

MM 11

- [Cc] = Clock Cycle
- [Sec] = Seconds
- [Ins] = Instructions
- [P] = Program
- [Mips] = Million instructions per second
- $1[\text{Ghz}] = 10^9[1/\text{sec}]$
- $1[\text{ns}] = 10^{-9}[\text{sec}]$

Question 1

Figure 1.16, in section 1.7 - The Power Wall, displays the frequency and power as from 1982 and until 2012 (8 generations of Intel processors).

Note that until 2004 there was a steady increase in frequency and power.

1. What is the reason for the decline in power starting at 2004?
2. Until 2004 both power and frequency had increased, however frequency growth rate was much faster. What was the reason for the difference?
3. What factors may influence processors frequency?

Answer

1. A key factor for enhancing processors performance was the increment of transistors per chip. The growing number of transistors allowed processors designers to introduce new complex and advanced architectures, which yield faster execution times.

With the growing number of transistors, the power consumption increased as well. The High power consumption exposed a basic flaw in the processor substance, which is the meltdown of silicon under high temperatures.

As long as processors designers were able to cope with the increasing power consumption, processors performance steadily improved. However, 2004 was the first year where the new designs failed to yield substantial performance improvements, due to the high thresholds of power consumption.

In essence, 2004 was the year where processor designers had hit the "Power Wall". Meaning, they failed to successfully design efficient processor with new, higher, thresholds of power consumption.

2. Technological improvements continuously lowered the voltage thresholds in processors architects. Since low voltage requires less power, the general raise in power was restrained.
3.
 - High frequency rises power consumption, which then yields more heat. Therefore higher frequency will mandate more cooling. Thus, frequency is bound by cooling measures capacity.
 - Advancements in optimization techniques may introduce more complex architectures, which requires longer cycles. Since frequency is directly correlated to cycle time, it may be affected by this.

Question 2

Let P1, P2 be two processors with the same instruction set.

Their instruction set is partitioned into 4 different groups, A, B, C and D.
We know:

	Clock Rate	A CPI	B CPI	C CPI	D CPI
P1	1.2GHz	1	2	2	3
P2	1.5GHz	2	4	2	3

1. In the sense of instructions per second, what is the fastest series of instruction for P1 and P2 possible?
2. Let T be some arbitrary program, comprised of instruction from A,B,C,D.
Say that the groups of instructions in T from B, C and D are equally represented, and the group of instructions in T from A is double the size of the others.
Which processor will run T faster? What is the MIPS rate?
3. Let V be some arbitrary program, comprised of instruction from A,B,C,D.
Say that the groups of instructions in V from A, B and C are equally represented, and the group of instructions in T from D is double the size of the others.
Which processor will run V faster? What is the MIPS rate?

Answer

1. MIPS (million instruction per second) is given by:

$$MIPS = \frac{Clock\ Rate}{CPI \times 10^6} \quad (1)$$

Group A has the lowest CPI for P1, therefore executing instruction of group A will yield the highest MIPS for P1, which is: $\frac{1.2 \times 10^9}{1 \times 10^6} = 1200$ MIPS.

Similarly, taking the lowest CPI value for P2 (e.g C CPI), we have:
 $\frac{1.5 \times 10^9}{2 \times 10^6} = 750$ MIPS.

Therefore, P1 is faster than P2 in MIPS metric.

2. Denote T the set of instructions from T, and:

$$A_T = \{x \in T \mid x \in A\}.$$

$$B_T = \{x \in T \mid x \in B\}.$$

$$C_T = \{x \in T \mid x \in C\}.$$

$$D_T = \{x \in T \mid x \in D\}.$$

Execution Time = Instruction Count x CPI / Clock Rate.

$$P1 \text{ Execution Time} = (|A_T| \times 1 + |B_T| \times 2 + |C_T| \times 2 + |D_T| \times 3) / 1.2 \text{ GHz}$$

$$= \frac{(9 \times |D_T|)}{1.2 \text{GHz}} = 7.5 \times 10^{-9} \times |D_T|.$$

$$\begin{aligned} \text{P2 Execution Time} &= (|A_T| \times 2 + |B_T| \times 4 + |C_T| \times 2 + |D_T| \times 3) / 1.5 \text{ GHz} \\ &= \frac{(13 \times |D_T|)}{1.5 \text{GHz}} = 8\frac{2}{3} \times 10^{-9} \times |D_T|. \end{aligned}$$

Therefore P1 runs T faster than P2.

$$\text{Acceleration rate} = \frac{P2 \text{ Execution Time}}{P1 \text{ Execution Time}} = \frac{8\frac{2}{3}}{7.5} \cong 1.1555.$$

3. Denote V the set of instructions from V , and:

$$A_V = \{x \in V \mid x \in A\}.$$

$$B_V = \{x \in V \mid x \in B\}.$$

$$C_V = \{x \in V \mid x \in C\}.$$

$$D_V = \{x \in V \mid x \in D\}.$$

$$\begin{aligned} \text{P1 Execution Time} &= (|A_V| \times 1 + |B_V| \times 2 + |C_V| \times 2 + |D_V| \times 3) / 1.2 \text{ GHz} \\ &= \frac{(11 \times |A_V|)}{1.2 \text{GHz}} = 9\frac{1}{6} \times 10^{-9} \times |A_V| \end{aligned}$$

$$\begin{aligned} \text{P2 Execution Time} &= (|A_V| \times 2 + |B_V| \times 4 + |C_V| \times 2 + |D_V| \times 3) / 1.5 \text{ GHz} \\ &= \frac{(14 \times |A_V|)}{1.5 \text{GHz}} = 9\frac{1}{3} \times 10^{-9} \times |A_V| \end{aligned}$$

Therefore P1 runs V faster than P2.

$$\text{Acceleration rate} = \frac{P2 \text{ Execution Time}}{P1 \text{ Execution Time}} = \frac{9\frac{2}{3}}{9\frac{1}{3}} \cong 1.035.$$

Question 3

P1 and P2 are two processors running some program T. The instructions are partitioned to groups A,B,C,D.

P1 clock rate is 3.5GHz.

P2 clock rate is 4.0GHz.

For the program T, and A,B,C,D instructions, we have:

P1	A	B	C	D
CPI	2	3	3	5
Frequency	45%	25%	25%	5%

P2	A	B	C	D
CPI	2	2	4	5
Frequency	45%	25%	25%	5%

1. What is the CPI of each processor?
2. What is the MIPS of each processor?
3. How much faster is P2 compared to P1?
4. What is the CPI of P1new?
5. How much faster is P1new compared to P1?
6. If we would have managed to combine the improved hardware and new compiler, how much faster will be the new processor compared to P1?

Answer

1. The average CPI is given by the taking the proportional CPI of each instruction group.

$$P1 \text{ CPI} = 2 \times 0.45 + 3 \times 0.25 + 3 \times 0.25 + 5 \times 0.05 = 2.65[\text{Cc/Ins}].$$

$$P2 \text{ CPI} = 2 \times 0.45 + 2 \times 0.25 + 4 \times 0.25 + 5 \times 0.05 = 2.65[\text{Cc/Ins}].$$

2. Instructions per second is given by: IPC x Clock Rate.

$$P1_MIPS = \frac{\text{Clock Rate}}{CPI \times 10^6} = \frac{3.5GHz}{2.65 \times 10^9} \cong 1.3207MIPS.$$

$$P2_MIPS = \frac{\text{Clock Rate}}{CPI \times 10^6} = \frac{4.0GHz}{2.65 \times 10^9} \cong 1.5094MIPS.$$

3. Since P1 and P2 have the same CPI values, we can deduce the result by considering the MIPS values alone. Then, P2 is $\frac{4}{3.5} \cong 1.1428$ times faster than P1.

4. We have changed the proportions of instructions executed from A,B,C,D.

In total, we have decreased the amount of instructions to 84.75%. And the new proportional amounts are:

Inst.	Calc	New Group Freq
A	$0.9 \times 45 / 84.75$	0.4778
B	$0.8 \times 25 / 84.75$	0.2359
C	$0.85 \times 25 / 84.75$	0.2507
D	$0.6 \times 5 / 84.75$	0.0353

Then, P1New CPI is

$$0.4778 \times 2 + 0.2359 \times 3 + 0.2507 \times 3 + 0.0353 \times 5 = 2.5919[\text{Cc/Ins}].$$

5. In general, Execution Time = Instruction Count x CPI / Clock Rate.

If $L = \text{Instructions Executed by } P1$, then:

$$P1 \text{ Time} = |L| \times 2.65 \times \frac{1}{3.5\text{GHz}}.$$

$$P1\text{New Time} = |L| \cdot \frac{84.75}{100} \times 2.5919 \times \frac{1}{3.5\text{GHz}}.$$

$$\text{Therefore } P1\text{New is } \frac{2.65}{0.8475 \times 2.5915} = 1.206 \text{ times faster than } P1.$$

6. The average CPI of P2New may be different. We find it first.

$$0.4778 \times 2 + 0.2359 \times 2 + 0.2507 \times 4 + 0.0353 \times 5 = 2.6067[\text{Cc/Ins}].$$

Then the new execution time is:

$$|L| \cdot \frac{84.75}{100} \times 2.6067 \times \frac{1}{3.5\text{GHz}}.$$

$$\text{And } P2\text{New is } \frac{2.65}{0.8475 \times 2.6067} = 1.199 \text{ times faster than } P1.$$

Question 4

1. Explain sign extension.
2. Represent 8574 in 2,4,8,16 bases. Then perform negation in two's complement in 15 bits.
3. What operations are required to perform division by 32 of a 16 bits number in two's complement form?

Answer

1. For a given binary number, a sign extension is the operation of extending the number of bits, while preserving the sign and value.
For example, the 4 bits binary numbers are limited to 16 values.
If 0110 is a 4 bits binary number, a sign extension to 8 bits will add 4 new bits, while preserving its sign and value, with respect to the binary form (e.g two's complement).

2. To find the binary form of a positive number n .

- Take mod 2
- Divide by 2, and round down.
- repeat.

$$8574 \equiv 0 \pmod{2}$$

$$4287 \equiv 1 \pmod{2}$$

$$2143 \equiv 1 \pmod{2}$$

...

$$2 \equiv 0 \pmod{2}$$

$$1 \equiv 1 \pmod{2}$$

$$8574 = \text{b}10000101111110.$$

$$8574 = \text{0x}217\text{E}.$$

$$8574 = (\text{base } 4)\text{b}2011332.$$

$$8574 = (\text{base } 8)\text{b}20676.$$

With 15 bits, $8745 = \text{b}010000101111110$.

The negation is $-8745 = \text{b}101111010000001 + \text{b}000000000000001 = \text{b}101111010000010$.

3. For x a 16 bits binary number in two's-complement form.

Here is a pseudo code for dividing x by 16.

```

divide_by_16(x) {

    if (x ^ b1000000000000000)
        if (x equals b111111111111111)
            return 0

    shift x bits to the left 4 places. pad with ones.
    x |= b1000000000000000

    else
        shift x bits to the right 4 places. pad with zeros.

    return x
}

```

4. For two's complement, both for subtraction, addition and multiplication the arithmetic operations don't differ between signed and unsigned integers.

This unique characteristic simplifies systems design and implementation, and thus a key factor choosing two's complement form above others.