Class Announcements

1. Programming Assignment 2 Due on Wednesday October 21st before 9AM

2. Collect your Midterm1 exams and Programming Assignment 1 feedback from me if you have not done so already.

Lecture Overview

1. Memory Hierarchy motivating example
2. Cache Organization
3. Direct Mapped Cache

Array copy

Changing the loop order while copying a 2 dimensional array alters the time taken by a factor of 37 on a CSL instructional machine
Memory Mountain

1. Read throughput decreases as the stride increases

2. Read throughput decreases as more data is copied.
Each main memory block maps to a specific block frame.

2 bits of the address define this mapping.

Address as used by the cache for a lookup.

Bits of SRAM cannot identify whether a block from memory has or has not been placed in a block frame.
So, keep 1 bit per frame to identify if data is valid or not.

Generic Cache Organization

Set 0:
Valid Tag 0 1 ... B-1
Valid Tag 0 1 ... B-1
Valid Tag 0 1 ... B-1

Set 1:
Valid Tag 0 1 ... B-1
Valid Tag 0 1 ... B-1
Valid Tag 0 1 ... B-1

Set S-1:
Valid Tag 0 1 ... B-1
Valid Tag 0 1 ... B-1
Valid Tag 0 1 ... B-1

Cache size: \( C = B \times E \times S \) data bytes
Looking up a memory address in Direct Mapped Cache

Three steps while determining whether a request is a hit or a miss:

- Set selection: Select the set where the address resides.
- Line matching: Select the cache line within the set.
- Word extraction: Extract the requested word from the right offset.
Looking up a memory address in Direct Mapped Cache

1. The valid bit must be set
2. The tag bits in the cache line must match the tag bits in the address
(3) If (1) and (2), then cache hit, and block offset selects starting byte.