

# 不均一クラスタにおける最適構成予測手法の適用と評価

指導教員：市川 周一

学籍番号：023719 高橋 翔

## 1 研究背景・目的

既存の並列応用の多くは、要素プロセッサ (PE) が均一な計算環境を前提としており、PE 性能が不均一なクラスタ上では低速 PE による性能低下が発生する。不均一クラスタ上での性能を改善するには、PE 性能に応じた不均一な負荷分散が必要である。

不均一な負荷分散手法の一つに、高速 PE に複数プロセスを割り当てる方法 (マルチプロセス法) がある。岸本ら [1] は HPL (High Performance Linpack) にマルチプロセス法を適用し、実測値から実行時間予測モデルを構築して (準) 最適な構成を予測することに成功している。しかし岸本らは HPL のみを検討しており、他の広範囲の応用に適用できるか否か実証されていない。

本研究では岸本の手法を 4 種のベンチマークプログラムに適用し、(準) 最適構成を予測できるか検証する。また、岸本らのモデルを改良して、より高精度なモデルを構築できることを示す。

## 2 実行時間予測モデル

### 2.1 岸本らによるモデル

ある応用の問題サイズを  $N$ 、不均一クラスタ内の同性能 PE の集合をサブクラスタ  $G_i$ 、 $G_i$  内で計算に利用する PE 数を  $P_i$  とする。 $G_i$  内の PE に等しく  $M_i$  プロセス起動するなら、プロセス総数  $P = \sum_i P_i M_i$  となる。 $G_i$  の実行時間  $T_i$  は  $N, P, M_i$  の関数となり、全体の実行時間  $T = \max_i T_i$  となる。この  $T$  の近似式を実行時間予測モデルと呼ぶ。不均一クラスタ上の全ての構成  $(P_i, M_i)$  のモデルが構築できれば、最適な構成を予測できる。

HPL を例にモデル構築を行う。HPL の実行時間  $T$  は式 (1) で表され、 $\exists(P_i, M_i)$  を定めると  $T_i$  は式 (2) となる。式に含まれる係数は実測値から最小二乗法で求める。このモデルを岸本らは N-T モデルと名づけた。

N-T モデルは  $(P_i, M_i)$  の全構成の測定が必要であり、大規模クラスタでは構築時間が膨大になる。そこで N-T モデルから  $P$  を変数に含むモデルを構築する。このとき  $T_i$  は通信相手に依存せず  $P$  にのみ依存すると仮定すると、式 (3) が得られる。これを岸本らは P-T モデルと呼んだ。P-T モデルは各  $G_i$  内の測定から構築できるため、N-T モデルより構築時間を削減できる。なお、PE 数=1 の N-T モデルには通信時間が含まれないため、P-T モデルは PE 数が 2 以上の N-T モデルより構築する。

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

$$T_i(N)|_{P, M_i} = k_0 N^3 + k_1 N^2 + k_2 N + k_3 \quad (2)$$

$$T_i(N, P)|_{M_i} = \frac{k_0}{P} \cdot T_i(N)|_{P, M_i} + k_1 P \cdot T_i(N)|_{P, M_i} + k_2 \quad (3)$$

### 2.2 本研究で提案するモデル

式 (1) を  $N, P$  で展開すると式 (4) となる。これを NP-T モデルと名づける。NP-T モデルは、N-T・P-T モデルより係数が多いため自由度が高く、構築に利用可能なデータ数が  $N$  の測定数  $\times$  (PE 数 - 1) と N-T・P-T モデルより多いため、正確なモデルが構築できると期待される。NP-T モデルも  $G_i$  内の測定から構築できるため、構築時間は P-T モデルと同等である。

$$T_i(N, P)|_{M_i} = \frac{1}{P} \cdot (k_0 N^3 + k_1 N^2 + k_2 N + k_3) + P \cdot (k_4 N^2 + k_5 N + k_6) + k_7 N^2 + k_8 N + k_9 \quad (4)$$

## 3 評価方法

本研究では表 1 の不均一クラスタを用い、次の 4 種のベンチマークを実験に用いた。各ベンチマークの測定範囲と評価範囲を表 2 に示す。各アプリケーションの N-T, P-T, NP-T モデルを構築し、各モデルで最適構成を予測する。

**HimenoBMT** Poisson 方程式を Jacobi 反復で解くカーネルの性能を測定する。 $N \times N \times N$  の問題を扱う。

**hpcmw-solver-test** 3次元有限要素解析を各種アーキテクチャで評価するためのベンチマーク。記憶容量等の関係上、 $N \times N \times 1$  の問題を扱う。

FFTE  $N = 2^{23} 5^7$  の FFT を行うプログラム。今回の評価では  $N$  を 2 のべき乗に固定する。プロセス数  $P$  に 3 か 5 の因数が含まれる場合と含まれない場合で領域分割が異なるため、場合分けして別々にフィッティングを行う。これは P-T, NP-T モデルに対して適用する。

**HPL** 連立一次方程式を解くベンチマーク。岸本の結果の追試として評価に加える。

表 1: 測定環境

	サブクラスタ $G_1$	サブクラスタ $G_2$
PE	Xeon 2.8 GHz	Celeron M 1.5 GHz
OS	Redhat Linux 9	FedoraCore 3
Compiler, Library	gcc 3.2.2, ifc 8.1, mpich-1.2.6 (Buffer 8KB)	
$P_i$	$1 \leq P_1 \leq 8$	$0 \leq P_2 \leq 8$
$M_i$	$1 \leq M_1 \leq 2$	$0 \leq M_2 \leq 1$

表 2: 測定・評価範囲

	測定範囲	評価範囲
HimenoBMT	32 ~ 192 9 点	32 ~ 256 10 点
hpcmw-solver-test	70 ~ 504 7 点	70 ~ 660 20 点
FFTE	$2^{12} \sim 2^{20}$ 9 点	$2^{16} \sim 2^{23}$ 8 点
HPL	400 ~ 6400 9 点	1600 ~ 9600 7 点

## 4 評価結果

各モデルによる予測最適構成の (実測) 実行時間、実測最適構成での実行時間 (Optimal) を、図 1 にまとめる。

HPL, hpcmw-solver-test では、NP-T モデルは (準) 最適構成の予測に成功している。N-T, P-T モデルも、内挿範囲では (準) 最適構成が予測できた。しかし外挿範囲では、フィッティングが失敗して高次項の係数が負になる構成が選択されるため、誤差が大きくなった。

HimenoBMT では  $N = 160$  で P-T モデル、 $N = 256$  で N-T モデルが実測最適時間から大きく外れている。NP-T モデルの予測構成の実行時間は最適実行時間とほぼ一致し、(準) 最適構成を予測できた。

FFTE では  $N$  の増加につれ N-T, P-T モデルの誤差が大きくなっている。NP-T モデルは予測構成の実行時間は最適構成とほぼ一致し、(準) 最適構成を予測できた。

このように、岸本らの N-T・P-T モデルは応用によっては精度が低下する。一方、本研究で提案する NP-T モデルは、P-T モデルと同等以上の構成を予測することができ、より広範囲の応用に適用できた。現在は 2 種類のプロセッサを用いて評価しているが、今後はプロセッサの種類を増やして、より不均一な環境での検討を進める。

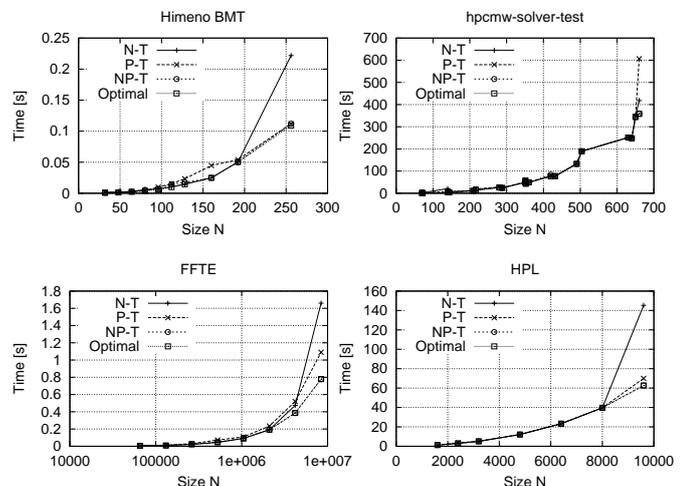


図 1: 各モデルの評価結果

## 参考文献

- [1] Kishimoto, Y. and Ichikawa, S.: Optimizing the Configuration of a Heterogeneous Cluster with Multiprocessing and Execution-Time Estimation, *Parallel Computing*, Vol. 32, No. 7, pp. 691-710 (2005).