

Architecture des ordinateurs haute performance

Soraya Zertal

Laboratoire Li-PaRAD, Université de Versailles

E-mail: soraya.zertal@uvsq.fr

Plan :

- Notions de performance
- Jeux d'instructions
- Pipeline
- Superscalaire et Vectoriel
- Hiérarchie mémoire
- Stockage de données

- Architecture des ordinateurs, John L. Hennessy et David A. Patterson, Thomson Editions.
- Architecture des ordinateurs", A. Tanenebaum. Editions Dunod.

Organisation du module

- Cours par Mme S. Zertal
- TDs par S. Zertal (gr1) et S. Gougeaud (gr2)
- Mode d'évaluation :
 - 2 Contrôles : épreuves écrites (en cours ou en TD) sur le semestre, **Annoncées ou PAS**
 - ET 1 examen : fin du semestre (voir date fixée par la scolarité)

Composants principaux d'un ordinateur

- Processeur ou unité de calcul
 - Hiérarchie mémoire (cache + mémoire principale)
 - Disques (stockage de masse)
 - Autres moyens (cartes) : réseau, graphique, son, lecteurs, ports, ...
- Chaque composant doit être performant.
- L'interaction des différents composants fonctionnant ensemble doit se faire au mieux pour assurer l'exploitation du potentiel de chacun.

Architecture à haute PERFORMANCE

La PERFORMANCE est liée à 3 facteurs :

- ① **Technologique** (l'électronique): augmentation de la vitesse et du degré d'intégration, qui va provoquer respectivement un temps de cycle très bas et de grand multiprocesseurs.
- ② **Architectural** : à plusieurs niveaux : CPU , hiérarchie mémoire et assemblage de multiprocesseurs.
- ③ **Logiciel** : Capacité à gérer le parallélisme sous différentes formes et à plusieurs niveaux (intra et inter processus).

Cours 1 : Notions de performance

Mesurer les performances : Métriques

La performance **se quantifie** par plusieurs métriques :

- ① **Temps de réponse** : temps séparant le début et la fin d'exécution.
- ② **Latence** : temps séparant deux exécutions successives.
- ③ **Débit** : quantité de travail effectué par unité de temps.
- ④ **Accélération** (speed up) : rapport du temps d'exécution initiale au temps d'exécution après l'amélioration.
- ⑤ **Efficacité** : rapport de l'accélération obtenue en considérant une amélioration au nombre d'éléments qui ont permis l'amélioration en question (coût de l'amélioration).
- ⑥ **MIPS et MFLOPS**

Mesurer les performances : Moyenne

Les métriques sont exprimés en moyenne. L'utilisation d'un type de moyenne dépend de la nature du métrique.

Soit (x_1, x_2, \dots, x_n) un échantillon des valeurs que peut prendre le paramètre X

- **Moyenne arithmétique** : (exple : temps, taille)

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$$

- **Moyenne géométrique** (exple : taux)

$$\bar{X} = (\prod_{i=1}^n x_i)^{1/n}$$

- **Moyenne harmonique** (exple : vitesse ou débit)

$$\bar{X} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Mesurer les performances : Moyenne

Avantage :

Une seule valeur pour représenter/caractériser un ensemble et simple à calculer

Inconvénient :

Sensible aux valeurs aberrantes, peut donner une idée erronée.

Exemple :

10 est la moyenne de $\{ 10,10,10,10 \}$ et de $\{20,20,0,0\}$

Mesurer les performances : Moyenne et variabilité

Afin de mieux interpréter la moyenne, on étudie la variabilité de l'échantillon avec :

- La variance : $var(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$
- L'écart type : $S = \sqrt{var(X)}$

Un écart type réduit montre que les valeurs de l'échantillon sont proches de la moyenne et améliore sa représentativité.

Mesurer les performances : Accélération

Loi d'Amdahl

T : le temps d'exécution du programme P .

N : le nombre d'instructions dans le programme P .

$t1_{moy}$: le temps moyen d'exécution d'une instruction.

$$T = N \times t1_{moy}$$

Exemple :

Soit un programme P de 10^3 instructions. Chaque instruction s'exécute en moyenne en 5 cycles.

$$T = 10^3 \times 5 \text{ cycles} = 5000 \text{ cycles}$$

Mesurer les performances : Accélération

Loi d'Amdahl

En exécutant P sur une autre machine, 30% des instructions s'exécutent comme sur la première machine.

Les 70% des instructions qui restent s'exécutent 5 fois plus rapidement que sur la première machine.

Le nouveau temps d'exécution de P est :

$$T_{new} = (30\% \times 10^3 \times 5 \text{ cycles}) + (70\% \times 10^3 \times \frac{5 \text{ cycles}}{5}) = 2200 \text{ cy}$$

Donc, en considérant P dans sa globalité, on va $\frac{5000}{2200}$ ($\approx 2,27$) plus vite avec le passage à la nouvelle machine.

Mesurer les performances : Accélération

Loi d'Amdahl : définition

L'amélioration par un facteur a d'un certain type d'instructions constituant une fraction F_a du nombre d'instructions total N , entraîne une accélération globale de :

$$Acc = \frac{1}{(1-F_a) + \frac{F_a}{a}}$$

Exemple précédent : $a = 5$ et $F_a = 70\% = 0.7$

$$Acc = S_p = \frac{1}{0.3 + \frac{0.7}{5}} = 2.27$$

Après l'amélioration, l'exécution est 2.27 fois plus rapide.

⇒ L'accélération globale de la machine est de 2.27

Mesurer les performances : Efficacité

Sachant que la nouvelle machine coûte 2.27 fois plus cher que la machine initiale, l'accélération obtenue est alors compensée par le coût.

L'investissement (la nouvelle machine) n'est pas rentable.

La nouvelle machine est aussi **EFFICACE** que celle initiale.

Mesurer les performances : Efficacité

Acc : l'accélération apportée par l'amélioration.

C : le coût de l'amélioration **par rapport à l'état initial**.

Exprimé en : coût financier, nombre de ressources, ... etc.

L'efficacité ou le rendement de l'amélioration est :

$$Eff = \frac{Acc}{C}$$

Remarque :

l'état considéré initial, doit l'être à la fois pour l'accélération et le coût

Mesurer les performances : Exemples

Loi d'Amdahl : exemple 1

La vitesse du CPU est augmentée par un facteur de 2. Le CPU est utilisé pendant 80% du temps d'exécution, le reste du temps étant consacré aux entrées/sorties.

Quelle est l'accélération totale obtenue ?

le facteur d'accélération $a = 2$, la fraction améliorée $F_a = 0.8$

$$\text{Acc} = S_p = \frac{1}{0.2 + \frac{0.8}{2}} = 1.66$$

L'accélération globale de la machine est de 1.66. l'exécution est 1.66 fois plus rapide avec le nouveau CPU.

Mesurer les performances : Exemples

Loi d'Amdahl : exemple 1

En sachant que le nouveau CPU est 3 fois plus coûteux que le premier,

Quelle est l'efficacité de cet investissement ?

$$\text{Efficacité} = \text{rendement} = \frac{Acc}{C}$$

$$\text{Efficacité} = \frac{1.66}{3} = 0.55$$

L'efficacité étant inférieure à 1, cet investissement n'est pas avantageux. L'accélération obtenue est contrebalancée par le coût élevé.

Mesurer les performances : Exemples

Loi d'Amdahl : exemple 2

On disposait d'une machine monoprocesseur sur laquelle l'exécution d'un programme P dure 13 s qu'on a remplacé par un biprocesseur (de même type) sur lequel le programme P s'exécute en 9 s.

Quelle est l'accélération obtenue ?

$$Acc = \frac{T_{mono}}{T_{biproc}} = \frac{13s}{9s} = 1.44$$

En doublant le nombre de processeurs, on va 1.44 fois plus vite et non 2 fois plus vite. Les deux processeurs ne peuvent être exploités simultanément tout le temps.

Mesurer les performances : Exemples

Loi d'Amdahl : exemple 2

Quelle est l'efficacité de ce biprocesseur ? L'investissement est-il justifié ?

$$Eff_{biProc} = \frac{1.44}{2} = 0.72$$

L'efficacité étant inférieure à 1, le coût l'emporte sur l'accélération obtenue. L'investissement n'est donc pas justifié.

La machine monoprocesseur est plus efficace (dans ce cas) que le biprocesseur.

Mesurer les performances CPU

But :

Fournir une base de comparaison pour le calcul.

La machine qui effectue une quantité de calcul en un minimum de temps est la plus performante.

Vitesse : nombre d'événements par unité de temps.

MIPS : Millions d'Instructions Par Seconde

MFLOPS : Millions d'Instructions Flottantes Par Seconde.

Remarque : Pour le calcul du MIPS et du MFLOPS, on considère que les programmes sont en mémoire centrale.

Mesurer les performances CPU

Temps CPU = temps d'exéc d'un pgme (cy) \times durée du cycle

Temps CPU = $\frac{\text{Nombre de cycles pour un programme}}{\text{frequence d'horloge}}$

Exemple :

Sachant qu'un cycle dure 5 ns sur la machine considérée, le temps CPU d'un programme P de 15 M cy est:

Temps CPU = $15 \times 10^6 \times 5 \times 10^{-9} s = 75ms$

La frequence d'horloge est de : $\frac{1}{5 \times 10^{-9}} Hz = 200MHz$

Mesurer les performances CPU

Les instructions sont différentes \Rightarrow ont des temps d'exécution différents.

C'est très rare (voire impossible) de connaître avec exactitude la composition des programmes et le temps d'exécution de chacune des instructions.

\Rightarrow CPI : nombre moyen de cycles pour exécuter une instruction quelconque d'un programme.

$$\text{Temps CPU} = \frac{\text{Nombre d'instructions} \times \text{CPI}}{\text{frequence d'horloge}}$$

Mesurer les performances CPU : Calcul du CPI

Méthode 1:

Le nombre de **C**ycles **P**ar **I**nstruction peut être calculé par :

$$CPI = \frac{\text{Nombre de cycles pour un programme } P}{\text{Nombre d'instructions dans le programme } P}$$

Exemple : P s'exécute en 15 Mcy et contient 6 Minstrs.

Chaque instruction (peu importe son type) s'exécute en

$$\frac{15 \times 10^6 \text{ cy}}{6 \times 10^6} = 2.5 \text{ cy}$$

⇒ CPI = 2.5 cycles

Mesurer les performances CPU : Calcul du CPI

Méthode 2:

Si la composition du programme en classes d'instructions et la durée de chaque classe sont connues :

$$CPI = \frac{\sum_i T_{inst_i} \times \text{pourcentage} - inst_i}{100}$$

La performance du CPU dépend de 3 paramètres:

- ① La fréquence de l'horloge qui dépend de la technologie du matériel
- ② Le nombre moyen de cycles par instruction qui dépend du jeu d'instructions
- ③ Le nombre d'instructions d'un programme qui dépend du jeu d'instructions et du compilateur.

Mesurer les performances CPU : Calcul du CPI

Exemple : La composition pour les programmes s'exécutant sur la machine M et le CPI de chaque classe des instructions

| Type d'instr | Fréquence d'utilisation en % | CPI |
|--------------|------------------------------|-----|
| Logique | 4 | 1 |
| Arith. | 64 | 2 |
| accès Mem. | 10 | 4 |
| Branchements | 22 | 3 |

$$CPI_{moy} = \frac{4*1+64*2+10*4+22*3}{100} = 2.38cy$$

Donc, **en moyenne** une instruction quelconque s'exécute en 2.38 cy sur la machine M.