

# 6J-05 情報発信機能の付加による自律負荷分散方式の改善

宮崎 慎一 小林 真也  
愛媛大学 工学部

## 1 はじめに

ネットワークによって接続された複数の計算機が同一の管理下におかれているマルチコンピュータ環境では、ランダムに投入されるユーザのタスクを効率良く処理するために負荷分散を行なう機能の必要性が高い。我々の研究グループでは負荷分散のアルゴリズムとして自律負荷分散方式を提案している。従来の研究により自律負荷分散方式がマルチコンピュータ環境において有効な負荷分散アルゴリズムであることが示されている。今回、自律負荷分散方式に低負荷ノードからの情報発信機能を付加することで、負荷分散の判断に用いる情報の信頼性を高め、性能の改善を実現する方式を提案する。またワークステーションクラスター上で従来の方式との比較評価を行なった。

## 2 自律負荷分散方式

自律負荷分散方式は、タスクを投入されたノードが自律的に他ノードと交渉を行ない、より応答の早いノードに実行を依頼する方式である。交渉は他ノードに依頼を行なうかどうかを決定する依頼判断と、受けた依頼を受諾するか否かを決定する諾否判断からなり、各ノードが個々に管理する局所的な情報を基にしてそれらの判断を行なう。依頼を受諾したノードは同様に交渉を繰り返しタスクを実行するのに最適なノードを探査する。また、交渉を通してノードは各々が持つノード情報を交換する。以下に依頼判断と諾否判断のアルゴリズムについて解説する。

### 依頼判断

タスクの投入のあったノードは依頼判断を行ない、他ノードと交渉を行なうかどうかを決定する。まず自ノードが保持するノード情報を基に自ノードおよび他ノードにおけるタスクの処理時間を推定し、候補ノードのリストを作成する。推定処理時間の最小であるノードから順次依頼を行なう。それが自ノードであった場合は依頼は行なわずに自ノードにおいてタスクを実行する。依頼が拒否された場合には次の候補ノードに対して依頼を繰り返す。依頼を受諾したノードは同様に依頼判断を繰り返す。

### 諾否判断

他ノードからタスクの実行を依頼された場合に諾否判断が行なわれる。ノード情報から自ノードでのタスク処理時間を推定し、これを依頼元ノードの推定処理時間より短かければ受諾する。自ノードの処理時間の方が大きい場合には、依頼元のノードに対して拒否メッセージを送信する。

---

Autonomous Load Distribution Algorithm with Receiver Initiated Informing Mechanism

Shin-ichi MIYAZAKI, Shin-ya KOBAYASHI  
Faculty of Engineering, Ehime University  
{miyazaki, kob}@koblab.cs.ehime-u.ac.jp

## 2.1 自律負荷分散方式の特徴

自律負荷分散方式の特徴としては以下のようなもののが挙げられる。

- ユーザが明示的に分散を指示する必要が無い(負荷分散の透過性)
- 必要な情報は各ノードが局所的に管理する(分散管理型)
- 各ノードはシステムの全てに関する情報を保持する必要が無い(不完全情報型)
- 一对一通信を用いる(非放送型)

これらの特徴により、自律負荷分散方式は大規模なマルチコンピュータ環境においても有効かつ実現性の高い方式であると言える。

## 3 情報発信機能

各ノードが保持する他ノード情報はそれが更新されない限り、時間の経過とともにその信頼性は低下する。効率的な負荷分散を実現するためには、それぞれのノードの所持する情報が新しいものである必要がある。従来の自律負荷分散方式では他ノードと交渉が行なわれない限り、他ノード情報の更新が行なわれなかった。そこで情報発信機能を付加することによって交渉時以外にもノード情報の交換が行なわれるようになり、ノード情報の信頼性を高める方式を提案する。この方式では各ノードが交渉時以外に自ノードが低負荷であるか否かを判断し、低負荷であれば高負荷であると予測されるノードに対して自ノードの情報を送信する。それにより高負荷ノードにおいて負荷の分散先候補について信頼性の高い情報が得られるため、より適切な交渉先の決定が期待できる。

### 3.1 アルゴリズム

各ノードは依頼判断も諾否判断もする必要がなくなった場合に情報発信判断を行なう。情報発信判断を行なうノードはまず保持するノード情報を元に各ノードの負荷状況  $L$  を以下の式により推定する。

$$L = \frac{l}{S}$$

ここではノード上で実行中のタスク数であり、 $S$  はノードの処理能力である。自ノードと他ノードとの負荷状況を比較し、自ノードの負荷状況が十分に小さければ負荷状況の値が最大であるノードに対して情報を発信する。これにより情報を受け取った高負荷ノードはより迅速に適切なノードを選択することが可能になる。また、この情報発信は新たな依頼判断や諾否判断が行なわれない限り、一定の間隔で繰り返し行なうことで、長期に渡って負荷のタスクの投入が行なわれない場合においても情報の有効性を保つ事が可能になる。

## 4 評価

情報発信型の自律負荷分散方式をワークステーションクラスター型並列計算機である FUJITSU AP3000 の各ノード WS 上に実装し平均応答時間を測定した。ノードは処理能力の等しい UNIX ワークステーションであり、それらは AP-Net と呼ばれるメッシュ状のネットワークによって接続されている。今回は 8 台のノード上に実装した。CPU のスケジューリング方式はラウンドロビン (プロセッサ・シェアリング) である。最大交渉回数は 5 回とし、全てのノードは全ての他ノードの情報を保持するものとする。

投入するタスク 1 つ当たりの処理量をタスクサイズと呼び、タスクを無負荷時に単独で実行した場合における実行時間 [sec.] とした。投入するタスクは、タスクサイズが平均 5[sec.] の指数分布に従う空ループプログラムである。またタスクの投入間隔も指数分布に従うものとする。ここで、あるノードに投入するタスクの平均タスクサイズ  $\bar{w}$  と平均到着間隔  $\bar{t}$  の比、

$$\lambda \equiv \frac{\bar{w}}{\bar{t}}$$

をノード負荷と定義する。また、システム全体のノード負荷の平均

$$\Lambda \equiv \frac{1}{N} \sum_{n=1}^{N_t} \lambda$$

をシステム負荷と定義する。ただし、 $N$  はシステム内の全ノード数であり、 $N_t$  ( $1 \leq N_t \leq N$ ) はタスクを投入するノード数 (投入ノード数) である。システム負荷が 1 を越えるとシステムは過負荷状態となるため、システム負荷の値は  $0 < \Lambda \leq 1$  の範囲内にある。今回の評価ではタスクを投入するノードを 1 台に限定した局所的な負荷投入により評価を行なった。また、平均タスクサイズ  $\bar{w}$  を 5[sec.] と固定しているため、タスクの平均到着間隔  $\bar{t}$  を変化させることによってシステム負荷を調節する。情報発信を決定する際には自ノードがノード情報を所持しているノードの内、負荷状態値が低い側の何パーセントの内にあるかを閾値として用いる。この閾値が何パーセントであれば最も有効な負荷分散が行なわれるかを検証した。

### 4.1 応答時間

従来の自律負荷分散方式と今回改良したものとの平均応答時間を図 1 に示す。1 回の実験あたり 10000 個のタスクを投入し、1000 個目から 9000 個目の区間で平均をとった。図 1 において ALD が従来の自律負荷分散方式であり、ALD-I が今回情報発信機能を付加したものである。情報発信型については閾値が 50% である場合、そのノードの負荷状況が推定されたシステム内の負荷状況の平均以下であれば情報発信が実行される。負荷状況が平均以下であれば低負荷ノードであると言えるので今回の実験では閾値を 30% から 60% まで 10% 毎に変化させ、どの値が適当であるかを調べた。M/M/1 は負荷分散を行なわない FCFS での応答時間、M/M/8 はシステムノードが 8 台の場合の理論限界応答時間である。なお、情報発信判断の間隔は通常の負荷分散処理に必要なオーバヘッドよりも十分に大きい 1[sec.] とした。この値についてはいずれ検証する予定である。

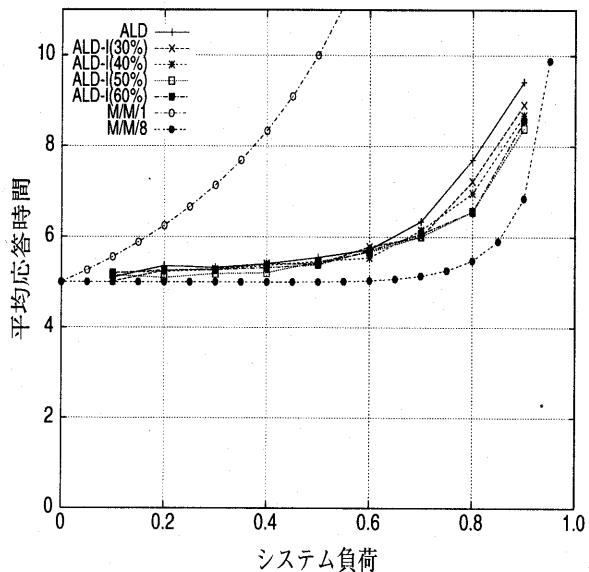


図 1: 平均応答時間

従来型、情報発信型ともにシステム負荷が低い状態ではほとんど差は見られないが、システム負荷が高まるにつれて情報発信型の方式では従来型よりも早い応答時間が得られるようになっている。また、情報発信判断の際の閾値を変化させた場合には、閾値を上げるより緩い条件でも情報発信が行なわれるようになるに従って、より早い応答時間が得られるようになっている。しかし、閾値が 50% の条件で得られる性能は最高となり、60% では 50% の時とほぼ変わらず、システム負荷が 0.9 の段階では逆に悪化してさえいる。これは情報発信機能によるオーバヘッドが性能に影響を及ぼしているためと考えられる。

## 5 まとめ

本稿では情報発信機能の追加による自律負荷分散方式の改善法を新たに提案した。この機能により負荷分散の交渉時以外にも負荷状態が一定の閾値以下である低負荷ノードから高負荷なノードに対してノード情報を発信することでノード情報の信頼性が高まる。実装システムを用いてこの方式を従来のものと比較評価を行ない、応答時間が改善される事を示した。また、閾値としては 50% が最も適当であることが分かった。

今後は情報の発信間隔を変化させた場合や、負荷の投入ノードが複数に渡る場合についての評価を行なう予定である。

## 参考文献

- [1] 小林真也、小川智之、渡辺尚、"マルチコンピュータシステムにおける自律負荷分散方式", 電子情報通信学会論文誌 D-I, VOL.J79-D-I NO.11(pp.903-915), Nov. 1996
- [2] 田中義幸、小林真也、"分散協調にもとづく自律負荷分散方式の実装と評価", 情報処理学会 マルチメディア・分散・協調とモバイルワークショップ論文集, (7)(pp.37-42), July 1997