

ヘテロな環境に対応したプログラミングのすすめ

SUPERCOMPUTING IN HETEROGENEOUS ENVIRONMENT

日置 慎治

Shinji Hioki

工博 帝塚山大学教授 経営情報学部 (〒631-8501 奈良市帝塚山7-1-1, hioki@tezukayama-u.ac.jp)

The way to achieve the high performance calculation in heterogeneous environment is discussed. If the code can run on any number of processors, you can do efficient partitioning of the processors and then the code will run fast. I will show one simple example to realize this situation. The status of the SuperComputing in the world will be also presented.

Key Words : SuperComputing, Heterogeneous Environment, PSE

1. Introduction

膨大な計算能力を必要とするアプリケーションは多い。そのような計算を遂行するためにはスーパーコンピュータに代表される並列計算機が必要であり、同時に、並列プログラミングが不可欠である。

はじめに、スーパーコンピュータについて触れておこう。スーパーコンピュータの歴史や現状を語るとき、2つの主な指標に注目するとよい。1つは俗にいう“Top500 list” [1]であり、他方は“ゴードンベル賞 (Gordon Bell Prize Winner)” [2]である。これら2つは、ともに計算能力を競うものであるが、以下の点が異なっている。

- (1) Top500ではLINPACK[3] 計算を対象とする
- (2) ゴードンベル賞は全アプリケーション対象

しかしながら、歴史を概観すると、これら2つのトップは、多くの場合と同じ顔で占められてきた。例えば、最近では、2009年のゴードンベル賞はORNLでのCray XT5を使った計算に与えられ、Top 500はORNLの“Jaguar - Cray XT5-HE Opteron Six Core 2.6 GHz” が受賞した。

図1は、Top 500に見る計算性能の推移である。(上のグラフから、1位から500位までの性能合計、1位の性能、500位の性能の順)

この図からは、とてもクリアなムーアの法則を見ることが出来る。

表1は、2010年6月におけるTOP500のベスト10である。特筆すべきは、中国が2つ入っていること。そして何より残念なのは、日本が見当たらないことである。

(2010年後半に、中国の天河1Aがトップに躍り出た)



図1 TOP500に見る性能の推移 (<http://www.top500.org>.)

Rank	Manufacturer	Country	RMax
1	Cray Inc.	United States	1759000
2	Dawning	China	1271000
3	IBM	United States	1042000
4	Cray Inc.	United States	831700
5	IBM	Germany	825500
6	SGI	United States	772700
7	NUDT	China	563100
8	IBM	United States	478200
9	IBM	United States	458611
10	Sun Microsystems	United States	433500

表1 2010年6月のTOP500 (<http://www.top500.org>.)

これら上位の性能をもつ計算機に共通するのは、ノードの性能が均一の「ホモジニアス」な並列計算機である、という点である。高性能を達成するために並列プログラムを作成するという立場からすれば、ホモジニアスな計算機は扱いやすく、最高性能を達成しやすい。

これは、すべてのノードで同じプログラムを走らせ、均一な負荷を与えるSIMD型のプログラミングが容易だからである。

一方、例えば大学や研究所等の計算センターのような環境に目を転じてみると、様々な計算機が混在している「ヘテロジニアス」な環境がむしろ一般的であり、そのような環境において、最高性能を達成しようとするならば、ホモジニアスな環境では理想的であったSIMD型のプログラミングスタイルは通用せず、我々はパラダイムの転換を必要とされることになる。

この稿では、既存のプログラムに少し変更を加えることで、パラダイムの交換を行う可能性および実証実験の結果について触れてみたい。

2. QCDMPI

QCDMPI [4] は、素粒子物理に関する、格子QCDモンテカルロシミュレーションを行うプログラムであり、筆者により開発され、これまで世界中で広く使われてきた。

QCDMPI はその名の通り、並列処理部分にMPIを利用しており、かつ、下記のようにポータブルな特徴を持っている。;

- 1) 何次元のQCDでも計算可能
- 2) 何次元分割にも対応
- 3) ノードの数は任意
- 4) 必要とする作業メモリが少ない

これらの特徴は密接に関係しているのであるが、今回は、3)に注目したい。

ある計算を遂行するために必要なノードの数を可変にしておくことで、ヘテロジニアスな環境に対応することが可能であると想定できる。つまり、計算能力に応じた負荷分散を行えばよい。ノードの数を可変にする方法はアプリケーション毎に異なると考えられるが、領域分割の問題に関してはQCDMPIの例が役に立つと思われる。

QCDMPIの場合の詳細については文献[4]を参照いただきたい。

3. ヘテロジニアスな環境での並列化

ヘテロジニアスな例として、2つのコンピュータAとBから構成されるシステムを考えよう。Aの性能はBの2倍であるとする。この場合、理想的な全体性能はBの3倍であるが、このシステムをホモジニアスな環境であると考えて、SIMD型のプログラム（全ノードで負荷が均一）を走らせた場合に得られる性能は、遅いBに制約されるためにBの2倍となる。（図2のCase1）

一方、ヘテロジニアスな環境に対応可能なプログラムの場合には、全体のノード数を3にしておいて、そのうち

2つをAで、残りの1つをBで実行することにより、全体性能は理想値であるBの3倍を達成することができる。

（図2のCase2）

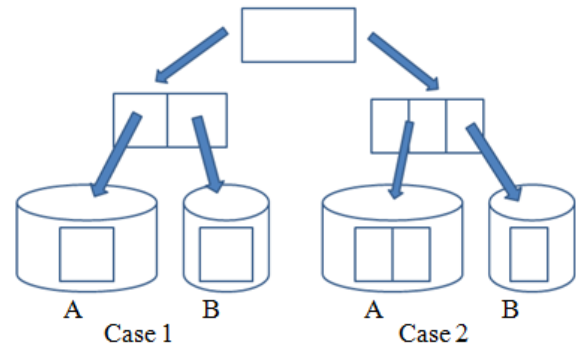


図2 並列化方法

性能比が1:Nの2台のコンピュータの場合であればノード数を(N+1)とし、高性能のコンピュータでNノードを分担させればよい。この場合には、ホモジニアスな環境と比べた時の、ヘテロジニアスな場合の計算時間の比の理想値は $2 / (1 + N)$ となる。

4. 実機による実測とまとめ

性能比 1 : Nの2台のコンピュータを使い、実測値と理想値の比較を行った。当然、通信時間が含まれるので、理想的な性能からずれる可能性も考えられるが、QCDMPIは通信時間も最小化するように設計されているため、大きな影響はないと予想される。

結果は、図3となり、2つはほぼ一致した。つまり、ノード数を可変にすることだけで、ヘテロジニアスな環境であっても理想的な並列化が実現できることが示された。

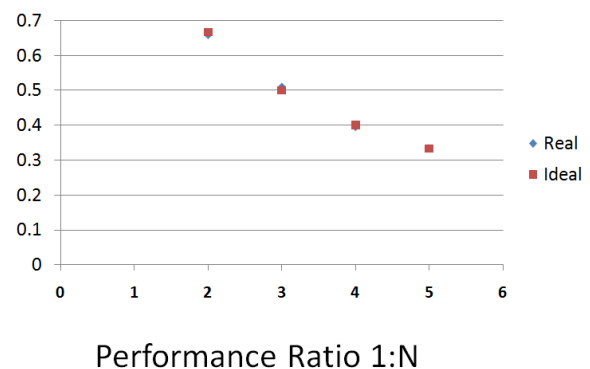


図3 ヘテロジニアスな環境における性能向上

5. References

- [1] <http://www.top500.org/>
- [2] for the past list, for example see <http://www.sc2000.org/bell/pastawrd.htm> and <http://awards.acm.org/homepage.cfm?awd=160>
- [3] <http://www.netlib.org/linpack/index.html>
- [4] S.Hioki, Nuclear Physics B63 (1998) 1000.