

TD1-M2: Métriques d'évaluation de performances

S. Zertal

1 Loi d'Amdhal

Les questions qui suivent sont indépendantes :

1. La fraction séquentielle d'une application représente 4% du temps d'exécution sur un processeur.

Calculer l'accélération et l'efficacité sur un multiprocessor de 2^n processeurs ($2 \leq n \leq 7$). Quelle est la borne supérieure de l'accélération ?

2. La mesure des accélérations des programmes A et B fournit les résultats suivants :

Procs	2	3	4	5	6	7	8
Acc_A	1.9	2.73	3.47	4.16	4.8	5.38	5.93
Acc_B	1.94	2.72	3.34	3.82	4.2	4.49	4.72

Calculer la fraction séquentielle pour chacune des exécutions. Comment peut-on décrire le comportement parallèle des deux programmes ?

3. On supposera qu'une machine parallèle (multiprocesseurs) à p processeurs permet de diviser par p le temps d'exécution d'une fraction de code parallélisé. On appellera T_1 le temps d'exécution d'un programme sur un processeur et T_p celui sur p processeurs.

(a) Sachant que seq est la fraction séquentielle du programme à paralléliser, exprimez l'accélération pour une machine parallèle à p processeurs. Quelle est l'accélération maximale envisageable ?

(b) Quelle est l'accélération pour une machine dotée d'un très grand nombre de processeurs ?

(c) Si seq est inversement proportionnelle à la taille n du code (nombre d'instructions). Calculez alors l'accélération. Quelle est l'accélération pour un très grand code ?

4. Mode scalaire/vectériel

Un mode d'exécution vectoriel est un mode dans lequel on effectue des opérations de haut niveau portant sur des vecteurs au même titre que l'on manipule habituellement des opérations sur de simples scalaires.

Pour cet exercice, on considérera l'amélioration d'une machine liée à l'ajout

d'un mode vectoriel. Quand un code s'exécute en mode vectoriel, il est 20 fois plus rapide qu'en mode scalaire.

Le pourcentage du nombre d'opérations effectuées en mode vectoriel est appelé pourcentage de vectorisation.

- (a) Tracez un graphe donnant l'accélération liée au calcul réalisé en mode vectoriel en fonction du pourcentage de vectorisation.
- (b) Quel est le pourcentage de vectorisation nécessaire pour obtenir une accélération de 2 ?
- (c) L'accélération est maximale lorsque la totalité des instructions est effectuée en mode vectoriel. Quel est le pourcentage de vectorisation nécessaire pour obtenir la moitié de cette accélération maximale ?
- (d) Le pourcentage de vectorisation mesuré pour les programmes est de 70%. On peut augmenter les performances en doublant la vitesse du mode vectoriel par l'intégration d'un nouveau dispositif matériel. En l'absence d'un tel dispositif, comment pourrait-on assurer les mêmes performances ?

2 Performance en MIPS/MFLOPS

Les deux questions suivantes sont indépendantes :

1. Un programme de test (*Whetstone*) contient 195578 opérations flottantes de base par itération. En considérant les types d'opérations flottantes, on obtient la décomposition suivante:

Type d'opération flottante	Nombre
Addition	82014
Soustraction	8229
Multiplication	73220
Division	21399
Conversion entier vers flottant	6006
Comparaison	4710
Total	195578

Whetstone a été exécuté sur une machine *M* cadencée par une horloge de 16.67 Mhz, dotée d'un coprocesseur flottant et utilisant un compilateur Fortran. Ce compilateur permet d'utiliser au choix le coprocesseur ou bien des routines logicielles, selon les options de compilation.

L'exécution d'une seule itération de *Whetstone* a duré 1.08 sec avec le coprocesseur et 13.6 sec avec l'option logicielle. On a mesuré un CPI moyen de 10 en utilisant le coprocesseur et de 6 en utilisant l'option logicielle.

- (a) Quel est le nombre de MIPS pour les deux exécutions
 - (b) Quel est le nombre total d'instructions exécutées dans les deux cas.
 - (c) En moyenne, combien faut-il d'instructions entières pour calculer chaque opération flottante par logiciel?
 - (d) Quel est le nombre de MFLOPS pour la machine M avec le coprocesseur flottant exécutant *Whetstone* (on suppose que toutes les opérations flottantes du tableau précédent comptent pour une opération)?
2. Votre entreprise a un programme de test qui est considéré comme représentatif de vos applications typiques. Une des stations de travail d'un modèle plus ancien n'a pas d'unité flottante, et doit émuler chaque instruction flottante par une séquence d'instructions entières. Cette station de modèle ancien a un débit de 120 MIPS sur ce programme test. Un vendeur tiers offre un coprocesseur qui est destiné à donner un "coup de fouet" à votre station. Ce coprocesseur exécute chaque instruction flottante sur un processeur dédié (aucune émulation n'est nécessaire). Le débit de l'ensemble station/coprocesseur est 80 MIPS sur le même programme test. Les symboles suivants sont utilisés pour répondre aux questions.

Symbole	Commentaire
I	Nombre d'instructions entières exécutées dans le programme
F	Nombre d'instructions flottantes exécutées dans le programme
Y	Nombre d'instructions entières pour émuler une instruction flottante
W	Temps pour exécuter le programme sur la station seule
B	Temps pour exécuter le programme sur l'ensemble station/coprocesseur

- (a) Ecrire la formule du débit MIPS de chaque configuration avec les symboles de la table. Justifiez votre réponse.
- (b) Pour la configuration sans coprocesseur, on mesure $F = 8 \times 10^6$, $Y = 50$, $W = 4$. Calculez I en conséquence.
- (c) Quelle est la valeur de B
- (d) Quel est le débit MFLOPS du système avec le coprocesseur.
- (e) Votre collègue veut acquérir le coprocesseur même si le débit MIPS pour la configuration l'utilisant est moindre que celui de la station de travail seule. Est ce que l'estimation de votre collègue est correcte? Justifiez votre réponse.

3 Loi de Gustafson-Barsis

Définition :

Le métrique de Gustafson-Barsis sert à calculer la capacité d'exécuter un code plus grand dans la même durée d'exécution que le problème initial en utilisant plus de processeurs. Ici, c'est le cas d'un passage du séquentiel (1 processeur) à P processeurs ($P > 1$) d'un code dont la fraction séquentielle est seq . Ce passage s'accompagne par une accélération de mise à l'échelle (Scaled speed up), tel que :

$$Acc(scaled) \leq p + (1 - p) \times seq$$

1. Quelle est l'accélération selon *Gustafson-Barsis* (scaled speedup) pour une application s'exécutant sur 8 processeurs et qui passe 14% de son temps à exécuter sa partie séquentielle.
2. Quelle est la fraction séquentielle maximale qu'on puisse tolérer pour garder au moins la même accélération (scaled speedup) quand le nombre de processeurs passe à 10.

4 Métrique de Karp-Flatt

Définition :

Le métrique de Karp-Flatt sert à inclure le coût de la parallélisation sur P processeurs en considérant davantage de sources de coût et d'inefficacité. Il se focalise sur la partie séquentielle qui prend plus d'importance avec l'augmentation du nombre de processeurs. Ce métrique se calcule par un facteur e

$$e = \frac{\frac{1}{Acc} - \frac{1}{P}}{1 - \frac{1}{P}}$$

1. Soit le tableau ci-dessous, regroupant les temps d'exécution mesurés pour un code P durant ses exécutions sur la machine Cray Y-MP/8 avec différent nombre de processeurs.

Nombre de processeurs P	Temps d'execution (s)
1	2.170
2	1.110
3	0.754
4	0.577
8	0.312

- (a) Calculez l'accélération, l'efficacité et le facteur e de Karp-Flatt pour chacun des quatres cas ($P = 2, 3, 4$ et 8)
 - (b) Que pensez vous des accélérations et efficacités obtenues ? Argumentez votre réponse.
 - (c) Qu'indique les valeurs du facteur e dans ce cas ?
2. Soit le tableau ci-dessous, regroupant les temps d'exécution mesurés pour un code P durant ses exécutions sur la machine Aliant FX/40 avec différent

nombre de processeurs.

Nombre de processeurs P	Temps d'exécution (s)
1	66.1
2	34.8
3	24.9
4	20.5

- (a) Calculez l'accélération, l'efficacité et le facteur e de Karp-Flatt pour chacun des trois cas ($P = 2, 3$ et 4).
- (b) Interpretez la variation de l'efficacité puis faites de même pour le facteur e .