Topic V34

Average Memory Access Time

Reading: (Section 5.4)

Average Access Time
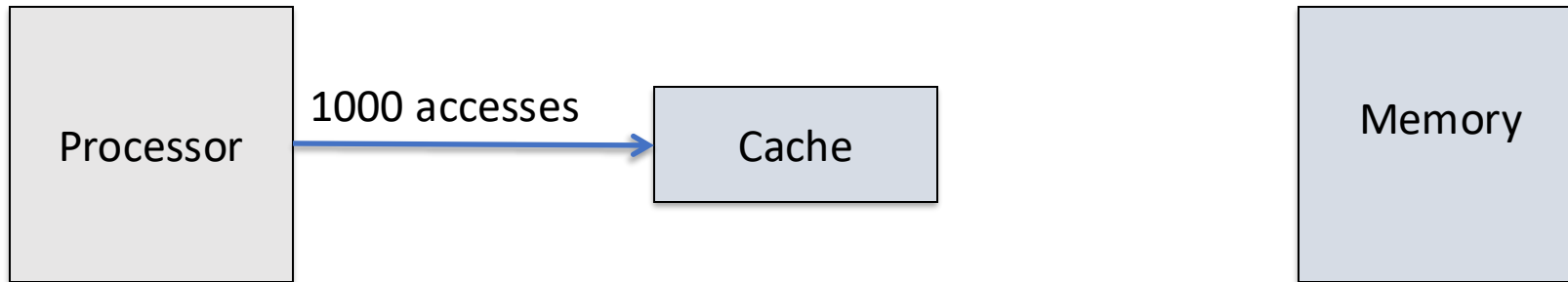
Hit time is also important for performance

Average memory access time (AMAT) = $\dfrac{\text{Total time to access memory}}{\text{Number of memory accesses}}$

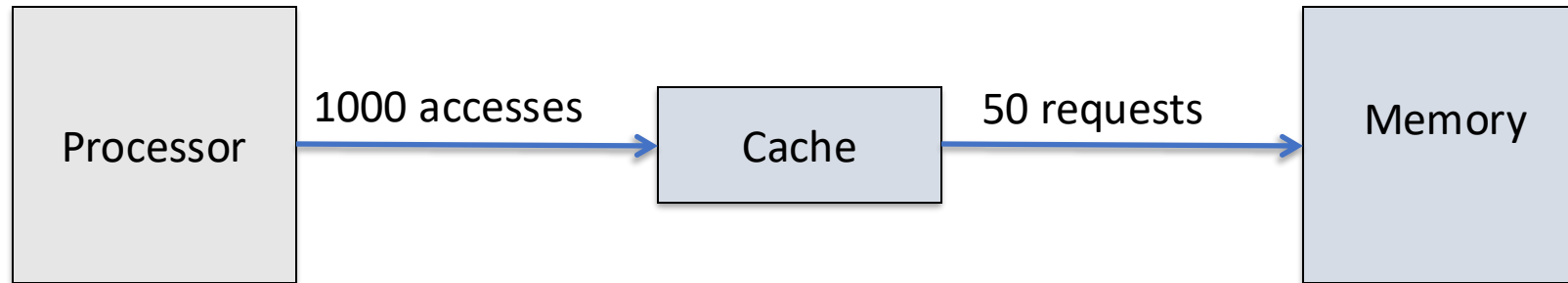Units: $\dfrac{ns}{\text{access}}$ $\dfrac{cycles}{\text{access}}$

Average Memory Access Time (example)

Find the AMAT for a processor with a 1 ns clock cycle time,
a miss penalty of 20 clock cycles,
a miss rate of 0.05 misses per access,
and a cache access time (including hit detection) of 1 clock cycle.
Assume that the read and write miss penalties are the same and ignore other write stalls.

Processor

1000 accesses

Cache

Memory

Average Memory Access Time (example)

Find the AMAT for a processor with a 1 ns clock cycle time,
a miss penalty of 20 clock cycles,
a miss rate of 0.05 misses per access,
and a cache access time (including hit detection) of 1 clock cycle.
Assume that the read and write miss penalties are the same and ignore
other write stalls.

Processor

1000 accesses

Cache
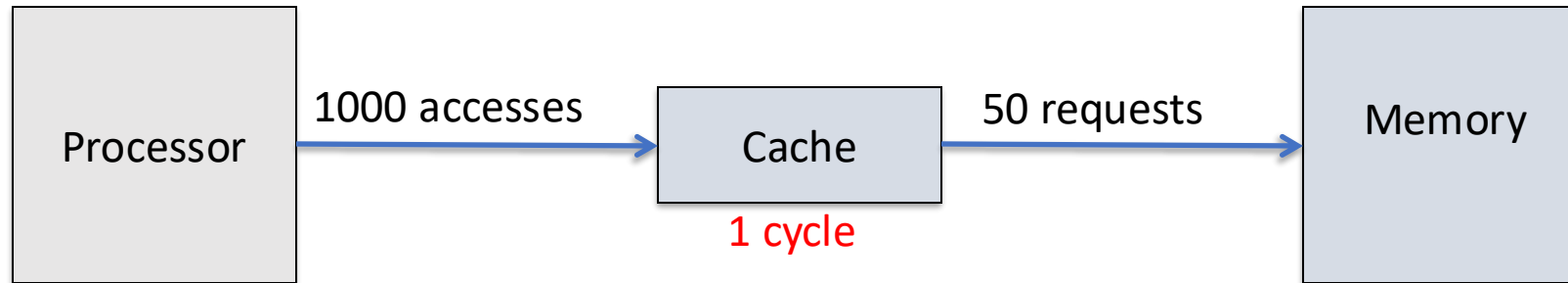
50 requests

Memory

Average Memory Access Time (example)

Find the AMAT for a processor with a 1 ns clock cycle time,
a miss penalty of 20 clock cycles,
a miss rate of 0.05 misses per access,
and a cache access time (including hit detection) of 1 clock cycle.
Assume that the read and write miss penalties are the same and ignore other write stalls.

| Processor | 1000 accesses → | Cache | 50 requests → | Memory |

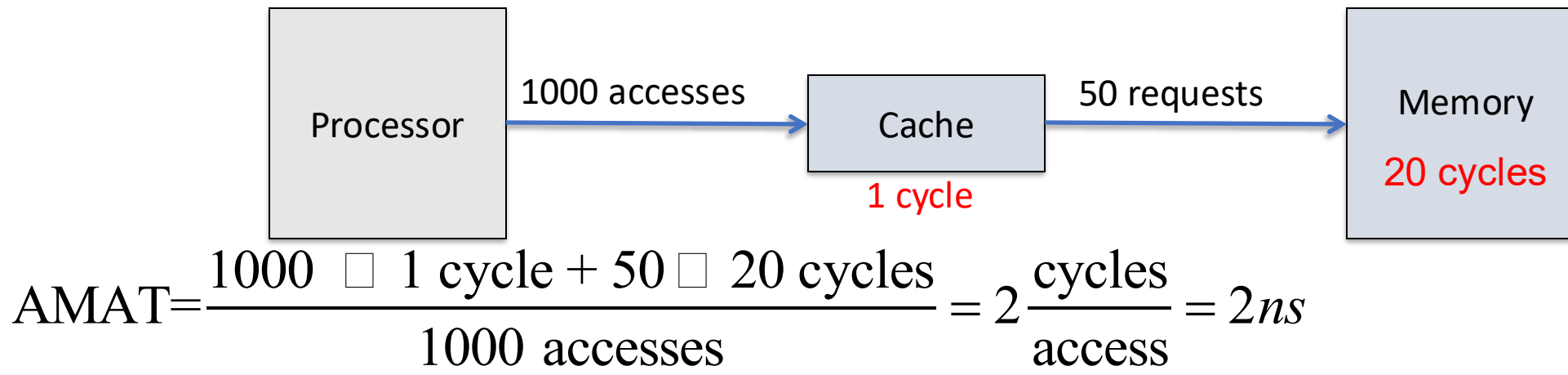1 cycle

Average Memory Access Time (example)

Find the AMAT for a processor with a 1 ns clock cycle time,
a <u>miss penalty of 20 clock cycles,</u>
a miss rate of 0.05 misses per access,
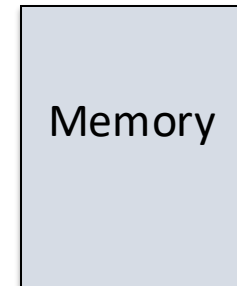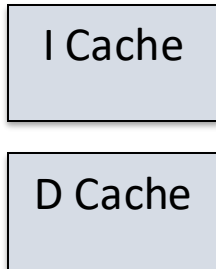and a cache access time (including hit detection) of 1 clock cycle.
Assume that the read and write miss penalties are the same and ignore
other write stalls.

Processor → 1000 accesses → Cache → 50 requests → Memory

Cache: 1 cycle

Memory: 20 cycles

$$AMAT = \frac{1000 \times 1\ cycle + 50 \times 20\ cycles}{1000\ accesses} = 2\frac{cycles}{access} = 2ns$$

Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store
running on a processor with a 1 ns clock cycle time,
a miss penalty of 100 clock cycles,
an I-cache miss rate of 0.03 misses per instruction,
and a D-cache miss rate of 0.1 misses per access.
The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

$$AMAT = \frac{\text{\# of clock cycle}}{\text{\# of accesses}}$$

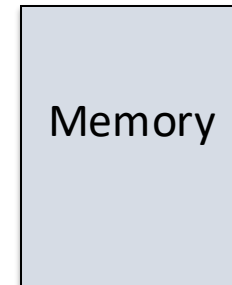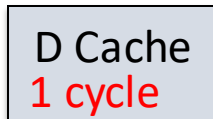| Processor		I Cache			Memory
		D Cache			

Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store
running on a processor with a 1 ns clock cycle time,
a miss penalty of 100 clock cycles,
an I-cache miss rate of 0.03 misses per instruction,
and a D-cache miss rate of 0.1 misses per access.
The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

$$AMAT = \frac{\text{\# of clock cycle}}{\text{\# of accesses}}$$

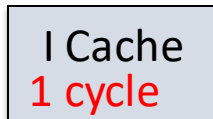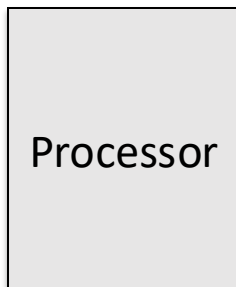Processor

I Cache
1 cycle

D Cache
1 cycle

Memory

Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store
running on a processor with a 1 ns clock cycle time,
a <u>miss penalty of 100 clock cycles,</u>
an I-cache miss rate of 0.03 misses per instruction,
and a D-cache miss rate of 0.1 misses per access.
The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

$$AMAT = \frac{\text{\# of clock cycle}}{\text{\# of accesses}}$$

Processor

1000 instruction fetches →

I Cache
1 cycle

D Cache
1 cycle

Memory
100 cycles

Average Memory Access Time (example)

Find the AMAT for a program where <u>20% of the instructions are load/store</u>
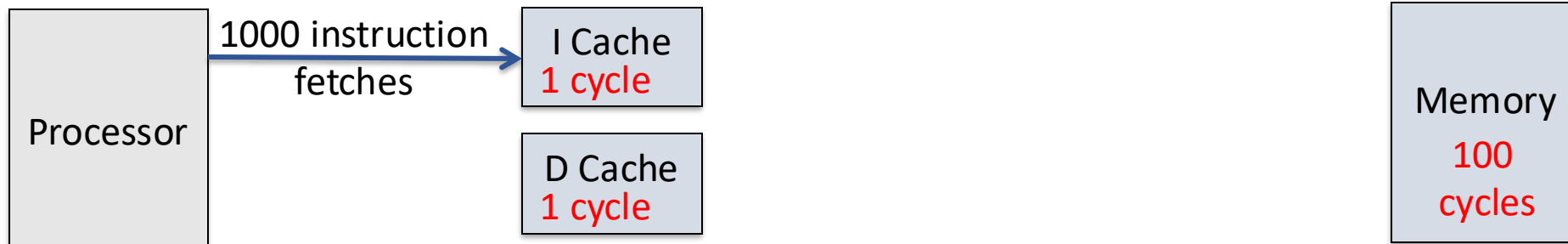running on a processor with a 1 ns clock cycle time,
a miss penalty of 100 clock cycles,
an I-cache miss rate of 0.03 misses per instruction,
and a D-cache miss rate of 0.1 misses per access.
The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

$$\text{AMAT} = \frac{\text{\# of clock cycle}}{\text{\# of accesses}} = \frac{1000 \times 1 \text{ cycle}}{1200 \text{ accesses}}$$

Parallel Accesses

Processor

1000 instruction fetches → I Cache 1 cycle

200 accesses → D Cache 1 cycle

Memory 100 cycles

Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store
running on a processor with a 1 ns clock cycle time,
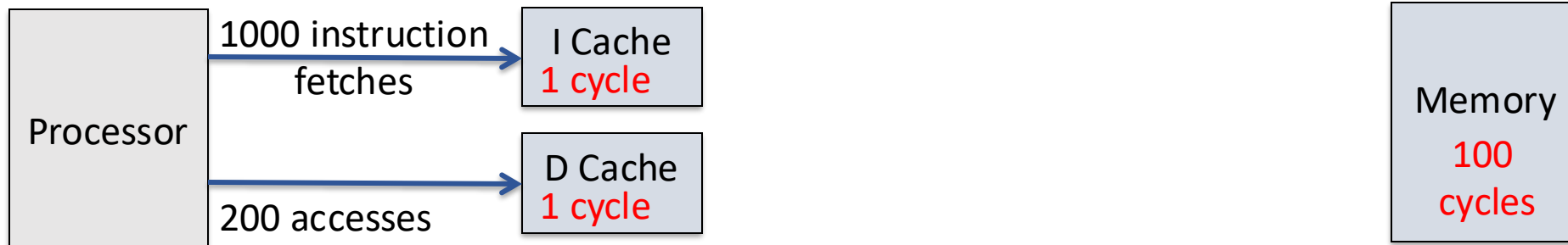a miss penalty of 100 clock cycles,
an I-cache miss rate of 0.03 misses per instruction,
and a D-cache miss rate of 0.1 misses per access.
The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

$$\text{AMAT} = \frac{\text{\# of clock cycle}}{\text{\# of accesses}} = \frac{1000 \times 1 \text{ cycle}}{1200 \text{ accesses}}$$
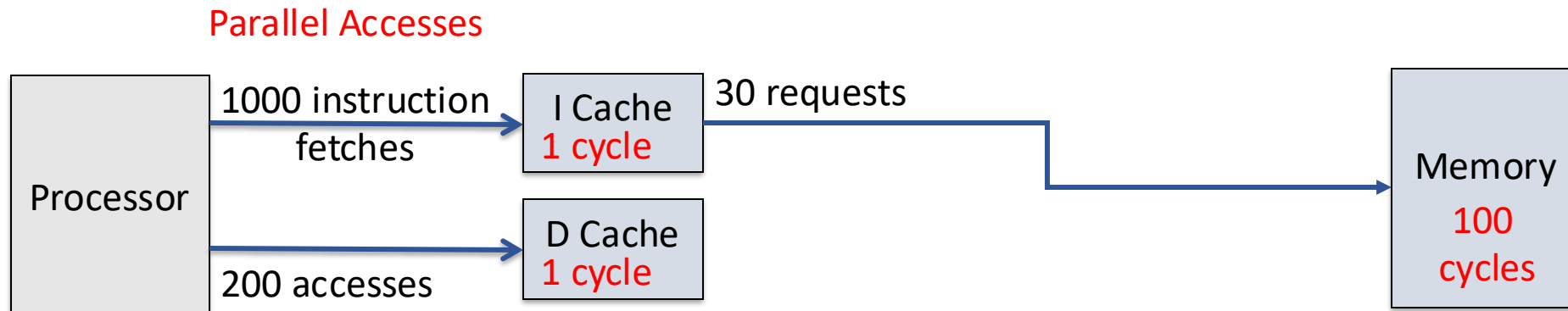
Parallel Accesses

Average Memory Access Time (example)

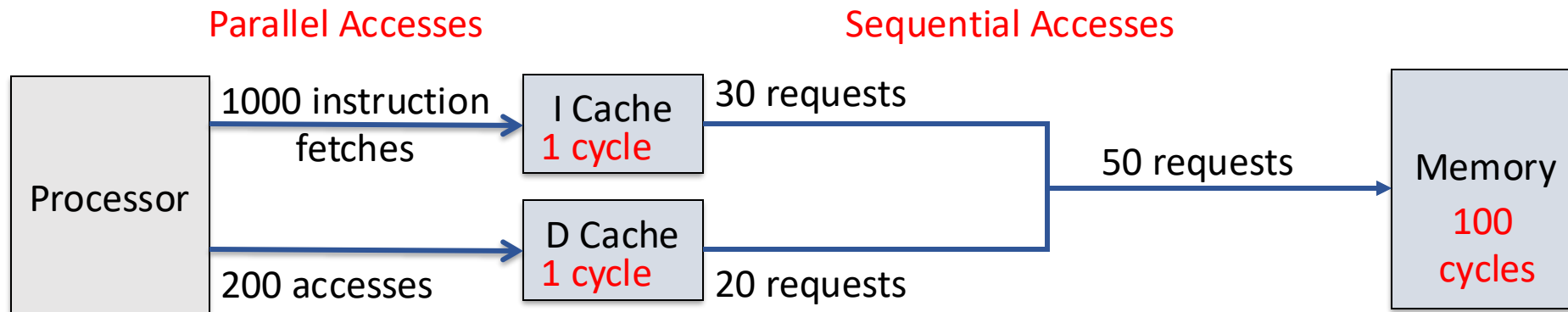Find the AMAT for a program where 20% of the instructions are load/store running on a processor with a 1 ns clock cycle time, a miss penalty of 100 clock cycles, an I-cache miss rate of 0.03 misses per instruction, and a D-cache miss rate of 0.1 misses per access. The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

$$\text{AMAT} = \frac{\text{\# of clock cycle}}{\text{\# of accesses}} = \frac{1000 \times 1 \text{ cycle} + 50 \times 100 \text{ cycles}}{1200 \text{ accesses}} = 5 \frac{\text{cycles}}{\text{access}}$$

Parallel Accesses   Sequential Accesses

Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store running on a processor with a 1 ns clock cycle time,
a miss penalty of 100 clock cycles,
an I-cache miss rate of 0.03 misses per instruction,
and a D-cache miss rate of 0.1 misses per access.
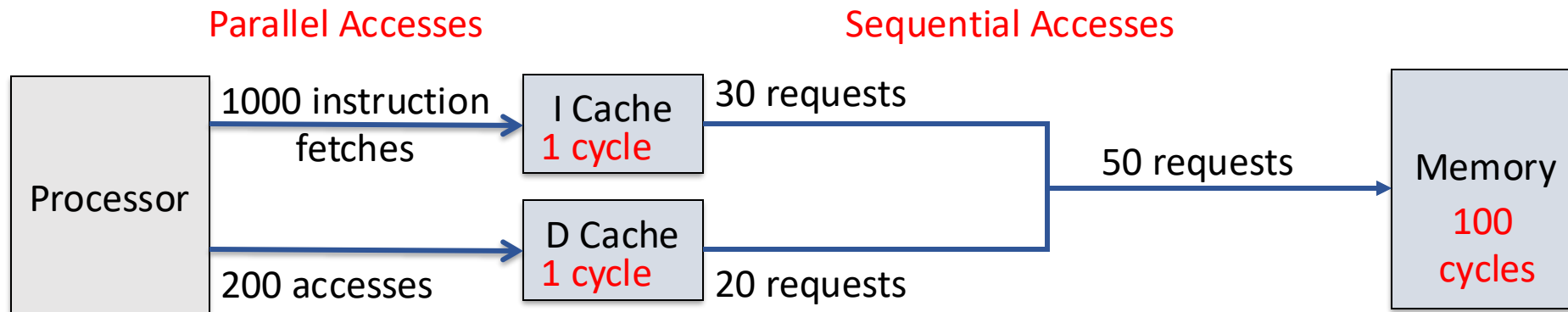The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

$$\text{AMAT} = \frac{\# \text{ of clock cycle}}{\# \text{ of accesses}} = \frac{1000 \times 1 \text{ cycle} + 50 \times 100 \text{ cycles}}{1200 \text{ accesses}} = 5 \frac{\text{cycles}}{\text{access}} = 5 \ ns$$

Parallel Accesses

Sequential Accesses
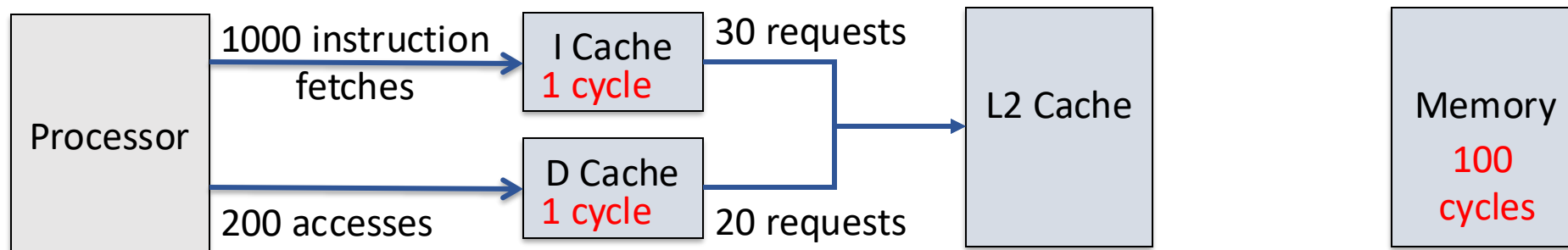
Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store running on a processor with a 1 ns clock cycle time, a miss penalty of 100 clock cycles, an I-cache miss rate of 0.03 misses per instruction, and a D-cache miss rate of 0.1 misses per access. The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

What is the new AMAT if an L2 cache is added to this memory hierarchy.

This L2 has an access time of 20 cycles and a local miss rate of 20%.

An access to the memory is only started after the access to the L2 (miss or hit) is completed.

$$AMAT = \frac{\# \text{ of clock cycle}}{\# \text{ of accesses}} = \frac{1000 \times 1 \text{ cycle}}{1200 \text{ accesses}}$$
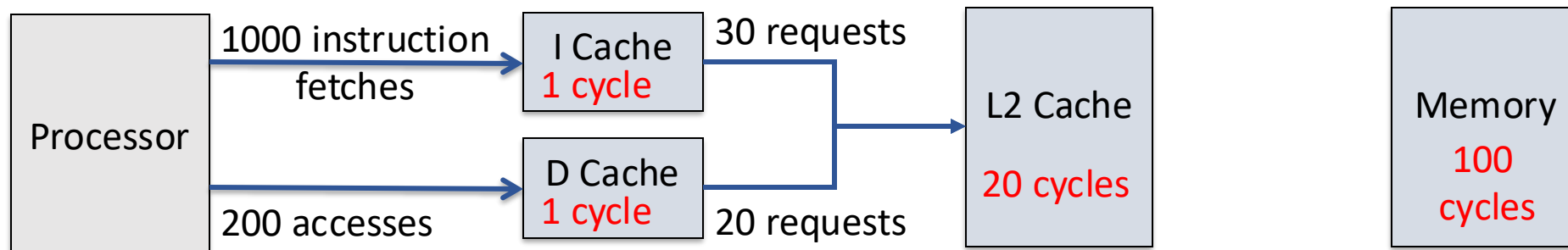
Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store running on a processor with a 1 ns clock cycle time, a miss penalty of 100 clock cycles, an I-cache miss rate of 0.03 misses per instruction, and a D-cache miss rate of 0.1 misses per access. The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

What is the new AMAT if an L2 cache is added to this memory hierarchy.

This L2 has an access time of 20 cycles and a local miss rate of 20%.

An access to the memory is only started after the access to the L2 (miss or hit) is completed.

$$\text{AMAT} = \frac{\text{\# of clock cycle}}{\text{\# of accesses}} = \frac{1000 \times 1\ \text{cycle}}{1200\ \text{accesses}}$$
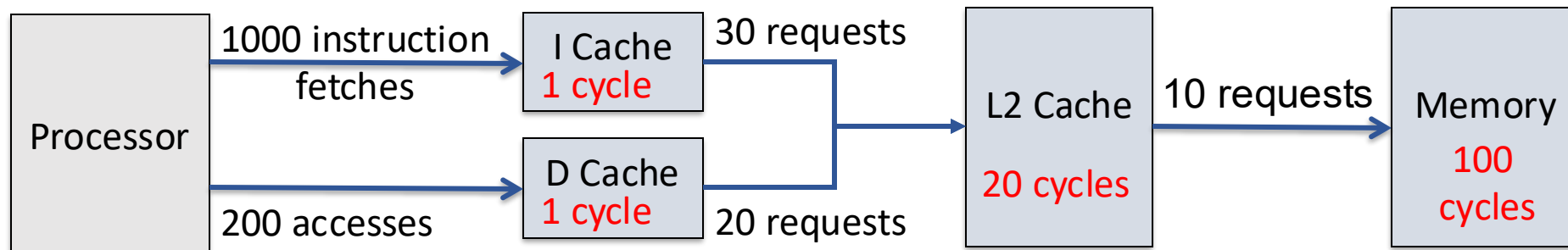
Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store running on a processor with a 1 ns clock cycle time, a miss penalty of 100 clock cycles, an I-cache miss rate of 0.03 misses per instruction, and a D-cache miss rate of 0.1 misses per access. The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.

What is the new AMAT if an L2 cache is added to this memory hierarchy.

This L2 has an access time of 20 cycles and a local miss rate of 20%.

An access to the memory is only started after the access to the L2 (miss or hit) is completed.

$$\text{AMAT} = \frac{\text{\# of clock cycle}}{\text{\# of accesses}} = \frac{1000 \times 1 \text{ cycle}}{1200 \text{ accesses}}$$

Average Memory Access Time (example)

Find the AMAT for a program where 20% of the instructions are load/store running on a processor with a 1 ns clock cycle time, a miss penalty of 100 clock cycles, an I-cache miss rate of 0.03 misses per instruction, and a D-cache miss rate of 0.1 misses per access. The access time (including hit detection) for both the I-cache and the D-cache is 1 clock cycle.
What is the new AMAT if an L2 cache is added to this memory hierarchy.
This L2 has an access time of 20 cycles and a local miss rate of 20%.
An access to the memory is only started after the access to the L2 (miss or hit) is completed.

$$\text{AMAT} = \frac{\text{\# of clock cycles}}{\text{\# of accesses}} = \frac{1000 \times 1 \text{ cycle} + 50 \times 20 \text{ cycles} + 10 \times 100 \text{ cycles}}{1200 \text{ accesses}}$$
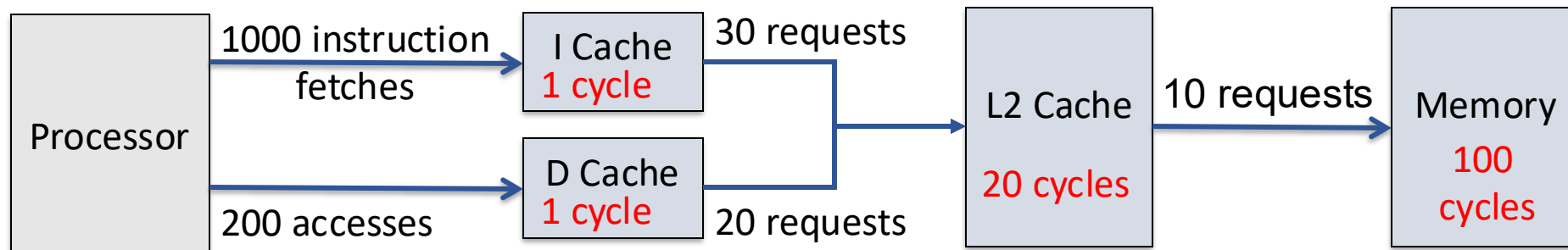
$$= 2.5 \ \frac{\text{cycles}}{\text{access}}$$

How much faster is the machine with L2 in comparison with the machine without L2?

$$AMAT_{no\ L2} = \frac{1000 \times 1\ cycle + 50 \times 100\ cycles}{1200\ accesses} = 5\ \frac{cycles}{access} = 5\ ns$$

$$AMAT_{with\ L2} = \frac{1000 \times 1\ cycle + 50 \times 20\ cycles + 10 \times 100\ cycles}{1200\ accesses} = 2.5\ \frac{cycles}{access}$$

Given that the basic CPI (with ideal caches) is the same, and the number of accesses is the same, we can simply divide the AMATs:

$$Speedup = \frac{AMAT_{no\ L2}}{AMAT_{with\ L2}} = \frac{5\ ns}{2.5\ ns} = 2$$
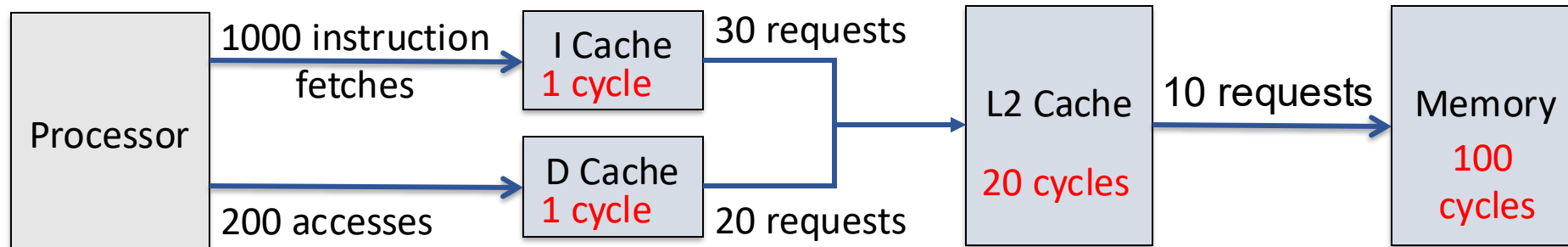
Design with L2 is two times faster

Local vs. Global Hit Rate

What is the *local hit rate* of the L2 cache?

$$\text{Local hit rate}_{L2} = \frac{\text{\# accesses that hit L2}}{\text{\# accesses that reach L2}} = \frac{40}{50} = 80\%$$

What is the *global hit rate* of the L2 cache?

$$\text{global hit rate}_{L2} = \frac{\text{\# accesses that hit L2}}{\text{Total \# of accesses}} = \frac{40}{1200} = 6.7\%$$

Performance Summary

When CPU performance increases

    Miss penalty becomes more significant

Decreasing base CPI

    Greater proportion of time spent on memory stalls

Increasing clock rate

    Memory stalls account for more CPU cycles