

# 確率的コーラムシステムに基づく

6J-07

## 負荷分散アルゴリズムとその実験的評価

吉村 英明

角川 裕次

阿江 忠

広島大学工学部

### 1 はじめに

近年のコンピュータの形態は、複数のコンピュータがネットワークを介して相互に接続されている分散システムへと変化を遂げて来ているが、このようなシステムの処理能力を十分に活用するには、それに適した計算機資源の管理が求められる。そのうちの重要な問題の一つである負荷分散は、分散システム内のコンピュータに負荷(仕事)を分散させて処理させることによって、システムの処理効率を向上させるものである。

本稿では分散システムにおける通信構造として、コートリ [2] の拡張である確率的コーラムシステム (Probabilistic Quorum System: 以下 PQS) [1] を用いて、システムの状態に影響を受けない安定した、比較的通信量の少ない負荷分散アルゴリズムを提案し、その効果をシミュレーションにより評価した結果について述べる。

### 2 確率的コーラムシステム

これまでの研究で分散システム上における様々な問題を、通信構造にコートリを用いて解決したものが示されている。ここで用いられているコートリは以下のように構築される。

**定義 1** (コートリ [2])

$U$  を大きさ  $n$  の集合とする。以下の二つの条件を満たす空でない  $U$  の部分集合  $Q_i$  の集合  $\{Q_1, Q_2, \dots, Q_m\}$  をコートリ (Coterie) と呼ぶ。

- 全ての  $i$  と  $j$  ( $1 \leq i, j \leq m$ ) に対して,  $Q_i \cap Q_j \neq \phi$ .
- 全ての  $i$  と  $j$  ( $1 \leq i, j \leq m, i \neq j$ ) に対して,  $Q_i \not\subseteq Q_j$ .

このとき、各  $Q_i$  をコーラム (quorum) と呼ぶ。 □

上記で定義されたコートリを以下のように拡張する。

**定義 2** (確率的コーラムシステム [1])

$U$  を大きさ  $n$  の集合、 $Q$  を  $U$  の部分集合の集合としたとき  $Q$  へのアクセス戦術  $w$  は、 $Q$  の構成要素への確率の分布を表し、

$$\sum_{Q \in \mathcal{Q}} w(Q) = 1$$

を満たす。

ここである定数  $\epsilon$  ( $0 < \epsilon < 1$ ) に対して、

$$\sum_{Q, Q': (Q \cap Q') \neq \phi} w(Q)w(Q') \geq 1 - \epsilon$$

が成立するとき  $\langle Q, w, \epsilon \rangle$  を確率的コーラムシステム (Probabilistic Quorum System) という。 □

本稿では、このように定義された PQS を用いる。

A load balancing algorithm based on probabilistic quorum system and its experimental evaluation  
Hideaki Yosimura, Hirotsugu Kakugawa, Tadashi Ae  
Faculty of Engineering, Hiroshima University

### 3 PQS を用いた負荷情報の通信

2章で定義された PQS  $\mathcal{Q} = \{Q_1, Q_2, \dots\}$  を用いて、以下のような手順にしたがって負荷情報を通信する。

- 各計算機  $p_i$  は、自身の負荷に変化が生じる度に  $Q_1, Q_2, \dots$  の中からランダムにあるコーラム  $Q_i$  を選び、そのコーラムに含まれる全ての計算機に自身の負荷情報を送信する。
- 他の計算機群の負荷情報を知りたい計算機  $p_j$  はあるコーラム  $Q_j$  を選び、そのコーラムに含まれる全ての計算機からシステム全体の負荷情報を受信する。
- このとき  $Q_i$  と  $Q_j$  が共通部分を持つならば、 $Q_i \cap Q_j$  の計算機を経由して  $p_i$  の情報が  $p_j$  に届く。
- $Q_i$  と  $Q_j$  が共通部分を持たないならば、その時点では  $p_i$  の情報は  $p_j$  には届かない。

図 1 と図 2 は、ランダムに選んだコーラムが共通部分を持つ場合と持たない場合の情報通信について図示したものである。

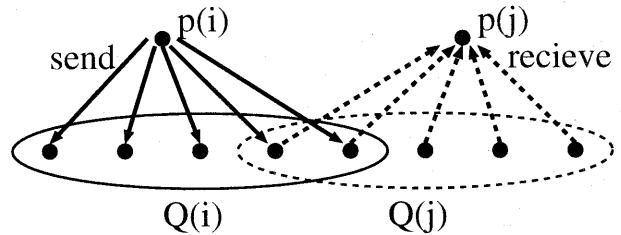


図 1: 共通部分を持つ場合

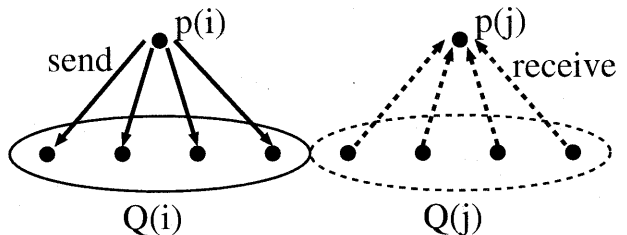


図 2: 共通部分を持たない場合

PQS を用いた通信構造では、コーラムの大きさ (全て一定) を変更することで通信量を変更することが可能である。しかし、コーラムの大きさを小さくした場合にはランダムに選ばれた 2 つのコーラムが共通部分を持つ確率も減少する。すなわち、入手できるシステム全体の負荷情報の正確性が損なわれることになる。

図3は、コーラムの大きさをシステム全体の大きさに対してある割合に設定したものがどのような確率で共通部分を持つかを表したグラフである。

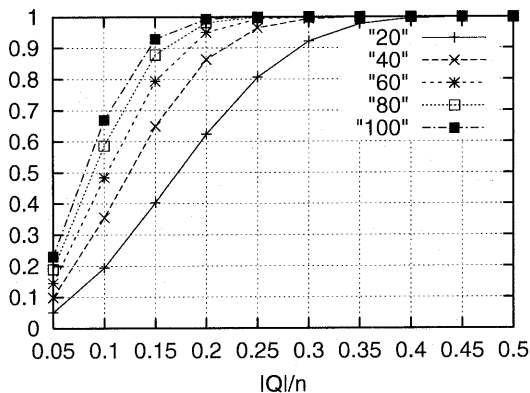


図3: コーラム同士が共通部分を持つ確率

本稿では、システム全体の正確な負荷状況を入力するために多大な通信量を払うことを避け、情報の正確性を多少犠牲にすることで、通信コストを下げながら十分な負荷分散が可能となるのではないかと考えた。

通信構造に PQS を用いた負荷分散の長所として、通信量の削減の他にまず計算機が負荷状態に変化が生じている度に通信を行っているため、各計算機が比較的新しいシステム全体の負荷状態を保持できるという点がある。また PQS の構築方法は、ある決められた個数の計算機をランダムに選ばば良いので、動的なネットワークへの対応が容易であるという点も挙げられる。

#### 4 実験結果と考察

通信構造として PQS を用いたものが、どの程度の効力を持っているかをシミュレーションを行って検証した。実験では、まずシステムの大きさに対してコーラムの大きさをどの程度に設定したものが、より効率的な負荷分散を行えるのかを検証した。

次に、この検証の結果を受けていくつかのシステムの大きさに対応した PQS を構築し、既存の負荷分散アルゴリズムと比較した。

今回は実験を行う環境を以下のように仮定した。

- 各計算機・計算機間の通信経路は故障しない。
- 各計算機の処理能力は全て等しい。
- 各計算機は局所時計を持つ。ここで各計算機の局所時計は、自身の負荷状態に変化が生じたとき、1増加するものとする。
- 各計算機間の通信は1対1である。
- 各計算機間の通信遅延時間は全て等しい。ここで通信遅延時間とは、ある計算機からメッセージが送信されてから受け手に受信されるまでの時間である。

このような仮定においてシミュレーションを行った結果、PQS を用いた負荷分散は既存の負荷分散アルゴリズムと比べて比較的良い評価を得ることを実証することができた。

図4は、仕事分散される毎の各計算機間の負荷の差の最大値を検出し、その平均値を検出した結果を示したグラフである。

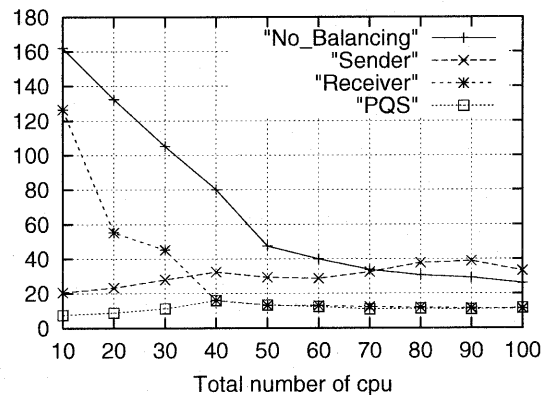


図4: <負荷の差の平均値>の実行結果

#### 5 おわりに

本稿では、分散システムにおいて負荷分散を行う際に必要な負荷情報の通信に確率的コーラムシステムを用いた新たな負荷分散アルゴリズムを提案し、その有効性をシミュレーションによって示した。

今後の課題としては、各計算機の処理能力に差が生じている場合に同一の仕事をもつ全ての計算機に処理させて各計算機の処理能力の差を検出することや各計算機が故障する可能性を持つ場合には通信の際にあるデータを付随させることで故障した計算機を検出するというような、実際の分散システムに対応したアルゴリズムを考えていくことが考えられる。

謝辞 本研究の一部は、文部省科学研究費補助金(奨励研究A, 課題番号11780229)の援助を受けている。

#### 参考文献

- [1] Malkhi, D., Reiter, M., Wright, R.: Probabilistic Quorum Systems, in *Proceedings of the 16th International Conference on Distributed Computing Systems*, (1997).
- [2] Barbara, D. and Garcia-Molina, H.: Mutual Exclusion in Partitioned Distributed Systems, *Distributed Computing*, Vol. 1, pp. 119-132 (1986).
- [3] Nakajima, A.: Fault-Tolerant Distributed Match-Making with Any Resiliency, *IEICE Transaction*, Vol. E 74, No. 2, pp. 427-434 (1991).
- [4] Ibaraki, T. and Kameda, T.: Theory of Coteries: Mutual Exclusion in Distributed Systems, *IEEE Transaction on Parallel and Distributed Systems*, Vol. 4, No. 7, pp. 770-794, (1993).
- [5] Maekawa, M.: A  $\sqrt{n}$  Algorithm for Mutual Exclusion in Decentralized Systems, *ACM Transactions on Computer Systems*, Vol. 3, No. 2, pp. 145-159 (1985).