

## 複合処理系について

大阪大学 工学部 岡田 博美

### 1. まえかき

複合交換系に代表される、到着時間间隔やサーバ占有時間の特性が大半に異なる2つのクラスのジョブを同一のサーバ（サーバ群）で処理する複合処理系を考える。サーバの処理容量配分としては、次の2形式がある。

- i) 各クラス毎に独立に一定容量を割当てる。(固定配分)
- ii) 各クラスに一定容量を割当てるが、一方に空き、他方に待ちがある場合、空き容量の時を一時的に再配分する。  
(動的配分)

ここではジョブクラス間の特性差より、ii)の形式、特に一方のクラスにのみ他方の空きを利用させる形式の複合処理系を対象とする。

ジョブクラスとその特性を概述する。

## a) クラス 1

- ・到着率は比較的小さく、サーバ占有時間加長。
- ・待ちは許されず、呼損する。
- ・クラス2に空きがあるても使用しない。

## b) クラス 2

- ・到着率は相当高く、サーバ占有時間はクラス1の10倍  
以下小さく、一定と考える。

- ・無限の待ちを許す。

- ・クラス1サーバに空きがあれば適宜これを利用する。

このよう複合処理系の主な例を列挙する。

- ・複合交換系（あるいは音声/データ総合通信系）

1) 音声 or 固線交換, 2) データ or パケット交換

- ・計算機通信網

1) ファイル転送情報, 2) 会話形メッセージ

- ・Resource Sharing 形分散処理システム

1) 長大ジョブ (ナンバ・クランゲージジョブ) 2) トランザクション・ジョブ

- ・計算機システム

1) Batch ジョブ, 2) TSS, 実時処理ジョブ。

本稿は、複合処理系のクラス2のジョブの振舞、特に処理遅延 (処理時間 + 待ち時間) の期待値の、クラス1ジョブに対する依存性について考察する。

## 2. モデル

- 1) クラス1のジョブは、単位時間当り到着率  $\lambda_1$  のポアソン到着、ジョブ当りの平均サーバ占有時間  $h = 1/\mu_1$  の指數処理とする。
- 2) クラス2のジョブは、単位時間当り到着率  $\lambda_2$  のポアソン到着、ジョブの処理時間は一定で単位時間とする。
- 3) サーバの構成を図1に示す。即ち、全サーバ数  $N = N_1 + N_2$  うち  $N_1$  はクラス1、 $N_2$  はクラス2に割当てられる。各クラスのサーバは等価であるものとし、クラス1に空きがある場合、クラス2が一時的にこれを使用する。

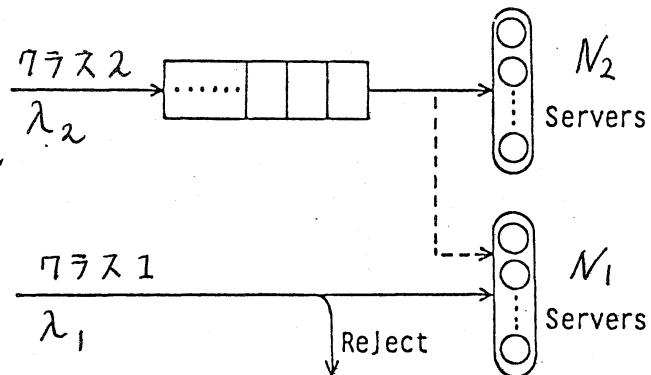


図1 複合処理系モデル

## 3. 二元マルコフ解析

前述のモデルに対し、次の仮定を用いて二元マルコフ過程として解析を行なう。

## [仮定]

$\lambda_1 \ll 1$ ,  $\mu_1 \ll 1$  より、単位時間当たりのクラス 1 の到着あるいは処理の終了はそれぞれ高々 1 とし、その確率を  $\lambda_1$ ,  $\mu_1$  とする。

また単位時間開始直後 n 級内にあるクラス 1, クラス 2 の個数をそれぞれ  $n_1^{(k)}$ ,  $n_2^{(k)}$  とし、系の状態を  $M^{(k)} = (n_1^{(k)}, n_2^{(k)})$  とする。状態系列  $\{M^{(k)}\}$  はマルコフ連鎖を成す。 $\lambda_1 < N_1 + N_2 - S_1$  のとき、 $k \rightarrow \infty$  に対する定常状態確率  $\pi_{ij}$  を得る。ここで  $S_1$  はクラス 1 の平均系内数を示す。また  $\pi_{ij}$  は次の状態方程式を満たす。

$$\begin{aligned} \pi_{ij} &= (1 - \delta_{i,0}) \cdot \lambda_1 \cdot \{ 1 - (i-1)\mu_1 \} \times \\ &\quad \left\{ \sum_{m=0}^{N_1+N_2-i} \pi_{i-1,m} \cdot P_j + \sum_{m=0}^j \pi_{i-1, N_1+N_2+m-i+1} \cdot P_{j-m} \right\} \\ &+ \{ i \cdot \lambda_1 \cdot \mu_1 + (1-\lambda_1)(1-i\mu_1) + \delta_{i,S_1} \cdot \lambda_1 (1-i\mu_1) \} \times \\ &\quad \left\{ \sum_{m=0}^{N_1+N_2-i-1} \pi_{i,m} \cdot P_j + \sum_{m=0}^j \pi_{i, N_1+N_2+m-i} \cdot P_{j-m} \right\} \\ &+ (1 - \delta_{i,S_1}) (i+1) (1-\lambda_1) \mu_1 \end{aligned}$$

$$\left\{ \sum_{m=0}^{N_1+N_2-i-2} \pi_{i+1,m} \cdot P_j + \sum_{m=0}^j \pi_{i+1,N_1+N_2+m-i-1} \cdot P_{j-m} \right\} \\ (i \geq 0, N_1 \geq j \geq 0) \quad (1)$$

ここで  $\delta_{ij}$  は、 $i = j \pm 1$ ,  $i \neq \pm 0$  となる。また、  
 $P_i$  はクラス 2 の単位時間当りの  $i$  ノード到着個数確率であり、  
平均入力のホアソン分布となる。

これよりクラス 2 の平均系内数  $S_2$ 、平均処理遅延  $W_2$  は  
次式となる。

$$S_2 = \sum_{i=0}^{N_1} \sum_{j=0}^{\infty} i \cdot \pi_{ij} \quad (2)$$

$$W_2 = \frac{S_2}{\lambda_2} + \frac{1}{2} \quad (3)$$

#### 4. 多段容量限界(MCL)効果

前章の解析の数値例を図2に示す。ここでサーバ数  $N_1 = 1$   
 $N_2 = 1$ 、クラス 1 の平均サーバ占有時間  $h = 500$ 、トランシット強度  $a = \lambda_1 \cdot h (= \frac{\lambda_1}{\mu_1})$  を  $0.1, 0.5, 2.0$  とし左場合の、クラス 2 の到着率  $\lambda_2$  に対する平均処理遅延  $W_2$  を  
図2に示す。

図2より次の2点が観測される。

- 1) クラス 2 の負荷  $\lambda_2$  をクラス 2 本来の容量  $N_2 = 1$  附近に設定すると仮想の容量限界が生じる。

2) 仮想の容量限界と真の容量限界の間に、処理遅延  $W_1$  の增加の傾きが緩慢となる、一種の高原状態が存在する。

この振舞を調べるために、2つの処理系を切り離し、クラス2の入力  $\lambda_2 + \lambda_1$  をクラス1のサーバで処理されたもの  $R_1$  とクラス2本来のサーバで処理されたもの  $R_2$  に行け、次に仮説を考える。ここで  $\lambda_2 = R_1 + R_2$  である。  $R_1, R_2$  は前述の二次元マルコフ解析より求められる。

[仮説1]  $R_1$  を無視し、 $R_2$  のみがクラス2サーバに加えられたものとして平均処理遅延を求める。

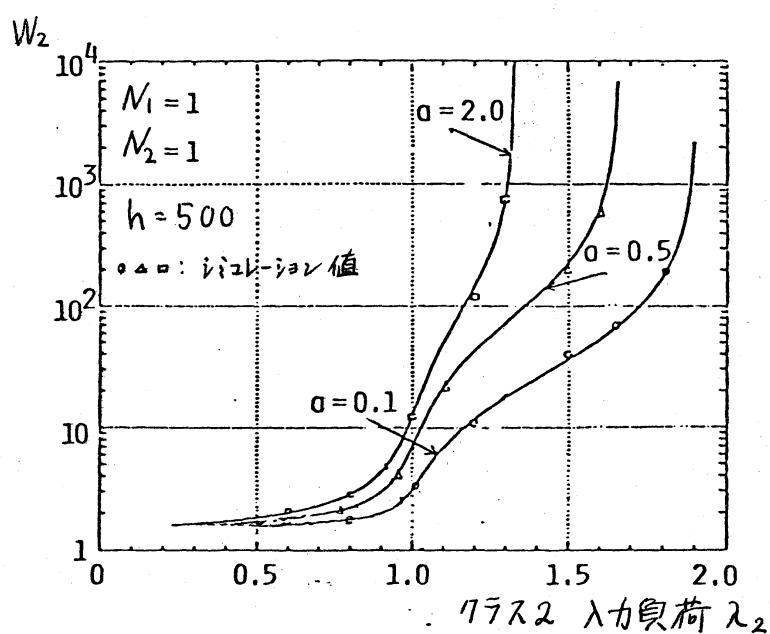


図2 入力負荷-平均処理遅延特性

[仮説2]  $R_1$ をクラス1サーカスに強制的ループ加し、 $R_2$ をクラス2サーカスで処理。2つの平均遅延を $R_1, R_2$ で加重平均。

[仮説3]  $R^* = \lambda_2 - N_2$ をクラス1サーカスに強制加し、残りをクラス2に加えた平均遅延は  $\lambda_2 = N_2$  の解析値とし、この両者を  $R^*, N_2$  で加重平均する。

図3に3つの仮説による特性、( $N_1=1, N_2=0$ ), ( $N_1=0, N_2=1$ ), および  $M$  ( $N_1=1, N_2=1$ ) の場合の特性を示す。仮説1, 2 が解析値と大きく異なっており、仮説3は解析値とよく一致している。仮説1は  $R_1$ を無視(そのため過小評価となる)、仮説2は  $R_1$ を単純なクラス1サーカスに強制加(そのため過大評価となる)である。仮説3は、クラス2の割当て容量内で処理仕切れる。

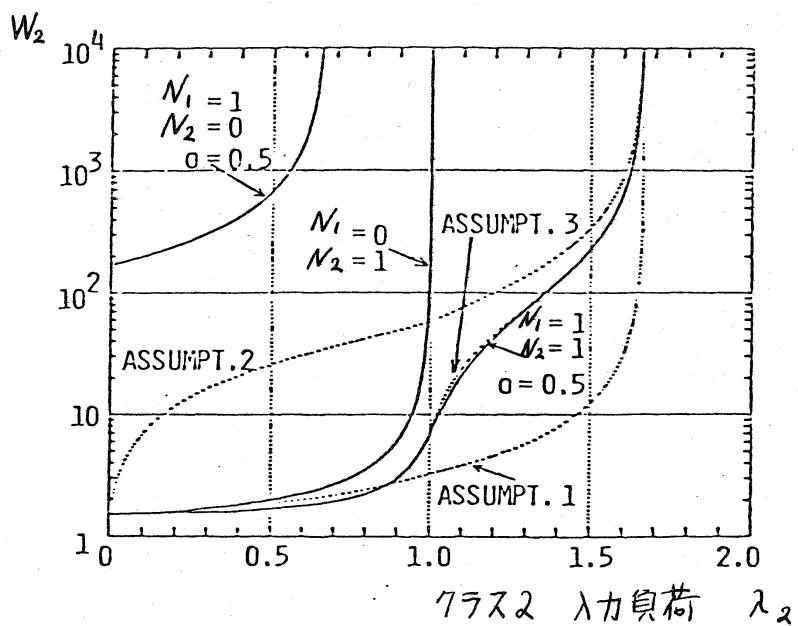


図3. 3つの仮説による入力負荷-処理遅延特性

・負荷  $R^*$  をクラス1 サーバル追加し評価しているため、問題の現象を捉えていけると考えられる。こゝよん前述の現象は次の様に考えられる。

処理時間が大半に異なる複合処理系では、クラス2の入力負荷入力加、1) クラス2 自体の処理容量内で処理しうる時の遅延特性 ( $M=0, N_2=1$ ) と、2) クラス2の外で処理せざれば、クラス1の空きを利用して最早得る時の特性 ( $N_1=1, N_2=0$ ) との間に大きさ差を感じる。従って  $\lambda_2$  が小さく時は1),  $\lambda_2$  が極めて高くなる時は主に2) の特性に支配される。 $\lambda_2$  がクラス2 本来の容量限界  $N_2$  よりも大きくなる附近より1) から2)への特性上の遷移が生じ、前記の現象が観測されたことになる。

この現象を多段容量限界 (Multi Capacity Limit; MCL) 効果と呼ぶ。

クラス1のサーバ数  $N_1$  が2以上の場合、クラス2の入力負荷  $\lambda_2$  の増加に対し、クラス1サーバへの依存性が "critical" に変化する各閾値 (ex.  $\lambda_2=N_2$ ,  $\lambda_2=N_2+1$ , ...) 每にこのMCL効果の生じる可能性がある。以下、MCL効果を中心として複合処理系の種々の特性を図示する。

図4は  $N_1=2, N_2=2$  の特性を示す。 $\lambda_2=N_2+1$  で MCL効果が観測される。図5は  $N_1$  の変化に対する MCL効果を示す。クラス1の負荷が比較的重くため、 $\lambda_2=N_2$  で MCL効果が顯

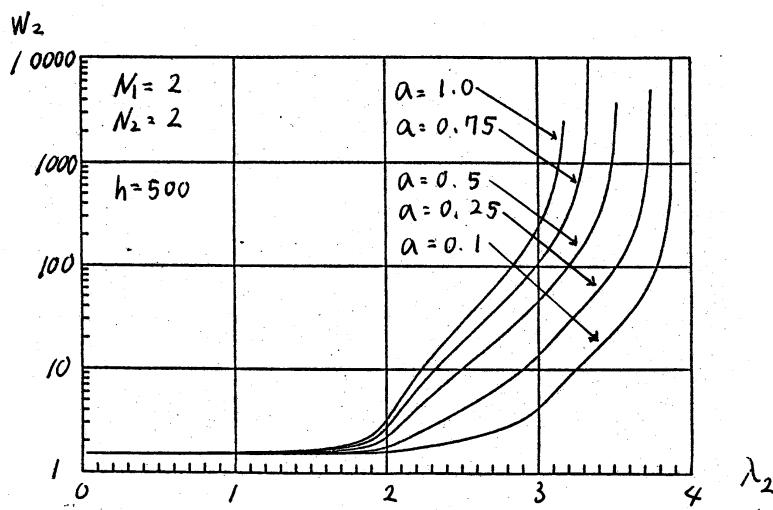


図.4

サ - バ数を増した場合の、入力負荷 - 处理遅延特性

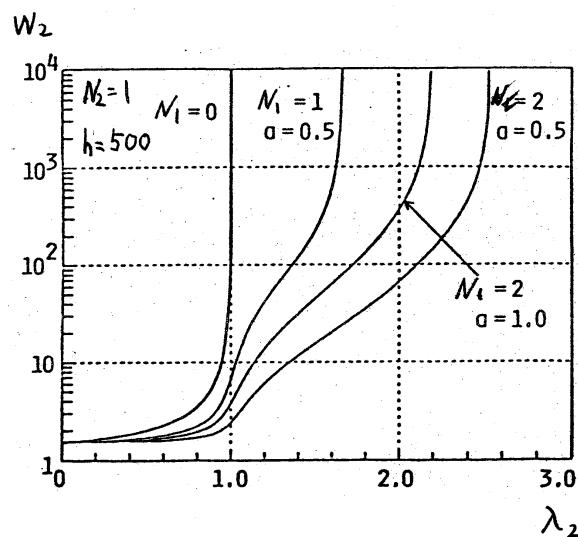


図.5

クラス1のサ - バ数を増した場合の、入力負荷 - 处理遅延特性

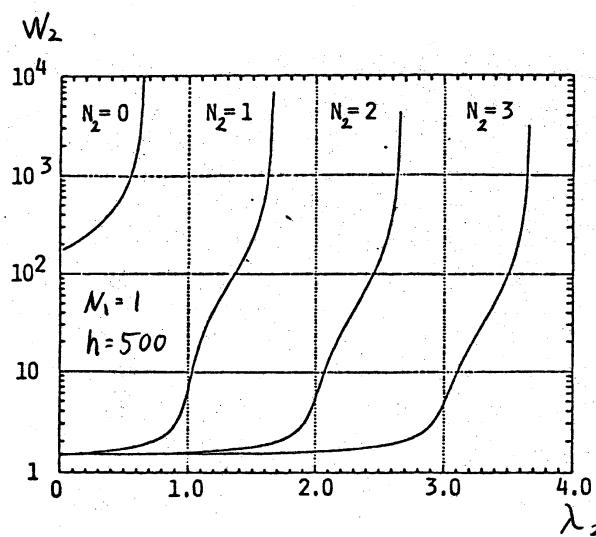


図.6

クラス2のサ - バ数を増した場合の、入力負荷 - 处理遅延特性

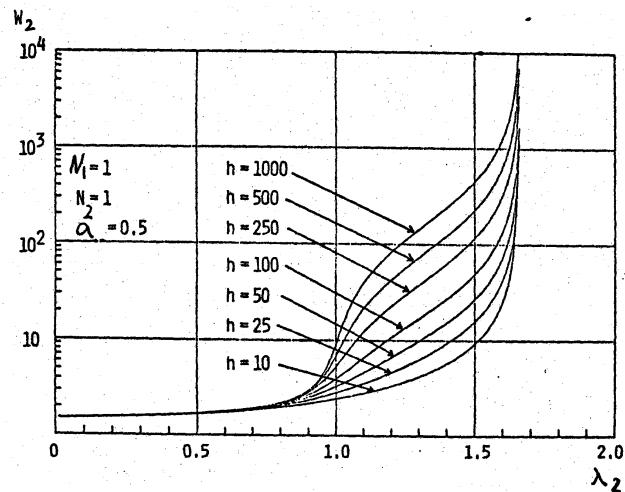


図.7

サーバ占有時間の変化  
に対する特性

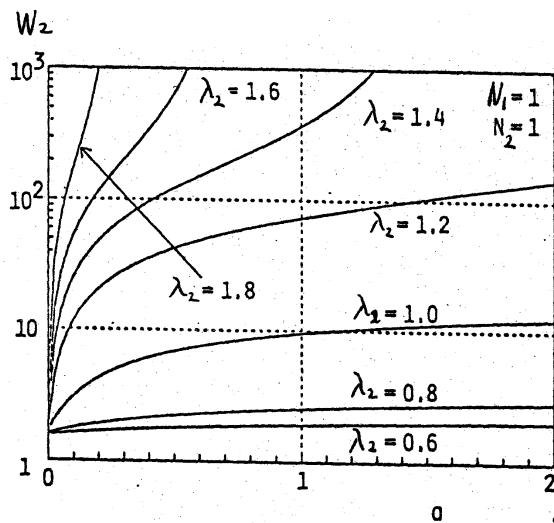


図.8

グラス1 入力負荷変化に  
対する特性

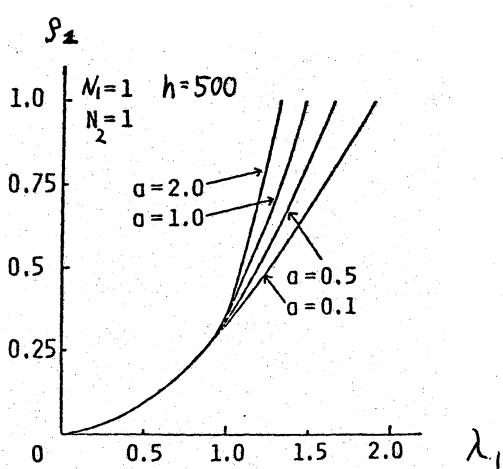


図.9

グラス2 レートに付ける  
ス1 空きサーバ利用率  $\rho_1$

著にしていい。図6に  $N_2$  を変化させた場合の特性を示す。  
 $N_2$  の変化に対し MCL効果の生じる率も変化していくことがわかる。

図7にクラス1の負荷αを一定とした時のサーカス占有時間ηに対する特性を示す。これより MCL効果のサーカス占有時間の差により生じていることが示される。図8にクラス1の負荷の変化に対するクラス2の処理遅延特性を示す。 $\lambda_2 < 1$  の時、 $\alpha$  の変化は  $W_2$  に大きな影響を与えないことがわかる。最後に、図9にクラス2の入力負荷  $\lambda_2$  に対する、クラス2のジョブのクラス1空き容量の利用率  $P_2$  の特性を示す。 $\lambda_2 < 1$  の時はクラス1の負荷αに係わりなく一定である。 $\lambda_2 > 1$  に対して  $\alpha$  の影響が生じてくることがわかる。

#### 4. むすび

処理時間加大中に異なる2種類の処理系を複合し、動的空き容量配分により効率改善を計る複合処理系のうえ、特に短時間処理系の振舞を検討した。この結果、短時間処理系の入力負荷が本来の処理容量を越える場合、異なる処理遅延特性間の遷移に起因する MCL(多段容量限界)効果が生じることを示した。

## 参考文献

- 1) I.Gitman, et al, "Issues in integrated network design", ICC, 1977.
- 2) K.Kummerle, "Multiplexor performance for integrated line and packet switched traffic", ICCC 1974.
- 3) H.Okada, Y.Tezuka, "Hybrid switching systems in computer communication networks", The 2-nd USA-Japan Comp. Conf., 1975.
- 4) G.Coviello, P.Vena, "Integration of circuit /packet switching by SENET concept", NTC 1975.
- 5) M.Fischer, T.Harris, "A model for evaluating the performance of an integrated circuit- and packet- switched multiplex structure", IEEE Trans., Vol.COM-24, No.2, 1976.
- 6) N.Jaiswal, "A bulk-service queueing problem with variable capacity", J.Roy. Stat. Soc., B23, pp143-148, 1961.
- 7) 中村, 村尾, "集団処理待ち行列の一般法", 通研実報 Vol. 17, No.8, 1968.
- 8) C.Weinstein, et al, "Data traffic performance of an integrated circuit and packet switched multiple structure", IEEE Trans., COM-28, No.6, pp873-878, 1980.

- 9) C. Weinstein, et al, "Efficient multiplexing of voice and data in integrated digital networks", ICC 1980.
- 10) M. Fischer , "Data performance in a system where data packets are transmitted during voice silence periods-single channel case", IEEE Trans., COM- 27, No.9, 1979.
- 11) 森村, 綱川 "複合交換方式の N.C.D. 近似 "
- 12) B. Maglaris , M. Schwartz , "Performance Evaluation of a Variable Frame Multiplexer for Integrated Switched Networks" IEEE Trans., COM-29, No. 6, 1981.
- 13) J.P. Lehoczky , D.P. Gaver, "Diffusion approximations for the cooperative service of voice and data messages" J. Appl. Prob. 18 PP 660-671, 1981.