# Intéraction Logiciel Architecture

## Les performances et leur mesure

Henri-Pierre Charles & Frédéric Rousseau

# Motivation

## Argumentation

- 

## Illustration

1. Motivation
   - Motivation
   - Partie1 Intro
   - Metrics How to Measure Performance ?
   - Amdahl2
   - Vocabulaire Amdahl
   - Vocabulaire : Cycle par seconds / Flops
   - Vocabulaire Peak / Sustained
   - Vocabulaire Lois
   - Vocabulaire Speedup

# Partie1 Intro

## Argumentation

- 

## Illustration

# Metrics How to Measure Performance ?

## What scale ?

- Application level (TPS, Frame/s, .../)
- Run to completion
- Method level
- Instruction level

## Tools

- Wall clock
- `gettimeofday`
- Performance counters

## Full application or function call ?

- Cold start / warmup
- Statistic / multiple calls
- How many calls

# Vocabulaire : Cycle par seconds / Flops

## Units

| | |
|---|---|
| Mips | Million operation per second |
| Flops | Floating point operation per second<br>`http://www.top500.org` |
| Flops/Watt | Floating point operation per watt per second<br>`http://www.green500.org` |
| IPC | : Instructions per Cycle |

## How to mesure

- Analytique (wall clock)
- Instrumentation
  - "Portable" (`gettimeofday()`)
  - Hardware (hardware performance counter)

# Vocabulaire Peak / Sustained

## Notions

**Peak** performance : maximal theoretical performance, assuming no bubble

**Sustain** performance : real acheived performance, on a real benchmark

## Cost

- What is the percentage you're ready to lose ? 90% 95% ?
- How many are you ready to pay (time, money) to minimise this loss ?

`http://www.top500.org`

# Vocabulaire Lois

- Moore `http://en.wikipedia.org/wiki/Moore's_law`
- Amdahl `http://en.wikipedia.org/wiki/Amdahl's_law`
- Memory bound `http://en.wikipedia.org/wiki/IO_bound`
- CPU bound `http://en.wikipedia.org/wiki/CPU_bound`

# Vocabulaire Speedup

## Notion

**Speedup** $S = 100 * \frac{T_{seq}}{T_{opt}}$
Can be between

- 1 and N processor
- 1 and vectorized
- non optimized versus optimized version

**Mesure Quality** what are the execution conditions : data set, computer workload, reproducibly, ...

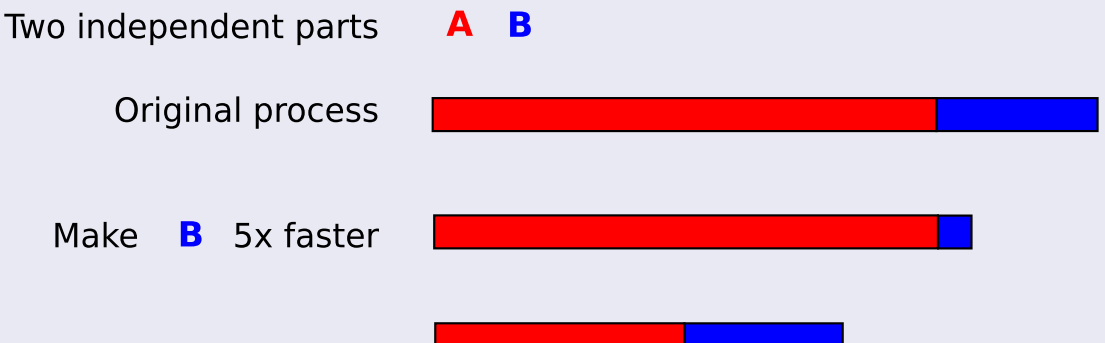
Computer science has to use “human science” tools & methodology

# Amdahl2

## Argumentation

Assume that a task has two independent parts, A and B. B takes roughly 25% of the time of the whole computation. By working very hard, one may be able to make this part 5 times faster, but this only reduces the time for the whole computation by a little. In contrast, one may need to perform less work to make part A be twice as fast. This will make the computation much faster than by optimizing part B, even though B's speed-up is greater by ratio, (5x versus 2x)

## Illustration



# Vocabulaire Amdahl

## Argumentation

The speedup of a program using multiple processors in parallel computing is limited by the sequential fraction of the program. For example, if 95% of the program can be parallelized, the theoretical maximum speedup using parallel computing would be 20x as shown in the diagram, no matter how many processors are used.

## Illustration