

不均ークラスタのための実行時間予測モデルの構築法の改良

指導教員：市川 周一

学籍番号：011011 河合 裕

1 はじめに

演算性能や通信性能が異なる要素プロセッサ (PE) で構成されるクラスタを不均ークラスタと呼ぶ。既存の並列応用の多くは各 PE に負荷を均等に分散するため、不均ークラスタ上で実行すると低速 PE がボトルネックとなり性能低下が生じる。PE の性能に応じた不均な負荷分散を行うことで性能向上が期待できる。負荷分散の手法の一つに、高速 PE に複数プロセスを起動するマルチプロセス法がある。この手法では PE に幾つプロセスを起動するかが問題となる。高橋ら [1] は実測値から実行時間予測モデルを構築し、4 つの科学技術応用について (準) 最適なプロセス構成を予測できることを示した。

高橋ら [1] の手法では均一 PE を使用して予測モデルを構築していたため、利用可能な状況は限られていた。また、実測値に誤差が含まれる場合に予測精度が著しく低下する問題があった。本研究では、不均ークラスタ自体から直接モデルを構築する方法を示す。さらに、予測精度の低下を抑える構築手法を検討する。

2 高橋らの実行時間予測モデル

応用の問題サイズを N 、不均ークラスタ中で等価な PE のグループ (サブクラスタ) を G_i 、 G_i 中で計算に使用する PE の台数を P_i とする。 G_i 内の PE には同数 (M_i 個) のプロセスを起動すれば、総プロセス数は $P = \sum P_i M_i$ となる。 G_i の実行時間 T_i を N, P, M_i の関数で近似すると、全体の実行時間 T は $\max_i T_i$ で見積もることができる。この T の近似式を実行時間予測モデルと呼ぶ。全ての可能な構成 (P_i, M_i) に対して予測モデルを構築できれば、 T が最小となる最適構成が予測できる。

HPL の実行時間 T は式 (1) で表され、ある M_i の実行時間 T_i を N と P の関数で表すと式 (2) になる。高橋らはこの式 (2) を NP-T モデルと名づけた。 k_0 から k_9 までの係数は実測値から最小二乗法で決定する。ただし通信時間を考慮して、均一 PE 2 台以上の実測値だけを使用する。

$$T(N, P) = \frac{1}{P} \cdot O(N^3) + P \cdot O(N^2) + O(N^2) \quad (1)$$

$$T_i(N, P)_{|M_i} = \frac{1}{P} \cdot (k_0 N^3 + k_1 N^2 + k_2 N + k_3) + P \cdot (k_4 N^2 + k_5 N + k_6) + k_7 N^2 + k_8 N + k_9 \quad (2)$$

式 (2) から P について最低でも 3 つの実測値が必要になる。 P の異なる実測値を得るには、PE の台数を変化させる必要があるため最低でも 4 台、応用の制限によってはそれ以上の台数の均一 PE が必要となる。また、実測値に誤差があるとパラメータ抽出に失敗し、負の実行時間を予測することがある (モデルの破綻)。その結果、最適でない構成が予測される場合がある。

3 改善方法

本研究ではモデル式自体は変更せず、モデル構築の方法を改良することを検討する。以下に示す 2 つの手法は互いに独立した手法であるため併用が可能である。

3.1 不均ークラスタからのモデル構築

低速な PE を 1 台以上含む不均ークラスタの実行時間は、同数の低速 PE で構成される均一クラスタの実行時間と同等だと仮定する。本研究では、この実行時間を使用してモデル構築を行う。同種の PE が 1 台しかない場合でも、それより高速な PE を利用してモデルを構築することが可能になる。

3.2 非負最小二乗法によるパラメータ抽出

負の実行時間が予測される原因は、 k_i ($i = 0 \sim 9$) が負の値を含むからである。負の実行時間が予測されないようにするためには $k_i \geq 0$ ($i = 0 \sim 9$) に制限すればよい。そこで本研究では、抽出されるパラメータに非負制約を付けた非負最小二乗法を使ってモデルを構築する。これによりモデルの破綻を回避して、精度低下を防ぐことが期待できる。

表 1: 評価環境

	サブクラスタ G_1	サブクラスタ G_2	サブクラスタ G_3
CPU	Pentium4 3.6GHz	Xeon 2.8GHz	CeleronM 1.5GHz
OS	FedoraCore4	RedHatLinux9	FedoraCore5
開発環境	icc 9.0, ifc 9.0, mpich 1.2.7p1		
P_i	$0 \leq P_1 \leq 2$	$0 \leq P_2 \leq 4$	$0 \leq P_3 \leq 2$
M_i	$0 \leq M_1 \leq 3$	$0 \leq M_2 \leq 2$	$0 \leq M_3 \leq 1$

表 2: 問題サイズ (N)

ベンチマーク	モデル構築時	評価時
HPL	400 ~ 6400 (9 点)	400 ~ 9600 (11 点)
Himeno BMT	32 ~ 192 (9 点)	32 ~ 256 (11 点)
hpcmw-solver-test	60 ~ 442 (7 点)	60 ~ 600 (12 点)
FFTE	$2^{12} \sim 2^{20}$ (9 点)	$2^{12} \sim 2^{23}$ (12 点)

4 評価

表 1 に示す不均ークラスタを用いて、HPL, Himeno BMT, hpcmw-solver-test, FFTE の 4 つのベンチマークに対して評価を行った。モデル構築の測定および評価は、表 2 に示す問題サイズで行った。 G_3 のモデルは G_1 (2 台), G_2 (4 台), G_3 (2 台) を使って構築し、 G_2 のモデルは G_1 (4 台) と G_2 (4 台) を使って構築する。最も高速な G_1 のモデルに限り G_1 (8 台) の均一サブクラスタから構築する。高速 PE のモデル構築法は今後の課題である。

各ベンチマークの予測最適構成の実測時間と実測最適構成の実測時間 (optimal) を図 1 にまとめる。HPL, hpcmw-solver-test では全ての手法で (準) 最適構成が予測できており、均一サブクラスタからモデル構築した場合 (homo) と不均ークラスタからモデル構築した場合 (hetero) で同程度の精度が得られた。Himeno BMT は通常の最小二乗法 (ls) では $N \leq 96$ および $N \geq 224$ で誤差が増大しているのに対し、非負最小二乗法 (nnls) では (準) 最適構成が予測できている。FFTE も Himeno BMT と同様、通常の最小二乗法では $N \leq 2^{16}$ および $N \geq 2^{21}$ で誤差が大きくなっているが、非負最小二乗法では改善されている。Himeno BMT と FFTE は、通常の最小二乗法では均一サブクラスタからモデル構築した場合と不均ークラスタからモデル構築した場合で予測精度が異なるが、非負最小二乗法では同程度の精度が得られた。

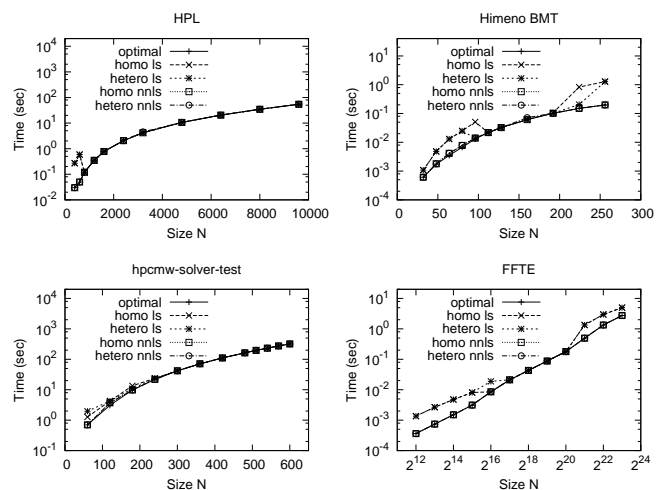


図 1: 評価結果

5 おわりに

本研究では、不均ークラスタ自体からモデル構築する手法を提案し、モデルの破綻が生じない場合については均一クラスタからモデルを構築した場合と同様の精度が得られることを示した。さらに、非負最小二乗法を用いることで予測精度の低下を防ぎ、(準) 最適構成を予測することに成功した。

参考文献

- [1] 高橋翔, 市川周一: 不均ークラスタの最適構成予測モデルの各応用への適用と評価, 情報処理学会研究報告 2006-HPC-105, pp. 97-102 (2006).