

複数待ち行列システムにおける負荷分散について

大阪大学 工学部 中西 暉
井上 健
手塚 慶一

1. まえがき

広域分散情報処理システムを構成する複数のローカルホストは、それぞれ独立して情報処理を行うことが可能であり、これは地域分散と呼ばれている。本システムでは、システム全体の処理状況を眺めた場合、ローカルシステムごとに負荷の不均衡が生じ易く、システム全体の処理効率が悪い。これに対処するのに負荷分散という考え方がある。これは、システムを構成する全処理装置がシステム全体の負荷を共有し、各処理装置を有効利用することによって応答特性を改善しようとするものである。

広域分散情報処理システムにおいては、各ローカルシステムは通信網を介して結合されているため、網内伝送遅延が原因となって緊密な通信による相互監視が望めない。したがっ

て、負荷分散システムにおいては、各処理装置への負荷配分法が重要な課題となる。本稿では、広域分散情報処理システムを複数待ち行列システムとして捉え、本システムにおける種々の負荷配分法について検討し、負荷分散効果を明らかにする。

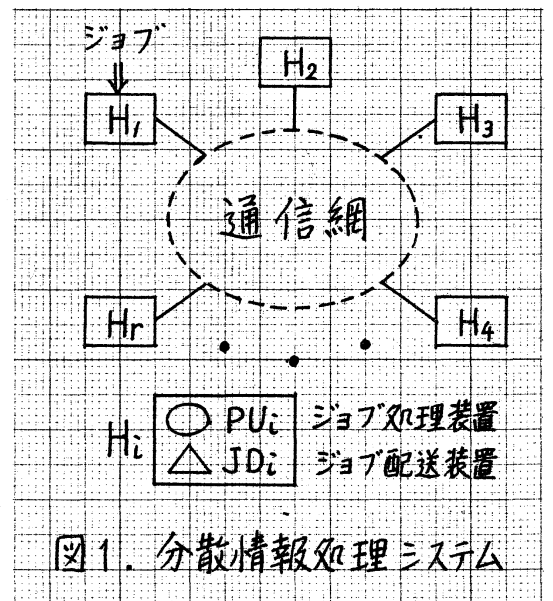
2. 分散情報処理システムとそのモデル化

負荷分散効果を評価する観点から、図1に示すような分散情報処理システムを考える。すなわち、

(i) 分散情報処理システムは、 r 箇所にある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。

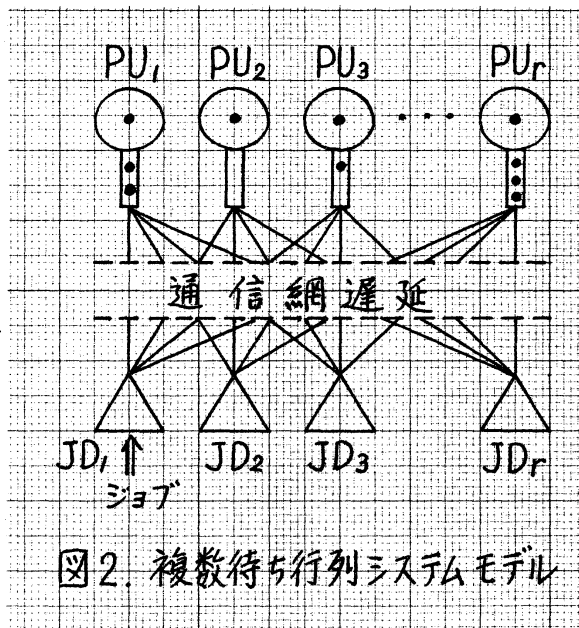
(ii) それぞれの H_i には端末

からジョブが入力され、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。



- (iii) 各 PU_i は網へ入力されるいか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, 先着順に処理される。
- (iv) 通信網は, 各 H_i 間で入力ジョブ及び処理結果を伝送するのに要する時間が d 秒以内であることを保証するものとする。
- (v) 端末から各 JD_i へのジョブ入力は, 平均 λ_i (個/秒) のポアソン到着をする。
- (vi) 各入力ジョブの処理要求量は, 平均 $1/\mu_i$ (命令) の指数分布に従う。
- (vii) PU_i の処理速度を C_i (命令/秒) とする。
- (viii) 負荷分散効果の評価基準として, PU での平均応答時間 T_p 又は端末での平均応答時間を用いる。

以上のような仮定のもとでは, 図1の分散情報処理システムは, 図2に示すような複数待ち行列システムとしてモデル化できる。



3. ジョブ配送方式

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方式（負荷配分法）について検討する。配送先を決定する際に利用する基礎情報が多様になるにつれて、より適確なジョブ配送が可能となり、応答特性が改善されると考えられるので、ジョブ配送方式を利用する基礎情報に応じて次の4段階に分けて考える。

レベル1：不変量，統計量だけを考慮した配送方式

レベル2：ローカル JD の配送履歴をも考慮した配送方式

レベル3：ローカル PU の瞬時情報をも考慮した配送方式

レベル4：システム全体の履歴，瞬時情報をも考慮した配送方式

各 JD_i が全 PU を実時間で監視できる集中制御システムとは異って、分散制御システムでは、各 JD_i が他ホスト内 PU_j の状況情報を得るのに伝送遅延を被るので、瞬時情報といえども過去の状況情報となる。このような遅延を伴う瞬時情報を用いるスケジュールにおいては、例えば、全 JD に対して特定の PU_j が透いているとの情報が伝えられた場合、各 JD_i が不用意な配送を行えば、特定の PU_j へのジョブの集中を引き起こし、システムを不安定にする。したがって、レベル4の配送法では、瞬時情報の利用に慎重な配慮が必要である。

本稿では、負荷分散効果を検討する第1段階として、上記の困難さを避け、各JD_iが容易に把握し得る基礎情報のみを利用するレベル1, 2の配送法について検討する。すなわち、ジョブ配送スケジュールを行うにあたり、次の基礎情報のみを用いる。

- (a) 各JD_iへのジョブの到着率 λ_i (個/秒)
- (b) ジョブの処理要求量 $1/\mu_i$ (命令)
- (c) 各PU_iの処理速度 C_i (命令/秒)
- (d) 各JD_iが各PU_jへ最後にジョブを配送してからの経過時間 t_{ij} (秒)

上記の基礎情報のみを用いたジョブ配送方式の基本的な特性を調べるために、さらにモデルを単純化する。すなわち、 $\lambda_i = \lambda$, $\mu_i = \mu$, $C_i = C$ ($i = 1, 2, \dots, r$) であるような均質な分散情報処理システムを考える。このような均質モデルにおける最適なジョブ配送方式は、各JDが新たなジョブを配送する際、配送後最も時間の経過したPUを選ぶことである。この場合、特定のJDは r 個存在するホストを順番(JDごとに任意)に配送先に選ぶという極めて単純な配送法となる。本方式をジョブ巡回配送方式と呼ぶことにする。

一方、不均質モデルに対しては、種々の配送法が考えられるが、ここでは以下に示す4つの配送方式について検討する。

(A) 処理能力比確率割当配送方式

各 JD_i は, 基礎情報 (b), (c) だけを用い, 確率 $P_{ij} = \mu_i C_j / \sum_{k=1}^r \mu_i C_k = C_j / \sum_{k=1}^r C_k$ に従ってジョブを PU_j へ配送する.

(B) 最適確率割当配送方式

各 JD_i は (a) ~ (c) を用い, 平均応答時間 T_p を最小にする最適確率 P_{ij}^* に従ってジョブを PU_j へ配送する.

(C) 平均処理量予測配送方式

各 JD_i は (b) ~ (d) を用い, 平均処理量 $PJ_{ij} = \mu_i C_j t_{ij}$ が最大である PU_j へジョブを配送する.

(D) 仮想平均処理量予測配送方式

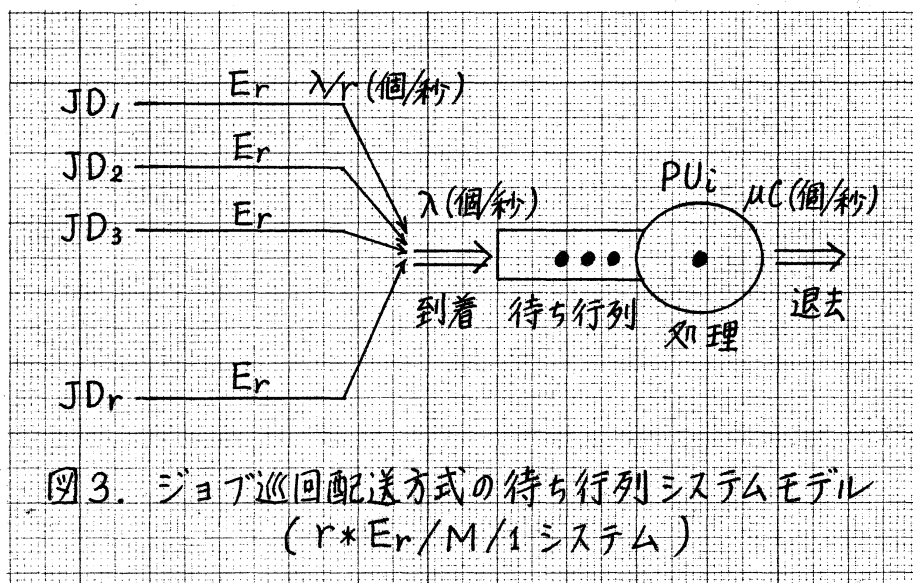
各 JD_i は (a) ~ (d) を用い, 仮想平均処理量 $VJ_{ij} = \mu_i C_{ij}^* t_{ij}$ (C_{ij}^* : 平均応答時間 T_p を最小にする定数) が最大である PU_j へジョブを配送する.

4. ジョブ巡回配送方式の近似解析

ジョブ巡回配送方式では, 特定の JD は入力されたジョブの r 個目ごとを特定の PU へ配送する. したがって, 特定の JD から特定の PU へのジョブの配送間隔は, 平均 r/λ (秒), 位相 r のアーラン分布 E_r をなす. 特定の PU へは全 JD が

ら同様の配送が行われるので、各PUへのジョブの到着は図3に示すような複合ア-ラン到着過程となる。又、各PUは配送された各ジョブを平均 $1/\mu C$ (秒) の指数分布に従う時間で順次処理する。このような待ち行列システムを $r * E_r / M / 1$ で表わすことにする。

このシステムのように個々の到着間隔分布が独立でない $G/M/1$ システムの解析は必ずしも容易ではない。そこで、ここでは到着過程を近似的に独立な $E_{r'}$ 到着 (r' は等価位相) と見なし、 $E_{r'}/M/1$ システムの解法を用いて近似解析を行う。すなわち、 $r * E_r$ 到着に関する分布の平均、分散に等しくなるような $E_{r'}$ 到着の位相 r' (r' は実数) を求め、等価位相 r' を用いて $E_{r'}/M/1$ システムを解き、平均応答時間 T_p を求める。



$$T_p(r', \rho) = \frac{1}{\mu C (1 - \sigma)} \quad (1)$$

ここで、 σ は次の方程式(2)の解である。

$$\sigma = \left(\frac{r' \rho}{1 - \sigma + r' \rho} \right)^{r'}, \quad \rho = \frac{\lambda}{\mu C} \quad (2)$$

等価位相 r' を求めるのに、次の3つの近似を用いる。

(A) 到着個数分布近似

一定時間 t に到着するジョブ数分布のZ変換を求めると、 $r * E_r$ 到着では

$$Pr^*(z, t) = \left\{ \sum_{k=0}^{\infty} \frac{1}{r} \sum_{i=1}^r \sum_{m=k-r-i+1}^{(k+i)r-i} \frac{(\lambda t)^m}{m!} e^{-\lambda t} z^k \right\}^r \quad (3)$$

$E_{r'}$ 到着では

$$P^*(z, t) = \sum_{k=0}^{\infty} \frac{1}{r'} \sum_{i=1}^{r'} \sum_{m=k-r'-i+1}^{(k+i)r'-i} \frac{(r'\lambda t)^m}{m!} e^{-r'\lambda t} z^k \quad (4)$$

$t = r/\lambda$ (PUへ平均 r 個のジョブが到着する時間) とおき、式(3)を用いて到着個数分布の分散を数値計算により求め、式(4)を用いて上で求めた分散と等しい分散を示す等価位相 $r' = r_a$ をグラフから求める。

(B) 複数到着間隔分布近似

あるジョブの到着時刻を0とし、その後複数個 (M 個) のジョブが到着するまでの時間分布 $G_M(t)$ を求めると、 $r * E_r$

到着では，時間 t 内に到着するジョブ数分布の z 変換は

$$\begin{aligned} M(z, t) &= \sum_{k=0}^{\infty} Q(k, t) z^k \\ &= \left(\sum_{k=0}^{\infty} \sum_{j=rk}^{r(k+1)-1} \frac{(\lambda t)^j}{j!} e^{-\lambda t} z^k \right) \left(\sum_{k=0}^{\infty} \frac{1}{r} \sum_{l=1}^r \sum_{j=r(k-l)+l}^{rk+l-1} \frac{(\lambda t)^j}{j!} e^{-\lambda t} z^k \right)^{r-1} \end{aligned} \quad (5)$$

であることより

$$G_M(t) = 1 - \sum_{k=0}^{M-1} Q(k, t) \quad (6)$$

Er' 到着では，

$$G(t) = 1 - \sum_{k=0}^{Mr'-1} \frac{(r'\lambda t)^k}{k!} e^{-r'\lambda t} \quad (7)$$

$M = r$ とおき，式 (6) より平均値 $\bar{t}$ 及び分散 σ^2 を数値的に求めると，等価位相 $r' = r_b$ は次式により求まる。

$$r_b = \frac{(\bar{t})^2}{M\sigma^2} \quad (8)$$

(C) 到着間隔分布近似

PU へのジョブ到着間隔分布は，式 (6) において $M = 1$ とおけば求まる。この分布の平均値 $\bar{t}$ ，分散 σ^2 を数値的に求めると，等価位相 $r' = r_c$ は次式により求まる。

$$r_c = \frac{(\bar{t})^2}{\sigma^2} \quad (9)$$

5. 数値計算例

まず，ジョブ巡回配送方式の負荷分散効果を調べるため，均質な分散情報処理システムを仮定し，シミュレーションを行った．図4～図5にこれらの結果を示す．

図4は，ホスト数 $r=4$ ， $1/\mu C=5$ (秒) のときの PU の平均応答時間特性 $T_p(r, \rho)$ を示したものである．同図には，比較対象として，各 JD に入力されたジョブをすべて自ホスト内の PU で処理する独立処理方式 (地域分散) の平均応答時間特性 $\hat{T}_p(\rho)$ を同時に示してある． $\hat{T}_p(\rho)$ は

$$\hat{T}_p(\rho) = \frac{1}{\mu C(1-\rho)} \quad (M/M/1) \quad (10)$$

により計算できる．これら両方式を比較すると，巡回配送方式を用いて負荷分散を行えば，応答時間特性が改善されることがわかる．これは，巡回配送によって，個々の JD へのジョブ到着間隔の時間的な不均衡が平滑化され，各 PU へのジョ

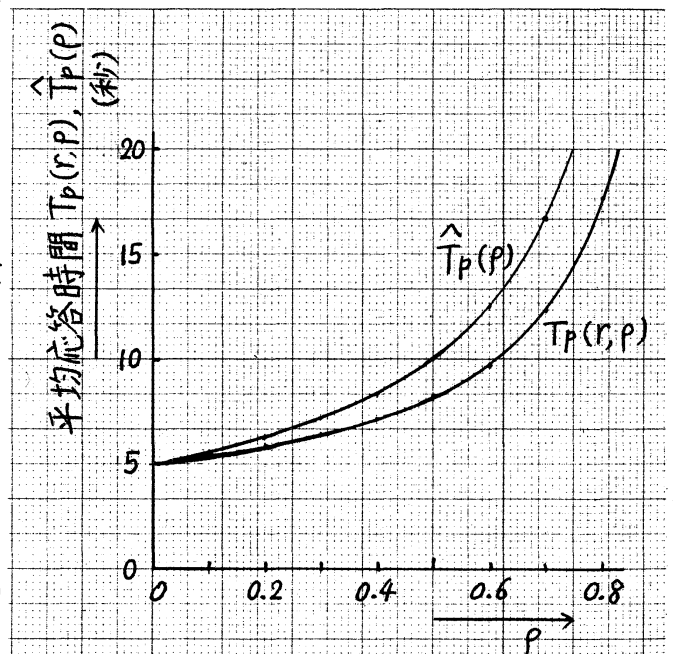


図4. 平均応答時間特性
($r=4$, $1/\mu C=5$ 秒)

ブの到着に規則性が生じるためである。

図5は、ホスト数 r の増加に対するシステム応答時間特性 $T_p(r, \rho)$ を ρ をパラメータとして示したものである。同図より、応答時間 $T_p(r, \rho)$ を最小にするという意味での最適ホスト r_{opt} は、

低負荷 ($\rho \leq 0.3$) では、 $r_{opt} = 2 \sim 3$

中負荷 ($0.4 \leq \rho \leq 0.7$) では、 $r_{opt} = 4$

高負荷 ($\rho \geq 0.8$) では、 $r_{opt} = 6$

となる。又、総合的に判断すると、 $r=4$ のとき全負荷にわたって応答時間特性の劣化が少ないことから、巡回配送方式の負荷分散効果を最も高めるのは4ホストからなる網構成であるといえる。

次に、前節で行った3つの近似解析について検討する。表1に各近似解析法によって求めた等価位相 r_a , r_b , r_c を示す。 r_a と r_b はほぼ等しい。これらは異なる分布に注目した到着過程近似により得

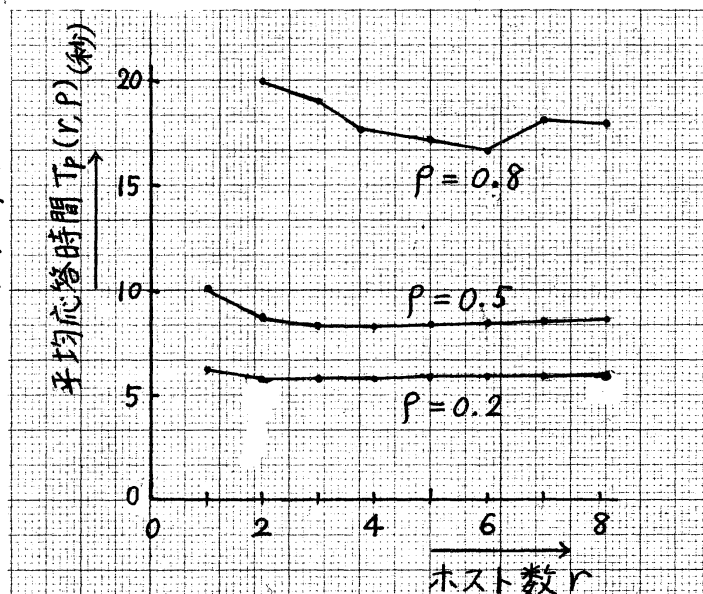


図5. ホスト数と平均応答時間特性の関係

た値であるが、 r_a と r_b はともに r 個のジョブ到着に関する分布を個数及び時間から眺めたものであるためだと考えられる。表1の等価位相を用いると、式(1)、(2)よりそれぞれの平均応答時間特性を計算することができる。 $r=5$ の場合の結果を図6に示す。同図には、近似精度を検討するため、シミュレーション結果を破線で示してある。図6より、利用率 ρ が低いときには(C)の到着間隔分布近似が有効であり、 ρ が高いときには(A)、(B)の複数ジョブ到着分布近似が有効であることがわかる。

この理由は次のように考えられる。 ρ が低いとき、ジョブ到着間隔に比べて処理時間が短いため、システムに待ち行列が生じがたい。

このような場合、待ち行列の形成による応答時間の増加は、主として個々の到着間隔のばらつきに依

表1. 各到着分布近似の等価位相

ホスト数 等価位相	2	3	4	5
r_a	1.76	2.28	2.67	3.03
r_b	1.78	2.31	2.73	3.08
r_c	1.60	1.65	1.56	1.46

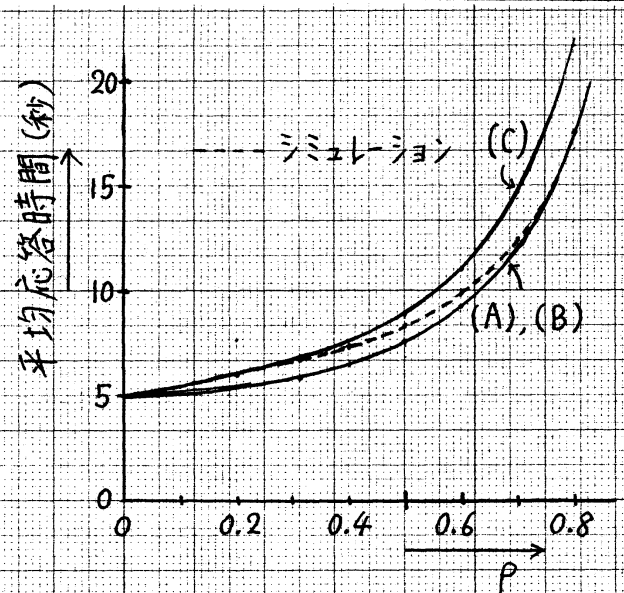


図6 近似解析結果とシミュレーション結果の比較

存する。一方、 ρ が高いとき、システムには比較的長い待ち行列が形成され易いため、システム稼働が継続する期間長が長くなる。このような場合、系内ジョブ滞留数の変化は個々の到着間隔のばらつきよりもむしろ一定期間内の到着ジョブ数のばらつきに依存するようになるためである。

最後に、第3節で示した4つのジョブ配送方式について検討するため、図7に示すような不均質モデルを設定する。このモデルを用いて求めた4つのジョブ配送方式

(A) 処理能力比確率割当配送方式

(B) 最適確率割当配送方式

(C) 平均処理量予測配送方式

(D) 仮想平均処理量予測配送方式

の負荷配分比特性を図8～図9に示す。図8は $C_{max} = 2C_{min}$ の場合であり、図9は $C_{max} = 3C_{min}$ の場合である。各PUへのジョブの配分率は(A)方式と(C)方式では一定であるのに対して、(B)方式と(D)方式では利用率 ρ の値により変化する。すなわち、(B)方式と(D)方式では、利用率 ρ が低いときすべての負荷が高速PUに配送さ

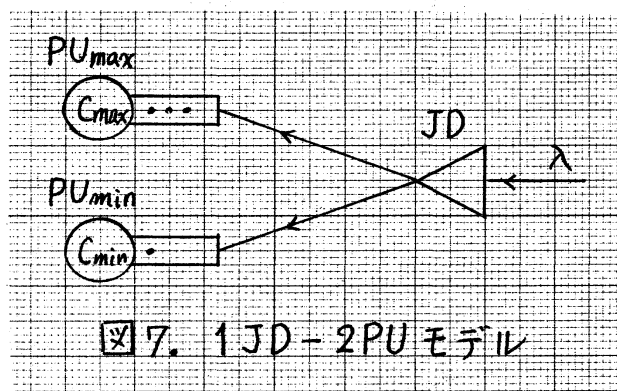


図7. 1JD-2PUモデル

れ、 ρ がある値に達すると ρ の増加につれて徐々に低速PUへ配分されるようになる。この傾向は処理速度の不均衡さが増すにつれてより顕著になる。以上のことより、低負荷のときにはあえて低速PUを使う必要のないこと、及びある程度以上の利用率のもとでは低速PUを適度に利用するのが有効であることがわかる。なお、 ρ の値を一定にした場合、高速

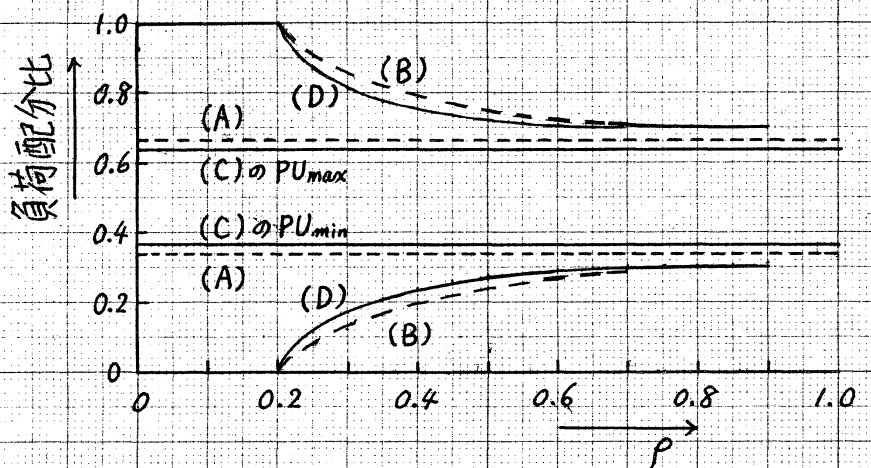


図8. 負荷配分比特性 ($C_{min}:C_{max}=1:2$)

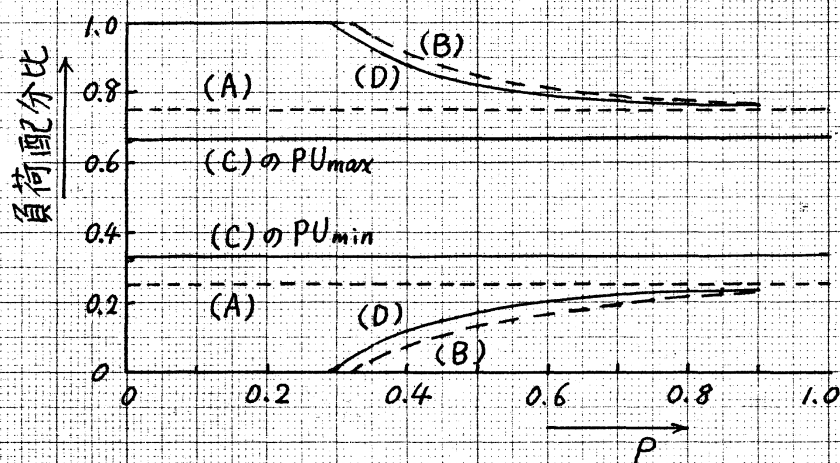


図9. 負荷配分比特性 ($C_{min}:C_{max}=1:3$)

PU への負荷の配分率は常に (B), (D), (A), (C) の順に高い。

図 10 及び図 11 に 4 つの配送方式の平均応答時間特性を示す。各配送方式とも利用率 ρ が増大すると平均応答時間は急激に増大する。ここには示していないが、処理速度が等しい場合には、(A) 方式と (B) 方式、(C) 方式と (D) 方式の応答特性がそれぞれ一致し、(C), (D) 方式の方が (A), (B) 方式より優れている。また、図 10, 11 より、処理速度比が 1:2 の場合には (D) 方式が、1:3 の場合には (B) 方式が最も

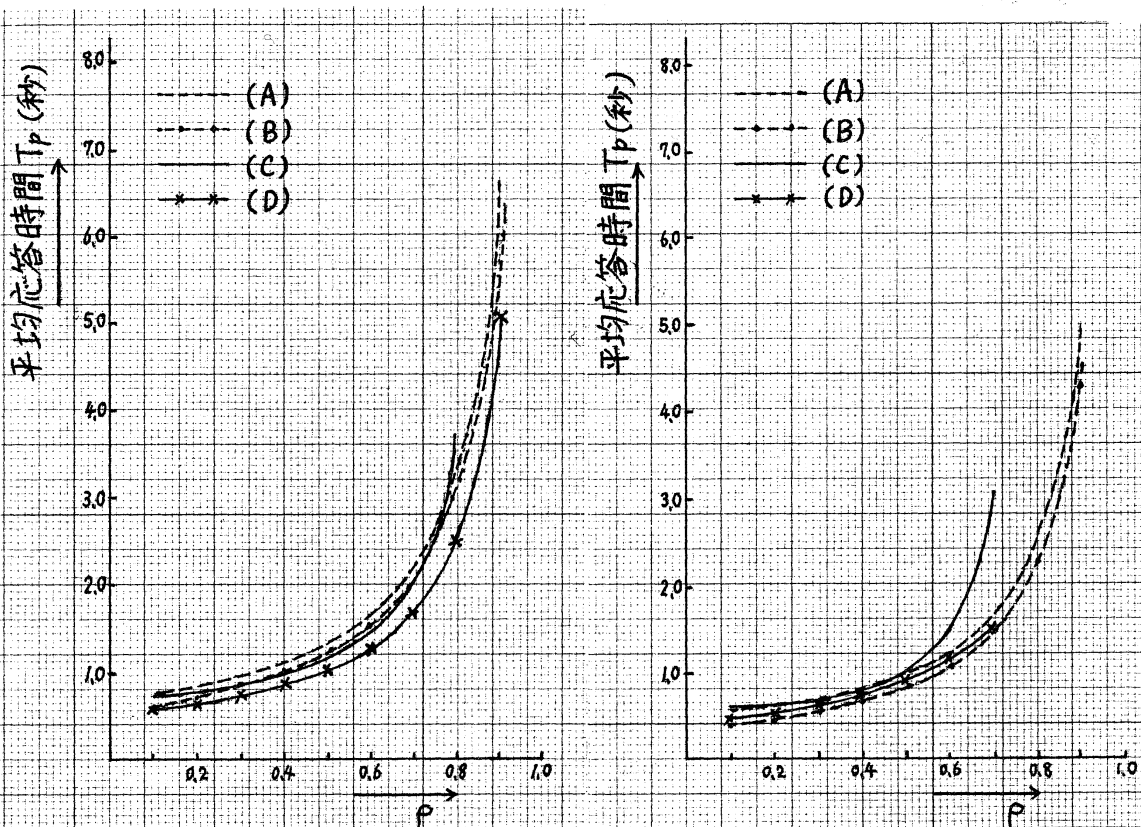


図 10. 平均応答時間特性
($C_{min}:C_{max}=1:2$)

図 11. 平均応答時間特性
($C_{min}:C_{max}=1:3$)

優れていることがわかる。以上のことより、処理速度比が倍程度までは(D)方式が優れ、それ以上の処理速度比では(B)方式が優れていることが明らかとなった。

6. むすび

本稿では、分散情報処理システムにおけるジョブ単位の負荷分散について検討し、不変情報、統計情報、ジョブ配送履歴だけを用いたいくつかのジョブ配送方式を示し、応答特性の改善に関する有効性を明らかにした。今後は、より多くの基礎情報を用いたレベル3, 4の配送方式による負荷分散効果の評価を行う予定である。