by byte blocks. What memory locations are now held?

A 32-bit CPU has this 2-byte DR cache with 4

Review Problem 40

it?
I/O or other processes invalidate the cache entry

Invalidation

Associativity

Cache size

Solutions

Intervening access may block the same location knock out

The block needed was in the cache, and there was enough room to hold it and all

Intervened

Solution is to increase cache size (but bigger is slower, more expensive)

Conflict

Conflict

For long-running programs this is a small fraction of misses

First access to a block - basically unavoidable (though bigger blocks help)

Compulsory/Colston

3 C's of cache design

Several different types of misses categorized based on problem/solution

Cache Miss Types
Fully Associative vs. Direct Mapped

A hit in cache before sensing the cache

send to CPU in parallel with hit miss check

On mapped can assume cache present in one location

slow - parallel lookup

Space for the components

DCL tags: Removing burn, no cache index

Significant, the choice one good

A hit or some structure would have

A miss at least, a miss is a miss is a miss
Fully Associative = $2^N$-way set associative (where $2^N$ is # of cache lines)

Direct Mapped = 1-way set associative

$N$ direct mapped caches working in parallel

$N$ lines are assigned to each cache index

$N$-way Set Associative
2-Way Set Associative Cache

Cache Index selects a "set", two tags compared in parallel.
Store the DM

N-way vs. Other Caches