Electronics and Communications in Japan, Part 3, Vol. 73, No. 3, 1990
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. 71-A, No. 5, May 1988, pp. 1193-1200

Analysis of a Single Server Model with Two Queues Having Different Service Disciplines

Toshihisa Ozawa, Member

NTT Communication Switching Laboratories, Musahino, Japan 180

SUMMARY

This paper analyzes a single server model with two queues using the generating-function method. Two service-discipline models are used: a model in which the service discipline of one queue is "exhaustive" and that of the other queue is "gated"; and a model in which the service discipline of one queue is "exhaustive" and that of the other queue is "1-limited." For each model, the Laplace-Stieltjes transform (LST) of waiting time distributions and mean waiting times are derived. By using the results, the difference of the mean waiting times in the two models is shown with numerical examples. It is possible to control the characteristics (e.g., the mean waiting times in each queue and their ratio, which depend on parameters such as the traffic density) of models with mixed service disciplines by changing the combination. Therefore, it is important to analyze these models for designing an integrated system which handles different traffic characteristics and required qualities.

1. Introduction

This paper analyzes a model in which two queues of customers are served alternately by one server. The service disciplines can be classified by the number of customers in a queue served during one visit of the server. The following three disciplines are considered in this paper:

Exhaustive service: The server serves customers in a queue until the queue is empty.

Gated service: The server serves continuously only those customers in a queue who arrived before the visit.

1-limited service: The server serves only one customer in a queue who arrived before the visit.

Time required for the server to move from one queue to another is called "walking time."

A head-of-line priority model with two classes of customers is one of two-queue models which have mixed service disciplines. In this model, the queue of priority customers is serviced according to the exhaustive service, and the queue of ordinary customers is serviced according to 1-limited service, without walking time.

The degree of the priority in each queue can be represented by the ratio of the mean waiting times and this ratio can be controlled by changing the combination of its service disciplines. Therefore, it is important for designing an integrated system, in which traffic characteristics and required qualities are processed, to combine different service disciplines for each queue.

In this paper, the Laplace-Stieltjes transforms for waiting times and mean waiting times are obtained for the following two models:

(a) Model combining exhaustive service and gated service;

(b) Model combining exhaustive service and 1-limited service.

These models are analyzed in the case of zero-walking times and the case of nonzero-walking times. In the former case, a model belonging to (b), has already been analyzed as a nonpreemptive priority queue, hence it is omitted from this paper.

These models with mixed service disciplines are multiqueue systems [1, 2], and their approximate analyses for waiting times have been done in [3, 4]. But exact solutions for waiting times of the models have not been obtained. In [5], model (b) with

ISSN1042-0967/90/0003-0018$7.50/0

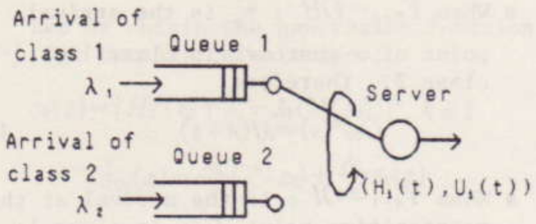


Fig. 1. A single server model with two queues.

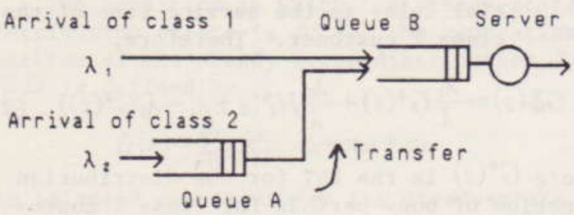


Fig. 2. A model for analysis without walking time.

walking times was analyzed by walking times from one queue served according to exhaustive service to the other queue are constant. These are variable with a general distribution in this paper.

Section 2 explains models and their notations. Section 3 analyzes models without walking times. Section 4 analyzes those with walking times, obtaining the LSTs for distribution functions of waiting times and mean waiting times. Section 5 gives numerical examples.

2. Models and Notations

For the models shown in Fig. 1, the following assumptions are made:

- There are only two queues.
- There is only one server who serves each queue alternately.
- Queue 1 is served according to exhaustive service, and queue 2 according to gated service or 1-limited service.
- Each queue has an infinity capacity.
- Customers arrive at the queues according to independent Poisson processes, and service times and walking times are independent distributed stochastic variables.

The notations are defined as follows:

$\lambda_i, i=1,2$: Arrival rate of class i customers,

$H_i(\cdot), i=1,2$: Distribution function of service times for class i customers, where h_i is the mean and $h_i^{(2)}$ is its second moment.

$U_1(\cdot)$: Distribution function of walking times from queue 1 to queue 2, where u_1 is its mean and $u_1^{(2)}$ is its second moment.

$U_2(\cdot)$: Distribution function of walking times from queue 2 to queue 1 where u_2 is its mean and $u_2^{(2)}$ is its second moment.

$$\lambda=\lambda_1+\lambda_2$$
$$u=u_1+u_2,\ u^{(2)}=u_1^{(2)}+2u_1u_2+u_2^{(2)}$$
$$\phi_i=u_i^{(2)}-(u_i)^2,\ i=1,2,$$
$$\rho_i=\lambda_i h_i,\ i=1,2,\ \rho=\rho_1+\rho_2,$$
$$w_0=(\lambda_1 h_1^{(2)}+\lambda_2 h_2^{(2)})/2$$

The Laplace-Stieltjes transform of a continuous-parameter distribution function $A(t)$ is represented by $A^*(s)$; and the generating function of a discrete-parameter distribution function $B(i)$ is represented by $\tilde{B}(z)$.

3. Analysis of a Model without Walking Times

3.1 Model

The model without walking times in which the service discipline of queue 1 is exhaustive and that of queue 2 is gated, is analyzed in this section.

3.2 A piecewise Markov process representation for the model

(1) setting of an equivalent model

Figure 2 shows an equivalent model for the analysis which consists of two queues and each queue has infinite capacity.

Class-1 customers arrive directly at queue B.

Class-2 customers arrive at queue A first and transfer to queue B. Transfers are performed by two methods: in one, immediately after all customers in queue B have been served, all customers in queue A transfer to queue B, and in the other, a class 2 customer arriving when no other customers are in the system is immediately transferred to queue B. It is assumed that no time is required for transferring the customers. The service discipline at queue B is first-in-first-out.

In this analytical model, the movement of the server is replaced by the transfers of class 2 customers. The order in which customers are served is not changed by this replacement, therefore, the equivalent model can be used to derive the mean waiting times for the original model.

(2) Representation as the piecewise Markov process

The number of customers in queue A are represented as the piecewise-Markov process [6].

Notations are defined as follows:

$N_A(t)$: Number of customers in queue A at time t

$S(t)$: State of the server at time t (= B : Busy, I : Idle)

$Y(t)=(N_A(t), S(t))$

$\Omega=\{OI, OB, 1, 2, \cdots, k, \cdots\}$: possible state space of $Y(t)$

$OI=(0, I), OB=(0, B), k=(k, B), k\geqq 1$

For the stochastic process $\{Y(t), t\geqq 0\}$ let $\{t_n, n\geqq 0\}$ be the times when the following state transitions occur, where t_0 is zero:

① $OB\to OI$ ② $OI\to OB$ ③ $k\to OB(k\geqq 1)$

From these definitions, the stochastic process $\{Y(t), t\geqq 0\}$ is a piecewise-Markov process, the time points $\{t_n, n\geqq 0\}$ are its regeneration points.

Now let us obtain the LSTs for the distribution functions of intervals between successive regeneration points. Let there be defined an embedded Markov chain for the process $\{Y(t), t\geqq 0\}$ as follows:

$$Y_n=Y(t_n-0), \quad n\geqq 0$$

The distribution functions of intervals t_n-t_{n-1}, $n=1, 2, \ldots$, which depend on Y_{n-1} are defined as follows:

$$G_y(t)=\Pr\{t_n-t_{n-1}\leqq t\,|\,Y_{n-1}=y\}, y\in\Omega$$

Let us obtain the LSTs $G_y^*(\cdot)$, $y\in\Omega$ for these functions are given as follows:

- When $Y_{n-1}=OB$: t_n is the arrival point of a customer in class 1 or class 2. Therefore,

$$G_{OB}^*(s)=\lambda/(\lambda+s) \tag{1}$$

- When $Y_{n-1}=OI$: If the arrival at the regeneration point t_{n-1} is class 1, then t_n-t_{n-1} is a busy period for class 1 customers or if that arrival is class 2 customer then that is the generalized busy period (see [7, p. 110]) of class 1 customers: its initial delay is the service time of the class 2 customer. Therefore,

$$G_{OI}^*(s)=\frac{\lambda_1}{\lambda}G^*(s)+\frac{\lambda_2}{\lambda}H_2^*(s+\lambda_1-\lambda_1G^*(s)) \tag{2}$$

where $G^*(s)$ is the LST for the distribution function of busy periods for class 1 customers, which is given by the solution of

$$G^*(s)=H_1^*(s+\lambda_1-\lambda_1G^*(s)) \tag{3}$$

- When $Y_{n-1}=k\geqq 1$: t_n-t_{n-1} is the generalized busy period for class 1 customers: k class 2 customers transfer to queue B at t_{n-1}, hence its initial delay is the sum of the service times for these class 2 customers:

$$G_k^*(s)=\{H_2^*(s+\lambda_1-\lambda_1G^*(s))\}^k, k\geqq 1 \tag{4}$$

3.3 An embedded Markov chain at regeneration points

(1) Transition probability matrix

Let us obtain the transition probability matrix $(p_{y1, y2})$.

$$p_{y1, y2}=\Pr\{Y_n=y_2|Y_{n-1}=y_1\}, y_1, y_2\in\Omega$$

o If $y_1=OB$, then $y_2=OI$; hence

$$P_{OB,OI}=1, p_{OB,y}=0, y\neq OI \tag{5}$$

o If $y_1=OI$ or $y_1=k\geqq 1$, then $N_A(t_n-0)$ is the number of class 2 customers which arrived during interval (t_{n-1}, t_n). Therefore, letting a_n be the number of class 2 customers which arrived during the interval and putting

$$\sigma_y(j)=\Pr\{a_n=j|Y_{n-1}=y\}, y\neq OB$$

the elements of the transition probability matrix are given by

$$\begin{aligned} p_{OI,OB}&=\sigma_{OI}(0)\\ p_{OI,k}&=\sigma_{OI}(k), k\geqq 1\\ p_{l,OB}&=\sigma_l(0), l\geqq 1\\ p_{l,k}&=\sigma_l(k), l,k\geqq 1 \end{aligned} \tag{6}$$

Let us obtain the generation function of $\sigma_y(\cdot)$ for further calculations:

$$\tilde{\sigma}_k(z)=[H_2^*(z_2+\lambda_1-\lambda_1 G^*(z_2))]^k,\ k\geq 1 \qquad (7)$$

$$\tilde{\sigma}_{0I}(z)=\frac{\lambda_1}{\lambda}G^*(z_2)+\frac{\lambda_2}{\lambda}\tilde{\sigma}_1(z) \qquad (8)$$

where $z_2=\lambda_2-\lambda_2 z$.

(2) Generating function of steady-state distribution for the embedded Markov chain

The condition in which the steady-state distribution exists is $\rho<1$. The generation function of the steady state distribution π_y, $y\in\Omega$ is defined by

$$\Pi(z)=\sum_{j=0}^{\infty}\pi_j z^j,\quad \pi_0=\pi_{0I}+\pi_{0B}$$

The balanced equations for the steady-state distribution are given by

$$\pi_{0I}=\pi_{0B}=\sigma_{0I}(0)\pi_{0I}+\sum_{j=1}^{\infty}\sigma_j(0)\pi_j$$

$$\pi_k=\sigma_{0I}(k)\pi_{0I}+\sum_{j=1}^{\infty}\sigma_j(k)\pi_j,\ k\geq 1 \qquad (9)$$

Using these equations, the generating function is given by

$$\Pi(z)=\{\tilde{\sigma}_{0I}(z)-1\}\pi_{0I}+\Pi\{\tilde{\sigma}_1(z)\} \qquad (10)$$

3.4 Distributions of waiting time

(1) Distribution of waiting time for class 1 customers

Let $Q_1(\cdot)$ be the generating function of the distribution for the number of class 1 customers in the system immediately after their departure. Using this, let us obtain the LST for the waiting time distribution function of class 1 customers.

$Q_1(\cdot)$ is represented by $\Pi(\cdot)$, which is given as Eq. (10). A steady-state process is assumed, and an interval $(t_n, t_{n+1}]$ for successive regeneration points is considered.

Notations are defined as follows:

- J : The number of class 1 customers departed during the interval $(t_n, t_{n+1}]$
- $t_n^{(0)}$: The time when all services for Y_n class 2 customers were done
- $t_n^{(j)}$: The departure times of class 1 customers during the interval $(t_n, t_{n+1}]$, where $t_n^{(J)}=t_{n+1}$
- v_j : The number of class 1 customers which arrived during an interval $(t_n^{(j-1)}, t_{n+1}^{(j)}]$
- X_j : The number of class 1 customers in the system immediately after time $t_n^{(j)}$

where $j=0, 1, 2, \cdots, J$

Let the distribution function of v_j be $v(i)=Pr\{v_j=i\}$; X_j satisfies $X_j=X_{j-1}-1+v_j$ (where $j=1, 2, \cdots, J$), therefore $\{X_j, j_z|\}$ is Markov chain having an absorbing state 0; J is the number of steps until the Markov chain reaches the absorbing state, hence $X_J = 0$.

From these results the following recursive expression for $\Pr\{X_j=m|X_0=k\}$ is obtained:

$$\Pr\{X_j=m|X_0=k\}=\begin{cases}0, & j\leq k \text{ and } m<k-j\\ \sum_{r=k-(j-1)}^{m+1} v(m-(r-1))\Pr\{X_{j-1}=r|X_0=k\}, & j\leq k \text{ and } m\geq k-j\\ \sum_{r=1}^{m+1} v(m-(r-1))\Pr\{X_{j-1}=r|X_0=k\}, & j>k\end{cases} \qquad (11)$$

$Q_1(\cdot)$ is given by

$$\begin{aligned}\frac{1}{C}Q_1(z)=\sum_{m=0}^{\infty}z^m[&\sum_{a=1}^{\infty}\pi_a\sum_{k=1}^{\infty}\Pr\{X_0=k|Y_n=a\}\\ &\cdot\sum_{i=1}^{\infty}\Pr\{X_i=m|X_0=k\}\\ &+\frac{\lambda_1\pi_{0I}}{\lambda}\sum_{i=1}^{\infty}\Pr\{X_i=m|X_0=1\}\\ &+\frac{\lambda_2\pi_{0I}}{\lambda}\sum_{k=1}^{\infty}\Pr\{X_0=k|Y_n=1\}\\ &\cdot\sum_{i=1}^{\infty}\Pr\{X_i=m|X_0=k\}]\end{aligned} \qquad (12)$$

where C is the normalized constant.

In the brackets of the forementioned equation: the first term represents the probability that the number of class 1 customers in the system immediately after their departure is m in the case of $Y_n=a\geq 1$; the second term represents that probability in the case where $Y_n=0I$ and a class 1 customer arrived at t_n; and the third term represents that probability in the case where $Y_n=0I$ and a class 2 customer arrived at t_n. Using Eq. (11), the following expression is obtained:

$$\begin{aligned}\frac{1}{C}Q_1(z)=\frac{H_1^*(\lambda_1-\lambda_1 z)}{z-H_1^*(\lambda_1-\lambda_1 z)}[&\Pi(H_2^*(\lambda_1-\lambda_1 z))\\ &+\frac{\lambda_1\pi_{0I}}{\lambda}z+\frac{\lambda_2\pi_{0I}}{\lambda}H_2^*(\lambda_1-\lambda_1 z)\\ &-(1+\pi_{0I})]\end{aligned} \qquad (13)$$

The value of the normalized constant C is obtained from the normalization condition for $Q_2(\cdot)$ as follows:

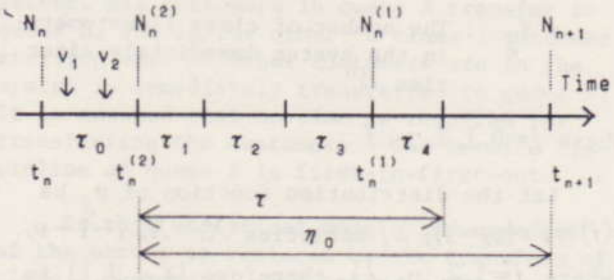


Fig. 3. Behavior of the server during $(t_n, t_{n+1}]$.

$$C=[\lambda(1-\rho)]/(\lambda_1\pi_{0I}) \tag{14}$$

From Eqs. (13) and (14), the LST $W_1^*(\cdot)$ for the waiting time distribution function for class 1 customers is obtained:

$$\begin{aligned}W_1^*(s)=&\frac{\lambda(1-\rho)}{\lambda_1\pi_{0I}\{H_1^*(s)-1+s/\lambda\}}\\&\cdot[1+\pi_{0I}-\Pi(H_2^*(s))\\&-(\pi_{0I}/\lambda)\{\lambda_1-s+\lambda_2H_2^*(s)\}]\end{aligned} \tag{15}$$

The mean waiting time w_1 is obtained by differentiating Eq. (15) with respect to s, and putting $s\to 0+$

$$w_1=[(1-\rho)w_0]/[(1-\rho)(1-\rho_1+\rho_2)] \tag{16}$$

(2) Distribution of waiting times for class 2 customers

Let $Q_2(\cdot)$ denote the generating function of the distribution for the number of class 1 customers in the system immediately after their departure. As in class 1 customers, the waiting time distribution function $W_2(\cdot)$ for class 2 customers is obtained by using $Q_2(\cdot)$.

Let us represent $Q_2(\cdot)$ by $\Pi(\cdot)$ in Eq. (10). In the following analysis, a steady-state process is assumed and an interval $(t_n, t_{n+1}]$ for successive regeneration points is considered.

Notations are defined as follows:

- t_n^j : The jth departure time of class 2 customers during interval $(t_n, t_{n+1}]$, where $t_n^{(0)}=t_n$.
- v_j : Number of class 2 customers who arrived during interval $(t_n^{(j-1)}, t_n^{(j)}]$
- X_j : Number of class 2 customers in the system just after time $t_n^{(j)}$
- $v(i)=\Pr\{v_j=i\}$.

Using these notations, $Q_2(\cdot)$ is given by

$$\begin{aligned}\frac{1}{C}Q_2(z)=&\sum_{m=0}^{\infty}z^m\Big[\frac{\lambda_2\pi_{0I}}{\lambda}v(m)\\&+\sum_{k=1}^{\infty}\pi_k\sum_{j=1}^{k}\Pr\{X_j=m|X_0=k\}\Big]\end{aligned} \tag{17}$$

where C is a normalized constant. In the brackets in the right-hand of Eq. (17); the first term represents the probability that the number of class 2 customers in the system immediately after their departure is m in the case where $Y_n=0I$ and a class 2 customer arrived at t_n; the second term represents that probability in the case of $Y_n=k\geqq 1$. Using the same procedure as for class 1, the LST $W_2^*(\cdot)$ for the waiting time distribution function for class 2 customers is obtained as follows:

$$\begin{aligned}W_2^*(s)=&(1-\rho)\\&+\frac{\lambda(1-\rho)}{\lambda_2\pi_{0I}}\frac{\Pi(H_2^*(s))-\Pi(1-s/\lambda)}{H_2^*(s)-1+s/\lambda_2}\end{aligned} \tag{18}$$

The mean waiting time is given by

$$w_2=[(1+\rho_2)w_0]/[(1-\rho)(1-\rho_1+\rho_2)] \tag{19}$$

w_1 and w_2 satisfy the following conservation law [7, p. 113]:

$$\rho_1w_1+\rho_2w_2=\rho w_0/(1-\rho) \tag{20}$$

4. Analyses of Models with Walking Times

Let $F_i(\cdot)$ denote the generating function of the distribution for the number of customers in queue i just before the time when the server arrived in this queue. Using $F_i(\cdot)$, the LST for the waiting time distribution for class i customers is given as follows:

- If queue i is serviced according to exhaustive service,

$$W_i^*(s)=\frac{1-\rho_i}{F_i'(1)}\frac{1-F_i(1-s/\lambda_i)}{H_i^*(s)-1+s/\lambda_i}$$

- If queue i is serviced according to gated service,

$$W_i^*(s)=\frac{F_i(H_i^*(s)-F_i(1-s/\lambda_i))}{F_i'(1)\{H_i^*(s)-1+s/\lambda_i\}}$$

- If queue i is serviced according to 1-limited service,

$$W_i^*(s)=\frac{F_i(1-s/\lambda_i)-F_i(0)}{(1-s/\lambda_i)\{1-F_i(0)\}}$$

where

$$F_i'(1)=\lim_{z\to 1-}(d/dz)F_i(z)$$

Therefore, $F_1(\cdot)$, and $F_2(\cdot)$ give the LST's for the waiting distributions. In the following sections, a model with mixed exhaustive and gated services is analyzed first. Then a model with mixed exhaustive and 1-limited services is analyzed.

4.1 Analysis of a model with mixed exhaustive and gated services

The procedure of the analysis is as follows: the generating function of the distribution for the number of class 2 customers in the system immediately after the server's departure from queue 1 is obtained first; $F_1(\cdot)$, and $F_2(\cdot)$ are represented by this generation function. Then substituting them into Eqs. (21) and (22), the LSTs of the waiting time distributions are obtained.

(1) Model

Exhausted service is applied to queue 1, and Gated service is applied to queue 2.

(2) Distribution of the number of customers immediately after departure of the server

Let us consider a time when the server departed from queue 1. The number of customers in queue 1 immediately after this time is always zero. Therefore, a one-dimensional Markov chain for the number of customers in queue 2 can be used for the analysis (see Fig. 3).

Notations are defined as follows:

$\{t_n\}$: Time when the server departed from queue 1

N_n : Number of customers in queue 2 immediately after t_n

$\{t_n^{(1)}\}$: Time when the server arrived at queue 1.

$N_n^{(1)}$: Number of customers in queue 1 at $t_n^{(1)}$

$\{t_n^{(2)}\}$: Time when the server arrived at queue 2

$N_n^{(2)}$: Numberof customers in queue 2 at $t_n^{(2)}$

υ_1 : Number of class 1 customers arrived during interval $(t_n, t_n^{(2)}]$

υ_2 : Number of class 2 customers arrived during interval $(t_n, t_n^{(2)}]$

τ_0 : Walking time from queue 1 to queue 2

τ_1 : Sum of service times for N_n class 2 customers

τ_2 : Sum of service times for υ_2 class 2 customers

τ_3 : Walking time from queue 2 to queue 1

τ_4 : Sum of service times for υ_1 class 1 customers

$$\tau=\tau_1+\tau_2+\tau_3+\tau_4$$

$$\eta_0=t_{n+1}-t_n^{(2)}$$

The condition that the stochastic process $\{N_n, n\geqq 0\}$ has a steady-state distribution is $\rho<1$. Assuming this condition, let us obtain the generating function $\Pi(\cdot)$ of the steady-state distribution $\{\Pi_j, j\geq 0\}$ of the stochastic process.

From Fig. (3), N_{n+1} is the number of class 2 customers arriving during interval $(t_n^{(2)}, t_{n+1}]$. Since it was assumed that the arrival processes were Poisson, the transition probability P_{ij} of the stochastic process can be represented using a conditions distribution for η_0 as follows:

$$G_k(t)=\Pr\{\eta_0\leqq t|N_n=k\}$$
$$p_{ij}=\int_0^\infty\frac{(\lambda_2 t)^j}{j!}\exp(-\lambda_2 t)dG_i(t),\quad i,j=0,1,2,\cdots \tag{24}$$

where η_0 is a generalized busy period for class 1 customers : its initial delay is τ. The LST for the conditional distribution function for τ is given as follows:

$$G_\tau(t;k)=\Pr\{\tau\leqq t|N_n=k\},\ k=0,1,\cdots$$
$$G_\tau^*(s;k)=U_1^*(\alpha)U_2^*(s)\{H_2^*(s)\}^k \tag{25}$$
$$\alpha=\lambda_1-\lambda_1H_1^*(s)+\lambda_2-\lambda_2H_2^*(s)$$

式(25)より $G_k^*(\cdot)$ が得られる.

$$G_k^*(s)=U_1^*(\beta-s+\lambda_2-\lambda_2H_2^*(\beta))\cdot U_2^*(\beta)\{H_2^*(\beta)\}^k \tag{26}$$
$$\beta=s+\lambda_1-\lambda_1G^*(s)$$

where $G^*(\cdot)$ is the LST for the distribution function of the generalized busy period for class 1 customers, and this is given by Eq. (3).

From Eq. (24), the generating function of p_{ij} with respect to j is given by

$$\sum_{j=0}^\infty z^jp_{ij}=G_i^*(\lambda_2-\lambda_2z),\quad i=0,1,2,\cdots \tag{27}$$

From Eqs. (26) and (27), and from balanced equations

$$\pi_j=\sum_{i=0}^{\infty}\pi_i p_{ij}, \quad j=0,1,2,\cdots$$

the following equation is obtained:

$$\Pi(z)=U_1^*(g+\lambda_2 z-\lambda_2 H_2^*(g))U_2^*(g)\Pi(H_2^*(g))$$
$$g=\lambda-\lambda_2 z-\lambda_1 G^*(\lambda_2-\lambda_2 z) \tag{28}$$

(3) Distributions of waiting times

The LST for the waiting time distribution function and the mean waiting time for each class are shown. In the following analysis, a steady state process is assumed.

(a) Distribution of waiting times for class 1 customers $F_1(\cdot)$ is represented by

$$F_1(z)=\sum_{j=0}^{\infty}z^j\Pr\{N_n^{(1)}=j\} \tag{29}$$

As shown in Fig. 3, $N_n^{(1)}$ is the number of class 1 customers arriving during interval $(t_n, t_n^{(1)}]$. By putting

$$\tau_a=t_n^{(1)}-t_n=\tau_0+\tau_1+\tau_2+\tau_3$$

let us obtain the LST for the distribution function of τ_a. Since $\tau_0+\tau_2$, τ_1, and τ_3 are mutually independent for given $N_n=j$, the following equation is obtained:

$$E[\exp(-s\tau_a)|N_n=j]$$
$$=U_1^*(s+\lambda_2-\lambda_2 H_2^*(s))U_2^*(s)\{H_2^*(s)\}^j$$
$$j=0,1,2,\cdots \tag{30}$$

By using this, the LST for the distribution function of τ_a is given as

$$E[\exp(-s\tau_a)]=U_1^*(s+\lambda_2-\lambda_2 H_2^*(s))\cdot U_2^*(s)\Pi(H_2^*(s)) \tag{31}$$

From this equation, $F_1(\cdot)$ is given as

$$F_1(z)=E[\exp(-(\lambda_1-\lambda_1 z)\tau_a]$$
$$=U_1^*(\lambda_1-\lambda_1 z+\lambda_2-\lambda_2 H_2^*(\lambda_1-\lambda_1 z))$$
$$\cdot U_2^*(\lambda_1-\lambda_1 z)\Pi(H_2^*(\lambda_1-\lambda_1 z)) \tag{32}$$

The LST $W_1^*(s)$ for the waiting time distribution function of class 1 customers is obtained by substituting Eq. (32) into Eq. (21). The mean waiting time w_1 is given by

$$w_1=\frac{(1-\rho_1)w_0}{(1-\rho)(1-\rho_1+\rho_2)}+\frac{(1-\rho_1)(\psi_1+\psi_2)}{2(1-\rho_1+\rho_2)u}$$
$$+\frac{(1-\rho_1)u}{2(1-\rho)}+\frac{\rho_2(1-\rho)(1+\rho_2)\psi_1}{(1-\rho_1+\rho_2)u} \tag{33}$$

(b) Distribution of waiting times for class 2 customers

$F_2(\cdot)$ is represented by

$$F_2(z)=\sum_{j=0}^{\infty}z^j\Pr\{N_n+v_2=j\} \tag{34}$$

Since N_n and v_2 are mutually independent, the following equation is obtained:

$$F_2(z)=U_1^*(\lambda_2-\lambda_2 z)\Pi(z) \tag{35}$$

By substituting Eq. (35) into Eq. (22), the LST $W_2^*(s)$ for the waiting time distribution function of class 2 customers is obtained. The mean waiting time w_2 is given by

$$w_2=\frac{(1+\rho_2)w_0}{(1-\rho)(1-\rho_1+\rho_2)}+\frac{(1+\rho_2)(\psi_1+\psi_2)}{2(1-\rho_1+\rho_2)u}$$
$$+\frac{(1+\rho_2)u}{2(1-\rho)}-\frac{\rho_1(1-\rho)(1+\rho_2)\psi_1}{(1-\rho_1+\rho_2)u} \tag{36}$$

w_1 and w_2 satisfy the following pseudoconvention law:

$$\rho_1 w_1+\rho_2 w_2=\frac{\rho w_0}{1-\rho}+\frac{\rho u^{(2)}}{2u}+\frac{\rho_2\rho u}{1-\rho} \tag{37}$$

4.2 Analyses of a model with mixed exhaustive and 1-limited services

Since the analyses procedure for this model is almost the same as before, the results only are shown below.

(1) Model

Queue 1 is served according to exhaustive service, and queue 2 according to the 1-limited service.

(2) Distribution of the number of customers immediately after departure of the server

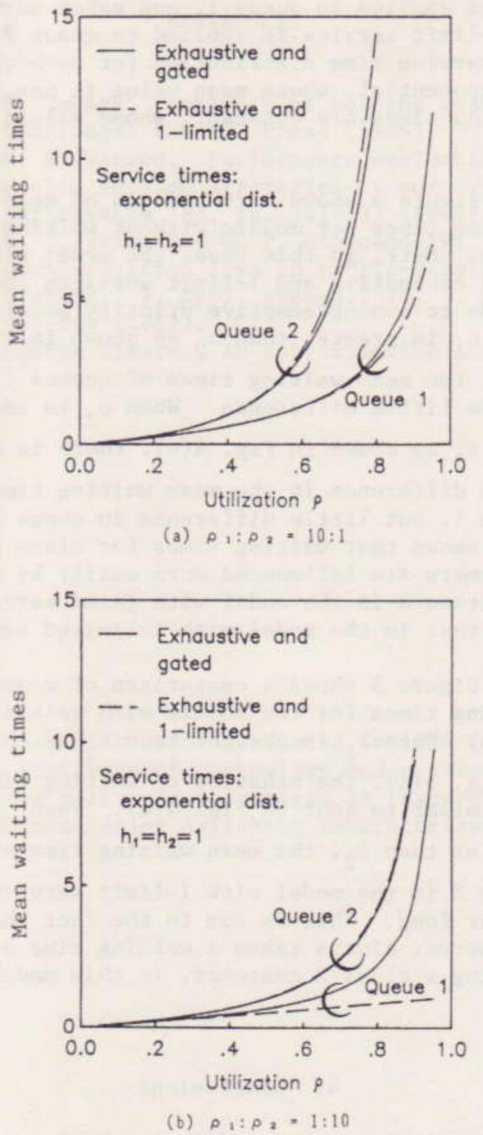
Since the number of customers in queue 1 immediately after a server's departure from this queue is always zero, the number of class 2 customers at these times can be represented as a one-dimensional Markov chain. The generating function $\Pi(z)$ of the steady-state distribution $\{\pi_j, j\geq\}$ for this Markov chain is obtained by the same method as in section 4.1. The condition that this steady-state distribution exists is given by

$$\rho+\lambda_2 u<1$$

$\Pi(z)$ is given by

$$\Pi(z)=\frac{U_1^*(g+\lambda_2 z)U_2^*(g)\{z-H_2^*(g)\}\pi_0}{z-U_1^*(g)U_2^*(g)H_2^*(g)}$$
$$g=\lambda-\lambda_2 z-\lambda_1 G^*(\lambda_2-\lambda_2 z) \tag{38}$$

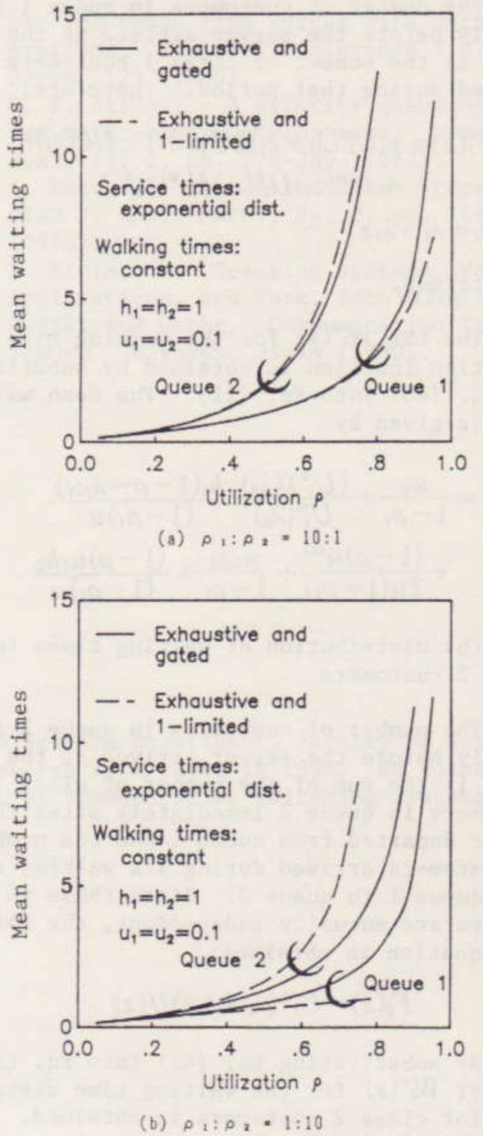
From this normalized condition for $\Pi(z)$ the value of π_0 is given by



(a) $\rho_1 : \rho_2 = 10:1$

(b) $\rho_1 : \rho_2 = 1:10$

Fig. 4. A comparison of mean waiting times (without walking times).



(a) $\rho_1 : \rho_2 = 10:1$

(b) $\rho_1 : \rho_2 = 1:10$

Fig. 5. A comparison of mean waiting times (with walking times).

$$\pi_0=(1-\rho-\lambda_2 u)/\{U_1^*(\lambda_2)(1-\rho)\} \qquad (39)$$

(3) Distributions of waiting times

By using the results of section (2), let us shown the LST for the waiting time distribution function and the mean waiting time for each class in steady state. The condition that steady state exists is $\rho_1<1$ for queue 1, and $\rho+\lambda_2 u<1$ for queue 2. Therefore, if $\rho_1<1$ and $\rho+\lambda_2 u\geqq 1$, after long run queue 1 is the steady state and queue 2 is saturated. In this case, the waiting time of queue 1 can be obtained by using the results of vacation models [8, p.302]. If $\rho+\lambda_2 u<1$, after a long run, both queue 1 and queue 2 are in steady state. In the following sections, this case is discussed.

(a) Distribution of waiting times for class 1 customers

Let us consider a period from the departure of the server from queue 1 to its return to the queue. This period depends on the number of customers in queue 2 immediately before the server arrived at the queue. If this number of customers is zero, the period is the sum of two successive walking times. Otherwise, the period is the sum of the two successive walking times and a service time for one class 2 customer.

The number of customers in queue 1 immediately before the server arrived at the queue is the number of class 1 customers who arrived during that period. Therefore,

$$F_1(z)=\pi_0 U_1^*(\lambda_2+z_1)U_2^*(z_1)\{1-H_2^*(z_1)\} + U_1^*(z_1)U_2^*(z_1)H_2^*(z_1) \quad (40)$$

$$z_1=\lambda_1-\lambda_1 z$$

is obtained.

The LST $W_1^*(s)$ for the waiting time distribution function is obtained by substituting Eq. (40) into Eq. (21). The mean waiting time is given by

$$w_1=\frac{w_0}{1-\rho_1}+\frac{(U_1^*)'(\lambda_2)}{U_1^*(\lambda_2)}\frac{h_2(1-\rho-\lambda_2 u)}{(1-\rho_1)u} +\frac{(1-\rho)u^{(2)}}{2u(1-\rho_1)}+\frac{u_2\rho_2}{1-\rho_1}+\frac{(1-\rho)u_1 h_2}{(1-\rho_1)u} \quad (41)$$

(b) Distribution of waiting times for class 2 customers

The number of customers in queue 2 immediately before the server arrived at the queue is the sum of the number of class 2 customers in queue 2 immediately after the server departed from queue 1 and the number of customers arrived during its walking time from queue 1 to queue 2. Since these two numbers are mutually independent, the following equation is obtained:

$$F_2(z)=U_1^*(\lambda_2-\lambda_2 z)\Pi(z) \quad (42)$$

By substituting Eq. (42) into Eq. (23), the LST $W_2^*(s)$ for the waiting time distribution for class 2 customers is obtained. The mean waiting time is given by

$$w_2=\frac{w_0}{(1-\rho_1)(1-\rho-\lambda_2 u)}-\frac{(U_1^*)'(\lambda_2)\rho_1(1-\rho)}{U_1^*(\lambda_2)\lambda_2 u(1-\rho_1)} +\frac{u\rho_2}{(1-\rho_1)(1-\rho-\lambda_2 u)}-\frac{\rho_1(1-\rho)u_1}{\lambda_2(1-\rho_1)u} +\frac{(1-\rho)u^{(2)}}{2(1-\rho_1)(1-\rho-\lambda_2 u)u} \quad (43)$$

w_1 and w_2 satisfy the following pseudoconvention law:

$$\rho_1 w_1+\rho_2\left(1-\frac{\lambda_2 u}{1-\rho}\right)w_2=\frac{\rho w_0}{1-\rho}+\frac{\rho u^{(2)}}{2u}+\frac{\rho_2\rho u}{1-\rho} \quad (44)$$

5. Numerical Examples

Using numerical examples, the dependence of mean waiting times on the combination of service disciplines is shown. In models discussed in this section exhaustive service is always applied to queue 1, and gated service or 1-limit service is applied to queue 2. The service time distribution for each class is exponential, whose mean value is one, and walking times are constant, whose values are 0.1.

Figure 4 shows a comparison of mean waiting times for models without walking times. Note, in this case, the model with mixed exhaustive and 1-limit services corresponds to a nonpreemptive priority model. When ρ_1 is greater than ρ_2 as shown in Fig. 4(a), the mean waiting times of queues 1 and 2 have little difference. When ρ_2 is smaller than ρ_2 as shown in Fig. 4(b), there is a large difference in the mean waiting time in queue 1, but little difference in queue 2. This shows that waiting times for class 1 customers are influenced more easily by class 2 customers in the model with gated service than that in the model with 1-limited service.

Figure 5 shows a comparison of mean waiting times for the models with walking times. When ρ_1 is greater than ρ_2 as shown in Fig. 5(a), the behavior of waiting times is similar to that in Fig. 4(a). When ρ_1 is smaller than ρ_2, the mean waiting time of queue 2 in the model with 1-limit service is rather long. This is due to the fact that the server always takes a walking time after serving a class 2 customer, in this model.

6. Conclusions

A single server model, in which queue 1 is served according to exhaustive service and queue 2 is served according to gated service or 1-limit service, was analyzed by using the generating function method; and the LSTs for waiting time distribution functions and mean waiting times were obtained. Differences between mean waiting time for class 1 customers and that for class 2 customers due to combinations of service disciplines were shown using numerical examples.

Analysis of a model in which the number of queues is expanded to N (i.e., a general case), and analysis of models combined with other service disciplines, such as K-limit service,which means that the server serves maximum K customers in a queue when it visits the queue, will be subjects of future study.

Acknowledgement. The author wishes to express his thanks to Mr. Konosuke Kawashima and Dr. Shuichi Sumita of NTT Communication Switching Laboratories for their advice.

REFERENCES

1. H. Takagi. Analysis of polling system. Cambridge, the MIT Press (1986).
2. K. S. Watson. Performance evaluation of cyclic service strategies—a survey. Performance '84, pp. 521-533 (1984).
3. O. J. Boxma and W. P. Groenendjik. Pseudoconservation laws in cyclic-service systems. J. Appl. Probab., 24, 4, pp. 949-964 (1987).
4. George Kimura. An approximation analysis for token ring system with priority stations. National Conference I.E.C.E., Japan, 125 (1986).
5. C. E. Skinner. A priority queueing system with server-walking time. Oper. Res., 15, 2, pp. 278-285 (1967).
6. A. Kuczura. Piecewise Markov processes. SIAM J. Appl. Math., 24, 2, pp. 169-181 (1973).
7. L. Kleinrock. Queueing systems, Vol. II: Applications, New York, John Wiley (1976).
8. Fujiki and Ganbe. Communication Traffic Theory. Maruzen, Japan (1980).

AUTHOR

Toshihisa Ozawa graduated in 1984 from Information Science Department, Faculty of Science, Tokyo Institute of Technology and obtained a Master's degree from there in 1986. He joined NTT the same year. He is engaged in research on evaluation of communication switching systems at NTT Communication Switching Laboratories.