Photon Netw Commun (2011) 22:209–220
DOI 10.1007/s11107-011-0320-4

An analytical model for input-buffered optical packet switches with reconfiguration overhead

Kuan-Hung Chou · Woei Lin

Received: 14 October 2010 / Accepted: 18 June 2011 / Published online: 7 July 2011

Abstract The overhead associated with reconfiguring a switch fabric in optical packet switches is an important issue in relation to the packet transmission time and can adversely affect switch performance. The reconfiguration overhead increases the mean waiting time of packets and reduces throughput. The scheduling of packets must take into account the reconfiguration frequency. This work proposes an analytical model for input-buffered optical packet switches with the reconfiguration overhead and analytically finds the optimal reconfiguration frequency that minimizes the mean waiting time of packets. The analytical model is suitable for several round-robin (RR) scheduling schemes in which only non-empty virtual output queues (VOQs) are served or all VOQs are served and is used to examine the effects of the RR scheduling schemes and various network parameters on the mean waiting time of packets. Quantitative examples demonstrate that properly balancing the reconfiguration frequency can effectively reduce the mean waiting time of packets.

Keywords Stochastic analysis · Stochastic decomposition · Discrete-time system · Batch arrival · Optical communication

K.-H. Chou · W. Lin (✉)
Department of Computer Science and Engineering, National Chung-Hsing University, No. 250 Guoguang Rd., South District, Taichung City 40227, Taiwan
e-mail: wlin@nchu.edu.tw

K.-H. Chou
e-mail: phd9604@cs.nchu.edu.tw

1 Introduction

Input-buffered packet switches using hybrid optical/electronic architecture are recently considered a promising approach for the implementation of high-speed scalable switches. The input-buffered architecture is a feasible solution to meet high-speed and large-scale switching requirements, because the architecture can internally operate at the same speed of input/output ports. Moreover, optical buffers using fiber delay lines are still expensive and bulky today and typically provide worse performance due to the dimension problem [5]. Therefore, the input-buffered packet switches using hybrid optical/electronic architecture are very interesting for future high-capacity routers and switches.

The input-buffered packet switches require high-speed and scalable optical switch fabrics for the packet transmission between buffers and output ports. Recent progress in optical switching technologies [9,16,17,20,24,29] has enabled these requirements. However, input-buffered packet switches with optical switch fabrics suffer from the overhead associated with reconfiguring a switch fabric [25–27]. Reconfiguring input and output connections in an optical switch fabric takes time, and aligning arriving packets and the optical switch fabric also takes time. Currently, optical switch fabrics in which the reconfiguration overhead is within the nanoseconds or picoseconds range are only available on a small scale, such as 2 × 2 [16,22] and 4 × 4 [24]. Most large-scale optical switch fabrics exploit technologies that are associated with a reconfiguration overhead within the milliseconds range, such as microelectromechanical systems (MEMS) [9,13,17,19,20,29]. Therefore, the reconfiguration overhead is an important issue in relation to the packet transmission time and throughput.

The reconfiguration overhead increases the mean waiting time of packets and reduces throughput, because packets cannot be transmitted across the switch fabric during the reconfiguration period. Scheduling schemes must take the reconfiguration frequency into account for improving the performance in terms of mean waiting time and throughput. Several scheduling schemes [2,14,25–27] that are based on time-slot assignment (TSA) have been presented to reduce the total reconfiguration overhead and guarantee 100% throughput. However, the time complexities of these scheduling schemes that are based on TSA fall in the range of $O(N^{2.5})$ to $O(N^{3.5})$, where N is the number of ports. At high N, these schemes give rise to prohibitively high overhead for scheduling packets in a batch.

This work uses a different method. It intends to find, through numerical analysis, the optimal reconfiguration frequency at which optical switches can transmit packets with the minimum waiting time and proposes an analytical model to facilitate the analysis on reconfiguration optimization for the input-buffered packet switches. It regards switch reconfiguration as a stochastic process and views the arrival time and the service time of each packet as a random variable. Switches are reconfigured cyclically, and packets are selected in each cycle for transmission from input buffers to outputs. The analytical model is based on Markovian analysis [1,3,4,6,11,28], and is suitable for round-robin scheduling schemes in which only non-empty VOQs [21] are served or all VOQs are served. The numerical analysis demonstrates how various scheduling schemes and network parameters affect the optimal reconfiguration frequency.

The rest of this paper is organized as follows. In Sect. 2, we introduce several scheduling schemes and the input-buffered switch architecture for the analysis on reconfiguration optimization. In Sect. 3, we propose the analytical model for the input-buffered packet switches with the reconfiguration overhead. In Sect. 4, we validate the analytical model by comparing the analytical results with the simulation results and discuss how various scheduling schemes and network parameters affect the optimal reconfiguration frequency. Finally, we conclude our discussion in Sect. 5.

2 Input-buffered switch architecture and scheduling schemes

This work considers a class of input-buffered switch architecture as shown in Fig. 1. The input-buffered switch architecture is composed of N VOQs at each input and a $N \times N$ buffer-less optical fabric. Each input and each output are interconnected by the optical fabric in a fully meshed manner. At each input, arriving packets are stored in the VOQ corresponding to the destined output. The optical fabric realizes any one-to-one mapping of inputs to outputs. We call such a mapping a switch configuration. With a configuration, the optical fabric is allowed to move packets from the specified VOQs to the corresponding outputs. The optical fabric is reconfigured periodically to vary the switch configuration.

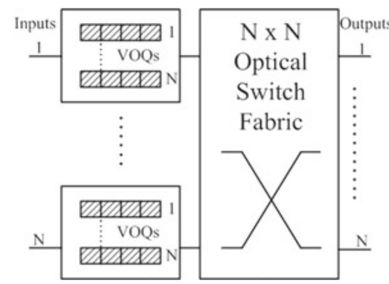


Fig. 1 Input-buffered switch architecture with VOQs

This work makes the following six assumptions to make our analytical model tractable:

1. All time durations are discretized and normalized so that the service time of a packet is a multiple of the time slot. The optical fabric forwards packets in a synchronous fashion and on the boundary of the time slot.
2. The optical fabric is periodically reconfigured and set at a regular time interval, T slots, referred to as a cycle. Each cycle comprises a reconfiguration period and a service period. Packets can be switched across the switch fabric during the service period.
3. Each reconfiguration period takes an overhead of δ slots, referred to as reconfiguration overhead. This assumption is intended to reflect all effects, such as synchronization overhead and mechanical settling times, that interrupt the transmission of packets as the optical fabric is reconfigured.
4. Each VOQ is served during a fixed-length service period of M slots and then waits during a variable-length vacation [7,10–12]. Each vacation comprises a reconfiguration period and several cycles, and the length of each vacation depends on the scheduling scheme used in the system.
5. A served packet will be preempted if the fixed-length service period of M slots ends and the packet transmission is still in progress.
6. Only those packets present in the served VOQ at the beginning of a service period can be served during the service period. The packets arriving during the service period will not be immediately served and they need to wait for the next service period.

Moreover, this work employs three round-robin (RR) scheduling scheme to study reconfiguration optimization of the input-buffered packet switches: a basic RR scheduling scheme (BRR), the iSLIP scheduling scheme [18], and a RR with empty queues skipped scheduling scheme (RREQS). In BRR, each output uses a RR counter to assign an input

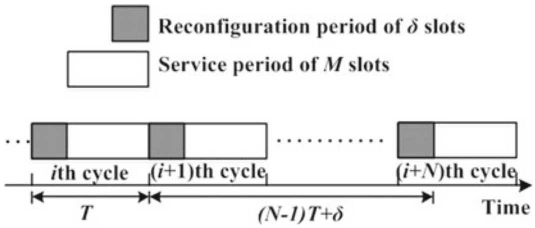


Fig. 2 A timing diagram for a basic RR scheduling scheme

to serve in circular order. The corresponding VOQ of the assigned input is served during a fixed-length service period of M slots and then waits during a vacation. The vacation is fixed-length, because the corresponding VOQ needs to wait for the other $(N-1)$ VOQs, that correspond to the other outputs, to be served during the vacation, as shown in Fig. 2.

In iSLIP and RREQS, the vacation is variable-length because only non-empty VOQs are served for reducing the unused bandwidth due to empty VOQs. However, iSLIP has a contention problem that several outputs may grant the requests of the same input but only an output will be accepted by the input. On the other hand, RREQS does not have the contention problem, because RREQS utilizes a $N^2 \times N$ buffer-less optical fabric to realize any many-to-one mapping of inputs to outputs. Simulation results of these RR scheduling schemes will be compared in Sect. 4.3.

3 Analytical model for optical packet switches with reconfiguration overhead

This section develops an analytical model for performance analysis of the input-buffered packet switches. The analytical model is based on a fixed-time service system with vacations. We employs stochastic decomposition [7,15,23] to derive equations of the fixed-time service system with vacations. We are particularly interested in calculating the mean waiting time of packets, denoted $E[W]$, and the average number of the packets in the system, denoted $E[P]$.

3.1 Basic service system with vacations

The key concept of the stochastic decomposition is that the number of packets present in the system at a random slot is distributed as the sum of two or more independent random variables, one of which is the number of packets present in the corresponding $Geo^X/G/1$ system at a random slot. Note that a $Geo^X/G/1$ system is a single-server discrete-time system with generally distributed service times and an arrival process in which more than one packet can arrive during a single slot. In this subsection, we begin with a $Geo^X/G/1$ system with vacations and use the stochastic decomposition to derive probability generating functions (PGFs) of steady-state probability distributions of the $Geo^X/G/1$ system with vacations. These PGFs and the stochastic decomposition will be used to derive PGFs of the fixed-time service system with vacations in the next subsection.

In the $Geo^X/G/1$ system with vacations, the length (measured in slots) of each service period is variable. A service period ends when there are no packets in the system and then a vacation begins. If there are packets in the system at the end of the vacation, another service period will begin. If there are still no packets in the system at the end of the vacation, another vacation will begin. We assume that the length V of each vacation is an independent and identically distributed random variable. Let $v(k)$ and $V(u)$ be the probability mass function and the PGF for V, respectively, and we have

$$v(l) \triangleq \Pr[V = l] \quad \text{for} \quad l = 1, 2, \ldots$$
$$V(u) \triangleq \sum_{l=1}^{\infty} v(l)u^l \quad \text{for} \quad |u| \le 1. \tag{1}$$

In order to derive equations of the $Geo^X/G/1$ system with vacations, we define B as the service time (measured in slots) of each packet, and each service is started and completed at exact slot boundaries. The probability distributions of B are

$$b(l) \triangleq \Pr[B = l] \quad \text{for} \quad l = 1, 2, \ldots$$
$$B(u) \triangleq \sum_{l=1}^{\infty} b(l)u^l \quad \text{for} \quad |u| \le 1, \tag{2}$$

where $b(i)$ is the probability mass function and $B(u)$ is the PGF for B. Let b and $b^{(i)}$ be the mean and the ith factorial moment of B, respectively, and we have

$$b \triangleq E[B] = B^{(1)}(1)$$
$$b^{(2)} \triangleq E[B^2] = B^{(2)}(1) + B^{(1)}(1) \tag{3}$$
$$b^{(i)} \triangleq E[B^i] \quad \text{for} \quad i = 3, 4 \ldots$$

Let Λ denotes the number of packets that arrive during a single slot. The probability distributions of Λ are defined as

$$\lambda(i) \triangleq \Pr[\Lambda = i] \quad \text{for} \quad i = 0, 1, 2, \ldots$$
$$\Lambda(z) \triangleq \sum_{i=0}^{\infty} \lambda(i)z^i, \tag{4}$$

where $\lambda(i)$ is the probability mass function and $\Lambda(z)$ is the PGF for Λ. Let λ and $\lambda^{(i)}$ be the mean and the ith factorial moment of Λ, respectively, and we have

$$\lambda \triangleq E[\Lambda] = \Lambda^{(1)}(1) = \sum_{k=0}^{\infty} \lambda(k)z^{k-1}k$$
$$\lambda^{(2)} \triangleq E[\Lambda(\Lambda-1)] = \Lambda^{(2)}(1) \tag{5}$$
$$\lambda^{(i)} \triangleq E[\Lambda(\Lambda-1)\ldots(\Lambda-i+1)] = \Lambda^{(i)}(1)$$
$$\text{for} \quad i = 3, 4 \ldots$$

Therefore, the traffic load ρ is equal to λb. Fut

Now, we take on the unfinished work U, which is defined as the remaining service time (measured in slots) of the served packets, immediately after an arbitrary slot in the $Geo^X/G/1$ system with vacations. According to the stochastic decomposition, U is distributed as the sum of the unfinished work $U_{Geo^X/G/1}$ immediately after an arbitrary slot in the corresponding $Geo^X/G/1$ system and an independent random variable Y. Boxma and Groenendijk [3] show that Y is the service time of the packets in the system at the beginning of an arbitrary slot during a vacation. Y is also the service time brought to the system by the packets that arrive until the arbitrary slot during a vacation because the system is empty at the beginning of a vacation. Let $U(z)$, $U_{Geo^X/G/1}(z)$, and $Y(z)$ be the PGFs of U, $U_{Geo^X/G/1}$, and Y, respectively, and we have

$$U(z) = Y(z)U_{Geo^X/G/1}(z),$$
$$\text{where } Y(z) = \frac{1 - V[\Lambda(B(z))]}{E[V][1 - \Lambda(B(z))]} \tag{6}$$
$$\text{and } U_{Geo^X/G/1}(z) = \frac{(1-\rho)(1-z)\Lambda[B(z)]}{\Lambda[B(z)] - z}.$$

Differentiating $U(z)$ with respect to z and taking the limit as $z \to 1$, we can find that the expected values of U, $U_{Geo^X/G/1}$, and Y are

$$E[U] = E[U_{Geo^X/G/1}] + E[Y]$$
$$E[Y] = \frac{\rho E[V(V-1)]}{2E[V]} \tag{7}$$
$$E[U_{Geo^X/G/1}] = \frac{\lambda b^{(2)} + \lambda^{(2)} b^2 + \rho(1-2\rho)}{2(1-\rho)}.$$

In a similar manner of computing $U(z)$, [23] shows that the PGF of the system size immediately after an arbitrary slot is

$$P(z) = \chi(z)P_{Geo^X/G/1}(z),$$
$$\text{where } \chi(z) = \frac{1 - V[\Lambda(z)]}{E[V][1 - \Lambda(z)]} \tag{8}$$
$$\text{and } P_{Geo^X/G/1}(z) = \frac{(1-\rho)(1-z)B[\Lambda(z)]}{B[\Lambda(z)] - z}.$$

Note that $P_{Geo^X/G/1}(z)$ is the $P(z)$ of the corresponding $Geo^X/G/1$ system and $\chi(z)$ is the PGF of the number of the packets present in the system at the beginning of an arbitrary slot during a vacation. Differentiating (8) with respect to z and taking the limit as $z \to 1$, we obtain the average number of packets in the system

$$E[P] = E[\chi] + E[P_{Geo^X/G/1}],$$
$$\text{where } E[\chi] = \frac{\lambda E[V(V-1)]}{2E[V]} \tag{9}$$
$$\text{and } E[P_{Geo^X/G/1}] = \frac{\lambda^2 b^{(2)} + \lambda^{(2)} b - \lambda\rho}{2(1-\rho)} + \rho.$$

From Little's theorem [8], we have $E[W] = E[P]/\lambda$.

Note that $E[Y] = bE[\chi]$ because Y is the service time and χ is the number of the packets in the system at the beginning of an arbitrary slot during a vacation, respectively. Takagi [23] shows a corresponding equation, $Y(z) = \chi[B(z)]$. From (7) and (9), we have

$$E[P] = \frac{E[Y]}{b} + E[P_{Geo^X/G/1}]. \tag{10}$$

3.2 Fixed-time service system with vacations

In this subsection, we deal with a fixed-time service system with vacations. We assume that the length of each service period is fixed of M slots even if packets in the system may not need M slots for transmission. In this service system, only those packets present in the system at the beginning of a service period are served during the service period. As the current service period ends and a packet transmission is not completed, the packets are preempted and the system is forced to take a vacation. After the vacation is over, the next service period begins and the system resumes the transmission of the preempted packet.

In the fixed-time service system with vacations, the unfinished work U immediately after an arbitrary slot is distributed as the sum of $U_{Geo^X/G/1}$ and Y. It is worth mentioning that the fixed-time service system with vacations may have an unfinished work, denoted as U^e, at the beginning of each vacation. Therefore, Y in the fixed-time service system with vacations is distributed as the sum of U^e and the service time brought to the system by the packets that arrive until an arbitrary slot during a vacation.

Therefore, according to the stochastic decomposition, the PGF $U(u)$ of the unfinished work immediately after an arbitrary slot boundary in the fixed-time service system with vacations can be computed by

$$U(u) = U_{Geo^X/G/1} \times U^e(u) \times \frac{1 - V[\Lambda[B(u)]]}{E[V](1 - \Lambda[B(u)])}, \tag{11}$$

where $U^e(u)$ denotes the PGF of the unfinished work U^e. The third factor of the r.h.s. of (11) is the service time brought to the system by the packets that arrive until an arbitrary slot during a vacation. From (11), the average unfinished work immediately after an arbitrary slot boundary can be computed by

$$E[U] = E[U_{Geo^X/G/1}] + E[U^e] + \frac{\rho E[V(V-1)]}{2E[V]} \tag{12}$$
$$\text{and } E[U_{Geo^X/G/1}] = \frac{\lambda b^{(2)} + \lambda^{(2)} b^2 + \rho(1-2\rho)}{2(1-\rho)},$$

where $E[U^e]$ is the mean unfinished work at the beginning of each vacation. From (10) and (12), the average number of packets can be computed by

$$E[P] = E\left[P_{Geo^X/G/1}\right] + \frac{E\left[U^e\right]}{b} + \frac{\lambda E[V(V-1)]}{2E[V]}. \quad (13)$$

Applying Little's theorem to (13), we have

$$E[W] = \frac{E\left[P_{Geo^X/G/1}\right]}{\lambda} + \frac{E\left[U^e\right]}{\lambda b} + \frac{E[V(V-1)]}{2E[V]}. \quad (14)$$

Now, we come to a point where $E\left[U^e\right]$ must be computed in order to solve $E[U]$, $E[P]$, and $E[W]$. This will be shown in the next subsection.

3.3 Computational method of the unfinished work $E[U^e]$

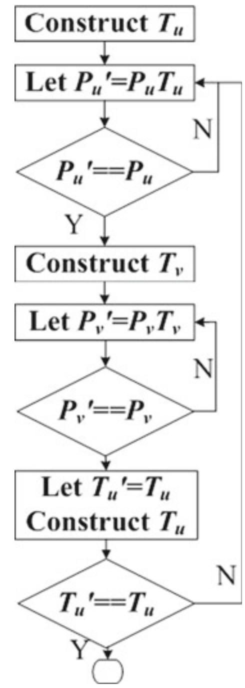
In this subsection, we present a computational method for computing $E[U^e]$ of the fixed-time service system with vacations. The computational method is based on the construction of two Markov chains described as follows.

Let us define u_k as the steady-state probability that the unfinished work U^e at the beginning of each vacation is equal to k slots in the first Markov chain. If the sum of U^e and the service time of the packets arriving during the vacation is less than or equal to M slots, U^e and the service time will be completely served during the service period. If the sum of U^e and the service time of the packets arriving during the vacation exceeds M slots, only M slots out of U^e and the service time will be served during the service period. Let $\lambda_v(n, i)$ be the probability that the service time of the packets arriving during a vacation is i slots and the length of the vacation is n slots, and let $\lambda_s(i)$ be the probability that the service time of the packets arriving during a service period of M slots is i slots. If the lengths of each vacation and each service period are independent and identically distributed, $\lambda_v(n, i)$ and $\lambda_s(i)$ can be calculated with the given $\lambda(i)$ and $b(i)$. The state transition probabilities from u_k to $u_{k'}$ can be calculated:

$$\begin{cases} \sum_{n=0}^{\infty} v(n)\{\lambda_s(0) \sum_{i=0}^{M-k} \lambda_v(n,i)\} & \text{if } k' = 0 \\ \sum_{n=0}^{\infty} v(n)\{\lambda_s(k') \sum_{i=0}^{M-k} \lambda_v(n,i) + \sum_{j=\max(0,M-k+1)}^{M+k'-k} \\ \quad \lambda_s(M+k'-k-j)\lambda_v(n,j)\} & \text{if } k' > 0. \end{cases} \quad (15)$$

Let P_u denotes the vector version of u_k and let T_u denotes the transition probability matrix constructed from the state transition probabilities. u_0 is initialized with one and u_k is initialized with zero for $k = 1, 2, 3, \ldots$. Essentially, a Markov chain is a stochastic process with the Markov property, meaning that the next state depends only on the current state but not on any past states. Therefore, our method recursively updates P_u through computing $P_u = P_u T_u$ until P_u converges. Finally, $E[U^e]$ can be calculated by $\sum_{k=0}^{\infty} u_k k$, and $E[P]$ and $E[W]$ can be calculated by (13) and $E[W] = E[P]/\lambda$, respectively.

Fig. 3 The flowchart of the computational method with the nested loop structure



Note that the probability distribution of $v(n)$ may depend on the packet scheduling scheme used in the system. The computational method considers the scheduling schemes in which the probability distributions of $v(n)$ and u_k depend on each other. Let $v(n)$ denotes the steady-state probability that the length of vacation is equal to n slots in the second Markov chain. Let P_v denotes the vector version of $v(n)$, and let T_v be the transition probability matrix that is determined by the packet scheduling scheme used in the system. $v(\delta)$ is initialized with one and $v(n)$ is initialized with zero for $n \neq \delta$. The computational method uses a nested loop structure to recursively updates P_u and P_v at the same time and consists of the steps as shown in Fig. 3. Since the probability distributions of $v(n)$ and u_k depend on each other, the computational method uses T_u to confirm if the two Markov chains converge at the same time.

We present a method to calculate the transition probabilities of $v(n)$ and construct the transition probability matrix T_v for RREQS. The method is also effective for similar RR scheduling schemes. The method constructs T_v by multiplying $(N-1)$ transition probability matrices because the transition probabilities of $v(n)$ are determined by u_k of the other $(N-1)$ VOQs in the same input. If all the other VOQs are non-empty, the length of the next vacation will be $(N-1)T + \delta$ slots; if one of the other VOQs is empty and there is no packet arriving during the vacation, the one will be skipped and the next vacation will be cut short by T slots. For example, a VOQ has waited during a vacation and the length of the vacation is $(N-1-l)T + \delta$ slots, where $l < (N-1)$. The length of the next vacation will be cut short to $(N-2-l)T + \delta$ if the first one of the other VOQs is empty and there is no packet arriving during the vacation; the length of the next vacation will be kept if the first one of the other VOQs is non-empty or there are packets arriving during the vacation. Therefore, the transition probability from $v(nT+\delta)$ to $v\left(n'T+\delta\right)$ in the first $(N-1-l)$ transition probability matrices can be computed by

$$\begin{cases} 1-u_0\lambda_v(nT+\delta,0) & \text{if } n'=n>0 \\ u_0\lambda_v(nT+\delta,0) & \text{if } n'=n-1\geqslant 0 \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

Since n' must be greater than or equal to zero, the transition probability from $v(nT+\delta)$ to $v(n'T+\delta)$ in the last l transition probability matrices is computed by

$$\begin{cases} 1-u_0\lambda_v(nT+\delta,0) & \text{if } n'=n+1\leqslant(N-1) \\ u_0\lambda_v(nT+\delta,0) & \text{if } n'=n<(N-1) \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

Let $D(l)$ denotes a diagonal matrix where all elements are equal to zero but the element, $d_{l,l}$, is equal to one, and let $T_v'^{(m)}(l)$ denote the mth transition probability matrix, where the targeted VOQ has waited during a vacation and the length of the vacation is $(N-1-l)T+\delta$ slots, constructed from (16) and (17). In summary, T_v can be computed by $\sum_{l=0}^{N-1}D(l)\prod_{m=1}^{N-1}T_v'^{(m)}(l)$ for RREQS, and then $E[U^e]$, $E[P]$, and $E[W]$ can be computed by the computational method with the nested loop structure.

4 Numerical analysis

In this section, we employ our analytical and simulation models to study reconfiguration optimization for the input-buffered optical packet switches. First, we present our simulation model. Second, we validate our analytical model by comparing the analytical and simulation results. Last, we examine how various network parameters affect the optimal choice of the reconfiguration frequency for minimizing the mean waiting time $E[W]$ of packets.

4.1 Simulation model

Our simulation model uses an input-buffered switch, as shown in Fig. 1 and a traffic model. The input-buffered switch has been presented in Sect. 2. The traffic model is composed of an arrival process and a service time distribution. The arrival process follows the binomial distribution in which the number of packets that arrive during a slot is i with probability $\binom{X}{i}(\frac{\lambda}{X})^i(1-\frac{\lambda}{X})^{X-i}$, and the service time distribution follows the geometric distribution in which the probability of success on each trial is p and the mean service time b of each packet is $\frac{1}{p}$.

Furthermore, our simulation model uses BRR, iSLIP [18], and RREQS to study reconfiguration optimization and uses RREQS to verify our analytical model. $v(n)$ and u_k of RREQS are computed by the computational method that is presented in Sect. 3.3.

4.2 Verification of the analytical model

In this subsection, we validate our analytical mode by comparing the analytical results with the simulation results. We assume $N=8$, $M=5$, and $\delta=1$ and calculate $E[U^e]$ and $E[W]$ for a series of ρ ranging from 0.1 to 0.8. Figures 4 and 5 illustrate $E[U^e]$ and $E[W]$ of the analytical and simulation results in RREQS with various X and b. They show that the analytical results well match the simulation results in terms of $E[U^e]$ and $E[W]$. Obviously, the analytical results also well match the simulation results in terms of $E[P]$ because $E[W]=E[P]/\lambda$. Therefore, the simulation results verify the accuracy of our analytical model.

Note that X does not directly affect $E[W]$, but b does, according to (14). Obviously, $E[P_{Geo^X/G/1}]$ increases with b and ρ and ρ also increases with b. Figure 5 shows the consistent results with this situation. Although the second factor of the r.h.s. of (14), $E[U^e]/b$, decreases as b increases, Figs. 4 and 5 also show that $E[U^e]$ is far smaller than $E[W]$. Therefore, $E[W]$ significantly increases with ρ and b.

Furthermore, we fix ρ and calculate $E[W]$ for a series of T ranging from 2 to 16 slots. Figures 6, 7, and 8 illustrate $E[W]$ of the analytical and simulation results in RREQS with various X, b, and ρ. They show that the analytical results still well match the simulation results.

4.3 Reconfiguration optimization

Now, we intend to find the optimal reconfiguration frequency F_{opt}, which minimizes $E[W]$, through numerical analysis. First, we fix ρ and $\delta=1$ and calculate $E[W]$ for a series of the reconfiguration frequency F ranging from 1/16 to 1/2 slots, where T ranges from 16 to 2 slots, to find F_{opt}, and then, we use surface charts of F_{opt} to further investigate how to determine F_{opt} under a wide range of network configurations. Second, we compare simulation results of BRR, iSLIP, and RREQS. This comparison helps us find out the difference between the effects of a RR scheduling scheme in which all VOQs are served and a RR scheduling scheme in which only non-empty VOQs are served on $E[W]$ and F_{opt}. Finally, we make several important observations on F_{opt}.

In the case of $\rho=0.3$ and $N=8$, Fig. 6 shows that $F_{\text{opt}}=1/4$ when $b=4$ and shows that $F_{\text{opt}}=1/5$ when $b=8$; X does not obviously affect F_{opt}. In the case of $\rho=0.5$ and $N=8$, Fig. 7 shows that $F_{\text{opt}}=1/4$ when $b=4$. It also shows that $F_{\text{opt}}=1/6$ when $b=8$; X still does not obviously affect F_{opt}. Figures 6 and 7 show that X slightly affects F_{opt} and F_{opt} decreases as b and ρ increase.

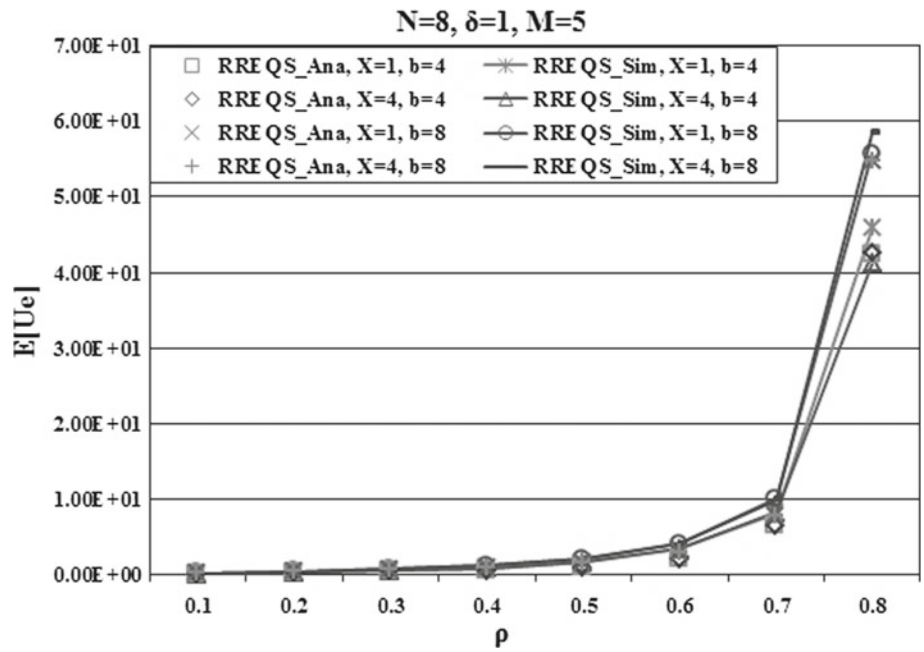


Fig. 4 $U^e(u)$ of analysis and simulation results in RREQS with various X and b

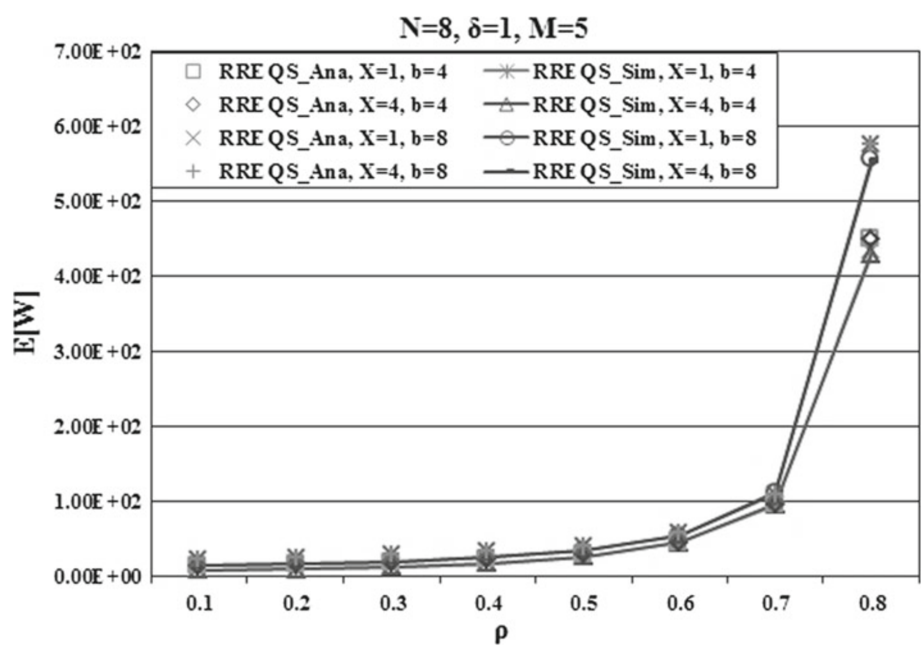


Fig. 5 $E[W]$ of analysis and simulation results in RREQS with various X and b

Figure 8 illustrates $E[W]$ of the analytical and simulation results in RREQS with various X and b under $\rho = 0.5$ and $N = 12$. Figure 7 shows that $F_{\text{opt}} = 1/4$ when $b = 4$ and $N = 8$, and Fig. 8 shows that $F_{\text{opt}} = 1/5$ when $b = 4$ and $N = 12$. Figures 7 and 8 show that F_{opt} decreases as N increases. This is because the length V of each vacation in RREQS is determined by u_k of the other $(N-1)$ VOQs in the same input. If all the other VOQs are non-empty, the length of the next vacation will be $(N-1)T + \delta$ slots. According to (13) and (14), $E[P]$ and $E[W]$ increase with V and V increases with N. When $E[W]$ increases, we need a larger T and a smaller F to reduce $E[U^e]$ for minimizing $E[W]$. However, V also increases with T. Thus, F_{opt} only slightly decreases as N increases.

We use surface charts of F_{opt} to further examine how various network parameters affect F_{opt} under a wide range of network configurations. Figure 9 is a surface chart of F_{opt} for various ρ and b. It shows that F_{opt} decreases as b and ρ increase and shows that F_{opt} is smaller than or equal to one-third. For example, when $b = 2$ and $\rho = 0.1$, F_{opt} are equal to $1/3$. When $b = 2$ and $\rho = 0.7$, F_{opt} is equal to $1/6$. When $b = 8$ and $\rho = 0.7$, F_{opt} are equal to $1/7$. On

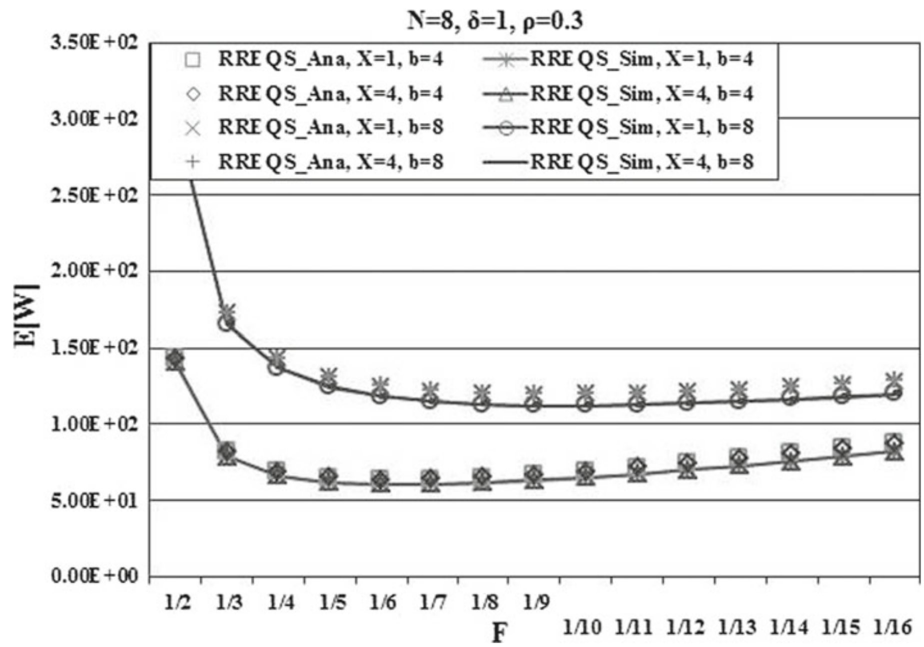


Fig. 6 $E[W]$ of analysis and simulation results in RREQS with various X and b under $\rho = 0.3$

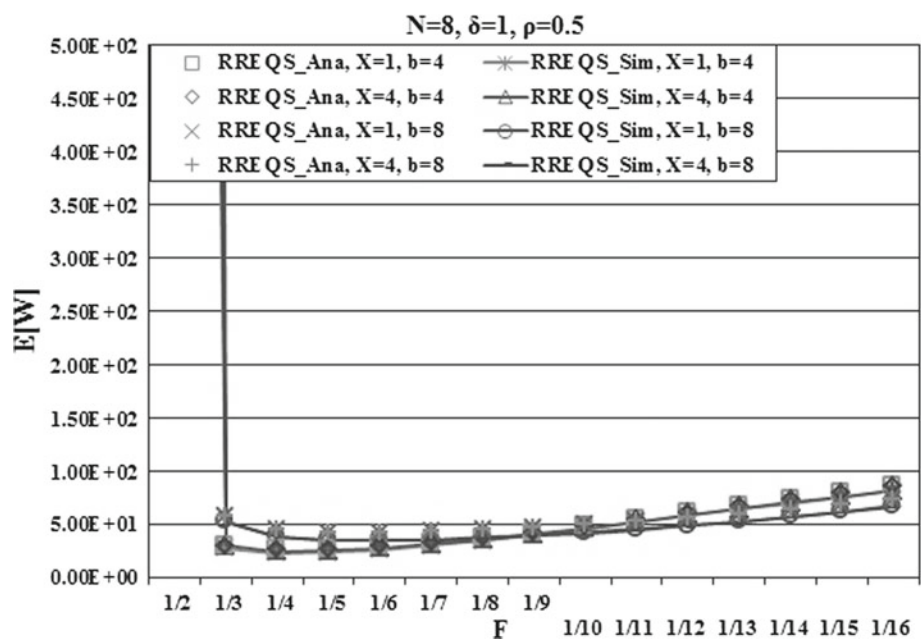


Fig. 7 $E[W]$ of analysis and simulation results in RREQS with various X and b under $\rho = 0.5$

the other hand, Fig. 10 is a surface chart of ρ_{opt} for various ρ and X. It shows that X has a minor effect on F_{opt}. These surface charts confirm our observations made in the previous two paragraphs.

Next, we compare the simulation results of BRR, iSLIP, and RREQS. In the case of $\rho = 0.3$ and $N = 8$, Fig. 11 shows that iSLIP and RREQS outperform BRR in the case of $E[W]$. This is because the RR scheduling schemes in which only non-empty VOQs are served can reduce unused bandwidth due to empty VOQs. Furthermore, Fig. 11 shows that F_{opt} of BRR is equal to $1/6$ when $b = 4$ and shows that F_{opt} of iSLIP and RREQS are equal to $1/4$. Therefore, F_{opt} in the RR scheduling schemes in which only non-empty VOQs are served is greater than the one of BRR in which all VOQs are served.

However, iSLIP has a contention problem that several outputs may grant the requests of the same input but only an output will be accepted by the input. In the case of $\rho = 0.5$ and $N = 8$, Fig. 12 shows that $E[W]$ in iSLIP may be larger than the one in BRR when ρ increases. By contrast, RREQS still outperforms BRR in the case of $E[W]$ because RREQS does not have the contention problem.

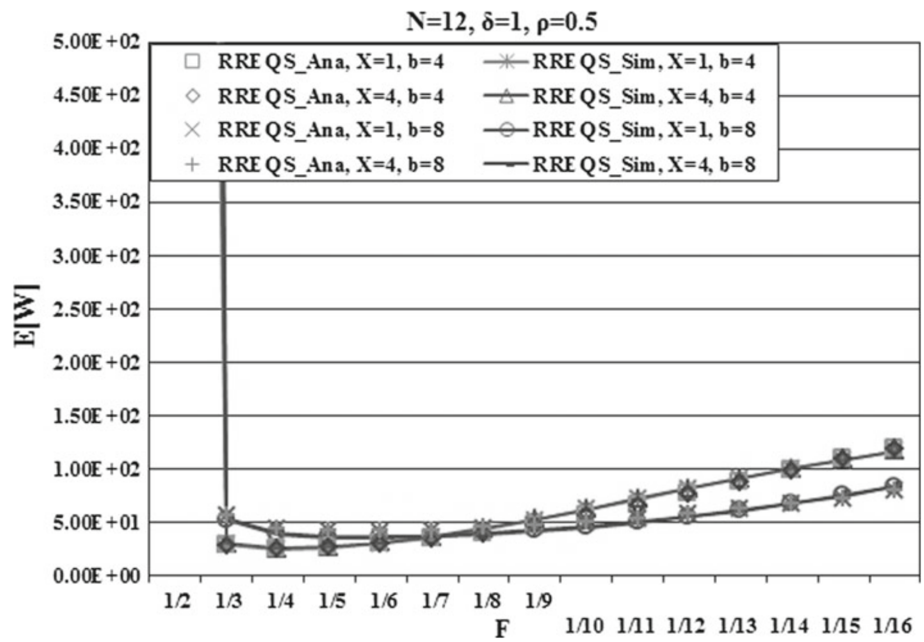


Fig. 8 $E[W]$ of analysis and simulation results in RREQS with various X and b under $\rho = 0.5$ and $N = 12$

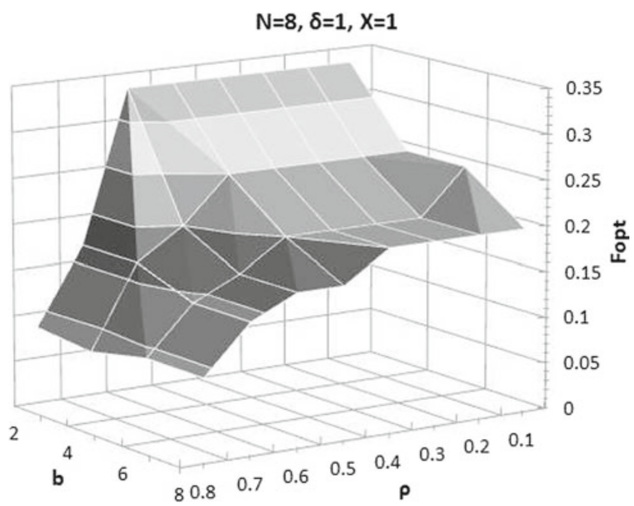


Fig. 9 A surface chart of F_{opt} for various ρ and b

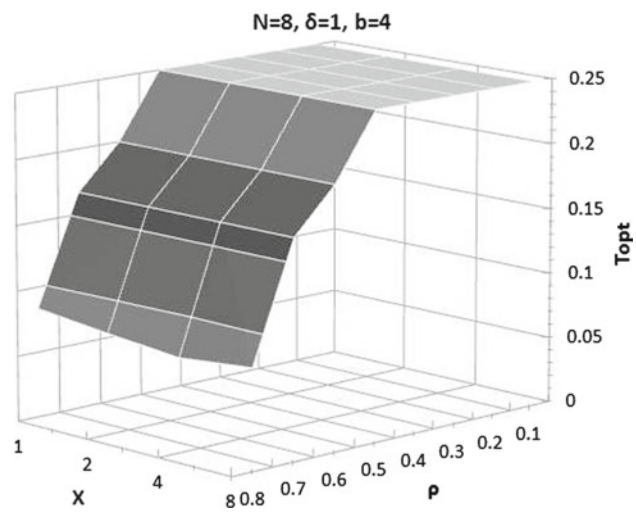


Fig. 10 A surface chart of F_{opt} for various ρ and X

In summary, the reconfiguration frequency can be optimized based on the following observations for minimizing $E[W]$

1. F_{opt} slightly decreases as N increases;
2. F_{opt} decreases as b and ρ increase;
3. F_{opt} in the RR scheduling scheme in which only non-empty VOQs are served is greater than that of BRR in which all VOQs are served.

We can dynamically adapt the value of F_{opt} based on the three observations listed above to minimize $E[W]$. Figure 13 shows that RREQS and iSLIP with F_{opt} outperform RREQS and iSLIP with $F = 1/10$ in the case of $E[W]$, respectively. Under a light traffic $\rho = 0.1$ and $b = 8$, RREQS with F_{opt} achieves the mean waiting time of packets 28% less than RREQS with $F = 1/10$. Even in the case of a heavy traffic load $\rho = 0.7$, the mean waiting time of packets of RREQS with F_{opt} is still 3% less than that of RREQS with $F = 1/10$.

5 Conclusions

This work has presented an analytical model to analyze the performance of the input-buffered packet switches for minimizing the mean waiting time of packets. The analytical model is based on a fixed-time service system with vacations

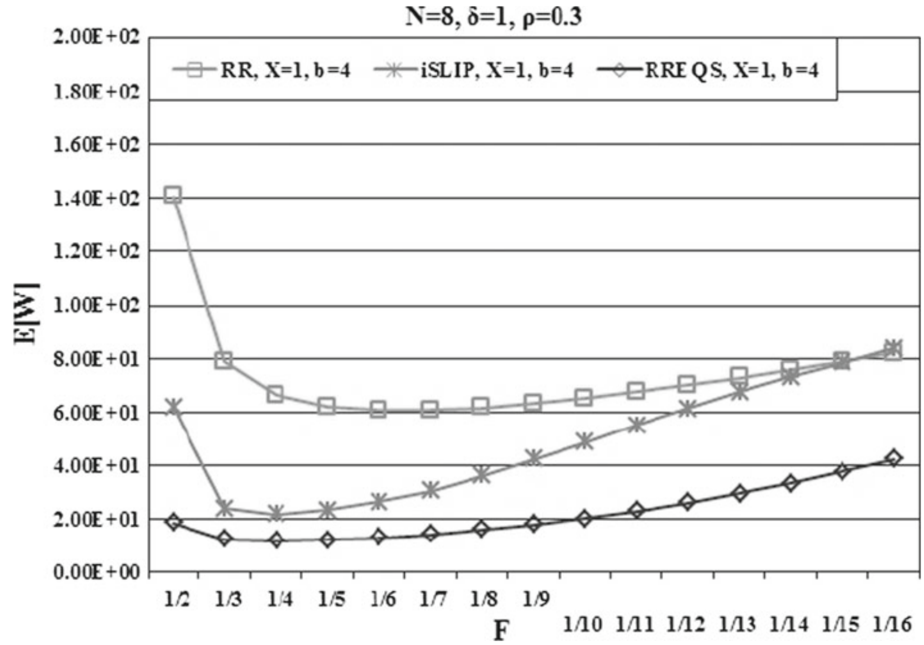


Fig. 11 $E[W]$ of analysis and simulation results in BRR, iSLIP, and RREQS with various X and b under $\rho = 0.3$

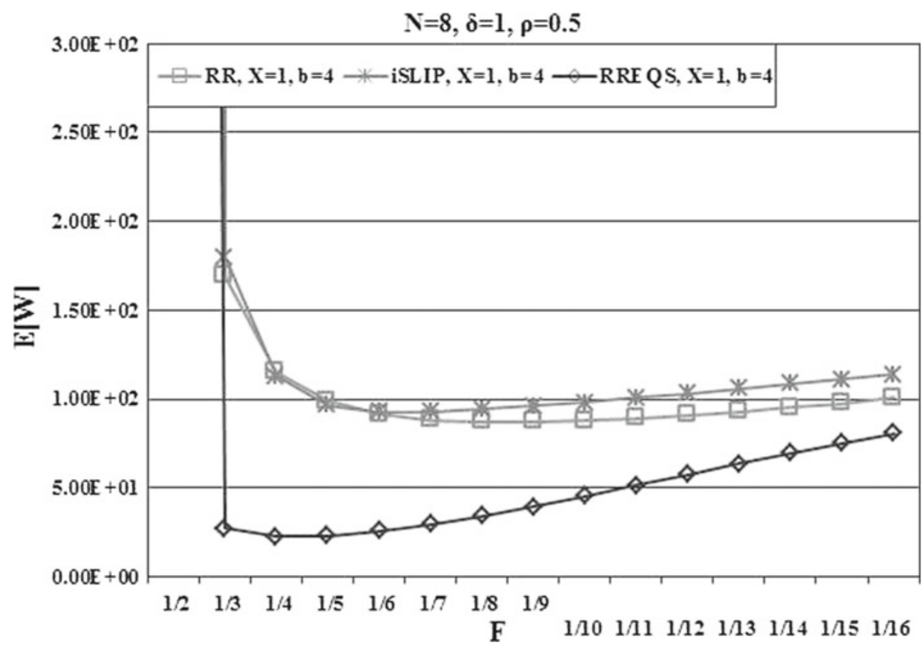


Fig. 12 $E[W]$ of analysis and simulation results in BRR, iSLIP, and RREQS with various X and b under $\rho = 0.5$

and employs a computational method with a nested loop structure to calculate the mean waiting time and the average number of the packets in the system. Our simulation results confirm that the proposed model is accurate and suitable for various scheduling schemes and network configurations.

Through numerical analysis, we find out the individual effects of various network parameters and several scheduling schemes on the performance and optimize the reconfiguration frequency to reduce the mean waiting time of packets. This quantitative work demonstrates that RR scheduling schemes with the optimal reconfiguration frequency can effectively reduce the mean waiting time of packets. The mean service time and the traffic load are key parameters. The optimal reconfiguration frequency decreases as the mean service time and the traffic load increase and slightly decreases as the number of ports increases. The optimal reconfiguration frequency in the RR scheduling schemes in which only non-empty VOQs are served is greater than that of BRR in which all VOQs are served.

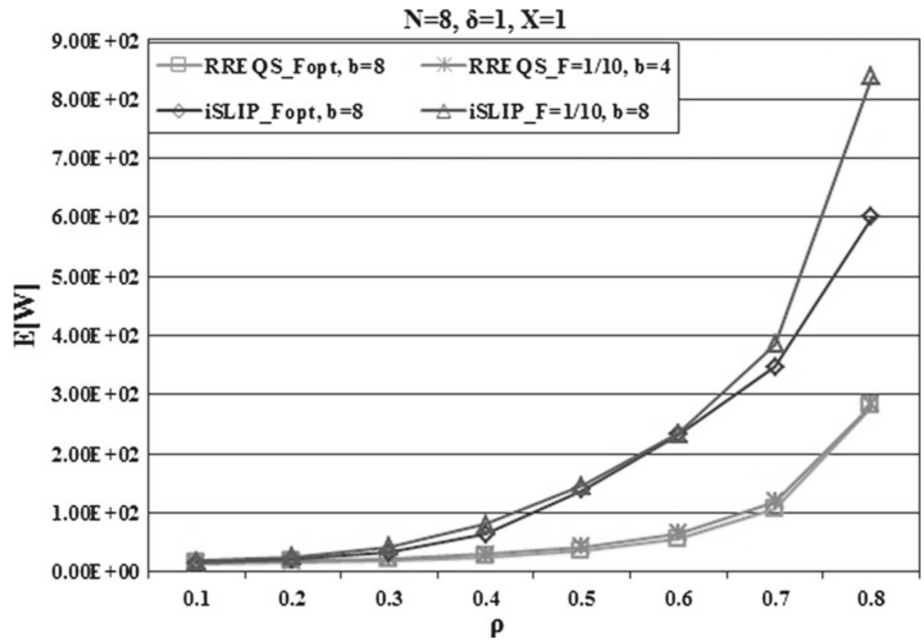


Fig. 13 $E[W]$ of analysis and simulation results in the RR $F = 1/5$

References

[1] Aissani, A.: Optimal control of an M/G/1 retrial queue with vacations. J. Syst. Sci. Syst. Eng. **17**, 487–502 (2008). doi:10.1007/s11518-008-5093-7

[2] Alaria, V., Bianco, A., Giaccone, P., Leonardi, E., Neri, F.: Multihop control schemes in switches with reconfiguration latency. IEEE/OSA J. Opt. Commun. Netw. **1**(3), B40–B55 (2009)

[3] Boxma, O., Groenendijk, W.: Waiting times in discrete-time cyclic-service systems. IEEE Trans. Commun. **36**(2), 164–170 (1988)

[4] Bruneel, H., Kim, B.G.: Discrete-Time Models for Communication Systems Including ATM. Kluwer Academic Publishers, Norwell, MA, (1992)

[5] Callegati, F.: Optical buffers for variable length packets. IEEE Commun. Lett. **4**(9), 292–294 (2000)

[6] Chou, K.H., Lin, W.: A latency-aware scheduling algorithm for all-optical packet switching networks with FDL buffers. Photonic Netw. Commun. **21**, 45–55 (2011). doi:10.1007/s11107-010-0279-6

[7] Fuhrmann, S., Cooper, R.: Stochastic decompositions in the M/G/1 queue with generalized vacations. Oper. Res. **33**(5), 1117–1129 (1985)

[8] Hluchyj, M., Karol, M.: Queueing in high-performance packet switching. IEEE J. Sel. Areas. Commun. **6**(9), 1587–1597 (1988)

[9] Kim, J., Nuzman, C., Kumar, B., Lieuwen, D., Kraus, J., Weiss, A., Lichtenwalner, C., Papazian, A., Frahm, R., Basavanhally, N., Ramsey, D., Aksyuk, V., Pardo, F., Simon, M., Lifton, V., Chan, H., Haueis, M., Gasparyan, A., Shea, H., Arney, S., Bolle, C., Kolodner, P., Ryf, R., Neilson, D., Gates, J.: 1100 × 1100 port MEMS-based optical crossconnect with 4-dB maximum loss. Photonics Technol. Lett., IEEE **15**(11), 1537–1539 (2003)

[10] Kleinrock, L., Levy, H.: The analysis of random polling systems. Oper. Res. **36**(5):716–732 (1988) http://or.journal.informs.org/cgi/reprint/36/5/716.pdf

[11] Leung, K., Eisenberg, M.: A single-server queue with vacations and gated time-limited service. In: INFOCOM ’89. Proceedings of the Eighth Annual Joint Conference of the IEEE Computer and Communications Societies. Technology: Emerging or Converging, IEEE, **3**:897–906 (1989)

[12] Leung, K.K., Eisenberg, M.: A single-server queue with vacations and non-gated time-limited service. Perform. Evaluation **12**(2), 115–125 (1991)

[13] Li, V., Li, C.Y., Wai, P.K.A.: Alternative structures for two-dimensional MEMS optical switches [invited]. J. Opt. Netw. **3**(10), 742–757 (2004)

[14] Li, X., Hamdi, M.: On scheduling optical packet switches with reconfiguration delay. IEEE J. Sel. Areas. Commun. **21**(7), 1156–1164 (2003)

[15] Liu, W., Xu, X., Tian, N.: Stochastic decompositions in the M/M/1 queue with working vacations. Oper. Res. Lett. **35**(5), 595–600 (2007)

[16] Ma, X., Kuo, G.S.: Optical switching technology comparison: optical MEMS vs. other technologies. Commun. Mag., IEEE **41**(11), S16–S23 (2003)

[17] Madamopoulos, N., Kaman, V., Yuan, S., Jerphagnon, O., Helkey, R., Bowers, J.: Applications of large-scale optical 3D-MEMS switches in fiber-based broadband-access networks. Photonic Netw. Commun. **19**(1), 62–73 (2010)

[18] McKeown, N.: The iSLIP scheduling algorithm for input-queued switches. IEEE/ACM Trans. Netw. **7**(2), 188–201 (1999)

[19] Mi, X., Soneda, H., Okuda, H., Tsuboi, O., Kouma, N., Mizuno, Y., Ueda, S., Sawaki, I.: A multi-chip directly mounted 512-MEMS-mirror array module with a hermetically sealed package for large optical cross-connects. J. Opt. A: Pure. Appl. Opt. **8**(7), S341 (2006)

[20] Neilson, D.T., Frahm, R., Kolodner, P., Bolle, C.A., Ryf, R., Kim, J., Papazian, A.R., Nuzman, C.J., Gasparyan, A., Basavanhally, N.R., Aksyuk, V.A., Gates, J.V.: 256 × 256 port optical cross-connect subsystem. J. Lightw. Technol. **22**(6), 1499 (2004)

[21] Pan, D., Yang, Y.: Bandwidth guaranteed multicast scheduling for virtual output queued packet switches. J. Parallel. Distrib. Comput. **69**(12), 939–949 (2009)

[22] Papadimitriou, G., Papazoglou, C., Pomportsis, A.: Optical switching: switch fabrics, techniques, and architectures. J. Lightw. Technol. **21**(2), 384–405 (2003)

[23] Takagi, H.: Queuing Analysis: A Foundation of Performance Evaluation. North-Holland, (1993)

[24] Wang, H., Aw, E., Williams, K., Wonfor, A., Penty, R., White, I.: Lossless multistage SOA switch fabric using high capacity mono-

lithic 4 × 4 SOA circuits. In: Optical Fiber Communication: incudes post deadline papers, 2009. OFC 2009, pp. 1–3 (2009)
[25] Wu, B, Yeung, K: Traffic scheduling in non-blocking optical packet switches with minimum delay. In: Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE 4:5-2045 (2005)
[26] Wu, B., Yeung, K.L., Hamdi, M., Li, X.: Minimizing internal speedup for performance guaranteed switches with optical fabrics. IEEE/ACM Trans. Netw. **17**(2), 632–645 (2009)
[27] Wu, B., Yeung, K.L., Ho, P.H., Jiang, X.: Minimum delay scheduling for performance guaranteed switches with optical fabrics. J. Lightw. Technol. **27**(16), 3453–3465 (2009)
[28] Xu, X., Zhang, Z.G.: Analysis of multi-server queue with a single vacation (e, d)-policy. Perform. Evaluation **63**(8), 825–838 (2006)
[29] Yuan, S., Lee, C.: Scaling optical switches to 100 Tb/s capacity. In: Photonics in Switching, Optical Society of America, p. PWB3 (2010)

Author Biographies

Kuan-Hung Chou received the BS degree in Information Engineering and Computer Science from Feng Chia University, Taichung, Taiwan, in 2005 and received the MS degree in Computer Science and Engineering from National Chung-Hsing University, Taichung, Taiwan, in 2007. Currently, he is working toward the PhD degree in Computer Science and Engineering from National Chung-Hsing University. His research interests include computer network, optical network, switching system, and interconnection network.

Woei Lin received the BS degree from National Chiao-Tung University, Hsing-Chu, Taiwan, in 1978 and the MS and PhD degrees in Electrical and Computer Engineering from University of Texas of Austin, USA, in 1982 and 1985, respectively. He is now a Professor in Institute of Computer Science, National Chung-Hsing University, during 1992–2005. His research interests include network switching system, network QoS, system performance evaluation, and parallel/distributed system.