

スター型分散コンピュータシステムにおける負荷分散

亀田 壽夫 , 張 勇兵
電気通信大学情報工学科

TantawiとTowsleyはスター型コンピュータネットワークにおける単一ジョブクラスの場合の静的負荷分散方式を考え、システム全体の平均応答時間を最小とするよう各ノードの負荷を決定する最適解を求めた。ここで、Tantawiらの方式を全体最適化方式と呼ぶ。本研究では、まずそれを複数クラスの場合へ拡張した。次に、ノードに到着する各ジョブがその応答時間の期待値を最小とするようにどこで処理を受けるかが決められる個別最適化方式を考え、そして、個別最適化方式による負荷分散の振舞いを調べ、全体最適化方式のそれと比較をした。それによると、個別最適化方式の場合、通信所要時間が増加するにもかかわらず、システム全体の平均応答時間が減少することがあるという一種の異常現象を発見した。

LOAD BALANCING IN STAR NETWORK CONFIGURATIONS

Hisao KAMEDA and Yong-Bing ZHANG

The University of Electro-Communications,
1-5-1 Chofugaoka, Chofu-shi, Tokyo 182, Japan

Tantawi and Towsley considered static load balancing for the model of a star network configuration wherein small computer systems are connected to a large central computer. They obtained an optimal strategy which minimizes the system-wide mean response time for the case of one job class. We call it the overall optimal policy. We first extend the results to the case of multiple job classes.

Furthermore, we consider an individually optimal policy whereby jobs are scheduled so that every job may feel that its own expected response time is minimized. We show that the individually optimal policy can be achievable, and obtained the characteristics of the policy. We compare both the policies, and observe that the individually optimal policy may show an anomalous behavior.

1. はじめに

パーソナルコンピュータやワークステーション等の小型コンピュータは、単独では、処理能力に限られ、他のシステムとのデータ共用等ができない。その解決法の1つとして、処理能力が十分に大きな計算機を中央に設置し（中央ノードと呼ぶ）、通信回線を通じて、多くの小型コンピュータ（衛星ノードと呼ぶ）を接続する方式がある。これはスター型ネットワークと呼ばれる。それによって、各ユーザからは、あたかも、身近にあるコンピュータの性能が向上され、アクセス能力が増えるように見える。ここで、各ジョブに対する応答時間の期待値を最小にするように衛星ノードと中央ノードに負荷を分散する方法が重要となる。

各衛星ノードに到着したジョブをその衛星ノードで処理させるか中央ノードに送るかを決定する負荷分散方式には静的方式と動的方式とがある。静的負荷分散はシステム全体の統計的な値に関する情報のみを使用し、（各ノードの待ち行列長などの）システムの現在の状態によらない方式であり、動的負荷分散はシステムの現在の状態に依存する方式である。刻一刻変化する各ノードの状態を別のノードが把握し続けるのは困難であるので、動的負荷分散の実現には多くの情報通信量を要する点で困難さがあると思われる。したがって、本論文では、静的負荷分散方式を追究する。

各ジョブに対する応答時間の期待値を最小にする方式には次の二通りが考えられる。一つはシステム全体にわたる平均応答時間を最小にしようとするものであり、これを**全体最適化方式**と呼ぶ。この方式によると、衛星ノードに到着したジョブのうち、局所的に衛星ノードで処理されるジョブに対する応答時間の期待値と、中央ノードで処理されるジョブに対する応答時間の期待値が必ずしも等しくなく、各ジョブに不公平感が残ることがある。もう一つは各ジョブの応答時間の期待値がもうこれ以上改善できないという意味で最適にしようとするものであり、これを**個別最適化方式**と呼ぶ。この方式によると、システム全体の平均応答時間は必ずしも最小ではないが、一つの衛星ノードに到着したジョブはどこで処理されるにしても同じ応答時間を受けるという点で公平である。本論文はこの二つの方式を比較することを目的とする。

Tantawiら [1]は、各衛星ノードや中央ノードへのジョブ到着率、各ノードの処理能力、通信路容量などの諸パラメータが与えられた単一ジョブクラスのスター型ネットワークモデルを考え、全体最適化方式を実現した。すなわち、システム全体の平均応答時間を最小にする最適解を求め、更にシステムパラメータによる影響を調べた。本研究では、まず、Tantawiらのモデルを複数ジョブ

クラスの場合に拡張し全体最適化方式の解を求める。さらに同等のスター型ネットワークモデルにおいて、個別最適化方式を実現する解を求める。すなわち、各ノードに到着する各ジョブの応答時間の期待値を最小にする負荷分散の解を求め、それを全体最適化方式の解と比較する。

2. モデルの記述

ネットワークは、 n 異機種衛星ノードと中央ノードからなり、各衛星ノードは通信線により、中央ノードに連結されている。各ノードのコンピュータはCPU、I/O装置などの資源からなる。

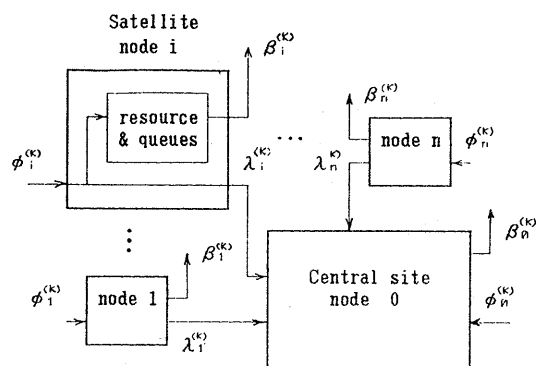


図1. クラス k ジョブの流れ

衛星ノードに1, 2, ..., n と番号を付け、中央ノードは0とする。ジョブは1, 2, ..., m のクラスに分類される。各ノード i へのクラス k のジョブ到着はポアソン分布とし、ジョブの到着率は $\phi_i^{(k)}$ とする。クラス k のジョブの全体の外部到着率は $\Phi^{(k)}$ で表す($\Phi^{(k)} = \phi_0^{(k)} + \phi_1^{(k)} + \phi_2^{(k)} + \dots + \phi_n^{(k)}$)。さらに $\phi_i = [\phi_i^{(1)}, \phi_i^{(2)}, \dots, \phi_i^{(m)}]$ とする。衛星ノード i に到着するジョブは衛星ノード i で処理されるか、または、中央ノードに転送されて処理されるとする。

ジョブが中央ノードに転送される際、及び中央ノードから返送する際に、遅延が生ずる。1つのジョブが一旦中央ノードに転送された場合、さらに別のノードに転送されることはないとする。（但し $\phi_0^{(k)} = 0$ とする。 $\phi_0^{(k)} > 0$ への拡張は容易である。）

ノード i で処理されるクラス k ジョブの割合を $\beta_i^{(k)}$ で表し、衛星ノード i と中央ノードとを結ぶ通信路を流れるクラス k ジョブのトラフィック量を $\lambda_i^{(k)}$ ($\lambda_i^{(k)} = \phi_i^{(k)} - \beta_i^{(k)}$)で表す。また、各ノードにおけるクラス k ジョブに対する遅延の期待値を $F_i^{(k)}(\beta_i)$ と定義し、各 $\beta_i^{(k)}$ に関して、微分可能で、凸型の増加関数とする。

クラスkジョブに対する通信による遅延の期待値を $G_i^{(k)}(\lambda_i)$ と表し、微分可能で、凸型の非減少関数とする。これは、中央ノードに転送するための遅延と返送の遅延とを合わせた時間の期待値を表す。

システム全体の平均応答時間 $D(\beta)$, $\beta = [\beta_0, \beta_1, \dots, \beta_n]$, $\beta_i = [\beta_i^{(1)}, \beta_i^{(2)}, \dots, \beta_i^{(m)}]$ は、各ノードによる平均のノード遅延と平均の通信遅延の和として表され、次のような形になる。

$$D(\beta) = \frac{1}{\Phi} \sum_{k=1}^m \sum_{i=0}^n \beta_i^{(k)} F_i^{(k)}(\beta_i) + \sum_{i=1}^n (\phi_i^{(k)} - \beta_i^{(k)}) G_i^{(k)}(\phi_i - \beta_i) \quad (2.1)$$

ただし

$$\begin{aligned} \sum_{i=0}^n \beta_i^{(k)} &= \Phi^{(k)}, & k=1, 2, \dots, m. \\ \beta_i^{(k)} &\leq \phi_i^{(k)}, & i=1, 2, \dots, n, k=1, 2, \dots, m. \\ \beta_i^{(k)} &\geq 0, & i=0, 1, \dots, n, k=1, 2, \dots, m. \end{aligned}$$

でなければならない。

衛星ノードは各クラスごとに3種類に分類される。

(1). クラスk idleノード(Rd^(k)): このノードでは全くクラスkジョブを処理せず、全てのクラスkジョブを中央ノードに転送する。

(2). クラスk activeノード(Ra^(k)): このノードでも一部のクラスkジョブを処理するが、残りのクラスkジョブを中央ノードに送る。

(3). クラスk neutralノード(N^(k)): クラスkジョブを中央ノードに送らず、すべてこのノードで処理する。

3. 負荷分散方式

1). 全体最適化方式

Tantawiらによる単一クラスの場合が次のように拡張される。

全体最適化方式における最適解を表現するために、下式の関数を使用する。

$$h_i^{(k)}(\beta_i) = f_i^{(k)}(\beta_i) - g_i^{(k)}(\phi_i - \beta_i) \quad (3.1)$$

ただし、 $f_i^{(k)}(\beta_i) = \partial(\sum_{k=1}^m \beta_i^{(k)} F_i^{(k)}(\beta_i)) / \partial \beta_i^{(k)}$,

$$g_i^{(k)}(\lambda_i) = \partial(\sum_{k=1}^m \lambda_i^{(k)} G_i^{(k)}(\lambda_i)) / \partial \lambda_i^{(k)}$$

である。

$f_i^{(k)}(\beta_i)$ は、衛星ノードiに滞在する平均ジョブ数 $\sum_{k=1}^m \beta_i^{(k)} F_i^{(k)}(\beta_i)$ が衛星ノードiでのクラスkジョブ処理率 $\beta_i^{(k)}$ に対する導関数であり、 $g_i^{(k)}(\lambda_i)$ は、通信線上にある平均ジョブ数 $\sum_{k=1}^m \lambda_i^{(k)} G_i^{(k)}(\lambda_i)$ が通信線i上のトラフィック量 $\lambda_i^{(k)}$ に対する導関数である。

$h_i^{(k)}(\beta_i)$ は上記の二導関数の差を表す。さらに任意のベクトル $\underline{x} = [x_1, x_2, \dots, x_l]$ に対して $\underline{x} |_{x_i=0} = [x_1, x_2, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_l]$ と定義する。

平均応答時間を最小にする(2.1)式中の各衛星ノードiで処理されるクラスkジョブの割合 $\beta_i^{(k)}$ は次の問題を解くことにより求められる。

$$D(\beta) = \frac{1}{\Phi} \sum_{k=1}^m \sum_{i=0}^n \beta_i^{(k)} F_i^{(k)}(\beta_i) + \sum_{i=0}^n (\phi_i^{(k)} - \beta_i^{(k)}) G_i^{(k)}(\phi_i - \beta_i) \quad (3.2)$$

を最小にする。ただし、制約条件は

$$\begin{aligned} \sum_{i=0}^n \beta_i^{(k)} &= \Phi^{(k)}, & k=1, 2, \dots, m. \\ \beta_i^{(k)} &\leq \phi_i^{(k)}, & i=1, 2, \dots, n, k=1, 2, \dots, m. \\ \beta_i^{(k)} &\geq 0, & i=0, 1, \dots, n, k=1, 2, \dots, m. \end{aligned}$$

[定理1] システム全体の平均応答時間で最小にする問題(3.2)の最適解は $k=1, 2, \dots, m$ に対し次の関係を満たす。

$$h_i^{(k)}(\beta_i) \geq \alpha^{(k)}, \beta_i^{(k)} = 0, \quad (i \in \text{Rd}^{(k)}), \quad (3.3a)$$

$$h_i^{(k)}(\beta_i) = \alpha^{(k)}, 0 < \beta_i^{(k)} < \phi_i^{(k)}, \quad (i \in \text{Ra}^{(k)}), \quad (3.3b)$$

$$h_i^{(k)}(\beta_i) \leq \alpha^{(k)}, \beta_i^{(k)} = \phi_i^{(k)}, \quad (i \in \text{N}^{(k)}), \quad i=1, 2, \dots, n, \quad (3.3c)$$

$$f_0^{(k)}(\beta_0) = \alpha^{(k)}. \quad (3.4)$$

ただし、次の制約条件が満たされなければならない。

$$\begin{aligned} f_0^{(k)-1}(\beta_0 |_{\beta_0^{(k)} = \alpha^{(k)}}) + \sum_{i \in \text{Ra}^{(k)}} h_i^{(k)-1}(\beta_i |_{\beta_i^{(k)} = \alpha^{(k)}}) \\ + \sum_{i \in \text{N}^{(k)}} \phi_i^{(k)} = \Phi^{(k)}. \quad (3.5) \end{aligned}$$

[証明] 付録A参照

2). 個別最適化方式

これは本研究で新たに求めたものである。各衛星ノードに到着する各ジョブはその応答時間の期待値が最小となるノードで処理を受けるように、意志決定を行う。ここで、全体最適化の解に対応するため、クラスkジョブに対する各衛星ノードにおける遅延と通信路における遅延の差を関数 $H_i^{(k)}(\beta_i)$ で表し、次式で示す。

$$H_i^{(k)}(\beta_i) = F_i^{(k)}(\beta_i) - G_i^{(k)}(\phi_i - \beta_i) \quad (3.6)$$

個別最適化方式においては、衛星ノードiの $H_i^{(k)}(\beta_i)$ が $F_0^{(k)}(\beta_0)$ より大きい場合は、クラスkジョブは全て中央ノードに送られ処理され、 $\beta_i^{(k)} = 0$ となるはずである。逆の場合は衛星ノードで処理され、 $\beta_i^{(k)} = \phi_i^{(k)}$ となるはずである。衛星ノードiの $H_i^{(k)}(\beta_i)$ と中央ノードの $F_0^{(k)}(\beta_0)$ とが等しい場合は、クラスkジョブはどちらで処理を受けてもよいので、 $0 \leq \beta_i^{(k)} \leq$

$\phi_i^{(k)}$ となるはずである。

このような条件を満たす β が存在するかどうかは必ずしも自明ではない。これに対してわれわれは次の定理を得る。すなわち、上述の β は存在することが得られる。われわれはこのような β を均衡解と呼ぶ。

[定理2] 個別最適化方式に対し均衡解が存在する。その均衡状態において、全ての衛星ノードは各 k について $Rd^{(k)}, Ra^{(k)}, N^{(k)}$ の3種類に分類され、 $k=1, 2, \dots, m$ について次の関係が成り立つ。

$$H_i^{(k)}(\beta_i) \geq R^{(k)}, \beta_i^{(k)} = 0 \quad (i \in Rd^{(k)}), \quad (3.7a)$$

$$H_i^{(k)}(\beta_i) = R^{(k)}, 0 < \beta_i^{(k)} < \phi_i^{(k)} \quad (i \in Ra^{(k)}), \quad (3.7b)$$

$$R^{(k)} \geq H_i^{(k)}(\beta_i), \beta_i^{(k)} = \phi_i^{(k)} \quad (i \in N^{(k)}), \quad (3.7c)$$

$$F_0^{(k)}(\beta_i) = R^{(k)}. \quad (3.8)$$

ただし次の制約条件を満たさなければならない。

$$F_0^{(k)-1}(\beta_0 |_{\beta_0 = R^{(k)}}) + \sum_{i \in Ra^{(k)}} H_i^{(k)-1}(\beta_i |_{\beta_i = R^{(k)}}) + \sum_{i \in N} \phi_i^{(k)} = \Phi^{(k)}. \quad (3.9)$$

[証明] 付録B参照。

4. 中央ノード及び通信線パラメータによる影響

前節では、負荷分散に関する解の条件が得られ、全ての衛星ノードが各クラス k について $Rd^{(k)}, Ra^{(k)}, N^{(k)}$ の三つに分けられる。一旦、衛星ノードの分割が決まると、 $R^{(k)}, \beta_i^{(k)}$ が得られる。(3.3), (3.4), (3.7), (3.8)式から分かるように、個別最適化方式が全体最適化方式に似ている形をしている。そのため、全体最適化方式のパラメータを少し調整することによって、個別最適化方式の場合に適應することができ、全体最適化方式との比較が簡単にできる。

この節では、個別最適化方式において、衛星ノードの分割におけるパラメータ t_i ($i=1, 2, \dots, n$), u_i, ϕ_i ($i=0, 1, \dots, n$)による影響を調べ、全体最適化方式の場合と比較検討する。ただし複数クラスの場合は非常に複雑になるので、単一クラス ($k=1$)の場合を示す。従って以下では上記の関係から (k) がすべて取り除いたものを用いる。 t_i はノード i と中央ノードとの通信線における通信所要時間を表し、 u_i はノード i におけるプロセッシングタイムを表す。 ϕ_i はノード i における外部到着率である。ノード遅延及び通信遅延の期待値はそれぞれ $F_i(\beta_i(P), u_i)$, $i=0, 1, \dots, n$, $G_i(\lambda_i(P), t_i)$, $i=1, 2, \dots, n$, $P=(t_1, t_2, \dots, t_n, u_0, u_1, \dots, u_n, \phi_0, \phi_1, \dots, \phi_n)$ で表す。

中央ノードについては次の関係がある。

$$F_0(\beta_0(P), u_0) = R(P) \quad (4.1)$$

$$\beta_0(P) = \sum_{i \in R_0} (\phi_i - \beta_i(P)) + \sum_{i \in R_d} \phi_i + \phi_0 \quad (4.2)$$

activeノードについては、(3.7b)により、次の関係がある。

$$F_i(\beta_i(P), u_i) - G_i(\lambda_i(P), t_i) = R(P) \quad (4.3)$$

λ_i がノード i と中央ノードとの通信線におけるトラフィック量を表しているので、下の式が成り立つ。

$$\lambda_i(P) = \phi_i - \beta_i(P) \quad (4.4)$$

個別最適化方式において中央ノードの遅延が $R(P)$ とする。neutralノードの場合、ジョブが到着したノードで処理されるので、ジョブの応答時間が $F_i(\phi_i)$ となる。activeノードの場合、ジョブが到着したノードで処理を受ける遅延と中央ノードに送り処理を受ける遅延と等しいので、ジョブの応答時間は中央ノードにおける遅延と通信線における遅延の和となり、 $G_i(\lambda_i(P), t_i) + R(P)$ で表される。idleノードの場合、全てのジョブが中央ノードに転送されるので、ジョブの応答時間が $G_i(\lambda_i(P), t_i) + R(P)$ となる。従って、システム全体の平均応答時間は次のような形に書ける。

$$T(P) = \frac{1}{\Phi} \left[\sum_{j \in R_0 \cup R_d} \phi_j (G_j(\lambda_j(P), t_j) + R(P)) + \sum_{j \in N} \phi_j F_j(\phi_j) + \phi_0 R(P) \right] \quad (4.5)$$

個別最適化方式において、次のような性質を持つことがわかる。

[性質4.1] $\frac{\partial R(P)}{\partial t_i} < 0, \quad i \in R_n,$
 $= 0, \quad i \in N, R_d.$
 $\frac{\partial R(P)}{\partial u_i} > 0, \quad i=0, i \in R_0,$
 $= 0, \quad i \in N, R_d.$
 $\frac{\partial R(P)}{\partial \phi_i} > 0, \quad i=0,$
 $> 0, \quad i \in R_0, R_d,$
 $= 0, \quad i \in N.$

[性質4.2] $\frac{\partial \beta_j(P)}{\partial t_i} > 0, \quad i \in R_0, j=i,$
 $< 0, \quad i \in R_0, j \neq i,$
 $= 0, \quad i \in N, R_d.$
 $\frac{\partial \beta_i(P)}{\partial u_i} > 0, \quad i=0,$
 $> 0, \quad i \in R_n, j \neq i,$
 $< 0, \quad i \in R_n, j=i,$
 $= 0, \quad i \in N, R_d.$
 $\frac{\partial \beta_i(P)}{\partial \phi_i} > 0, \quad i=0, i \in R_0, R_d,$
 $= 0, \quad i \in N.$

$$\begin{aligned}
 \text{[性質4.3]} \quad \frac{\partial R_i(P)}{\partial t_i} &< 0, \quad i \in R_0, j \neq i, \\
 &> 0, \quad i \in R_0, j = i, \\
 &= 0, \quad i \in N, R_0.
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial R_i(P)}{\partial u_i} &> 0, \quad i \in R_0, i=0, \\
 &= 0, \quad i \in N, R_0.
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial R_i(P)}{\partial \phi_i} &> 0, \quad i \in R_0, R_0, i=0, \\
 &= 0, \quad i \in N.
 \end{aligned}$$

$$\begin{aligned}
 \text{[性質4.4]} \quad \frac{\partial I(P)}{\partial t_i} &> 0, \quad i \in R_0, \\
 &= 0, \quad i \in N, \\
 &\text{uncertain, } i \in R_0.
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial I(P)}{\partial u_i} &> 0, \quad i \in R_0, N, i=0, \\
 &= 0, \quad i \in R_0.
 \end{aligned}$$

$$\frac{\partial I(P)}{\partial \phi_i} \quad \text{uncertain.}$$

これらの証明は付録Cに示す。

【性質4.1-4.3】は、 R を α , R_j を $\alpha_j = \alpha + G_j(\lambda_j)$ と置き換えると、全体最適化方式についても成立する。

5. 数値実験

この節では、上記各パラメータを変化させた場合、全体最適化方式と個別最適化方式による負荷分散の様子を調べる。

ノードモデルはセントラルサーバモデルを考える。各ノードには、1台のCPUと m 台のI/O装置がある。1つのジョブが処理が終わり、システムから出て行くと、直ちに新しいジョブがシステムに入る。ここで、ジョブがCPUから出ていった後、終了する確率を p_0 とし、I/O装置 j でサービスを受ける確率を $p_j, j=1, 2, \dots, m$ とすると、ジョブの平均応答時間は、

$$F(\beta) = \sum_{j=0}^m \frac{q_j}{\mu_j - q_j} \beta_j < \mu_j / q_j, j=0, 1, \dots, m$$

となる。但し、 $q_0=1/p_0, q_j=p_j/p_0$ で、 μ_j はサーバ j のサービス率である。

通信ネットワークモデルは、 $M/D/\infty$ を用いる。すなわち、平均通信遅延は、ネットワークを流れるトラフィック量に依存せず、 $G(\lambda_j, t_j) = t_j$ とする。

6. 実験結果

図2は中央ノードの処理能力を変化させた場合であり、図3はノード1と中央ノードとの通信時間を変化させた場合である。図4は全てのノードと中央ノードとの通信所要時間を同一にして変化させた場合である。図5は両

方式による応答時間の関係を示す。

7. 議論

定理1、2より全体最適化方式と個別最適化方式の解に類似性があることが分かり、個別最適化方式におけるシステム全体の平均応答時間が大体において全体最適化方式の値に近いことが見られる。しかし、性質[4.4]と図5から分かるように個別最適化方式の場合、通信所要時間または外部到着を増加させるとき、システム全体の平均応答時間がかえって減少することがある。

付録A. 定理1の証明

次のようなラグランジュ関数を考える。

$$L(\beta, \alpha, \gamma, \psi) = \Phi D(\beta) + \sum_{k=1}^m [\alpha^{(k)} (\Phi^{(k)} - \sum_{i=1}^n \beta_i^{(k)})] \quad (A1)$$

$$+ \sum_{i=1}^n \gamma_i^{(k)} (\phi_i^{(k)} - \beta_i^{(k)}) + \sum_{i=0}^n \psi_i^{(k)} \beta_i^{(k)} \quad (A1)$$

最適解は次の条件を $k=1, 2, \dots, m$ に対して満たす。

$$\begin{aligned}
 \frac{\partial L}{\partial \beta_i^{(k)}} = h_i^{(k)}(\beta_i) - \alpha^{(k)} - \gamma_i^{(k)} + \psi_i^{(k)} = 0, \\
 i=1, 2, \dots, n, \quad (A2)
 \end{aligned}$$

$$\frac{\partial L}{\partial \beta^{(k)}} = f_0^{(k)}(\beta_0) - \alpha^{(k)} + \psi_0^{(k)} = 0, \quad (A3)$$

$$\frac{\partial L}{\partial \alpha^{(k)}} = \Phi^{(k)} - \sum_{i=0}^n \beta_i^{(k)} = 0, \quad (A4)$$

$$\begin{aligned}
 (\phi_i^{(k)} - \beta_i^{(k)}) \geq 0, \gamma_i^{(k)} (\phi_i^{(k)} - \beta_i^{(k)}) = 0, \\
 \gamma_i^{(k)} \leq 0, \quad i=1, 2, \dots, n \quad (A5)
 \end{aligned}$$

$$\begin{aligned}
 \beta_i^{(k)} \geq 0, \psi_i^{(k)} \beta_i^{(k)} = 0, \psi_i^{(k)} \leq 0, \\
 i=0, 1, \dots, n. \quad (A6)
 \end{aligned}$$

条件(A5)より $\phi_i^{(k)} \geq \beta_i^{(k)}$, Tantawi&Towsley[1]と同様に場合に分けて調べる。

case 1. $\phi_i^{(k)} = \beta_i^{(k)}$ ($i \in N^{(k)}$) (A5)より $\gamma_i^{(k)} \leq 0$

(A6)より $\psi_i^{(k)} = 0$

故に(A2)より $h_i^{(k)}(\beta_i) \leq \alpha^{(k)}$

$\therefore h_i^{(k)}(\beta_i) \leq f_0^{(k)}(\beta_0)$

Case 2. $\phi_i^{(k)} > \beta_i^{(k)}$ (A5)より $\gamma_i^{(k)} = 0$, また、

$\beta_0^{(k)} > 0$ より $\psi_i^{(k)} = 0$. さらに次の場合に分かれる。

Case 2.1. $\beta_i^{(k)} > 0$ ($i \in R_0^{(k)}$) (A6)より $\psi_i^{(k)} = 0$

従って(A2)より $h_i^{(k)}(\beta_i) = \alpha^{(k)}$

$\therefore h_i^{(k)}(\beta_i) = f_0^{(k)}(\beta_0)$

Case 2.2. $\beta_i^{(k)} = 0$ ($i \in R_0^{(k)}$) (A6)より $\psi_i^{(k)} \leq 0$, よ

って(A2)より $h_i^{(k)}(\beta_i) \geq \alpha^{(k)}$

$\therefore h_i^{(k)}(\beta_i) \geq f_0^{(k)}(\beta_0)$

条件(A6)と(A3)より $f_0^{(k)}(\beta_0) \geq \alpha^{(k)}, \beta_0 = 0$
 $f_0^{(k)}(\beta_0) = \alpha^{(k)}, \beta_0 \geq 0.$

付録B. 定理2の証明

均衡解の存在. ショブスケジューリング決定が行われると、各ノードの負荷 β が決定される。次の定義をする。

$$\underline{\beta}^{(k)} = [\beta_1^{(k)}, \beta_2^{(k)}, \dots, \beta_n^{(k)}]$$

$$B^{(k)} = \{ \underline{\beta}^{(k)} \mid \sum_{i=0}^n \beta_i^{(k)} = \Phi^{(k)}, \phi_0^{(k)} \leq \beta_0^{(k)} \leq \Phi^{(k)}, \\ 0 \leq \beta_i^{(k)} \leq \phi_i^{(k)}, i=1, 2, \dots, n \}$$

$$B = \prod_{k=1}^m B^{(k)}$$

Bは $\underline{\beta}$ が取りうる値の集合で、mn-次元ユークリッド空間の、閉じた、凸で、有界の部分集合である。個別最適化状態であるためにはスケジューリング決定は次の条件を満たさなければならない。

中央ノードでの遅れの期待値は次のようであるとする。

$$F_0^{(k)}(\underline{\beta}_i) = R^{(k)}$$

- (1) $H_i^{(k)}(\underline{\beta}_i) \geq R^{(k)}$ ならばノードi到着のクラスkジョブは中央ノードで処理
- (2) $H_i^{(k)}(\underline{\beta}_i) = R^{(k)}$ ならばノードi到着のクラスkジョブはノードiか中央ノードで処理
- (3) $H_i^{(k)}(\underline{\beta}_i) \leq R^{(k)}$ ならばノードi到着のクラスkジョブはノードiで処理

このような決定により決まる。各ノードの負荷を $\underline{\beta}'$ と表すと $\beta_i^{(k)'}$ が取りうる領域 $\psi_i^{(k)}(\underline{\beta})$ は次のようになる。

- (1)の場合 $\beta_i^{(k)'} = 0$
すなわち $\psi_i^{(k)}(\underline{\beta}) = \{0\}$
- (2)の場合 $0 \leq \beta_i^{(k)'} \leq \phi_i^{(k)}$
すなわち $\psi_i^{(k)}(\underline{\beta}) = [0, \phi_i^{(k)}]$
- (3)の場合 $\beta_i^{(k)'} = \phi_i^{(k)}$
すなわち $\psi_i^{(k)}(\underline{\beta}) = \{\phi_i^{(k)}\}$

このようにして決まる $\underline{\beta}'$ の領域を $\psi(\underline{\beta})$ と表すと

$$\psi(\underline{\beta}) = \prod_{k=1}^m \prod_{i=1}^n \psi_i^{(k)}(\underline{\beta}_i) \cap B^{(k)}$$

である。明らかに $\psi(\underline{\beta})$ は凸で閉集合である。これに対して次の補題が成り立つ。

補題1. $\psi(\underline{\beta})$ はupper semicontinuous(上半連続)である。

証明 文献[3]の付録の補題と同様に証明できる。

対応 $\psi(\underline{\beta})$ は upper semicontinuousであり、Bは閉じた凸で有界なので角谷[2]のFixed Point Theoremが成り立つ。従って、 $\underline{\beta} \in \psi(\underline{\beta})$ なる $\underline{\beta}$ が存在する。すなわち、Bの均衡解が存在する。

均衡状態において満たされる条件も文献[3]の付録と同

様に導かれる。

付録C. 性質の証明

[性質1]の証明

$i \in R_0$ の場合、

(4.1)-(4.4)より

$$\frac{\partial R(P)}{\partial t_i} = \frac{\partial F_0}{\partial \beta_0} \frac{\partial \beta_0(P)}{\partial t_i} \quad (C1)$$

$j \in R_a$ の場合

$$\frac{\partial R(P)}{\partial t_i} = \frac{\partial F_j}{\partial \beta_j} \frac{\partial \beta_j(P)}{\partial t_i} - \frac{\partial G_j}{\partial \lambda_j} \frac{\partial \lambda_j(P)}{\partial t_i} - \frac{\partial G_i}{\partial t_i} \delta_{ij} \quad (C2)$$

ここで、 $i=j$ なら $\delta_{ij}=1$ そうでなければ $\delta_{ij}=0$

$$\frac{\partial \beta_0(P)}{\partial t_i} = - \sum_{j \in R_a} \frac{\partial \beta_j(P)}{\partial t_i} \quad (C3)$$

$$\frac{\partial \lambda_j(P)}{\partial t_i} = - \frac{\partial \beta_j(P)}{\partial t_i} \quad (C4)$$

(C1)-(C4)から

$$\frac{\partial R(P)}{\partial t_i} = - \frac{\partial F_0}{\partial \beta_0} \frac{\partial G}{\partial t_i} / [1 + \frac{\partial F_0}{\partial \beta_0} \sum_{j \in R_a} \frac{1}{A_j}]$$

が得られる。ここで $A_j = \frac{\partial F_j}{\partial \beta_j} + \frac{\partial G_j}{\partial \lambda_j}$

$\frac{\partial R(P)}{\partial u_i}, \frac{\partial R(P)}{\partial \phi_i}$ についても同じ様な手法で導かれる。

[性質2]と[性質3]の証明は[性質1]の証明を参照されたい。

[性質4]の証明

$$i \in R_0 \text{の場合} \quad \frac{\partial I(P)}{\partial t_i} = \frac{\phi_i}{\Phi} \frac{\partial G_i}{\partial t_i} > 0,$$

$$i \in N \text{の場合} \quad \frac{\partial I(P)}{\partial t_i} = 0,$$

$i \in R_a$ の場合

$$\frac{\partial I(P)}{\partial t_i} = \frac{1}{\Phi} [\sum_{j \in R_a \cup R_d} \phi_j \frac{\partial R_j(P)}{\partial t_i} + \phi_0 \frac{\partial R(P)}{\partial t_i}]$$

$$= \frac{1}{\Phi} \frac{\partial G_i}{\partial t_i} \frac{1}{A_i} \{ \phi_i \frac{\partial F_i}{\partial \beta_i} (1 + \frac{\partial F_0}{\partial \beta_0} \sum_{j \in R_a} \frac{1}{A_j})$$

$$- (\sum_{j \in R_a \cup R_d} \phi_j \frac{\partial F_j}{\partial \beta_j} + \phi_0) \frac{\partial F_0}{\partial \beta_0} / [1 + \sum_{j \in R_a} \frac{1}{A_j}] \}$$

ϕ_0 が大きく、かつ通信線の遅延が小さい場合、ノードiに到着するジョブの多くは中央ノードに送るため、 β_i が小さく、 λ_i が大きい。従って、このようなとき、 $\partial F_0 / \partial \beta_0$ が大きく、かつ $\partial F_i / \partial \beta_i$ が小さくなり、

$$\frac{\partial I(P)}{\partial t_i} < 0$$

となることがある。

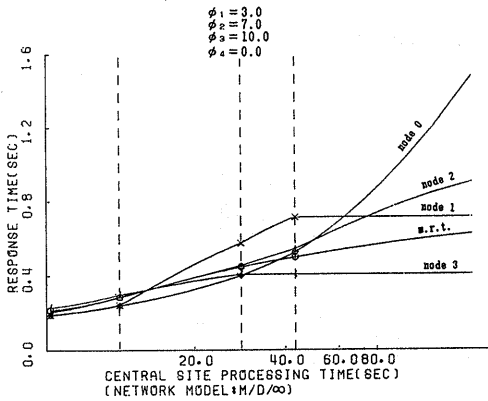
逆の場合は、 $\frac{\partial I(P)}{\partial t_i} > 0$

となる。

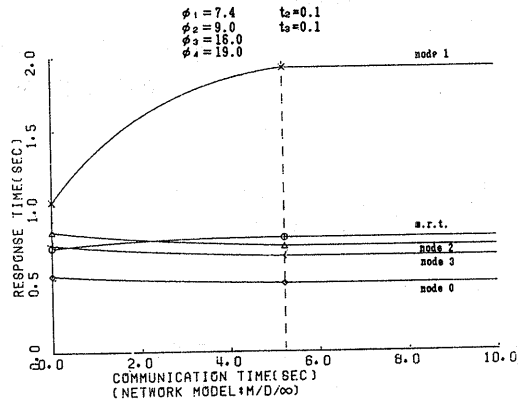
参考文献

- [1] Tantawi, A.N., and Towsley, D. "A general model for optimal static load balancing in star network configurations", PERFORMANCE '84 (Paris, Dec. 19-21, 1984) North-Holland, New York, 277-291.
- [2] Kakutani, S. "A generalization of Brouwer's fixed point theorem", Duke Mathematical Journal 8(1941), 457-459.

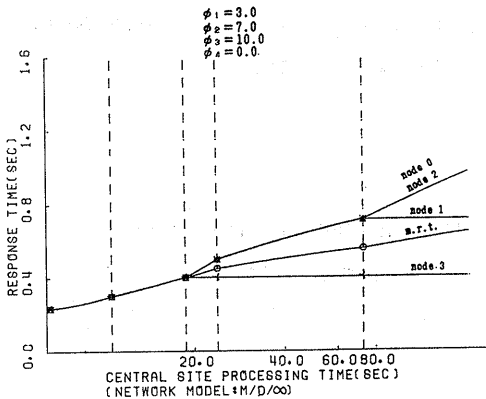
- [3] 亀田壽夫、榎山淳雄 “コンピュータネットワークにおける負荷分散について”、情報処理学会「オペレーティングシステム研究会」研究報告 34-7 (昭和62年2月26日)



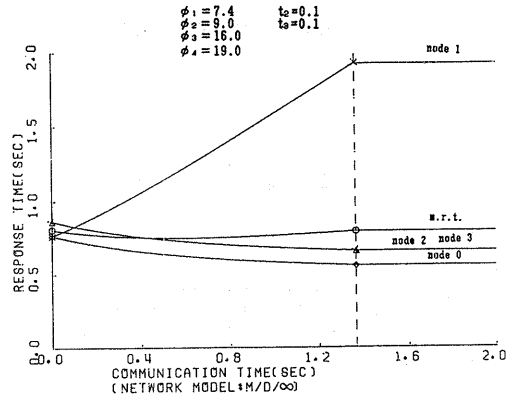
(全体最適化方式)



(全体最適化方式)



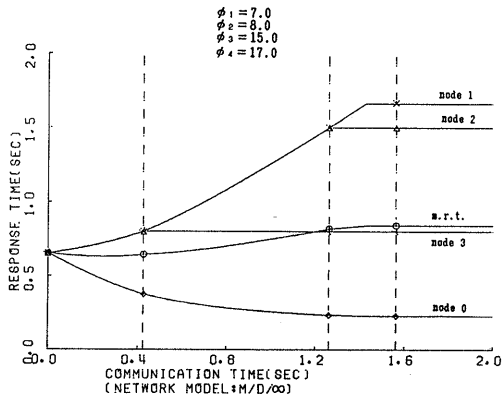
(個別最適化方式)



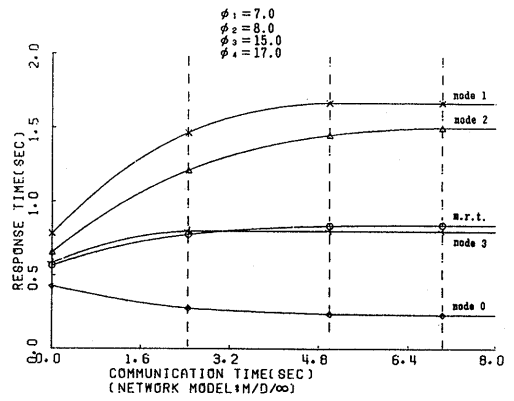
(個別最適化方式)

図2. 中央ノードの処理時間と応答時間の関係

図3. ノード1と中央ノードとの通信所要時間と応答時間の関係

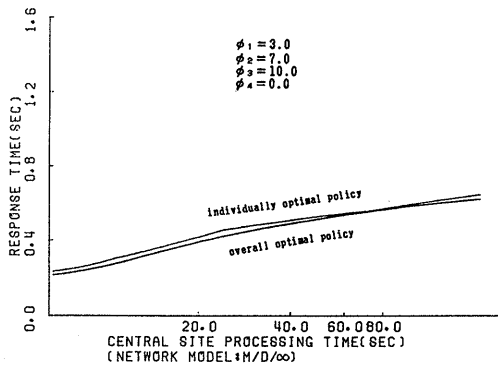


(個別最適化方式)

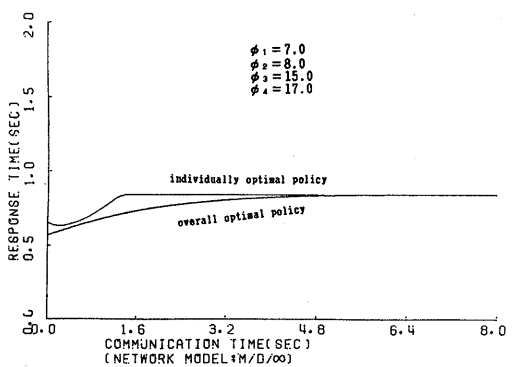


(全体最適化方式)

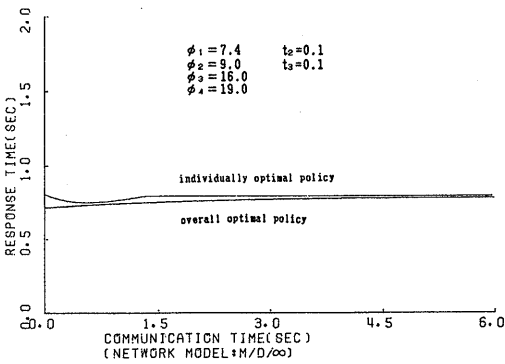
図4. 全てのノードと中央ノードとの通信所要
時間と応答時間の関係



(中央ノードの処理時間を変化させる場合)



(全てのノードと中央ノードとの通信所要時間を
同一に変化させた場合)



(ノード1と中央ノードとの通信時間を
変化させた場合)

図5. 全体最適化方式と個別最適化方式の
平均応答時間の比較