Bus型コンピュータネットワークにおける複数ジョブクラスの負荷分散

7H-6

金 宗根　　亀田 壽夫
(電気通信大学情報工学科)

1 はじめに

ワークステーション等のコンピュータ (以下ノードと呼ぶ) をbus型通信回線を通じて接続させた、分散型コンピュータシステムが考えられる。このシステムにおいて、各ジョブに対する応答時間の期待値を最小にする負荷分散法が重要となる。

Tantawi と Towsley[1] は、ネットワークの各ノードへのジョブの到着率、各ノードの処理能力などの諸パラメータが与えられ、また通信による遅延が発信元、着信先の違いによらないという仮定 (この仮定はLANや衛星通信等の場合に成り立つ) の下で単一ジョブクラスの全体最適化方式の解を示した。

Kameda と Hazeyama[2] は Tantawi らのモデルと同じ枠組において、複数ジョブクラスの場合の全体最適化方式の解を示した。さらに、各ジョブにとって各々の応答時間の期待値が最小になる個別最適化方式についての解を示した。

本研究の主な目的は複数ジョブクラスの場合の全体最適化方式及び個別最適化方式での最適負荷分散問題を解決する最適アルゴリズムの提案である。また数値実験を行うことにより二つの最適化方式での応答時間、最適負荷などの振舞いを比較検討する。

2 モデルの記述

分散コンピュータシステムのモデルを考える (図1)。このモデルではn個のノードが Bus形通信ネットワークに任意の形態で接続されており各ノードでは複数クラスのジョブが処理できる。モデルで使用される記号等を次に示す。

- n: ノード数
- m: クラス数
- $\phi_i^{(k)}$: ノード i におけるクラス k ジョブの到着率
- $\beta_i^{(k)}$: ノード i におけるクラス k ジョブの処理率
- $x_{ij}^{(k)}$: ノードiからノードjへ転送されるクラスkジョブの割合
- $\lambda^{(k)}$: 通信線におけるクラス k ジョブの通信量
- $\phi^{(k)}$: クラスkジョブのネットワーク全体への到着率
- Φ: ネットワーク全体の到着率
- $\boldsymbol{\phi}_i$: $[\phi_i^{(1)}, \phi_i^{(2)}, \ldots, \phi_i^{(m)}]$
- $\boldsymbol{\beta}$: $[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n]$
- $\boldsymbol{\beta}_i$: $[\beta_i^{(1)}, \beta_i^{(2)}, \ldots, \beta_i^{(m)}]$
- $\boldsymbol{\lambda}$: $[\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(m)}]$
- $F_i^{(k)}(\boldsymbol{\beta}_i)$: 到着ノード$i$で処理を受ける場合の平均遅延 (微分可能で凸形の増加関数)
- $F_j^{(k)}(\boldsymbol{\beta}_j) + G^{(k)}(\boldsymbol{\lambda})$: 他のノード$j$へ伝送され 処理を受けるクラスkジョブの応答時間
- $G^{(k)}(\boldsymbol{\lambda})$: 通信によるクラス$k$の平均遅延 (発信元、着信先の違いによらない微分可能で 凸形の$\boldsymbol{\lambda}$の非減少関数)

各ノードiへのクラス k ジョブの到着はポアソン過程により到着する。ノードiに到着するジョブはノードiで処理されるか、あるいは、他のノードjに転送されて処理されるかになる。一つのクラスkジョブが一旦他のノードに伝送された場合、さらに別のノードに転送されることはないとする。

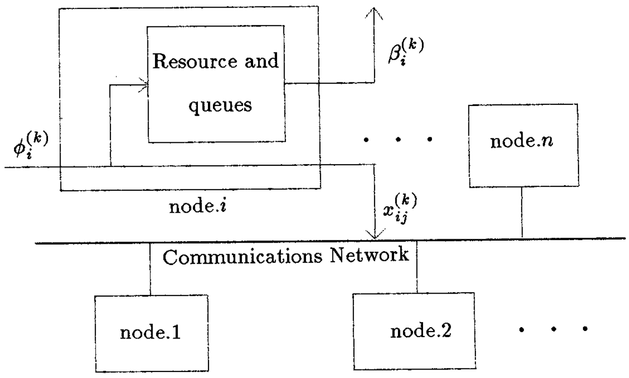

図1.Bus型コンピュータネットワーク

システム全体の平均応答時間 ($D(\boldsymbol{\beta})$) は平均のノード遅延と平均の通信遅延の和として表される。

$$D(\boldsymbol{\beta}) = \frac{1}{\Phi}\sum_{k=1}^{m}[\sum_{i=1}^{n}\beta_i^{(k)}F_i^{(k)}(\boldsymbol{\beta}_i) + \lambda^{(k)}G^{(k)}(\boldsymbol{\lambda})]. \quad (1)$$

ただし

$$\sum_{i=1}^{n}\beta_i^{(k)} = \phi^{(k)}, \quad k = 1, 2, \ldots, m,$$

$$\beta_i^{(k)} \geq 0, \quad i = 1, 2, \ldots, n, k = 1, 2, \ldots, m.$$

各ノードをジョブを処理する性質により各クラスごとに次のように呼ぶ。(1)source($R^{(k)}$): 到着したジョブの全部あるいは一部を他ノードに送る。(2)neutral($N^{(k)}$): 到着したジョブを全部自分で処理する。(3)sink($S^{(k)}$): 到着したジョブと他ノードから送ってきたジョブを処理する。

3 全体最適化方式と個別最適化方式の解

全体最適化方式に置ける最適解を表現するためにノードiのクラスkの限界ノード遅延 $f_i^{(k)}(\boldsymbol{\beta}_i)$ とクラスkの限界通信遅延 $g^{(k)}(\boldsymbol{\lambda})$ 定義を 与える。

$$f_i^{(k)}(\boldsymbol{\beta}_i) = \frac{\partial}{\partial \beta_i^{(k)}}\sum_{l=1}^{m}\beta_i^{(l)}F_i^{(l)}(\boldsymbol{\beta}_i), \quad (2)$$

$$g^{(k)}(\boldsymbol{\lambda}) = \frac{\partial}{\partial \lambda^{(k)}}\sum_{l=1}^{m}\lambda^{(l)}G^{(l)}(\boldsymbol{\lambda}). \quad (3)$$

Load balancing of multi-class jobs in bus computer networks.
Chonggun KIM and Hisao KAMEDA
The University of Electro-Communications

[定理 1] システム全体の平均応答時間を最小にする全体最適化の解は $k=1,2,..,m$ に対し次の関係を満たす。

$$\begin{array}{lll} f_i^{(k)}(\boldsymbol{\phi}_i) > \alpha^{(k)} + g^{(k)}(\boldsymbol{\lambda}), & 0 \le \beta_i^{(k)} < \phi_i^{(k)} & (i \in R^{(k)}),\\ \alpha^{(k)} \le f_i^{(k)}(\boldsymbol{\phi}_i) \le \alpha^{(k)} + g^{(k)}(\boldsymbol{\lambda}), & \beta_i^{(k)} = \phi_i^{(k)} & (i \in N^{(k)}),\\ f_i^{(k)}(\boldsymbol{\phi}_i) < \alpha^{(k)}, & \beta_i^{(k)} > \phi_i^{(k)} & (i \in S^{(k)}). \end{array}$$

ただし、次の制約条件が満たさなければならない。

$$\lambda_{sen}^{(k)} = \lambda_{rec}^{(k)},$$

$$\lambda_{sen}^{(k)} = \sum_{i \in R^{(k)}} \min\{\phi_i^{(k)}, \phi_i^{(k)} - f_i^{(k)^{-1}}(\boldsymbol{\beta}_i|_{f_i^{(k)}(\beta_i) > \alpha^{(k)} + g^{(k)}(\lambda)})\},$$

$$\lambda_{rec}^{(k)} = \sum_{i \in S^{(k)}} (f_i^{(k)^{-1}}(\alpha^{(k)}) - \phi_i^{(k)}).$$

ここで、$\alpha^{(k)} = \min_i f_i^{(k)}(\boldsymbol{\beta}_i)$である。

証明は[4] 参照。

個別最適化方式に置ける最適解を表現するためにノードiのクラスkのノード遅延 $F_i^{(k)}(\boldsymbol{\beta}_i)$ とクラスkの通信遅延 $G^{(k)}(\boldsymbol{\lambda})$ を使用する。

[定理 2] 各ノードのクラスkの平均応答時間を最小にする問題の解は $k=1,2,..,m$ に対し次の関係を満たす。

$$\begin{array}{lll} F_i^{(k)}(\boldsymbol{\phi}_i) > R^{(k)} + G^{(k)}(\boldsymbol{\lambda}), & 0 \le \beta_i^{(k)} < \phi_i^{(k)} & (i \in R^{(k)}),\\ R^{(k)} \le F_i^{(k)}(\boldsymbol{\phi}_i) \le R^{(k)} + G^{(k)}(\boldsymbol{\lambda}), & \beta_i^{(k)} = \phi_i^{(k)} & (i \in N^{(k)}),\\ F_i^{(k)}(\boldsymbol{\phi}_i) < R^{(k)}, & \beta_i^{(k)} > \phi_i^{(k)} & (i \in S^{(k)}). \end{array}$$

ただし、次の制約条件が満たさなければならない。

$$\lambda_{sen}^{(k)} = \lambda_{rec}^{(k)},$$

$$\lambda_{sen}^{(k)} = \sum_{i \in R^{(k)}} \min\{\phi_i^{(k)}, \phi_i^{(k)} - F_i^{(k)^{-1}}(\boldsymbol{\beta}_i|_{F_i^{(k)}(\beta_i) > R^{(k)} + G^{(k)}(\lambda)})\},$$

$$\lambda_{rec}^{(k)} = \sum_{i \in S^{(k)}} (F_i^{(k)^{-1}}(R^{(k)}) - \phi_i^{(k)}).$$

ここで、$R^{(k)} = \min_i F_i^{(k)}(\boldsymbol{\beta}_i)$である。

証明は[4] 参照。

上記の解に基づく両方式の詳細アルゴリズムは[4] で与える。

4 数値実験

本研究で提案したアルゴリズムにより全体最適化方式による負荷分散と個別最適化方式による負荷分散の振舞がどの様になるかを調べるためにシステムモデルの具体例を与えて調べる。ノードモデルとしてはプロセッサシェアリングモデルを考える。通信ネットワークモデルとしては M/G/∞ を考える。図 2. から図 4. に各最適化方式で通信所要時間を変化させた場合の実験結果を示す。

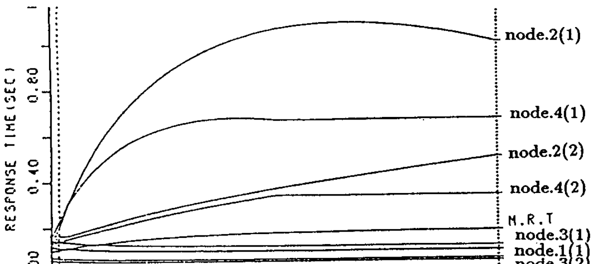

図 2. 全体最適化方式

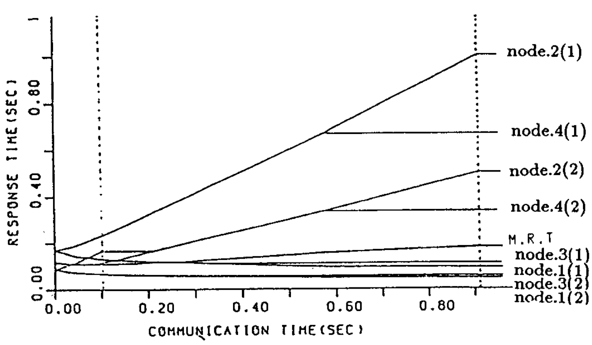

図 3. 個別最適化方式

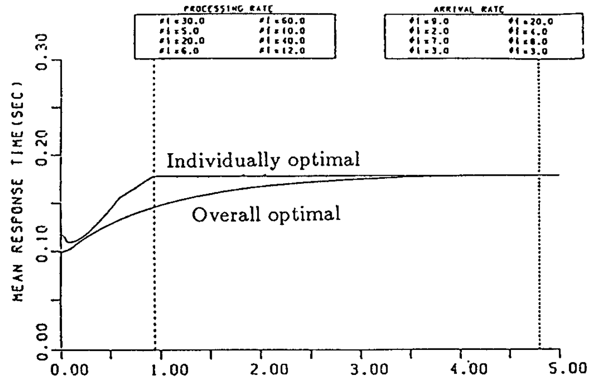

図 4. 各々の最適化方式での応答時間

5 まとめ

本研究では複数クラスジョブの場合の両方式の解を示した。これは Tantawi ら [1] の単一クラスについて示した解より簡単化されている。この解に基づいて全体最適化方式と個別最適化方式で決定される負荷分散を比較検討するため、各々の負荷分散アルゴリズムを求め、このアルゴリズムによる数値実験を行った。数値実験の結果、単一クラスの場合[3] と同じいくつかの性質と異常現象が起こるのを観測した。また、複数クラスの場合のみに起こるクラス間の不平等のいくつかの異常現象が見出された。

参考文献

[1] A.N.Tantawi and D.Towsley. Optimal static load balancing in distributed computer systems, *J.ACM* 32,2(April), pp.445-465 (1985).

[2] H.Kameda and A.Hazeyama. Individual vs. Overall optimization for static load balancing in distributed computer systems, Computer Science Report, The University of Electro-Communications (1988).

[3] 亀田 壽夫，櫨山 淳雄．コンピュータネットワークに置ける負荷分散について。 情報処理学会オペレーティングシステム研究会研究報告, 34-7(1987).

[4] 金 宗根，亀田 壽夫．分散コンピュータシステムにおける複数クラスジョブの負荷分散 。 情報処理学会オペレーティングシステム研究会研究報告, 42-4(1989).