

NAS Parallel Benchmark の実行時間予測モデルの構築

指導教員：市川周一

学籍番号：043710 作元佑輔

1 背景と研究目的

今日、PC クラスタを用いた並列処理は広く用いられている。様々な PC が混在するクラスタ (不均一クラスタ) 環境では、様々な構成で応用プログラムを実行することが可能であり、その中から最適な構成を選択して応用の実行時間を最小化することは重要なテーマである。岸本 [1] は、HPL を例として不均一クラスタ上での実行時間予測モデルを構築し、実際に (準) 最適構成が予測可能であることを示した。しかし岸本の手法は、HPL 以外の応用では検証されていない。

本研究では、NAS Parallel Benchmark (NPB)[2] の実行時間予測モデルを構築し、その精度を検討することを目的とする。NPB は並列計算機の性能評価を目的として開発されたベンチマークプログラムで、計算流体力学のための 5 つのカーネルと 3 つの擬似応用から構成されている。NPB には 5 つの問題サイズ (クラス S, A, B, C, D) だけが規定されているが、実行時間予測モデルを構築するには多くの問題サイズに対して実行時間を測定する必要がある。そこで本研究では、NPB (ver.2.4) のうち計算サイズが変更できる 4 つのプログラム (MG, CG, IS, LU) について実行時間予測モデルを構築し、最適なクラスタ構成を予測できるか検証する。

2 実行時間予測モデル

本研究では、問題を単純化するために均一なサブクラスタだけを扱うこととし、岸本 [1] の N-T モデルを簡略化して用いる。N-T モデルとは、問題サイズ N に対する実行時間 $T(N)$ を見積もる式である。実行時間 $T(N)$ は計算時間 $Ta(N)$ と通信時間 $Tc(N)$ からなり、 $T(N) = Ta(N) + Tc(N)$ であると仮定する。

以下、MG を例として説明する。MG の計算量は $O(N^3)$ [3] なので、 $Ta(N)$ は N の 3 次式で見積もることができる。MG のソースコードによれば通信量も $O(N^3)$ なので、 $Tc(N)$ も N の 3 次式となる。従って、あるクラスタ構成における MG の $T(N)$ は以下の式で見積もることができる。

$$T(N) = k_0 \cdot N^3 + k_1 \cdot N^2 + k_2 \cdot N + k_3 \quad (1)$$

式 (1) に含まれる定数 ($k_0 \sim k_3$) は、テストケースの実測値から最小二乗法で決定する。未知パラメータは 4 つなので、各構成の $T(N)$ を決定するには 4 つ以上の N について実測する必要がある。

CG, LU, IS で同様に $T(N)$ を検討したところ、CG と LU は N の 3 次式、IS は N の 1 次式になることがわかった。

3 評価結果

本研究で用いる測定環境は、表 2 に示すとおりである。このクラスタから、Pentium2 $\times$ 1, 2, 4, 8 (CPU), Pentium3 $\times$ 1, 2, 4 (CPU) の計 7 構成についてモデルを構築する。各 CPU に割り当てるプロセス数は 1 である。

MG の問題サイズは 2^n ($n = 1, \dots, 8$) とした。各構成に関して実行時間を実測し、実測時間からパラメータ抽出を行って 7 つのモデルを構築した。パラメータ抽出には gnuplot の fit コマンドを用いた。

表 2: 測定環境

CPU	Pentium3 866MHz	Pentium2 400MHz
CPU 数	4 (dual $\times$ 2 台)	8 (dual $\times$ 4 台)
メモリ	768MB	
NIC	100BASE-TX	
Compiler	gcc 3.2.2	
OS	Red Hat Linux 9	
通信ライブラリ	mpich-1.2.7	

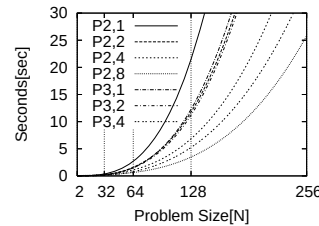


図 1: 各構成のモデル

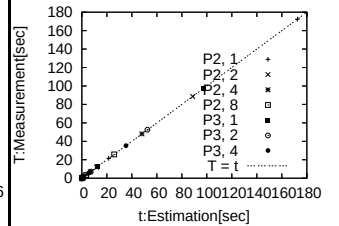


図 2: 予測値と実測値

各構成の $T(N)$ をグラフ化したものを図 1 に示す。各モデルの予測値と実測値の相関関係は図 2 の通りであり、予測値と実測値は非常に良く一致している。

構築した実行時間予測モデルを用いて、任意のサイズ N に対する最適構成を予測することができる。表 1 に、MG に関する評価結果をまとめる。表中、 t は予測最良構成の予測実行時間、 t' は予測最良構成の実測実行時間、 T' は実測最良構成の実測時間をあらわす。問題サイズ 32 以下で予測実行時間の相対誤差 $(t - T')/T'$ が大きくなっているが、絶対誤差 (予測実行時間と実測実行時間の差) は 0.1 秒未満である。サイズ 8~256 では、予測最良構成と実測最良構成の実行時間差は 6% に留まっており、準最適~最適構成が実際に予測できたことがわかる。

CG について $N = 500 \sim 75000$ で評価を行った結果、 $N = 1000 \sim 75000$ で最適構成の予測に成功した。このとき予測時間の最大誤差は 6% 程度であった。LU では、 $N = 22 \sim 64$ で評価した結果、予測実行時間の誤差は最大 13%、実測実行時間の誤差は 14% となり、十分に正確な (準) 最適構成を予測することができた。IS では $N = 2^{16} \sim 2^{25}$ の範囲で評価し、全ての N で最適構成の予測に成功した。このときの予測実行時間の絶対誤差は 1 秒以下であった。

参考文献

- [1] 岸本 芳典: “不均一クラスタ上での実行時間予測モデルとその評価,” 豊橋技術科学大学 知識情報工学課程 修士論文 (2003).
- [2] “The NAS Parallel Benchmarks,” <http://www.nas.nasa.gov/Software/NPB/> (2005).
- [3] 笹生 健, 松岡 聡: “ヘテロなクラスタ環境における NAS Parallel Benchmarks の適用化,” 情処研報 2003-HPC-93, pp.1-6 (2003).

表 1: 最良構成の予測結果 (MG)

サイズ N	予測最良構成			実測最良構成			誤差	
	CPU, 数	t	t'	CPU, 数	T'		$(t - T')/T'$	$(t' - T')/T'$
2	P2, 1	-4.36E-04	5.71E-04	P3, 1	2.47E-04		-2.76	1.31
4	P2, 1	2.50E-03	1.70E-03	P3, 1	7.66E-04		2.27	1.22
8	P3, 1	2.12E-03	3.90E-03	P3, 1	3.90E-03		-0.46	0.00
16	P3, 1	7.95E-03	2.34E-02	P3, 1	2.34E-02		-0.66	0.00
32	P3, 4	1.09E-01	1.83E-01	P3, 2	1.72E-01		-0.37	0.06
64	P2, 8	6.31E-01	5.93E-01	P2, 8	5.93E-01		0.06	0.00
128	P2, 8	3.48E+00	3.49E+00	P2, 8	3.49E+00		0.00	0.00
256	P2, 8	2.58E+01	2.58E+01	P2, 8	2.58E+01		0.00	0.00