#12 : Performance

Computer Architecture 2019/2020

João Soares & Ricardo Rocha

Computer Science Department, Faculty of Sciences, University of Porto

CPU Clocking

The duration of a complete clock cycle is the **clock period** and the number of cycles per second is the **clock rate (or clock frequency)**, which is the inverse of the clock period.

$$ClockRate\,(Hz) = \frac{1}{ClockPeriod\,(s)}$$

Clock period (duration of a clock cycle)

- e.g., 250ps (picoseconds) = 0.25ns (nanoseconds) = 250×10^{-12}s (seconds)

Clock rate (cycles per second)

- e.g., 4.0GHz = 4000MHz = 4.0×10^{9}Hz = $1 / (250 \times 10^{-12}\text{s})$

Performance Equation

$$CPUTime = InstructionCount \times CPI \times ClockPeriod$$

$$= \frac{InstructionCount \times CPI}{ClockRate}$$

Components of performance	Units of measure
CPU execution time for a program	Seconds for the program
Instruction count	Instructions executed for the program
Clock cycles per instruction (CPI)	Average number of clock cycles per instruction
Clock cycle time	Seconds per clock cycle

Cache Performance

CPU time can be divided into the clock cycles that the CPU spends executing the instructions with no misses (**CPIPerfect**) and the clock cycles that the CPU spends waiting for the memory system (**CPIStall**).

$$CPI = CPIPerfect + CPIStall$$

Memory-stall clock cycles can be defined as the sum of the stall cycles coming from reads plus those coming from writes. For simplicity, let's assume that the read/write miss rates and miss penalties are the same:

$$CPIStall = \frac{MemoryAccesses}{Instructions} \times MissRate \times MissPenalty$$

Cache Performance

If we consider separate caches/memories for instructions and data then:

$$CPIStall = CPIStallInstructionAccess + CPIStallDataAccess$$

In more detail:

$$CPIStall = 1 \times MissRateInstructionAccess \times MissPenalty + \frac{LoadStores}{Instructions} \times MissRateDataAccess \times MissPenalty$$

Multilevel Caches

With multilevel caches, memory-stall clock cycles can be defined as the sum of the stall cycles coming from the several cache levels (L1, L2, …). For simplicity, let's assume single instruction/data caches:

$$\begin{aligned} CPIStall &= MissRate \times MissPenalty \\ &= MissRateL1 \times MissPenaltyL1 \\ &\quad + GlobalMissRateL2 \times MissPenaltyL2 + \ldots \end{aligned}$$

The global miss rate for a level L represents the miss rate for the set of levels up to L:

$$GlobalMissRateL2 = MissRateL1 \times MissRateL2$$

$$GlobalMissRateL3 = MissRateL1 \times MissRateL2 \times MissRateL3$$