Assignment 4 Solutions
Pipelining and Hazards

Alice Liang

May 3, 2013

1 Processor Performance

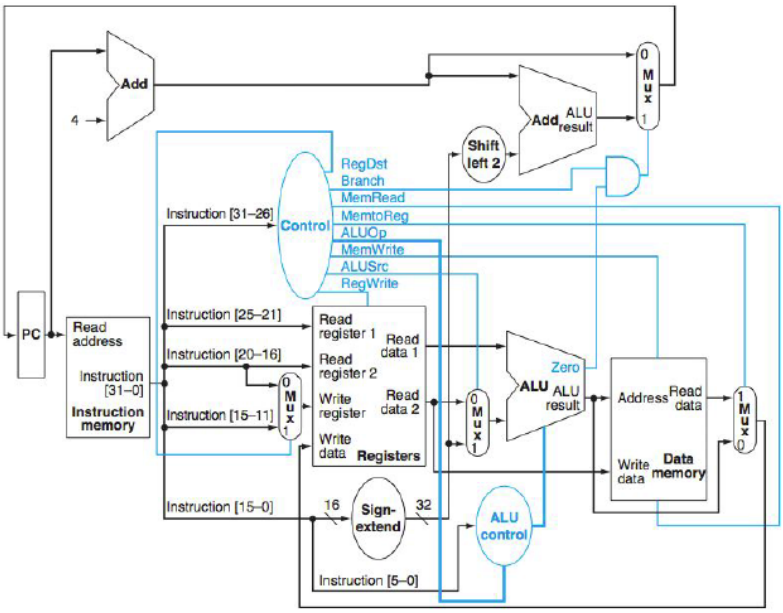
The critical path latencies for the 7 major blocks in a simple processor are given below.

	I-Mem	Add	Mux	ALU	Regs	D-Mem	Control
a.	400ps	100ps	30ps	140ps	200ps	350ps	100ps
b.	500ps	150ps	100ps	180ps	320ps	1000ps	65ps

For each part, answer the following questions:

1.1

What is the critical path for a MIPS ADD instruction? Explain your break-up.



`Using the above schematic, some paths of note are the following:`

```
1. PC increment = Add + Add + Mux
2. Control Unit = I-Mem + Control + ALUControl + ALU + Mux
3. Write Regs = I-Mem + Regs + Mux + ALU + Mux

After comparing latencies, 3 should be the longest for both as Regs takes longer than
Control.

a. 400 + 200 + 30 + 140 + 30 = 800ps
b. 500 + 320 + 100 + 180 + 100 = 1200ps
```

1.2

If the number of registers is doubled, this increases Regs by 100ps and Control by 20ps. This results in 10% fewer instructions due to fewer load/stores. What is the new critical path for a MIPS ADD instruction?

```
10\% fewer instructions doesn't affect the critical path at all. Regs still takes longer
than Control and so does not change the critical path from 1.1.

a. 800 + 100 = 900ps
b. 1200 + 100 = 1300ps
```

2 Pipelining

The 5 stages of the processor have the following latencies:

	Fetch	Decode	Execute	Memory	Writeback
a.	300ps	400ps	350ps	550ps	100ps
b.	200ps	150ps	100ps	190ps	140ps

Assume that when pipelining, each pipeline stage costs 20ps extra for the registers between pipeline stages.

2.1

Non-pipelined processor: what is the cycle time? What is the latency of an instruction? What is the throughput?

```
Because there is no pipelining, the cycle time must allow an instruction to go through all
stages in one cycle. The latency is the same as cycle time since it takes the instruction
one cycle to go from the beginning of fetch to the end of writeback. The throughput is
defined as 1/CT inst/s.

a. CT = 300 + 400 + 350 + 550 + 100 = 1700ps
   Latency = 1700ps
   Throughput = 1/1700 inst/ps

b. CT = 200 + 150 + 100 + 190 + 140 = 780ps
   Latency = 780ps
   Throughput = 1/780 inst/ps
```

2.2

Pipelined processor: What is the cycle time? What is the latency of an instruction? What is the throughput?

Pipelining reduces the cycle time to the length of the longest stage plus the register delay. Latency becomes CT*N where N is the number of stages as one instruction will need to go through each of the stages and each stage takes one cycle. The throughput formula remains the same.

a. CT = 550 + 20 = 570 ps
   Latency = 5 * 570 = 2850ps
   Throughput = 1/570 inst/ps

b. CT = 200 + 20 = 220 ps
   Latency = 5 * 220 = 1100ps
   Throughput = 1/220 inst/ps

2.3

If you could split one of the pipeline stages into 2 equal halves, which one would you choose? What is the new cycle time? What is the new latency? What is the new throughput?

We would want to choose the longest stage to split in half. The new cycle time becomes the originally 2nd longest stage length. Calculate latency and throughput correspondingly, but remember there are now 6 stages instead of 5.

a. CT = 400 + 20 = 420 ps
   Latency = 6 * 420 = 2520 ps
   Throughput = 1/420 inst/ps

b. CT = 190 + 20 = 210 ps
   Latency = 6 * 210 = 1260 ps
   Throughput = 1/210 inst/ps

2.4

Assume the distribution of instructions that run on the processor is:

50%: ALU
25%: BEQ
15%: LW
10%: SW

Assuming there are no stalls or hazards, what is the utilization of the data memory? What is the utilization of the register block's write port? (Utilization in percentage of clock cycles used)

LW and SW instructions use the data memory. As a result, the utilization of the data memory is 15% + 10% = 25%.

Similarly, ALU and LW instructions use the register block's write port. As a result, the utilization of the register block's write port is 50% + 15% = 65%.

3 Data Hazards

Skipped.

4 More Pipelines

You are given a non-pipelined processor design which has a cycle time of 10ns and average CPI of 1.4. Calculate the latency speedup in the following questions.

Note: The solutions given assume the base CPI = 1.4 throughput. Since the question is ambiguous, you could assume pipelining changes the CPI to 1. The method for computing the answers still apply.

4.1

What is the best speedup you can get by pipelining it into 5 stages?

5x speedup.

The new latency would be 10ns/5 = 2ns.

4.2

If the 5 stages are 1ns, 1.5ns, 4ns, 3ns, and 0.5ns, what is the best speedup you can get compared to the original processor?

The cycle time is limited by the slowest stage, so CT = 4ns.

Speedup = old CT / new CT = 10ns/4ns = 2.5x

4.3

If each pipeline stage added also adds 20ps due to register setup delay, what is the best speedup you can get compared to the original processor?

Adding the register delay, the new CT = 4.02ns.

Speedup = 10ns/4.02ns = 2.488x

4.4

The pipeline from Q4.3 stalls 20% of the time for 1 cycle and 5% of the time for 2 cycles (these occurences are disjoint). What is the new CPI? What is the speedup compared to the original processor?

New CPI = 0.2(2.4) + 0.05(3.4) + 0.75(1.4) = 1.7
Old performance = old CT * old CPI = 10 * 1.4 = 14
New performance = new CT * new CPI = 4.02 * 1.7 = 6.834
Speedup = 14/6.834 = 2.049x