

分散情報処理網におけるジョブ配送法について

阪大工学部 中西 暉 (Hikaru Nakanishi)

阪大工学部 加茂博史 (Hiroshi Kamo)

阪大工学部 真田英彦 (Hidehiko Sanada)

阪大工学部 手塚慶一 (Yoshikazu Tezuka)

1. まえがき

分散情報処理システムにおいては、処理装置の有効利用を図る上で負荷分散が不可欠である。負荷分散システムにおいては、入力負荷を各処理装置に如何に配分するかが重要な課題であり、システム全体の処理状況によく整合する負荷配分法を確立する必要がある。本稿では、負荷分散型情報処理網におけるジョブ単位の負荷配分法であるジョブ配送法について検討する。

まず、ジョブ配送法を配送スケジュールに用いる基礎情報に関して4つのレベルに分類し、各レベルに属する具体的な配送法を示す。次に、シミュレーション並びに近似解析により平均システム時間特性を求め、各レベルで用いる基礎情報

が特性改善にどのような効果をもたらすかについて検討し、負荷分散の有効性を明らかにする。

2. 負荷分散型情報処理網

負荷分散型情報処理網として、図1に示すような複数の独立した情報処理システムが通信回線によって結合された横形分配網を考える。これは、広域に分散した大型計算機からなる広域情報処理網と考えることもでき、また地域的に密集したマイクロプロセッサからなる狭域情報処理網とも考えることができる。同図において以下に示すような仮定を設ける。

- (1) 本システムは r 個のローカルホストシステム H_i ($i = 1, 2, \dots, r$)から構成され、これらは通信網で結ばれている。

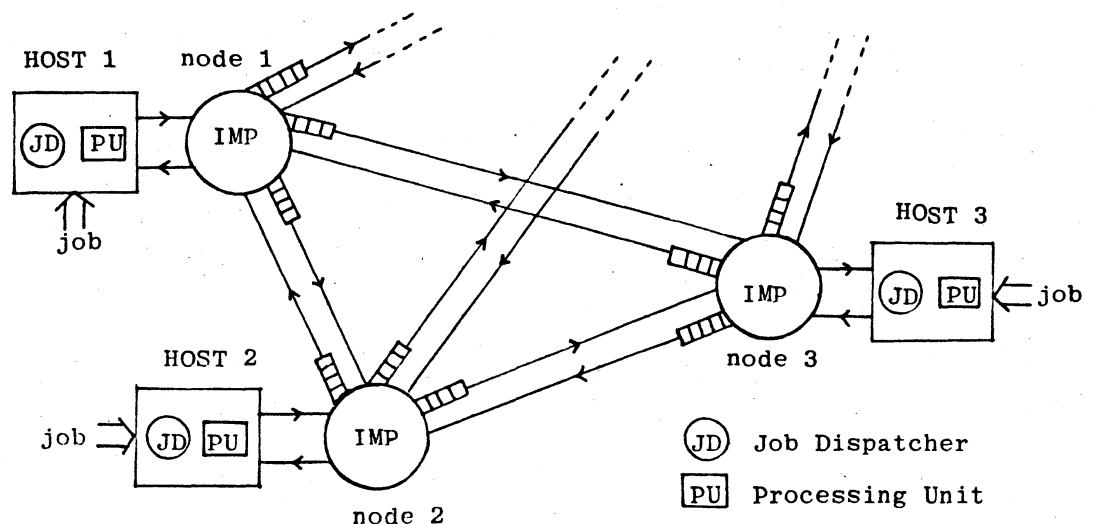


図1 負荷分散型情報処理網

- (2) 端末からの入力ジョブは、各 H_i 内にあるジョブ配送装置 JD_i によりジョブ単位で各 H_j 内にあるジョブ処理装置 PU_j に割当てられ、システム全体で負荷を分担することにより入力ジョブを処理する。
- (3) 各 PU_i は網へ入力される如何なるジョブも処理可能であるとし、 JD_i により割当てられたジョブは PU_i ごとに待ち行列を形成し、先着順に処理する。
- (4) 端末から JD_i への入力ジョブは平均 λ 個/秒のポアソン到着をする。
- (5) 入力ジョブの処理要求量は平均 $1/\mu$ 命令の指数分布をなす。
- (6) PU_i の処理速度はすべて等しく C 命令/秒である。
- (7) 各 H_i は IMP を介して全二重回線で接続され、網は完全網である。
- (8) 網内ではメッセージはパケット交換で伝送され、制御パケットはデータパケットより高い優先権をもつ。

3. ジョブ配送法

本節では、ジョブ配送装置 JD によりジョブ配送先 PU を決定するジョブ配送法について論じる。配送先 PU を決定するのに利用する基礎情報（状況情報）が多くなるにつれてよ

り適確なジョブ配送が可能であると考えられるので、配送スケジュールに利用する基礎情報に応じてジョブ配送法を次の4つのレベルに分類する。

Level 1 : 不変量・統計量だけを考慮した配送法

Level 2 : 配送履歴をも考慮した配送法

Level 3 : 自ホストの瞬時情報をも考慮した配送法

Level 4 : 全ホストの輻輳情報をも考慮した配送法

Level 1 及び Level 2 の配送法については既に報告しているので^{(1),(2)}、ここでは Level 3 及び Level 4 の配送法を中心に述べる。

3.1 Level 3 の配送法

Level 3 の配送法は、瞬時情報として自ホストの PU のシステム内容数を用い、各ホストの PU でのジョブの待ち合わせ遅延を最小にするよう制御する方式である。

具体的には、自ホストのシステム内容数 N が閾値 m に対して、 $N \leq m$ のとき JD はジョブを自ホストの PU に割当て、 $N > m$ のとき JD はジョブを基本配送法 (Level 2 以下の配送法) に従って適切なホストの PU に割当てる方式である。本配送法は、用いる基本配送法に応じて更に次の2つに分類することができる。

L3-1 : 自ホストの瞬時情報を用いたランダム配送法

L3-2 : 自ホストの瞬時情報を用いた巡回配送法

3.2 Level 4の配送法

Level 4の配送法は、全ホストの輻輳情報を用いることにより、各ホストのPUでのジョブの待ち合わせ遅延を最小にするよう制御する方式である。

具体的には、以下に示す配送スケジュールに従う方式である。

- (i) ジョブの配送は基本的には各基本配送法に従う。
- (ii) あるホスト H_i のシステム内容数 N_i が閾値 S に対して、 $N_i \geq S$ となったとき (Set状態), 全ホストに対し輻輳情報パケットを送出する。この輻輳情報パケットを受信した各ホストは、ホスト H_i へのジョブの配送を停止する。
- (iii) Set状態になったホスト H_j は、そのシステム内容数 N_j が閾値 R に対して、 $N_j \leq R$ となったとき (Reset状態), 全ホストに対し輻輳解除パケットを送出する。この輻輳解除パケットを受信した各ホストは、ホスト H_j へのジョブの配送を再開する。
- (iv) 全ホストが Set状態になったとき、各基本配送法に従

ってジョブを配送する。

本配送法は、用いる基本配送法により更に次の4つに分類することができる。

L4(L1) : 全ホストの輻輳情報を用いたランダム配送法

L4(L2) : 全ホストの輻輳情報を用いた巡回配送法

L4(L3-1) : 全ホストの輻輳情報を用いたL3-1配送法

L4(L3-2) : 全ホストの輻輳情報を用いたL3-2配送法

4. シミュレーション結果及び検討

図1に示した負荷分散型情報処理網において、Level 3及びLevel 4の6つの配送法を用いてジョブ配送を行ったときの平均システム時間特性及び平均応答時間特性をシミュレーションにより求めた。これらを図2～図7及び表1に示す。

まずLevel 3の配送法について検討する。本配送法では、閾値 m に平均システム時間 T_s を最小にするという意味での最適値 m_{opt} が存在し、

$$m_{opt} = \begin{cases} 0 & (0 < \rho \leq 0.6) \\ 1 & (0.6 < \rho \leq 0.85) \\ 2 & (0.85 < \rho < 1) \end{cases} \quad (1)$$

図2はホスト数を $r=5$ に固定し利用率 ρ を変化させたときの、図3はホスト数 r を変化させたときのLevel 3の配送法

の平均システム時間特性である。これらの図より、自ホストの瞬時情報を用いて負荷分散を行えば、各ホストで独自に処理する地域分散システム（M/M/1 システム）に比べ、平均システム時間特性がかなり改善されること、及び利用率が高くなるにつれて、またホスト数が多くなるにつれて改善率はより大きくなることがわかる。これらのことより、高負荷、大規模網ほど自ホストの瞬時情報による負荷制御が効果的に行われると言える。また L3-2の方が L3-1 より常に良好な特性を示していることから、基本配送法のレベルが高いほどより適切な負荷配分が行われると言える。

次に Level 4 の配送法について検討する。本配送法の Set 値 S は、オーバーヘッドと平均システム時間の兼ね合いから $S=3$ が最適であり、Reset 値 R の最適値は

$$R_{opt} = \begin{cases} 0 & (0 < p < 0.65) \\ 1 & (0.65 < p < 1) \end{cases} \quad (2)$$

図4は利用率 p を変化させたときの、図5はホスト数 r を変化させたときの Level 4 の配送法の平均システム時間特性である。図4より配送スケジュールに用いる基礎情報が多いほどより効果的な負荷配分が行われることがわかる。また利用率が高いほどホスト数の増加に対する改善効果が大きい。これは高利用率では網内輻輳が起こり易く輻輳情報が有効に利

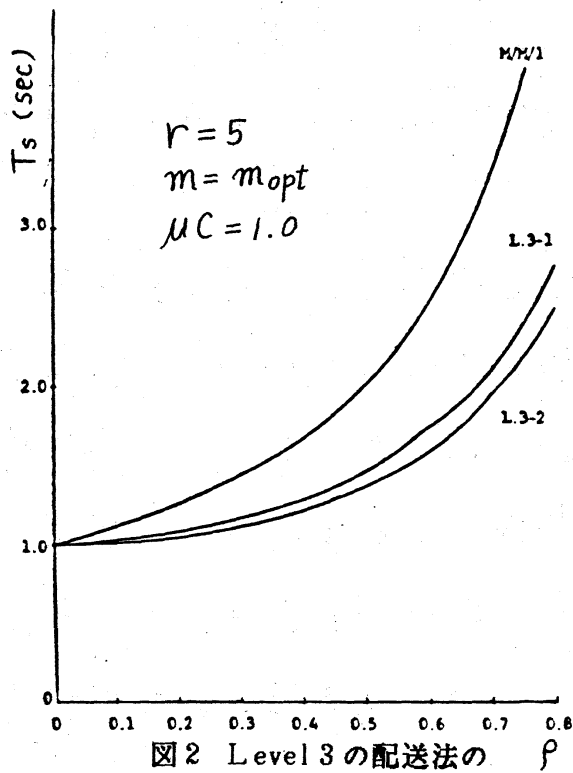


図2 Level 3の配送法の
平均システム時間特性

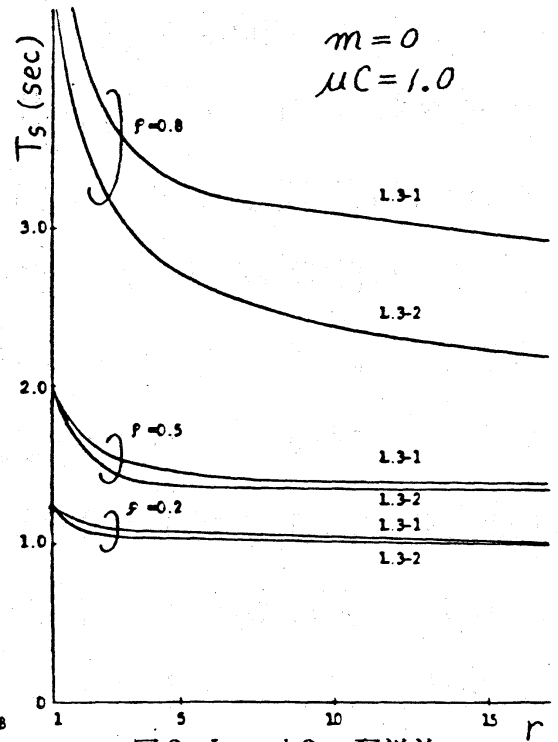


図3 Level 3の配送法の
平均システム時間特性

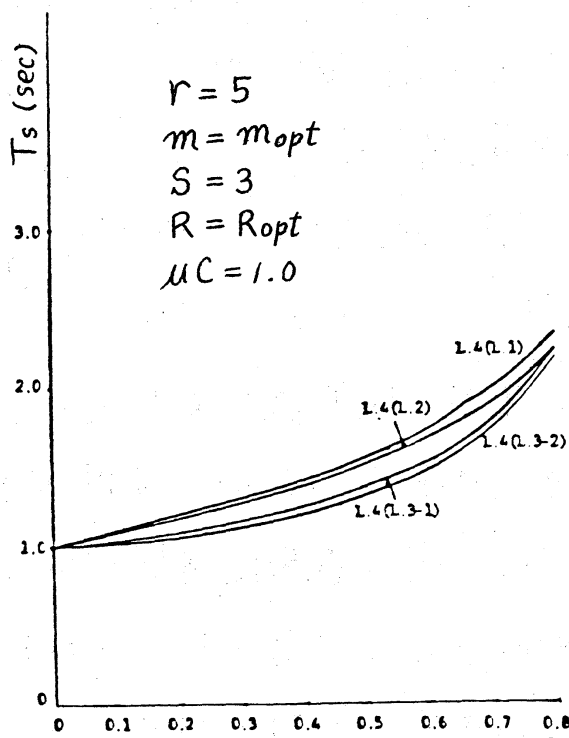


図4 Level 4の配送法の
平均システム時間特性

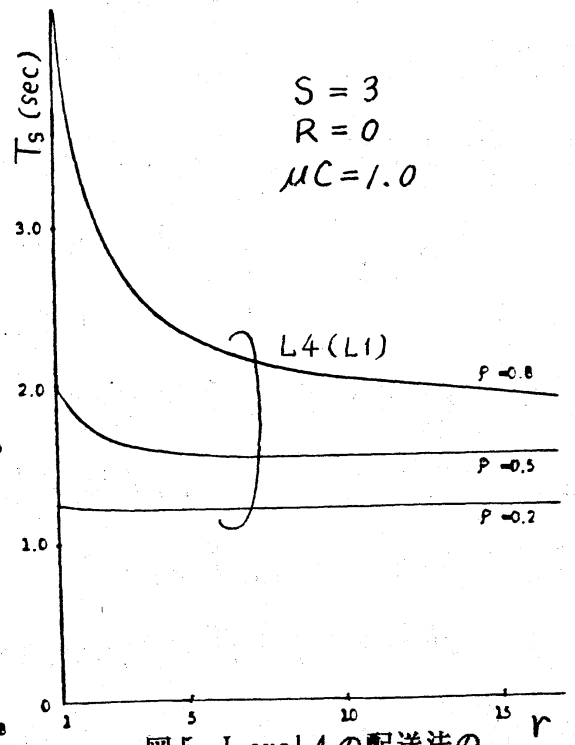


図5 Level 4の配送法の
平均システム時間特性

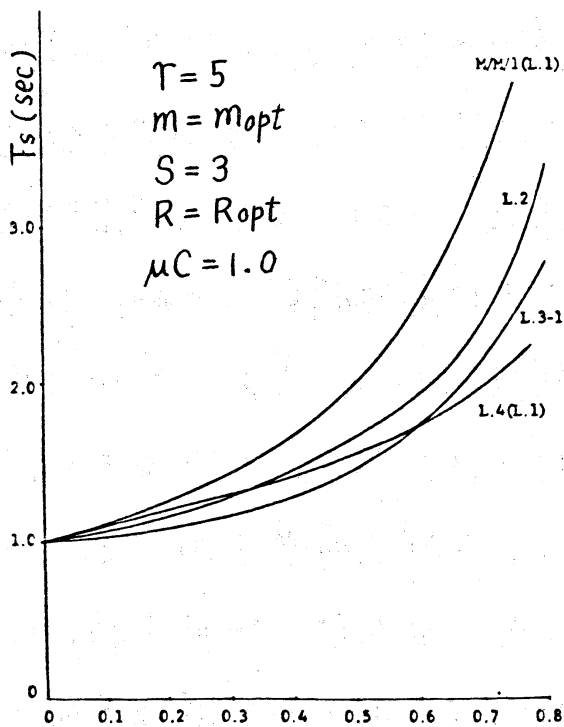


図6 各種ジョブ配送法の
平均システム時間特性

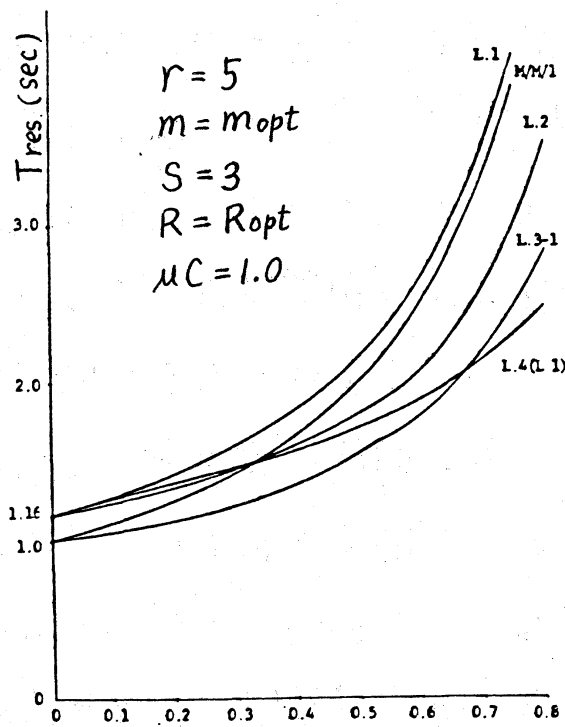


図7 各種ジョブ配送法の
平均応答時間特性

表 1 ジョブ配送法の比較

ジョブ配送法	基礎情報			平均システム時間		
	LEVEL2	LEVEL3	LEVEL4	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$
	自ホストの配送履歴	自ホストの瞬時情報	全ホストの輻輳情報	低利用率	中利用率	高利用率
L 1				8	8	8
L 2	○			6	7	7
L 3 - 1		○		4	4	6
L 3 - 2	○	○		2	3	5
L 4 (L 1)			○	7	6	4
L 4 (L 2)	○		○	5	5	3
L 4 (L 3 - 1)		○	○	3	2	2
L 4 (L 3 - 2)	○	○	○	1	1	1

用できるためである。

次に各レベルで用いる基礎情報の効果についてより詳細に検討する。図6及び図7は基本配送法としてランダム配送法を用いたときの各レベルの配送法の平均システム時間特性及び平均応答時間特性であり、表1は基礎情報の各種組み合わせに対する平均システム時間特性である。なお、表中の平均システム時間の欄の番号は特性の良好順に付した番号である。これらの図及び表より、自ホストの配送履歴を用いるだけでかなりの改善効果が得られること、及び低・中負荷時には自ホストの瞬時情報が、高負荷時には全ホストの輻輳情報が特性改善に有効であることがわかる。

5. 近似解析

ここでは、Level 3 及び Level 4 の配送法として最も単純な L3- / 配送法（自ホストの瞬時情報のみを利用する配送法）及び L4 (L/) 配送法（全ホストの輻輳情報のみを利用する配送法）を取り上げ、それぞれの平均システム時間の近似解析を行う。

5. 1 L3- / 配送法の近似解析

解析を行うにあたり次のような仮定を設ける。

〔仮定〕本配送法では自ホストからあふれたジョブは他ホストにランダムに配送されるが，このあふれジョブを含む各ホストへのジョブの到着はポアソン到着である。

上記の仮定の下では，L3-1 配送を行った場合の各ホストのシステム内容数に対する状態遷移率図は図8のようになる。ここで，

λ ：各ホストへのジョブの平均到着率

ν ： $N > m$ である他ホストからのジョブの平均到着率

$\mu (= \mu_C)$ ：各ホストのPUの平均処理率

システム内容数が k 人である平衡状態確率を p_k とすると，

$$\begin{cases} (\lambda + \nu) p_k = \mu p_{k+1} & (0 \leq k \leq m) \\ (\lambda/r + \nu) p_k = \mu p_{k+1} & (k > m) \end{cases} \quad (3)$$

簡単化のため $\mu = 1$ とおき， p_k を求めると，

$$p_k = \begin{cases} (\rho + \nu)^k p_0 & (0 \leq k \leq m) \\ \left(\frac{\rho}{r} + \nu\right)^k \left(\frac{\rho + \nu}{\rho/r + \nu}\right)^{m+1} p_0 & (k > m) \end{cases} \quad (4)$$

ここで， $\rho = \lambda / \mu$

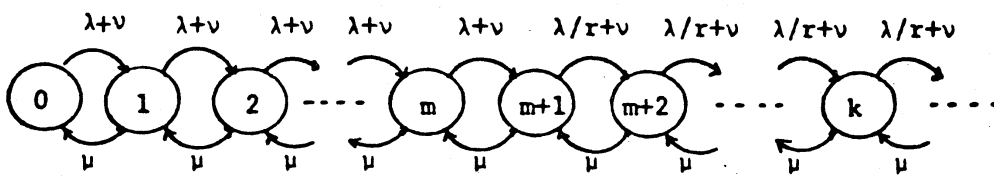


図8 Level 3 の配送法の状態遷移率図

確率の保存則より, p_0 は

$$p_0 = \frac{\{1 - (\rho + \nu)\} \{1 - (\frac{\rho}{r} + \nu)\}}{\{1 - (\frac{\rho}{r} + \nu)\} \{1 - (\rho + \nu)^{m+1}\} + \{1 - (\rho + \nu)\} (\rho + \nu)^{m+1}} \quad (5)$$

式(4), (5) より平均システム内容数 $\bar{N}$ 及び平均システム時間は容易に求まる。

$$\bar{N} = \sum_{k=0}^{\infty} k p_k \quad (6)$$

$$T_s = \bar{N} / (1 - p_0) \quad (7)$$

また未知数 ν は次式より計算できる。

$$\nu = \sum_{\substack{i=1 \\ i \neq j}}^r P[N_i > m] \frac{\lambda}{r} = \frac{r-1}{r} \left(1 - \sum_{k=0}^m p_k\right) \lambda \quad (8)$$

式(4) を式(6) に代入し, 簡略化すると,

$$\left(1 - \frac{\rho}{r} - \nu\right) \nu - \frac{r-1}{r} \rho (1 - \rho) (\rho + \nu)^{m+1} = 0 \quad (9)$$

任意の m に対して, 方程式(7) を数値計算により解けば, ν が求まる。

5. 2 L4(L1)配送法の近似解析

解析を行うにあたり次のような仮定を設ける。

[仮定] 他ホストからのあふれジョブを含むジョブの各ホストへの到着はポアソンであり, 制御情報パケットの伝送遅延は0とする。

上記の仮定の下では、本配送法の状態遷移率図は図9のようになる。ここで、

ν : 自ホストが Reset 状態にあるとき、自ホストに到着するジョブの平均到着率

κ : 全ホストが Set 状態 (Congestion 状態) にあるとき、自ホストに到着するジョブの平均到着率

$\mu (= \mu_C)$: 各ホストの PU の平均処理率

Reset 状態における定常確率を p_k , Set 状態における定常確率を p'_k で表すと、図9より次のような平衡状態方程式が得られる。

$$\left\{ \begin{array}{ll} \nu p_0 = \mu p_1 & (k=0) \\ (\nu + \mu) p_k = \nu p_{k-1} + \mu p_{k+1} & (1 \leq k \leq R-1) \\ (\nu + \mu) p_R = \nu p_{R-1} + \mu p_{R+1} + \mu p'_{R+1} & (k=R) \\ (\nu + \mu) p_k = \nu p_{k-1} + \mu p_{k+1} & (R+1 \leq k \leq S-2) \\ (\nu + \mu) p_{S-1} = \nu p_{S-2} & (k=S-1) \end{array} \right. \quad (10)$$

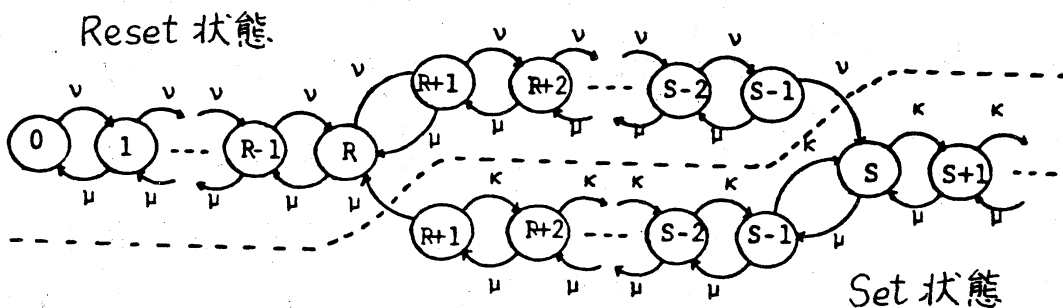


図9 Level 4 の配送法の状態遷移率図

$$\begin{cases}
 (\kappa + \mu) p'_S = \kappa p'_{S-1} + \mu p'_{S+1} + \nu p_{S-1} & (k=S) \\
 (\kappa + \mu) p'_{R+1} = \mu p'_{R+2} & (k=R+1) \\
 (\kappa + \mu) p'_k = \kappa p'_{k-1} + \mu p'_{k+1} & (R+2 \leq k \leq S-1) \\
 (\kappa + \mu) p'_k = \kappa p'_{k-1} + \mu p'_{k+1} & (k \geq S+1)
 \end{cases}$$

式(10)より p_k 及び p'_k を求めると,

$$p_k = \begin{cases}
 \left(\frac{\nu}{\mu}\right)^k p_0 & (0 \leq k \leq R-1) \\
 \frac{\sum_{i=0}^{S-k-1} \left(\frac{\mu}{\nu}\right)^i}{\sum_{i=0}^{S-R-1} \left(\frac{\mu}{\nu}\right)^i} \left(\frac{\nu}{\mu}\right)^R p_0 & (R \leq k \leq S-1) \\
 0 & (k \geq S)
 \end{cases} \quad (11)$$

$$p'_k = \begin{cases}
 0 & (0 \leq k \leq R) \\
 \frac{\sum_{i=0}^{k-R-1} \left(\frac{\kappa}{\mu}\right)^i}{\sum_{i=0}^{S-R-1} \left(\frac{\mu}{\nu}\right)^i} \left(\frac{\nu}{\mu}\right)^{R+1} p_0 & (R+1 \leq k \leq S-1) \\
 \frac{\sum_{i=0}^{S-k-1} \left(\frac{\kappa}{\mu}\right)^i}{\sum_{i=0}^{S-R-1} \left(\frac{\mu}{\nu}\right)^i} \left(\frac{\nu}{\mu}\right)^{R+1} \left(\frac{\kappa}{\mu}\right)^{k-S} p_0 & (k \geq S)
 \end{cases} \quad (12)$$

簡単化のため, $\mu = 1$ とおくと

$$p_k = \begin{cases}
 \nu^k p_0 & (0 \leq k \leq R-1) \\
 \frac{\nu^{k-S-1}}{\nu^{-S}-\nu^{-R}} p_0 & (R \leq k \leq S-1) \\
 0 & (k \geq S)
 \end{cases} \quad (13)$$

$$p'_k = \begin{cases}
 0 & (0 \leq k \leq R) \\
 \frac{1-\nu}{1-\kappa} \frac{1-\kappa^{k-R}}{\nu^{-S}-\nu^{-R}} p_0 & (R+1 \leq k \leq S-1) \\
 \frac{1-\nu}{1-\kappa} \frac{1-\kappa^{S-k}}{\nu^{-S}-\nu^{-R}} \kappa^{k-S} p_0 & (k \geq S)
 \end{cases} \quad (14)$$

確率の保存則より, p_0 は

$$p_0 = \frac{1}{\frac{1}{1-\nu} - \frac{\nu-k}{1-k} \frac{S-R}{\nu^{-S}-\nu^{-R}}} \quad (15)$$

式 (13), (14), (15) より, 平均システム内容数 $\bar{N}$ 及び平均システム時間 T_s は容易に求まる。

$$\bar{N} = \sum_{k=0}^{S-1} k p_k + \sum_{k=R+1}^{\infty} k p'_k \quad (16)$$

$$T_s = \frac{\bar{N}}{1-p_0} \quad (17)$$

また未知数 ν , k は次式より計算できる。

$$\begin{cases} \nu = p \sum_{i=0}^{r-1} \binom{r-1}{i} p^i q^{(r-1)-i} \frac{r\lambda}{i+1} \\ k = \lambda q^r \end{cases} \quad (18)$$

ここで

$$\begin{cases} p = P[\text{Reset}] = \sum_{k=0}^{S-1} p_k \\ q = P[\text{Set}] = \sum_{k=R+1}^{\infty} p'_k \end{cases} \quad (19)$$

式 (13), (14), (15) 及び (18), (19) より

$$\begin{cases} \nu = p - k \\ \left(\frac{k}{p}\right)^{\frac{1}{r}} \left\{ \frac{1}{1-p+k} - \frac{p-2k}{1-k} \frac{S-R}{(p-k)^{-S}-(p-k)^{-R}} \right\} = \frac{1-p+k}{1-k} \frac{S-R}{(p-k)^{-S}-(p-k)^{-R}} \end{cases} \quad (20)$$

6. 近似解析の精度の検討

前節で行った近似解析の精度について検討するため、種々のパラメータのもとに数値計算を行った。これらを図10～図15に示す。なお、これらの図には、比較のためにシミュレーション値を点線で併記してある。

図10, 11は $m=0, 2$ としたときの L3-1 配送法の利用率 ρ に対する平均システム時間特性であり、図12はホスト数 r に対する平均システム時間特性である。これらの図より、全般にわたって近似精度は良好で、利用率が低いほど、閾値 m が大きいほど、またホスト数が多いほど誤差が少ない

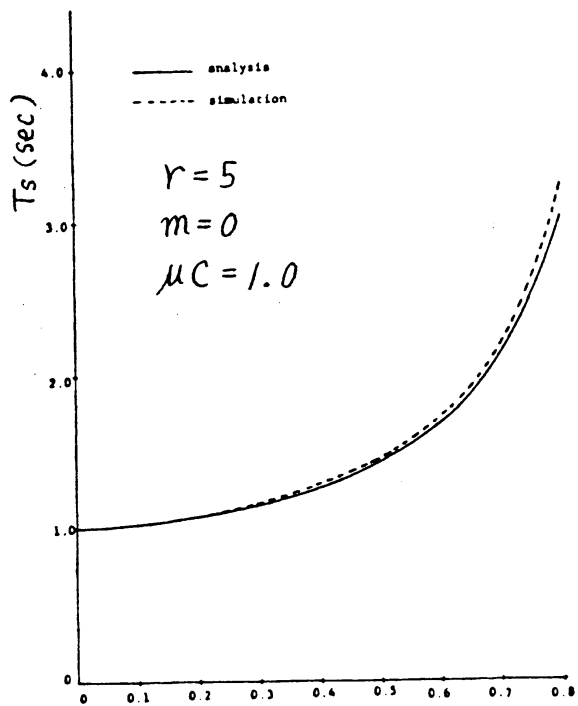


図10 Level 3の配送法の
平均システム時間特性

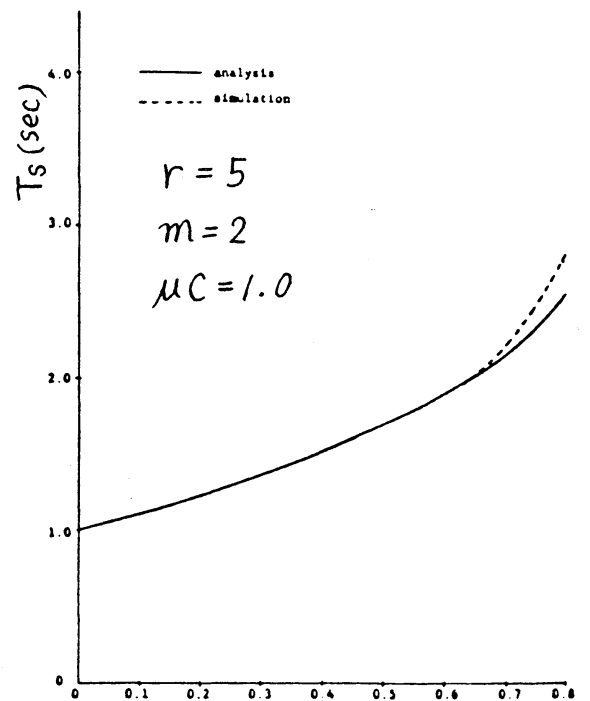


図11 Level 3の配送法の
平均システム時間特性

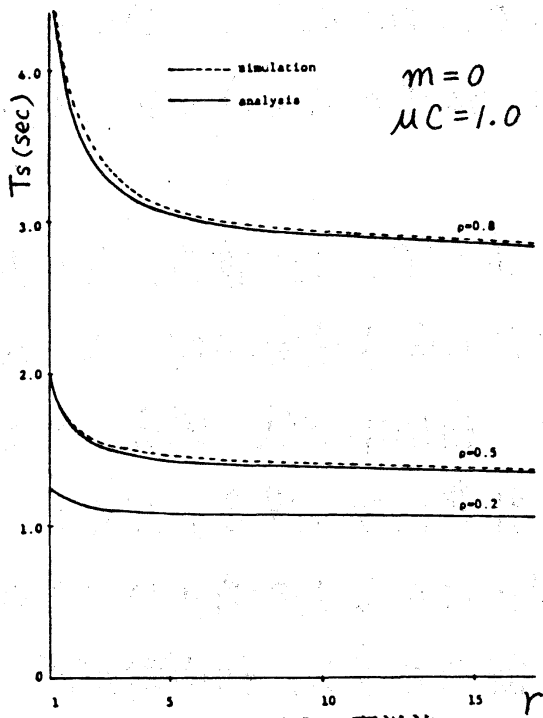


図12 Level 3の配送法の
平均システム時間特性

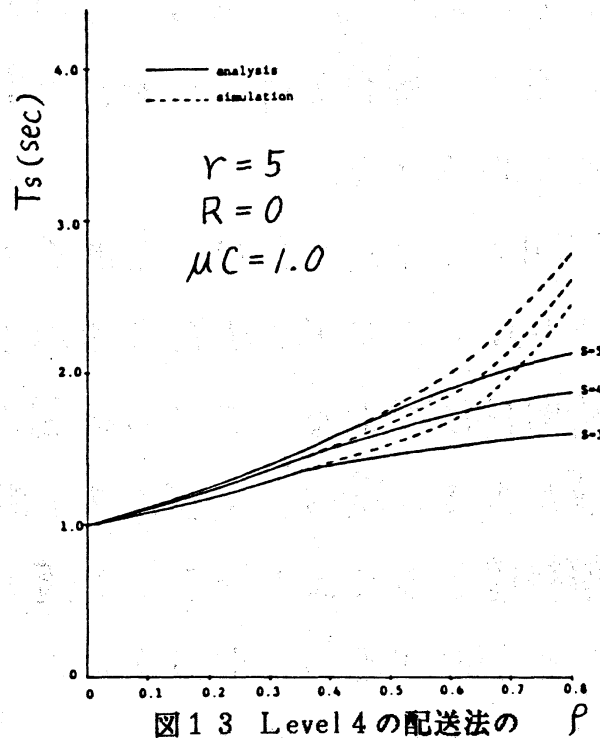


図13 Level 4の配送法の
平均システム時間特性

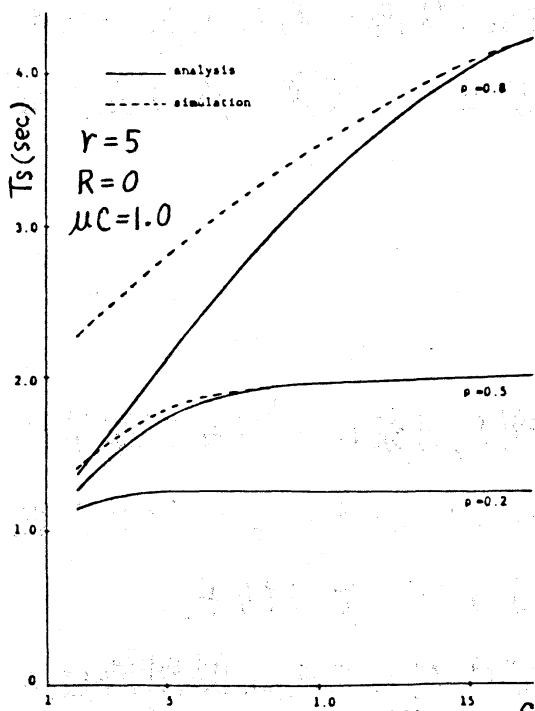


図14 Level 4の配送法の
平均システム時間特性

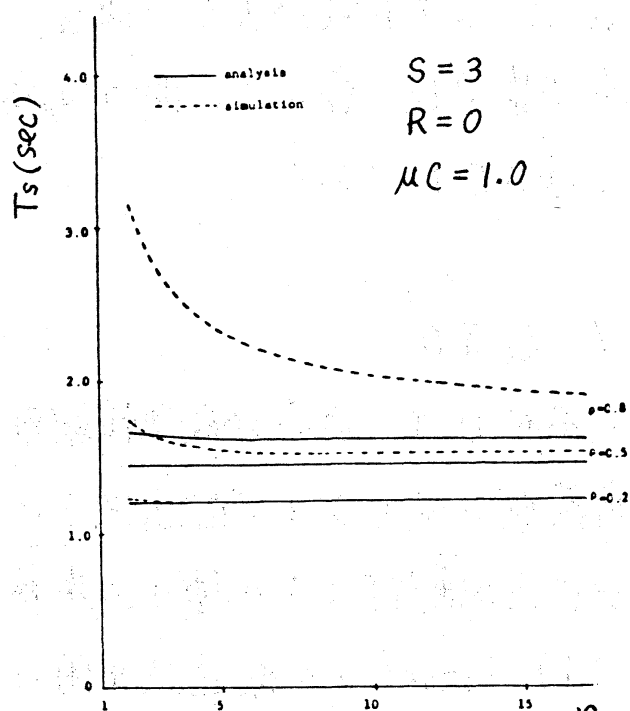


図15 Level 4の配送法の
平均システム時間特性

ことがわかる。これは、この領域ではあふれジョブがほとんどなく、解析を行うにあたり設けた仮定がよく成り立っているためである。

図13, 14, 15は、それぞれ利用率 ρ , Set値 S , 及びホスト数 r を変化させたときのL4(L1)配送法の平均システム時間特性である。これらの図より、利用率が低い場合、Set値が大きい場合、及びホスト数が非常に多い場合を除くと、解析値とシミュレーション値の間にはかなりの隔りができることがわかる。これは次のように説明できる。利用率が高くなるにつれて、各ホストのシステム内容数が増加し、Set値が低い場合、Set状態になるホスト数が増加する。それにつれて、ホスト数が少ない場合、Reset状態にあるホストへ負荷が集中し、各ホスト間の従属性が増し、仮定が成り立たなくなるためである。

7. むすび

本稿では、負荷分散型情報処理網を対象に、ジョブ単位の負荷配分法(ジョブ配送法)について、シミュレーション並びに近似解析により種々の考察を行った。その結果

- (1) Level 2の配送法を用いるだけでシステム時間特性のかなりの改善が得られる,

- (2) 低・中負荷時には自ホストの瞬時情報を, 高負荷時には全ホストの輻輳情報を用いるのが特性改善に最も効果的である,
- (3) ホスト数が多いほど, また配送スケジュールに用いる基礎情報が多いほど負荷分散効果を高め得る, ことなどを明らかにし, 負荷分散の有効性を示した。

参考文献

- (1) 井上, 加茂, 中西, 真田, 手塚, “分散情報処理網における統計的予測を用いた負荷制御について”, 信学技報, SE 81-18 (1981年5月).
- (2) 中西, 井上, 手塚, “複数待ち行列システムにおける負荷分散について”, 数理解析研究所講究録452, 待ち行列理論とその応用(Ⅱ) (1982年2月).