Cache Performance

Lecture Objectives:

1) Explain the relationship between miss rate and block size in a cache.
2) Construct a flowchart explaining how a cache miss is handled.
3) Compare and contrast write through and write back caching schemes.
4) Define the term write buffer.
5) Perform cache related calculations to analyze the impact of having a cache on a computer system.

Partner Discussion

- If you are in an odd row, discuss part 1. Otherwise, discuss part 2 with the person sitting beside you.
  - An engineer is designing a computer system and is designing a system with either 64KB of cache or 256KB of cache. Which will perform better?
  - An engineer is designing a system with 16KB of cache, and is trying to decide whether to use a block size of 16 or 64. Which will perform better?

Most likely

More typically is better

It depends on the program running...

MSOE

Block size considerations

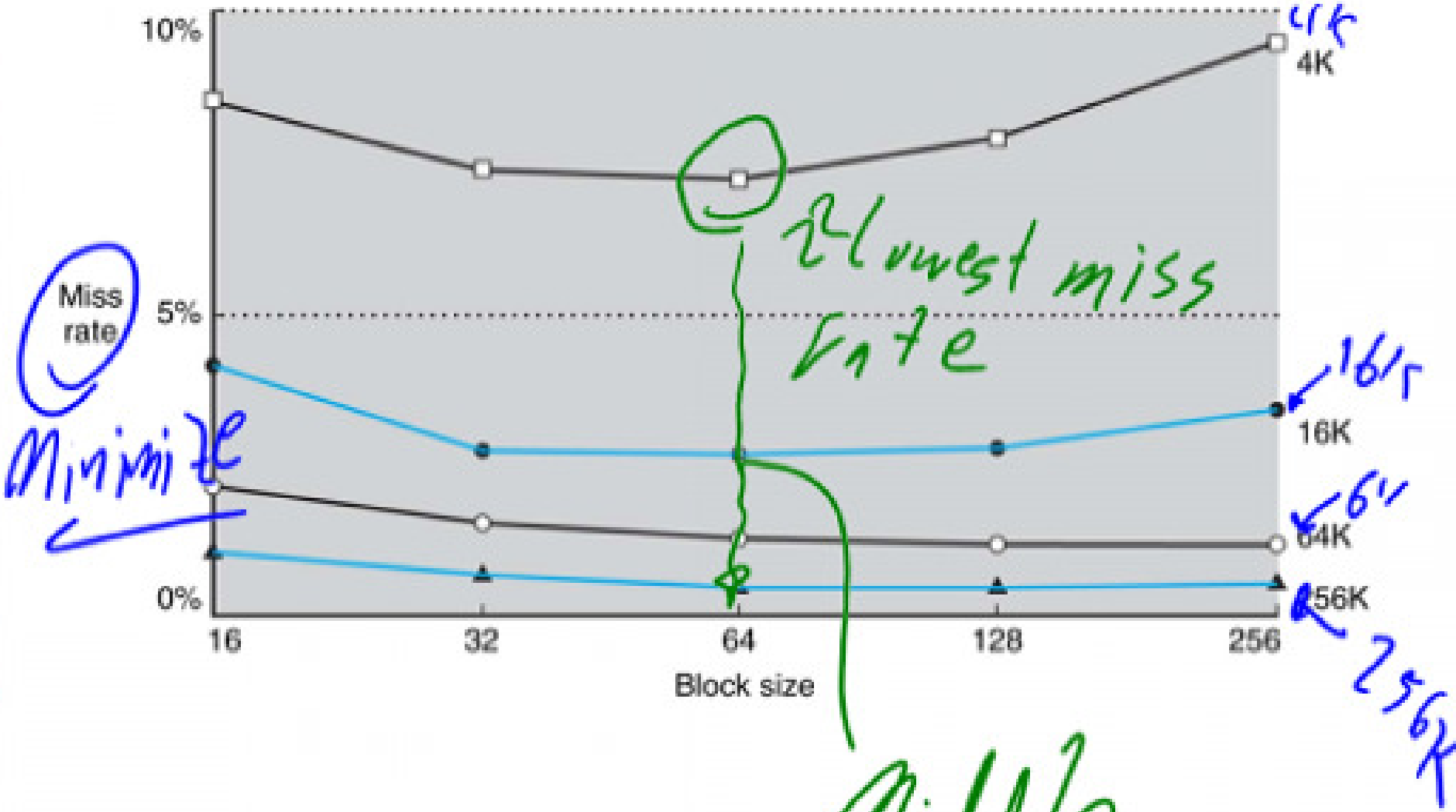
- Larger blocks? – Reduce miss rate due to spatial Locality.
- Fixed size cache

In a fixed size cache, fewer of them if Blocks are Large.
⇒ Increased miss rate.

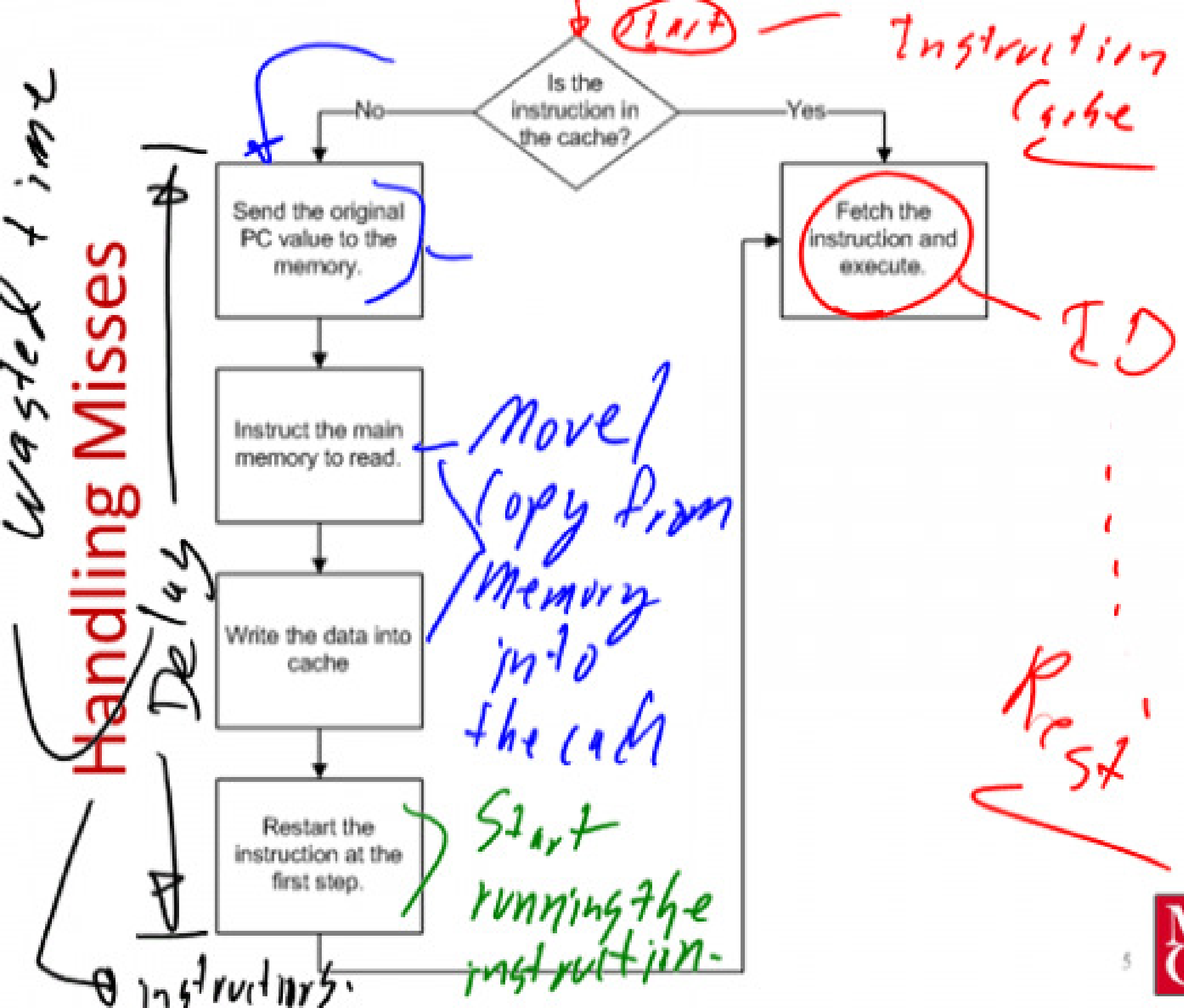
- Larger Miss penalty

can override benefit of reduced cache miss

Experimental Results



Handling Misses



Writing Data to memory

- Update the block in cache
  - Write through
    - Update cache plus block in memory
  - Write buffer

On data hit, we just update the value in cache.
⇒ Result is in cache Memory being different. Must store later.

⇒ Slower, increases CPI
⇒ But a faster way moving of cache values from

1,000,000 waiting time

500,000

CPU execution cycles: 1,000,000

100,000 memory accesses

5%, 000 cache misses

⇒ miss penalty is 10 clock cycles.

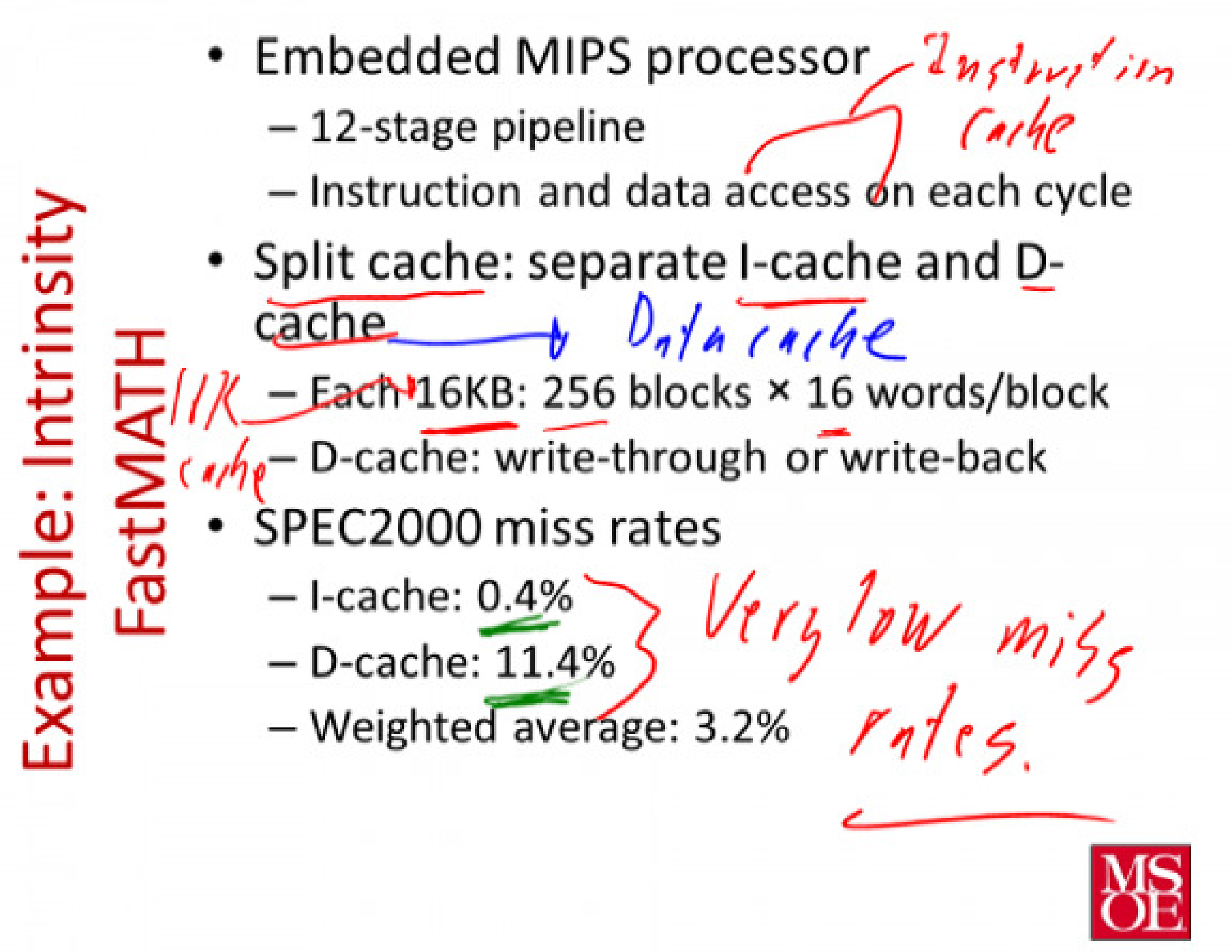
50,000 X 10

$$\text{CPU time} = \frac{1000\,000 + 500\,000}{1\,\text{M}}$$

$$= 1.5\,\text{s}$$

Example: Intrinsity FastMATH

- Embedded MIPS processor
  - 12-stage pipeline
  - Instruction and data access on each cycle
- Split cache: separate I-cache and D-cache
  - Each 16KB: 256 blocks × 16 words/block
  - D-cache: write-through or write-back
- SPEC2000 miss rates
  - I-cache: 0.4%
  - D-cache: 11.4%
  - Weighted average: 3.2%



Processor Performance

1MHz

Same as usual

$$\text{CPU Time} = (\text{CPU execution clock cycles} + \text{Memory-stall clock cycles}) \times \text{Clock cycle time}$$

$$\text{Memory-stall clock cycles} = \text{Read-stall cycles} + \text{Write-stall cycles}$$

10000 instructions

10% are memory accesses

50% have a cache miss

Instruction cache is perfect

~~100~~

Time it takes to read from RAM

MSOE

Cache Performance

- Working with your neighbor, solve the following problem:
  - Two different configurations for a cache system are being developed. In both cases, the miss penalty is 100 clock cycles, and the processor has an average CPI of 2 if no memory stalls occur.
  - The first cache system has a miss rate of 2% in the instruction cache and 4% in the data cache.
  - The second system has a miss rate of 4% in the instruction cache and 2% in the data cache.
  - 36% of instructions are loads and stores.
  - Determine the performance relationship between processor A and processor B.

Problem solution

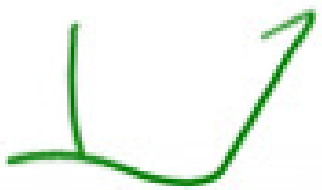
$$CPI_A = \text{Instruction miss cycles}_A + \text{Data miss cycles}_A + CPI_{Perfect}$$
$$= 2\% \times 100 + 4\% \times 36\% \times 100 + 2$$
$$= 2.00 + 1.44 + 2$$
$$= 5.44$$

$$CPI_B = \text{Instruction miss cycles}_B + \text{Data miss cycles}_B + CPI_{Perfect}$$
$$= 4\% \times 100 + 2\% \times 36\% \times 100 + 2$$
$$= 4.00I + 0.72I + 2$$
$$= 6.72$$

$$PerformanceDifference = \frac{6.72}{5.44} = 1.23$$

Average Memory Access Time

- Average Memory Access Time (AMAT)
  - The average of the time it takes to access memory considering both hits and misses.



Relative to the cache hit and miss rates

MSOE

AMAT Problem

- With your neighbor, solve the following:
  - A processor has a 1ns clock cycle time, a miss penalty of 50 clock cycles, a miss rate of 0.1 misses per instruction, and a cache access time of 1 clock cycle.
  - Determine the average memory access time per instruction.

50 clock cycles * 1ns / clock cycle

MSOE

Solution

$$AMAT = \text{Time for a hit} + \text{Miss Rate} \times \text{Miss Penalty}$$

$$= 1 + .1 \times 50$$

$$= 1 + 5$$

$$= 6$$