a cache (256 kbyte, 4-way set associative cache using 1 kbyte blocks). Your first design has some cache performance problems. Your colleagues made the following suggestions. For each suggestion, first state whether or not the idea will work, and then *briefly* explain why. If the idea works explain under what conditions.

- (a) (2 points) Alain: "We have too many cache misses due to *conflicts*. We need to reduce the degree of associativity, so that we reduce conflict misses in the cache":

Solution:

This idea **will not work**. Just the opposite: increasing set associativity gives data more possibilities to be stored in the cache without replacing other data. This reduces conflict misses.

- (b) (2 points) Beatrice: "There are many *compulsory* cache misses. To combat this, we should increase our block size"

Solution: This idea **could work**. A larger block size will take advantage of spatial locality and assume that nearby data items will also be accessed by the program. If the program has such accesses, the first data access will result in a compulsory miss, but the subsequent accesses will find data in the cache.

- (c) (2 points) Cathy: "Our cache has many *capacity* misses. Instead of using a set associative cache, we should convert it to a direct mapped cache of the same size. This will allow more sets to be stored in the cache, hence reducing capacity misses"

Solution: This idea **will not work**. The organization of the cache does not change its capacity. The capacity miss occurs because data that is needed can not fit into the cache.