

複数種類の呼に対する待ち行列システムの性能比較

岩 瀬 雅 治*

Abstract

A fundamental problem in the design of stochastic service system is that of pooling, namely, the replacement of several ingredients by a functionally equivalent single ingredient. In the queueing system this problem takes one of the two forms: pooling queues or pooling servers. In this paper three queueing models are considered. In Model 1 there are several servers and every task has a server dedicated to it, and in front of each server there is an individual queue. In Model 2, queues in Model 1 are pooled into a single queue and every server can serve any task. In Model 3 servers in Model 2 are pooled into a single server with high ability. The effect of pooling in various situation is assessed with steady state analysis.

Key Words : Queueing system, Service Facility Design

1. はじめに

確率的なサービスシステムの設計段階における基本的な問題として、システムを構成する複数個の要素を、同じ働きを持つ1つの要素に置き換えることが適切であるかを検討する問題がある。待ち行列システムにおいては、これは、到着した呼が入る queue を server 毎に設ける場

Performance of Queueing System with Multiple Class Customers

* Masaharu Iwase

Correspondence Address : Department of Business Administration, Bunkyo Women's University, 1196 Kamekubo, Oimachi, Iruma-gun, Saitama 356-8533, Japan.

Accepted October 7, 1999.

Published December 20, 1999.

合と共通な queue を設ける場合を比較する問題と、複数の server を 1 個の能力の高い server で置き換えることが適切かを検討する問題となる。[1], [4] はサービス段階が 1 段階のシステムについて、次の 3 つの基本的なモデルを考えている。サービス時間分布が指数分布で service 率 μ/m を持つ m 個の複数 server を持ち、server 毎に無限個の呼を収容できる queue があり、呼は各 server に到着率 λ/m でポアソン到着し FCFS 規則で service が行われる分散方式、分散方式において m 個の queue を共通化し 1 つの queue にした多重処理方式、多重処理方式で m 個の server をサービス時間分布が指数分布で service 率 μ の 1 個の server で置き換えた集中方式である (図 1)。そして、集中方式の平均系内数 L_C 、多重処理方式の平均系内数 L_M 、分散方式の平均系内数 L_D には、一般に $L_C \leq L_M \leq L_D$ の関係があることを示している。しかし [1], [4] では、到着する呼は 1 種類であり、複数種類の呼に対する service は考えられていない。本稿では、異なるサービスを要求する複数種類の呼が到着し、各 server は特定種類の呼に対するサービスのみを行い、各 server 毎に queue を持つ待ち行列システムについて、各 server がすべての種類の呼に対する service 能力を獲得し queue を 1 つにして並列システムになった場合と、この並列システムにおいて各 server の service 能力を 1 個の server に集積した場合について、システム能力の比較を行う。計算機システムにおいて CPU やネットワーク等の資源に対する負荷分散を考える場合や工場の製造工程を設計する場合に、性能評価の基礎となることを目的とする。

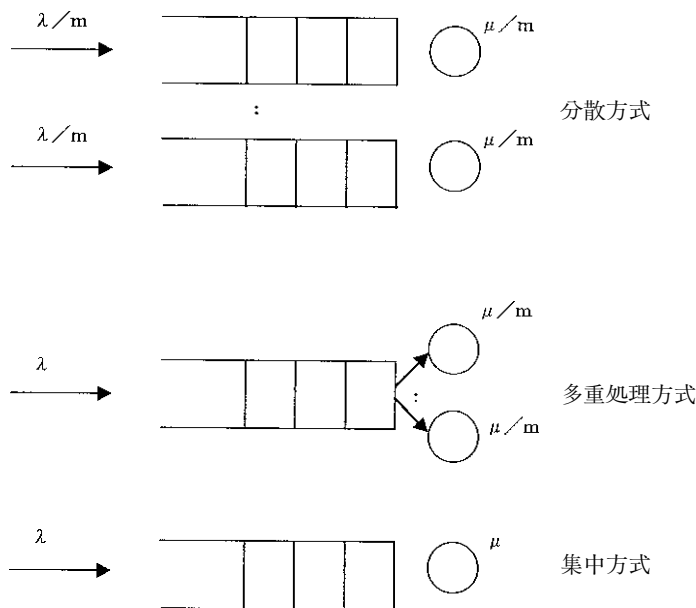


図 1 3 種の処理方式 (呼種が 1 の場合)

2. モデル

2-1 モデル 1

システム全体への呼の到着は、平均到着間隔 $1/\lambda$ のポアソン到着とする。到着した呼は、class 1から class c までのいずれかの class の呼である。到着した呼は確率 α_i で class i の呼であるとする。server 数は c とする。c 個の server を第 1 server から第 c server とする。class i の呼の service 時間分布は、平均 $1/\mu_i$ の指数分布である。第 i server は、class i の呼のみの service を行う。class i の呼は、FCFS 規則によって、第 i server の service を受け、service 終了時にシステムより退出する。モデル 1 で $\alpha_i=1/c(1\leq i\leq c)$ で、 $\mu_i(1\leq i\leq c)$ がすべて等しい場合、1. に記した分散方式になる。

2-2 モデル 2

システム全体への呼の到着は、平均到着間隔 $1/\lambda$ のポアソン到着とする。到着した呼は、class 1から class c までのいずれかの class の呼である。到着した呼は確率 α_i で class i の呼であるとする。server 数は c とする。c 個の server を第 1 server から第 c server とする。class i の呼の service 時間分布は、平均 $1/\mu_i$ の指数分布である。各 server は、すべての class の呼の service を行うことができる。到着した呼は、FCFS 規則によって、いずれかの server の service を受け、service 終了時にシステムより退出する。ここで、service 時間分布は、mixed exponential distribution であり、その確率密度関数 $g(x)$ は、 $g(x) = \sum_{i=1}^c \alpha_i \mu_i \exp(-\mu_i x)$ である。モデル 2 は、モデル 1 で、各 server がすべての class の呼に対する service 能力を獲得し、並列システムになった場合と考えられる。モデル 2 で $\alpha_i=1/c(1\leq i\leq c)$ で、 $\mu_i(1\leq i\leq c)$ がすべて等しい場合、1. に記した多重処理方式になる。

2-3 モデル 3

システム全体への呼の到着は、平均到着間隔 $1/\lambda$ のポアソン到着とする。到着した呼は、class 1から class c までのいずれかの class の呼である。到着した呼は確率 α_i で class i の呼であるとする。server 数は 1 とする。class i の呼の service 時間分布は、平均 $1/(c\mu_i)$ の指数分布である。到着した呼は、FCFS 規則によって service を受け、service 終了時にシステムより退出する。モデル 3 は、モデル 2 で、server 数を c から 1 にして、各 server の service 能力を、1 個の server に集積したシステムと考えられる。モデル 3 で $\alpha_i=1/c(1\leq i\leq c)$ で、 $\mu_i(1\leq i\leq c)$ がすべて等しい場合、1. に記した集中方式になる。

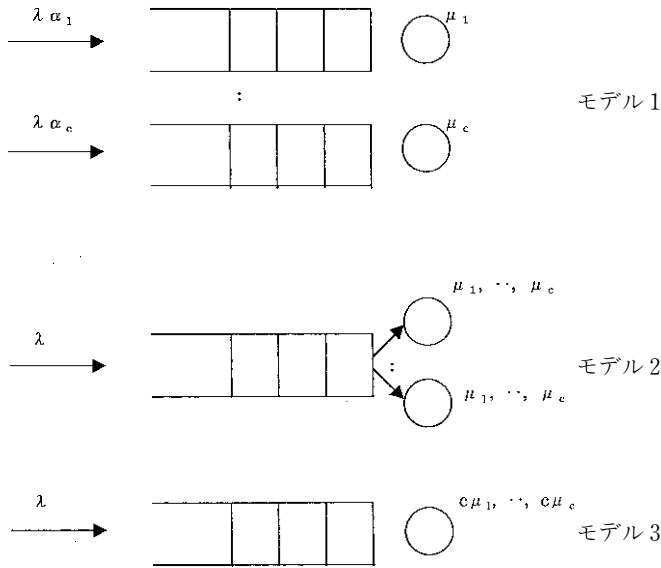


図2 3種のモデル

3. 定常状態における解析

3-1 モデル1

$\rho_i = \lambda \alpha_i / \mu_i$ とする。第*i* serverは、トラフィック密度 ρ_i の M/M/1 システムである。モデル1の平均系内数を L_1 、平均待ち数を L_1q とする。 L_1 、 L_1q は(1)、(2)で与えられる。

$$L_1 = \sum_{i=1}^c \frac{\rho_i}{1 - \rho_i} \quad (1)$$

$$L_1q = \sum_{i=1}^c \frac{\rho_i^2}{1 - \rho_i} \quad (2)$$

3-2 モデル2

系内にある呼数を i 、service を受けている class m の呼数を j_m とすると系内状態は

$$\{(i, j_1, \dots, j_c) \mid i \geq 0, \sum_{m=1}^c j_m = \min\{i, c\}\} \text{ で表される。}$$

系内状態を系内数によって可算個に分割し、各状態間の遷移確率を考えることにより、平均系内数 L_2 、平均待ち数 L_2q は、Neuts[2]のアルゴリズムを使って、数値的に求められる。以下に $c=2$ の場合の L_2 、 L_2q の求め方を示す。システム状態を次のように定める。

(i, j) : システムの系内数が i 、service 中の class 1 の呼数が j である状態。

$$(0 \leq i, 0 \leq j \leq \min\{i, 2\})$$

システムの状態を以下のように共通部分のない部分集合に分割する。

$$\text{level } 0 = \{(0, 0), (1, 0), (1, 1)\}$$

$$\text{level } i = \{(i+1, j) \mid 0 \leq j \leq 2\}, \quad i \geq 1$$

level を level 番号の昇順に並べ、level 0 に属する状態を $((0, 0), (1, 0), (1, 1))$ 、level $i (i \geq 1)$ に属する状態を $((i, 0), (i, 1), (i, 2))$ と並べることにより、すべての状態を一列に並べる。この状態の並べ方に対し、状態遷移を表す状態遷移行列を Q とする。

$$Q = \begin{pmatrix} B_{00} & B_{01} & 0 & \cdots & & & & & \\ B_{10} & A_1 & A_0 & 0 & \cdots & & & & \\ 0 & A_2 & A_1 & A_0 & 0 & \cdots & & & \\ 0 & 0 & A_2 & A_1 & A_0 & 0 & \cdots & & \\ 0 & 0 & 0 & A_2 & A_1 & A_0 & 0 & \cdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

ここで 0 は 3×3 零行列、 B_{00} , B_{01} , B_{10} , A_0 , A_1 , A_2 は $2 - 2$ より定まり、以下に示す 3×3 行列である。ここで、Neuts[2] の記法に従い $(D)_{i, j}$ によって行列 D の i 行 j 列の要素を表す。

$$A_0 = \lambda I_3 \tag{3}$$

$$(A_1)_{i, i} = -(\lambda + (i-1)\mu_1 + (2-i+1)\mu_2) \tag{4}$$

$$(A_1)_{i, j} = 0 (i \neq j) \tag{5}$$

$$(A_2)_{1, 1} = 2\mu_2\alpha_2 \tag{6}$$

$$(A_2)_{1, 2} = 2\mu_2\alpha_1 \tag{7}$$

$$(A_2)_{2, 1} = \mu_1\alpha_2 \tag{8}$$

$$(A_2)_{2, 2} = \mu_1\alpha_1 + \mu_2\alpha_2 \tag{9}$$

$$(A_2)_{2, 3} = \mu_2\alpha_1 \tag{10}$$

$$(A_2)_{3, 2} = 2\mu_1\alpha_2 \tag{11}$$

$$(A_2)_{3, 3} = 2\mu_1\alpha_1 \tag{12}$$

$$(B_{00})_{1, 1} = -\lambda \tag{13}$$

$$(B_{00})_{1, 2} = \lambda\alpha_1 \tag{14}$$

$$(B_{00})_{1, 3} = \lambda\alpha_2 \tag{15}$$

$$(B_{00})_{2, 1} = -\mu_2 \tag{16}$$

$$(B_{00})_{2, 2} = -(\lambda + \mu_2) \tag{17}$$

$$(B_{00})_{3, 1} = \mu_1 \tag{18}$$

$$(B_{00})_{3, 3} = -(\lambda + \mu_1) \tag{19}$$

$$(B_{01})3, 2 = \lambda \alpha_2 \quad (20)$$

$$(B_{01})3, 3 = \lambda \alpha_1 \quad (21)$$

$$(B_{01})2, 1 = \lambda \alpha_2 \quad (22)$$

$$(B_{01})2, 2 = \lambda \alpha_1 \quad (23)$$

$$(B_{10})1, 2 = 2\mu_2 \quad (24)$$

$$(B_{10})2, 2 = \mu_2 \quad (25)$$

$$(B_{10})2, 3 = \mu_1 \quad (26)$$

$$(B_{10})3, 3 = 2\mu_1 \quad (27)$$

(3)~(27)と Neuts[2]の方法により, iteration method を使い, システムの平均系内数 L_2 , 平均待ち数 L_2q を求められる。

3-3 モデル3

到着率は λ , service 時間分布は, mixed exponential distribution で, service 時間分布の確率密度関数は, $\sum_{i=1}^c \alpha_i \mu_i \exp(-c\mu_i x)$ の M/G/1 待ち行列システムである。Pollaczek-Khintchine の式より, 平均系内数 L_3 , 平均待ち数 L_3q は (28), (29) となる。

$$L_3 = \frac{\sum_{i=1}^c \rho_i}{c} \left(1 + \frac{\frac{\sum_{i=1}^c \rho_i}{c}}{1 - \frac{\sum_{i=1}^c \rho_i}{c}} \frac{1 + \frac{\sum_{i=1}^c \rho_i^2}{\sum_{i=1}^c \alpha_i}}{\frac{(\sum_{i=1}^c \rho_i)^2}{2}} \right) \quad (28)$$

$$L_3q = L_3 - \frac{\sum_{i=1}^c \rho_i}{c} \quad (29)$$

4. $c = 2$ の場合のモデル比較

server 数 $c = 2$ の場合について, 平均系内数, 平均待ち数の比較を行う。 $c = 2$ の場合, $\alpha_1 + \alpha_2 = 1$ より, モデル1における平均系内数 L_1 , 平均待ち数 L_1q , モデル3における平均系内数 L_3 , 平均待ち数 L_3q は (30) (31) (32) (33) となる。

$$L_1 = \frac{\rho_1}{1 - \rho_1} + \frac{\rho_2}{1 - \rho_2} \quad (30)$$

$$L_1q = \frac{\rho_1^2}{1 - \rho_1} + \frac{\rho_2^2}{1 - \rho_2} \quad (31)$$

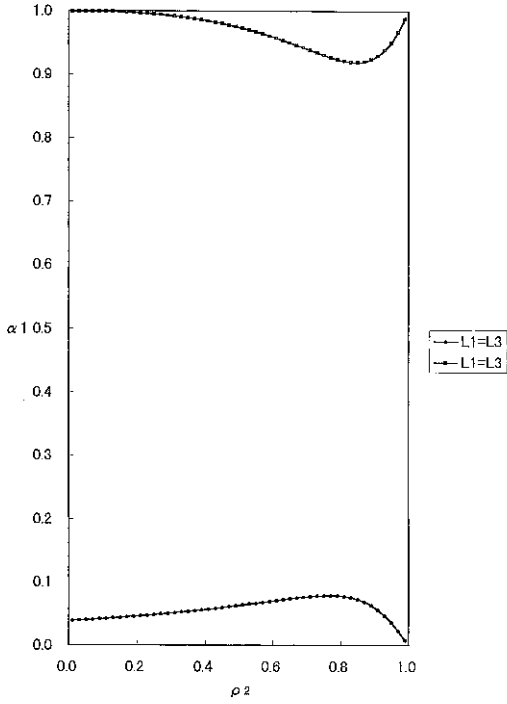


図3 最小平均系内数を与えるモデルの境界 ($\rho_1=0.8$)

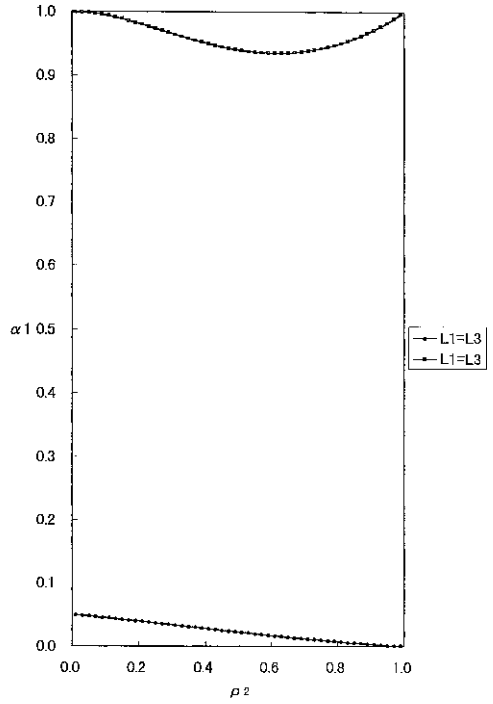


図4 最小平均系内数を与えるモデルの境界 ($\rho_2=0.3$)

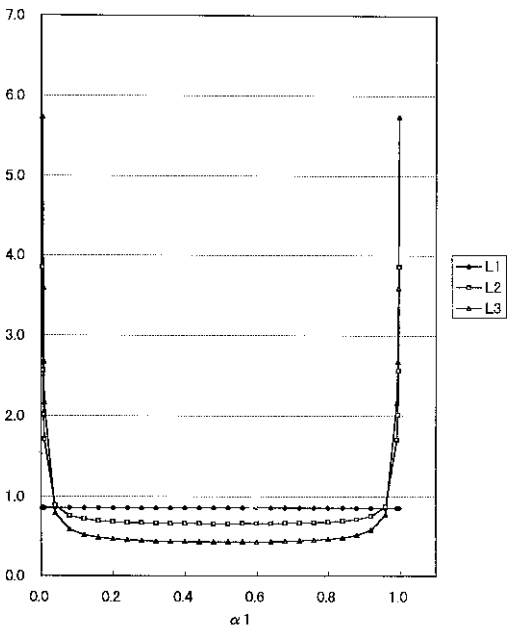


図5 平均系内数の変化 ($\rho_1=0.3, \rho_2=0.3$)

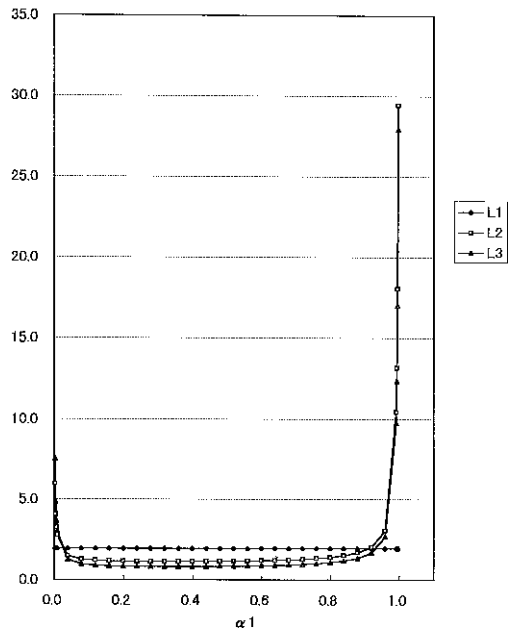


図6 平均系内数の変化 ($\rho_1=0.3, \rho_2=0.6$)

$$L_3 = \frac{\rho_1 + \rho_2}{2} \left(1 + \frac{\frac{\rho_1 + \rho_2}{2}}{1 - \frac{\rho_1 + \rho_2}{2}} \frac{1 + \frac{\rho_1^2 + \rho_2^2}{1 - \alpha_1}}{(\rho_1 + \rho_2)^2} \right) \quad (32)$$

$$L_3q = L_3 - \frac{\rho_1 + \rho_2}{2} \quad (33)$$

図3, 4に $\rho_1 = 0.3, 0.8$ の場合, $L_1 = L_3$ となる α_1 の値を示した。2つの曲線に挟まれた領域では $L_1 > L_3$ でありモデル3がモデル1より平均系内数が小さい。図3, 4で2つの曲線の上側と下側の領域では $L_1 < L_3$ でありモデル1がモデル3より平均系内数が小さい。図3, 4に示したように, モデル1とモデル3のどちらが平均系内数が小さいかは, ρ_1, ρ_2, α_1 に依存して定まる。図5, 6に ρ_1, ρ_2 が一定のとき, α_1 の変化に伴う L_1, L_2, L_3 の変化を示した。 L_1, L_2, L_3 について次の(i)(ii)(iii)が成り立つ。

- (i) 各 ρ_1, ρ_2 に対し, L_1 は α_1 の値によらず一定の値となる。
- (ii) 各 ρ_1, ρ_2 に対し, $0 < a_{(\rho_1, \rho_2)} < b_{(\rho_1, \rho_2)} < 1$ をみたす $a_{(\rho_1, \rho_2)}, b_{(\rho_1, \rho_2)}$ が存在し,
 $0 < \alpha_1 < a_{(\rho_1, \rho_2)}$ または $b_{(\rho_1, \rho_2)} < \alpha_1 < 1$ のとき $L_2 > L_1$,
 $a_{(\rho_1, \rho_2)} < \alpha_1 < b_{(\rho_1, \rho_2)}$ のとき $L_2 < L_1$ となる。
- (iii) 各 ρ_1, ρ_2 に対し, $0 < c_{(\rho_1, \rho_2)} < d_{(\rho_1, \rho_2)} < 1$ をみたす $c_{(\rho_1, \rho_2)}, d_{(\rho_1, \rho_2)}$ が存在し,
 $0 < \alpha_1 < c_{(\rho_1, \rho_2)}$ または $d_{(\rho_1, \rho_2)} < \alpha_1 < 1$ のとき $L_3 > L_1$,
 $c_{(\rho_1, \rho_2)} < \alpha_1 < d_{(\rho_1, \rho_2)}$ のとき $L_3 < L_1$ となる。

図7, 8に ρ_1, ρ_2 が一定のとき, α_1 の変化に伴う L_1q, L_2q, L_3q の変化を示した。 L_1q, L_2q, L_3q について次の(iv)(v)(vi)が成り立つ。

- (iv) 各 ρ_1, ρ_2 に対し, L_1q は α_1 の値によらず一定の値となる。
- (v) 各 ρ_1, ρ_2 に対し, $0 < e_{(\rho_1, \rho_2)} < f_{(\rho_1, \rho_2)} < 1$ をみたす $e_{(\rho_1, \rho_2)}, f_{(\rho_1, \rho_2)}$ が存在し,
 $0 < \alpha_1 < e_{(\rho_1, \rho_2)}$ または $f_{(\rho_1, \rho_2)} < \alpha_1 < 1$ のとき $L_2q > L_1q$,
 $e_{(\rho_1, \rho_2)} < \alpha_1 < f_{(\rho_1, \rho_2)}$ のとき $L_2q < L_1q$ となる。
- (vi) 各 ρ_1, ρ_2 に対し, $0 < g_{(\rho_1, \rho_2)} < h_{(\rho_1, \rho_2)} < 1$ をみたす $g_{(\rho_1, \rho_2)}, h_{(\rho_1, \rho_2)}$ が存在し,
 $0 < \alpha_1 < g_{(\rho_1, \rho_2)}$ または $h_{(\rho_1, \rho_2)} < \alpha_1 < 1$ のとき $L_3q > L_1q$,
 $g_{(\rho_1, \rho_2)} < \alpha_1 < h_{(\rho_1, \rho_2)}$ のとき $L_3q < L_1q$ となる。

定常状態確率が存在する parameter (ρ_1, ρ_2) の範囲は, モデル1では $0 < \rho_1 < 1, 0 < \rho_2 < 1$, モデル2とモデル3ではモデル1より広い領域で, $0 < \rho_1 + \rho_2 < 2, 0 < \rho_1, 0 < \rho_2$ である。単位時間に class i の呼を service する平均時間は各モデルで ρ_i , class 1 の呼に対する平均 service 時間と class 2 の呼に対する平均 service 時間の比は, $\mu_2/\mu_1 = (\alpha_2\rho_1)/(\alpha_1\rho_2)$ である。モデル2, モデル3では, (ρ_1, ρ_2) が一定の場合, α_1 または α_2 が 0 に近づくと平均系内数,

平均待ち数が大きくなる。これは2種類の呼に対するサービス時間の割合は変化せずに、それぞれの種類の呼1つに対する平均 service 時間が大きく異なった場合である。

5. 結 言

異なるサービスを要求する複数種類の呼が到着し、各 server は特定種類の呼に対するサービスのみを行い、server 毎に queue を持つ待ち行列システムについて次のようなシステムの変更を考えた。

- (1) 各 server がすべての class の呼に対する service 能力を獲得し queue を1つにして並列システムに変更した場合。
- (2) (1)の並列システムにおいて、各 server の service 能力を1個の server に集積したシステムにした場合。

そして server 数2の場合、元のシステムとこれらのシステムを定常状態において数値的に比較した。その結果、それぞれの呼の利用率と各呼の到着割合によって並列化、集積化による影響に違いがあること、各種類の呼に対する平均 service 時間が大きく異なる場合、並列化、集積化によって平均系内数、平均待ち数を小さくする効果が得られない場合があることを示した。

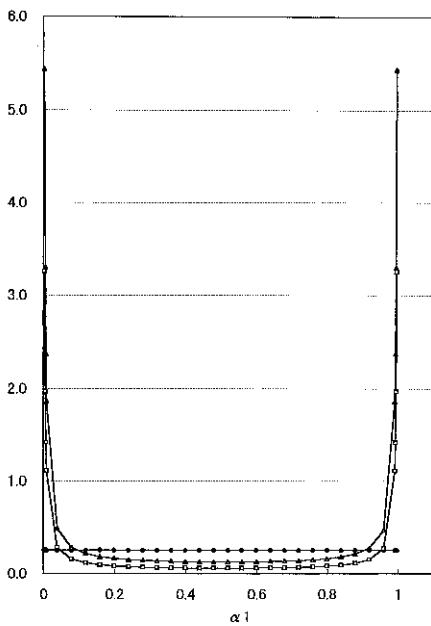


図7 平均待ち数の変化
 ($\rho_1=0.3, \rho_2=0.3$)

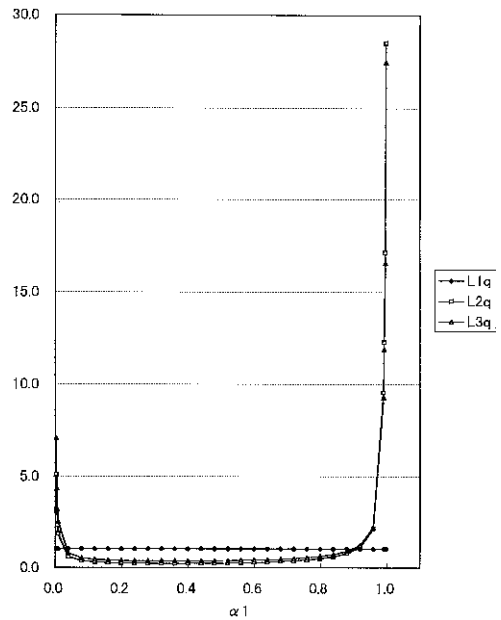


図8 平均待ち数の変化
 ($\rho_1=0.3, \rho_2=0.6$)

文 献

- [1] 亀田, 紀, 李, “性能評価の基礎と応用”, 共立出版 (1998)
- [2] Neuts, M.F., “Matrix-Geometric Solutions in Stochastic Models”, John Hopkins Univ. Press (1981)
- [3] Mandelbaum, A. and Reiman, M.I., “On Pooling Queueing Networks”, Management Sci., 44, pp971-981 (1998)
- [4] 村田, 宮原, “通信トラヒック理論とその応用(II)”, 電子情報通信学会誌, 77, pp1043-1051 (1994)
- [5] Wolff, R.W. “Stochastic Modeling and the Theory of Queues”, Prentice-Hall (1989)
- [6] 竹田, 岩瀬, “多重・並列型転送方式の性能比較”, 日本経営数学会誌, 18, pp38-47 (1996)