Performance Analysis of a Photonic Single-Hop ATM Switch Architecture with Bursty and Correlated Arrivals

Martin W. McKinnon George N. Rouskas Harry G. Perros

TR-96-11

June 4, 1996

Abstract

We consider a photonic ATM switch based on the single-hop WDM architecture. The switch operates under schedules that mask the transceiver tuning latency. We develop an *exact* queueing-based decomposition algorithm to obtain the queue-length distribution at the input and output ports of the switch. The analysis is carried out using arrival models that capture the notion of burstiness and correlation, two important characteristics of ATM type of traffic, and which permit non-uniform destinations. We then derive analytic expressions for the cell-loss probability at the input and output ports, and the delay distribution to traverse the switch. To the best of our knowledge, such a comprehensive performance analysis of an optical switch architecture has not been done before.

Keywords: Optical networks, performance evaluation, photonic ATM switch, Markov modulated Bernoulli process (MMBP), Wavelength Division Multiplexing (WDM)

Department of Computer Science
North Carolina State University
Raleigh, NC 27695

1 Introduction

One of the issues in Asynchronous Transfer Mode (ATM) networks is that of developing switch architectures that can effectively switch cells at very high data rates (currently, data rates on the order of a few tens of Gigabits per second per port are envisioned). Over the last decade, a great deal of research effort has been devoted to the design of fast cell switches suitable to a broadband, integrated services environment; surveys of some of these architectures may be found in [1, 2]. Mainstream research and development activities in the area of broadband switching are focused exclusively on electronics-based technologies which have attained a high level of maturity. On the other hand, the deployment of optics is limited to mere point-to-point transmission where the technology has proven successful in a short time span.

Given the continued rapid progress in lightwave technology (including the demonstration of fast tunable transceivers [3, 4], the development of erbium-doped fiber amplifiers [5], and guided-wave optical switching [6]), and the anticipated total dominance of optical fiber in the wired network, the issue of deeper penetration of optics naturally arises. Given the potential of optical solutions to cell switching, the possibility of employing photonics to implement switching functions hitherto reserved for electronics is currently being explored (see [7] and references thereof). However, there remain at least two major technical challenges to be overcome before one can contemplate the design of *all-optical* switches. First, there is the difficulty of "controlling light by light", and secondly, the technologies for implementing buffering in the optical domain are not yet mature enough. Consequently, the most likely scenarios for near-term photonic cell switching will involve an optical switching fabric with electronic control and buffering.

It has long been recognized that Wavelength Division Multiplexing (WDM) will be instrumental in bridging the gap between the speed of electronics and the virtually unlimited bandwidth available within the optical medium. The wavelength domain adds a significant new degree of freedom to network design, allowing new network concepts to be developed. With a few exceptions (e.g., [8, 9, 10]), however, most broadcast WDM architectures that have appeared in the literature require a large number of wavelengths and/or very fast tunable transceivers [11]-[12]. Furthermore, the performance analysis of these architectures has been typically carried out assuming uniform traffic and memoryless arrival processes (see most of the above references, as well as [13]-[14]). Similar traffic assumptions were made in performance studies of electronic ATM switches in the late 80s. However, as it was later shown, such assumptions lead to overwhelmingly erroneous results regarding the performance of a switch. In order to study correctly the performance of a switch, one needs to use traffic models that capture the notion of burstiness and correlation, and which permit non-uniform output port destinations [15].

In this paper we revisit the well known and widely studied single-hop broadcast-and-select WDM architecture [16]. Unlike previous papers, however, we develop an *exact* queueing-based decomposition algorithm to study the performance of a single-hop ATM switch architecture operating under schedules that mask the transceiver tuning latency [9]. The analysis is carried out using arrival models that capture the important characteristics of ATM type of traffic, and non-uniform destinations. To the best of our knowledge, such a comprehensive performance analysis of an optical switch architecture has not been done before.

The next section presents our system model and provides some background information. The performance analysis of the switch is presented in Sections 3 and 4, and we conclude the paper in Section 5.

2 The ATM Switch Under Study

2.1 The Switch Architecture

We consider an optical switch architecture with N input ports and N output ports interconnected through a broadcast passive star (the switch fabric) that can support $C \leq N$ wavelengths $\lambda_1, \cdots, \lambda_C$ (see Figure 1). Each input port is equipped with a laser that enables it to inject signals into the optical medium. Similarly, each output port is capable of receiving optical signals through an optical filter. The laser at each input port is assumed to be tunable over all available wavelengths. The optical filters, on the other hand, are fixed to a given wavelength. Let $\lambda(j)$ denote the receive wavelength of output port j. Since $C \leq N$, a set $\mathcal{R}_c$ of output ports may be sharing a single receive wavelength λ_c:

$$\mathcal{R}_c \quad = \quad \{j \mid \lambda(j) = \lambda_c\}, \qquad c = 1, \cdots, C \tag{1}$$

The switch operates in a slotted mode. Since there are N ports but $C \leq N$ channels, each channel must run at a rate $\frac{N}{C}$ times faster than the rate of the input or output links ($\frac{N}{C}$ need not be an integer). Thus, we distinguish between *arrival* slots (which correspond to the ATM cell transmission time at the input/output link rate) and *service* slots (which are equal to the cell transmission time at the channel rate within the switch). Obviously, the duration of a service slot is equal to $\frac{C}{N}$ times that of an arrival slot. Without loss of generality, we assume that all input links are synchronized at arrival slot boundaries; similarly for output links. On the other hand, all C channels internal to the switch are synchronized at service slot boundaries.

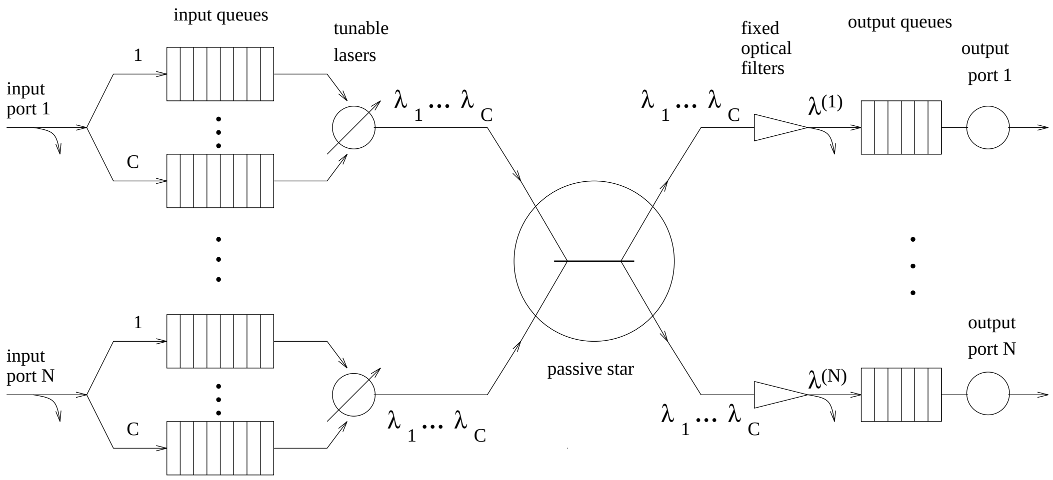

Figure 1: Queueing model of a switch architecture with N ports and C wavelengths

The switch employs electronic queueing at both the input and output ports, as Figure 1 illustrates. Cells arrive at an input port i and are buffered at a finite capacity queue, if the queue is not full. Otherwise, they are dropped. As Figure 1 indicates, the buffer space at each input port is assumed to be partitioned into C independent queues. Each queue c at input port i contains cells destined for the output ports which listen to a particular wavelength λ_c, $c = 1, \cdots, C$. This arrangement eliminates the head-of-line problem, and permits an input port to send a number of cells back-to-back when tuned to a certain wavelength. We let $B_{ic}^{(in)}$ denote the capacity of the queue at input port i corresponding to wavelength λ_c.

Cells buffered at an input port are transmitted on a FIFO basis onto the optical medium by the port's laser. This transmission takes place on an appropriate service slot which guarantees that the cell will be correctly received by its destination output port (more on this in the next subsection). Upon arriving at the output port, the cell is once again placed at a finite capacity buffer. Let $B_j^{(out)}$ denote the buffer capacity of output port j. Cells arriving at an output port to find a full buffer are lost. Cells in an output buffer are also served on a FIFO basis.

Interest in such a photonic switch architecture arises from several observations:

- it is highly modular, allowing the switch to grow relatively easily by adding ports and wavelengths;
- it is scalable, since the number of wavelengths need not be equal to the number of ports, and since the data rate within the switch needs only be $\frac{N}{C}$ times the rate of the input/output links;
- it provides end-to-end optical paths;

- its hardware requirements, in terms of the number of transceivers per port is minimum;
- it can be reconfigured to adapt to changing traffic patterns or to overcome failures of ports or transceivers; and
- it does not require extremely fast tunable transmitters (as explained below), and thus can be built using *currently available* tunable optical devices.

2.2 Transmission Schedules

One of the potentially difficult issues that arise in a WDM environment, such as the one described above, is that of coordinating the various transmitters/receivers. Some form of coordination is necessary because (a) a transmitter and a receiver must both be tuned to the same channel for the duration of a cell's transmission, and (b) a simultaneous transmission by one or more input ports on the same channel will result in a *collision*. The issue of coordination is further complicated by the fact that tunable transceivers need a non-negligible amount of time to switch between wavelengths. For the Gigabit per second rates envisioned here, and for 53-byte ATM cells, the tuning latency of state-of-the-art tunable lasers or filters can be as long as several times the size of a service slot [3]. Consequently, approaches that require each tunable transmitter to send a single cell and then switch to a new wavelength, will suffer a high tuning overhead and will result in a very low throughput.

In a recent paper [9] it was shown that careful scheduling can mask the effects of arbitrarily long tuning latencies, making it possible to build high-throughput photonic ATM switches using *currently available* lightwave technology. The key idea is to have each tunable transmitter send a *block* of cells on a wavelength before switching to another one. The main result of [9] was a set of new algorithms for constructing near-optimal (and, under certain conditions, optimal) schedules for transmitting a set of traffic demands $\{a_{ic}\}$. Quantity a_{ic} represents the number of cells to be transmitted by input port i onto channel λ_c per frame. The schedules are such that no collisions ever occur. Furthermore, they are easy to implement in a high speed environment, since the order in which the various input ports transmit is the same for all channels [9].

Quantity a_{ic}, $i = 1, \cdots, N$, $c = 1, \cdots, C$, can be seen as the number of service slots per frame allocated to input port i, so that the port can satisfy the required QoS of its incoming traffic intended for wavelength λ_c. By fixing a_{ic}, we indirectly allocate a certain amount of the bandwidth of wavelength λ_c to port i. This bandwidth could be an upper bound of the effective bandwidth of the total traffic carried by input port i on wavelength λ_c. In general, the estimation of the quantities a_{ic}, $i = 1, \cdots, N$, $c = 1, \cdots, C$, is part of the call admission algorithms, and is beyond the scope of this paper. We note that as the traffic varies, a_{ic} may vary as well. In this paper,

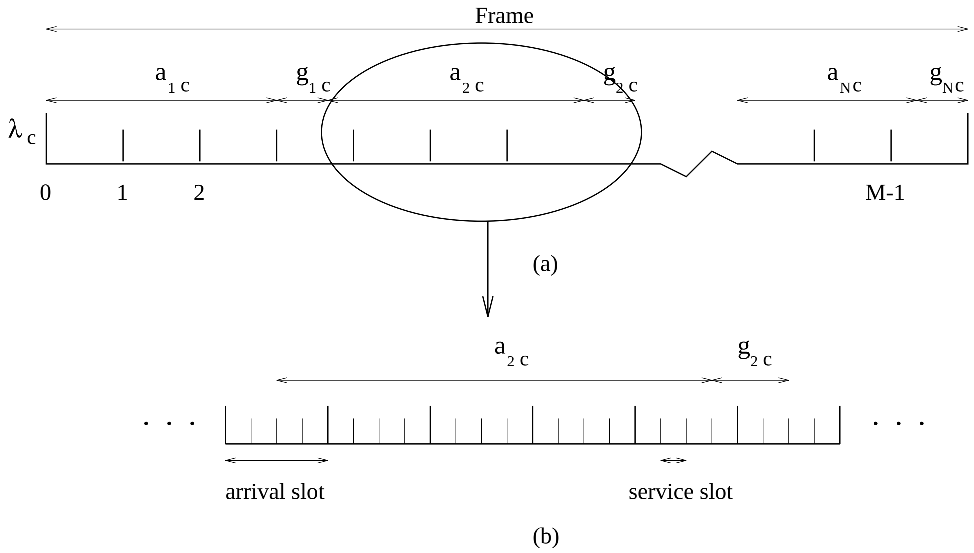

Figure 2: (a) Schedule for channel λ_c, and (b) detail corresponding to input port 2

we assume that quantities a_{ic} are fixed, since this variation will more likely take place over larger scales in time.

We assume that transmissions by the input ports onto wavelength λ_c follow a schedule as shown in Figure 2. This schedule repeats over time. Each frame of the schedule consists of M arrival slots. Within each frame, input port i is assigned a_{ic} *contiguous* service slots for transmitting cells on channel λ_c. These a_{ic} slots are followed by a *gap* of $g_{ic} \geq 0$ slots during which no port can transmit on λ_c. This gap may be necessary to ensure that input port $i+1$ has sufficient time to tune from wavelength λ_{c-1} to λ_c before it starts transmission. The algorithms in [9] are such that the number of slots in most of the gaps is equal to either zero or a small integer. Thus, the length of the schedule is very close to the lower bound $\max_i\{\sum_{c=1}^{C} a_{ic}\}$. Note that in Figure 2 we have assumed that an arrival slot is an integer multiple of service slots, but this need not be true in general, and it is not a necessary assumption for our model. Observe also that, although the frame begins and ends on *arrival* slot boundaries, the beginning or end of transmissions by a port does not necessarily coincide with the beginning or end of an *arrival* slot (although it is, obviously, synchronized with *service* slots).

2.3 Traffic Model

The arrival process to each input port of the switch is characterized by a two-state Markov Modulated Bernoulli Process (MMBP), hereafter referred to as 2-MMBP. This is a Bernoulli process whose arrival rate varies according to a two-state Markov chain. It captures the notion of burstiness and the correlation of successive interarrival times, two important characteristics of ATM type of traffic. For details on the properties of the 2-MMBP, the reader is referred to [17]. (We note that the algorithm for analyzing the switch was developed so that it can be readily extended to MMBPs with more than two states.)

We assume that the arrival process to port i, $i = 1, \cdots, N$, is given by a 2-MMBP characterized by the transition probability matrix $\mathbf{Q}_i$, and by $\mathbf{A}_i$ as follows:

$$\mathbf{Q}_i = \begin{bmatrix} q_i^{(00)} & q_i^{(01)} \\ q_i^{(10)} & q_i^{(11)} \end{bmatrix} \quad \text{and} \quad \mathbf{A}_i = \begin{bmatrix} \alpha_i^{(0)} & 0 \\ 0 & \alpha_i^{(1)} \end{bmatrix} \tag{2}$$

In (2), $q_i^{(kl)}, k, l = 0, 1$, is the probability that the 2-MMBP will make a transition to state l, given that it is currently at state k. Obviously, $q_i^{(k0)} + q_i^{(k1)} = 1$, $k = 0, 1$. Also, $a_i^{(0)}$ and $a_i^{(1)}$ are the arrival rates of the Bernoulli process at states 0 and 1, respectively. Transitions between states of the 2-MMBP occur only at the boundaries of *arrival* slots. We assume that the arrival process to each input port is given by a different 2-MMBP.

Let r_{ij} denote the probability that a cell arriving to input port i will have j as its destination output port. We will refer to $\{r_{ij}\}$ as the *routing* probabilities; this description implies that the routing probabilities can be input port dependent and non-uniformly distributed. The destination probabilities of successive cells are not correlated. That is, in an input port, the destination of one cell does not affect the destination of the cell behind it. This is a reasonable assumption when the switch is used as part of a backbone network. Given these assumptions, the probability that a cell arriving to port i will have to be transmitted on wavelength λ_c is:

$$r_{ic} = \sum_{j \in \mathcal{R}_c} r_{ij}, \qquad i = 1, \cdots, N \tag{3}$$

3 Exact Queueing Analysis

In this section we analyze exactly the queueing network shown in Figure 1. This queueing network represents the tunable-transmitter, fixed-receiver switch under study. The arrival process to each input port is assumed to be a 2-MMBP, and the access of the input ports to the wavelengths is governed by the schedule described in Section 2.2. The objective of the analysis of this queueing network is to obtain the queue-length distribution in an input or output port, from which performance measures such as cell-loss probability and delay distribution can be obtained. As will be seen below, the analysis of this queueing network is exact.

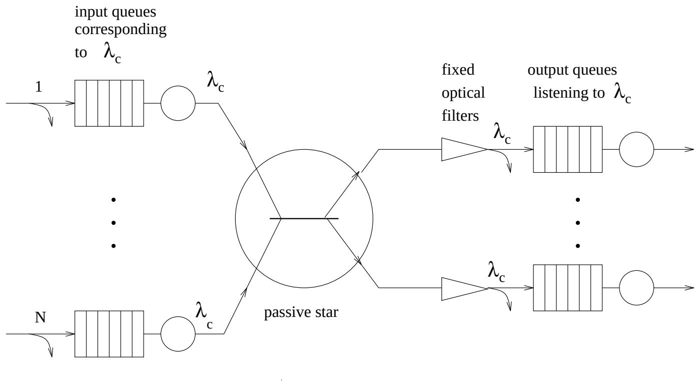

Figure 3: Queueing sub-network for wavelength λ_c

We first note that the original queueing network can be decomposed into C sub-networks, one per wavelength, as in Figure 3. For each wavelength λ_c, the corresponding sub-system consists of N input queues, and all the output queues that listen to wavelength λ_c. Each input queue i of the sub-system, is the one associated with wavelength λ_c in the i-th input port. These input queues will transmit to the output queues of the sub-system over wavelength λ_c. Note that, due to the independence among the C queues at each input port, the transmission schedule, and the fact that each output port listens to a specific wavelength, this is decomposition is *exact*. In view of this decomposition, it suffices to analyze a single sub-network, since the same analysis can be applied to all other sub-networks.

Consider now the sub-network for wavelength λ_c. We will analyze this sub-network by decomposing it into individual input and output ports. The decomposition algorithm gives an exact solution. As discussed in the previous section, each input queue i of the sub-network is only served for a_{ic} consecutive service slots per frame. During that time, no other input port is served. Input queue i is not served in the remaining slots of the frame. In view of this, there is no dependence among the input queues of the sub-system, and consequently each one can be analyzed in isolation in order to obtain its queue-length distribution. Each output port will also be considered in isolation. However, in order to consider each output port in isolation, we need to characterize the departure process from the input queues, which is offered as the arrival process to the output queues. This characterization is the most critical component of the decomposition algorithm, and is discussed in detail in Section 3.2.

From the queueing point of view, the above queueing network can be seen as a polling system in discrete time. Despite the fact that polling systems have been extensively analyzed, we note that very little work has been done within the context of discrete time (see for example [18]). In addition, this particular problem differs from the typical polling system since we consider output queues, not typically analyzed in polling systems.

We now proceed to obtain the queue-length distribution at an input queue. Following that, we obtain the queue-length distribution at an output queue.

3.1 The Queue-Length Distribution of an Input Queue

We now turn our attention to the sub-network for wavelength λ_c. The i-th input queue of this sub-network is the c-th queue of input port i. Since throughout this section we only consider the sub-network corresponding to λ_c, we will simply refer to this queue as "input queue i".

Consider input queue i of the sub-network in isolation. This input queue receives exactly a_{ic} service slots on wavelength λ_c, as shown in Figure 4(a). The block of a_{ic} service slots may not be aligned with the boundaries of the arrival slots. For instance, in the example shown in Figure 4(a), the block of a_{ic} service slots begins at the second service slot of arrival slot $x-1$, and it ends at the end of the second service slot in arrival slot $x+1$. Here, $x-1$, x, and $x+1$ represent the arrival slot number within a frame.

For each arrival slot, define $v_{ic}(x)$ as the number of service slots allocated to input queue i, that lie within arrival slot x [1]. Then, in the example in Figure 4(a), we have: $v_{ic}(x-1) = 3, v_{ic}(x) = 4, v_{ic}(x+1) = 2$, and $v_{ic}(x') = 0$ for all other x'. Obviously we have,

$$\sum_{x=0}^{M-1} v_{ic}(x) = a_{ic} \tag{4}$$

We analyze input queue i by constructing its underlying Markov chain embedded at arrival slot boundaries. The order of events is as follows. The service (i.e., transmission) completion of a cell occurs at an instant just before the end of a service slot. An arrival may occur at an instant just before the end of an arrival slot, but after the service completion instant of a service slot whose end is aligned with the end of an arrival slot. The 2-MMBP describing the arrival process to the queue makes a state transition immediately after the arrival instant. Finally, the Markov chain is observed at the boundary of each arrival slot, *after* the state transition by the 2-MMBP. The order of occurrence of these events is shown in Figure 4(b).

The state of the input queue is described by the tuple (x, y, z), where:

[1] In Figure 4 we assume that each arrival slot contains an integral number of service slots. If this is not the case, $v_{ic}(x)$ is defined as the number of service slots that lie completely within arrival slot x, plus one if there is a service slot that lies partially within slots $x-1$ and x. If there is a service slot that lies partially within arrival slots x and $x+1$, it will be counted in $v_{ic}(x+1)$.

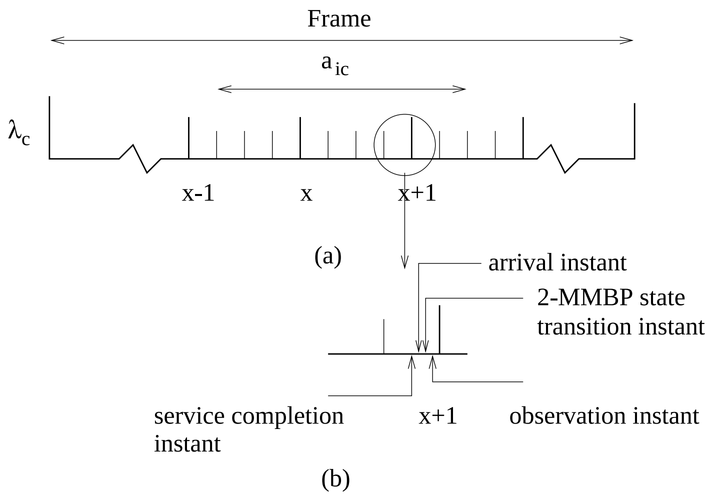

Figure 4: (a) Service period of input port i on channel λ_c, and (b) detail showing the relationship among service completion, arrival, 2-MMBP state transition, and observation instants within a service and an arrival slot

Table 1: Transition probabilities out of state (x, y, z) of the Markov chain

Current State	**Next State**	**Transition Probability**
(x, y, z)	$(x \oplus 1, \max\{0, y - v_{ic}(x)\}, z')$	$q_i^{(zz')}(1 - \alpha_i^{(z)} r_{ic})$
(x, y, z)	$(x \oplus 1, \min\{B_{ic}^{(in)}, \max\{0, y - v_{ic}(x)\} + 1\}, z')$	$q_i^{(zz')} \alpha_i^{(z)} r_{ic}$

- x represents the arrival slot number within a frame ($x = 0, 1, \cdots, M - 1$),
- y indicates the number of cells in the input queue ($y = 0, 1, \cdots, B_{ic}^{(in)}$), and
- z indicates the state of the 2-MMBP describing the arrival process to this queue, that is, $z = 0, 1$.

It is straightforward to verify that, as the state of the queue evolves in time, it defines a Markov chain. Let $\oplus$ denote modulo-M addition, where M is the number of arrival slots per frame. Then, the transition probabilities out of state (x, y, z) are given in Table 1. Note that, the next state after (x, y, z) always has an arrival slot number equal to $x \oplus 1$. In the first row of Table 1 we assume that the 2-MMBP makes a transition from state z to state z' (from (2), this event has a probability $q_i^{(zz')}$ of occurring), and that no cell arrives to this queue during the current slot (from (2) and (3), this occurs with probability $1 - \alpha_i^{(z)} r_{ic}$). Since at most $v_{ic}(x)$ cells are serviced during this arrival slot, and since no cell arrives, the queue length at the beginning of the next slot is equal to

$\max\{0, y - v_{ic}(x)\}$. In the second row of Table 1 we assume that the 2-MMBP makes a transition from state z to state z' and a cell arrives to the queue. This arriving cell cannot be serviced during this slot, and has to be added to the queue. Finally, the expression for the new queue length ensures that it will not exceed the capacity $B_{ic}^{(in)}$ of the input queue.

We observe that the probability transition matrix of this Markov chain has the following block form:

$$\mathbf{S}_{ic} \quad = \quad \left[\begin{array}{cccccc} 0 & \mathbf{R}_{ic}(0) & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{R}_{ic}(1) & 0 & \cdots & 0 \\ 0 & 0 & 0 & \mathbf{R}_{ic}(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \mathbf{R}_{ic}(M-2) \\ \mathbf{R}_{ic}(M-1) & 0 & 0 & 0 & \cdots & 0 \end{array}\right] \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ M-2 \\ M-1 \end{array} \tag{5}$$

Note that the only non-zero elements of the $M \times M$ matrix $\mathbf{S}_{ic}$ are those of the upper diagonal. This is due to the fact that at each transition instant, i.e., at each arrival slot boundary, the random variable x changes to $x \oplus 1$. Changes in the other two random variables, y and z, of the state of the queue are governed by the $(B_{ic}^{(in)} + 1) \times (B_{ic}^{(in)} + 1)$ matrices $\mathbf{R}_{ic}(x)$. There are M different $\mathbf{R}_{ic}$ matrices, one for each arrival slot x in the frame.

Define matrices $\mathbf{X}_{ic}$ and $\mathbf{Y}_{ic}$ as follows:

$$\mathbf{X}_{ic} \;=\; r_{ic}\, \mathbf{A}_i\, \mathbf{Q}_i \qquad \text{and} \qquad \mathbf{Y}_{ic} \;=\; (\mathbf{I} - r_{ic}\, \mathbf{A}_i)\, \mathbf{Q}_i, \tag{6}$$

where $\mathbf{I}$ is the identity matrix. Then, the transition matrix $\mathbf{R}_{ic}(x)$ associated with arrival slot x can be written as:

$$\mathbf{R}_{ic}(x) \quad = \quad \left[\begin{array}{ccccccccc} \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & \cdots & 0 \end{array}\right] \begin{array}{c} 0 \\ \vdots \\ v_{ic}(x) \\ v_{ic}(x)+1 \\ v_{ic}(x)+2 \\ \vdots \\ B_{ic}^{(in)} \end{array} \tag{7}$$

The structure of matrix $\mathbf{R}_{ic}(x)$ given in (7) can be explained as follows. Suppose that the number of cells y in the queue at the beginning of slot x is at most $v_{ic}(x)$. Since up to $v_{ic}(x)$ cells can be served within slot x, the number in the queue at the end of the slot will be 1 or 0, depending on whether an arrival occurred or not. This is indicated by the transitions in rows 0 through $v_{ic}(x)$ of matrix $\mathbf{R}_{ic}(x)$. However, if at the beginning of slot x we have $y > v_{ic}(x)$, then the number in the queue at the next transition will be $y - v_{ic}(x)$ (plus one if an arrival occurred). This is indicated by the transitions in rows $v_{ic}(x) + 1$ and higher of $\mathbf{R}_{ic}(x)$. Of course, y cannot exceed the queue capacity $B_{ic}^{(in)}$. Since the number of service slots $v_{ic}(x)$ depends on the particular slot x within the frame, $\mathbf{R}_{ic}(x)$ is a function of x.

Matrix $\mathbf{R}_{ic}(x)$ is slightly different when $v_{ic}(x) = 0$. This is because, in this case, if the state of the input queue is such that $y = B_{ic}^{(in)}$, a new arrival will be discarded. So when $y = B_{ic}^{(in)}$, the 2-MMBP is allowed to make a transition, but regardless of whether or not an arrival is generated, the number of cells in the queue will remain equal to $B_{ic}^{(in)}$. Thus, the the last row of $\mathbf{R}_{ic}(x)$ will be: $[0\ 0\ \cdots\ 0\ \mathbf{Q}_i]$.

It is now straightforward to verify that the Markov chain with transition matrix $\mathbf{S}_{ic}$ is irreducible, and therefore a steady-state distribution exists. Transition matrix $\mathbf{S}_{ic}$ defines a *p-cyclic* Markov chain [19], and therefore it can be solved using any of the techniques for p-cyclic Markov chains in [19, ch. 7]. We have used the power method in [19, ch. 7] to obtain the steady state probability $\pi_{ic}(x, y, z)$ that at the beginning of arrival slot x, the 2-MMBP is in state z and the input queue has y cells. The steady-state probability that the queue has y cells at the beginning of slot x, independent of the state of the 2-MMBP is:

$$\pi_{ic}(x, y) = \sum_{z=0,1} \pi_{ic}(x, y, z) \tag{8}$$

Finally, we note that all the results obtained in this subsection can be readily extended to MMBP-type of arrival processes with more than two states. For this, it would suffice to appropriately modify matrices $\mathbf{X}_{ic}$ and $\mathbf{Y}_{ic}$.

3.2 The Queue-Length Distribution of an Output Queue

We now proceed to obtain the queue-length distribution of an output queue. Let j be an output port listening to wavelength λ_c. Below, we analyze this queue in isolation. As will be seen, the analysis is exact.

In order to analyze queue j in isolation, we need to characterize the arrival process to the queue. This arrival process is, in fact, the departure process from the input queues. An interesting aspect of the departure process from the input queues is that for each frame, during the sub-period a_{ic} we only have departures from the i-th input queue. This period is then followed by a gap g_{ic}

during which no departure occurs. This cycle repeats for the next input queue. Thus, in order to characterize the overall departure process offered as the arrival process to output queue j, it suffices to characterize the departure process from each input queue, and then combine them together. (We note that this overall departure process is quite different from the typical process of constructing the superposition of a number of departures into a single stream, where at each slot more than one cell may be departing.)

As in the previous section, we obtain the queue-length distribution of output port j at arrival slot boundaries. Recall that an arrival slot to an input queue is equal to a departure slot from an output queue. Also, arrival and departure slots are synchronized. Therefore, during an arrival slot x a cell may be transmitted out from the output queue. However, during slot x there may be several arrivals to the output queue from the input queues.

Let (x, w) be the state associated with output port j, where

- x indicates the arrival slot number within the frame ($x = 0, 1, \cdots, M-1$), and
- w indicates the number of cells at the output queue ($w = 0, 1, \cdots, B_j^{(out)}$).

We assume the following order of events. A cell will depart from the output queue at an instant immediately after the beginning of an arrival slot and the departure will be completed at the end of the slot. A cell from an input port arrives at an instant just before the end of a service slot. Finally, the state of the queue is observed just before the end of an arrival slot and after the arrival associated with the last service slot in the arrival slot has occurred (see Figure 5(b)).

Let $u_j(x)$ be the number of service slots of any input queue on wavelength λ_c within arrival slot x. We have that

$$u_j(x) \quad = \quad \sum_{i=1}^{N} v_{ic}(x) \tag{9}$$

where $v_{ic}(x)$ is as defined in (4). Quantity $u_j(x)$ represents the maximum number of cells that may arrive to output port j within slot x. In the example of Figure 5(a) where we show the arrival slots during which cells from input ports i and $i+1$ may arrive to output port j, we have: $u_j(x-1) = v_{ic}(x-1) = 4$, $u_j(x) = v_{ic}(x) + v_{i+1,c}(x) = 1 + 2 = 3$, and $u_j(x+1) = v_{i+1,c}(x+1) = 4$.

Observe now that (a) at each state transition x advances by one (modulo-M), (b) at most one cell departs from the queue as long as the queue is not empty, (c) a number $s \leq u_j(x)$ of cells may be transmitted from the input ports to output port j within arrival slot x, and that (d) the queue capacity is $B_j^{(out)}$. Then, the transition probabilities out of state (x, w) for this Markov chain are given in Table 2.

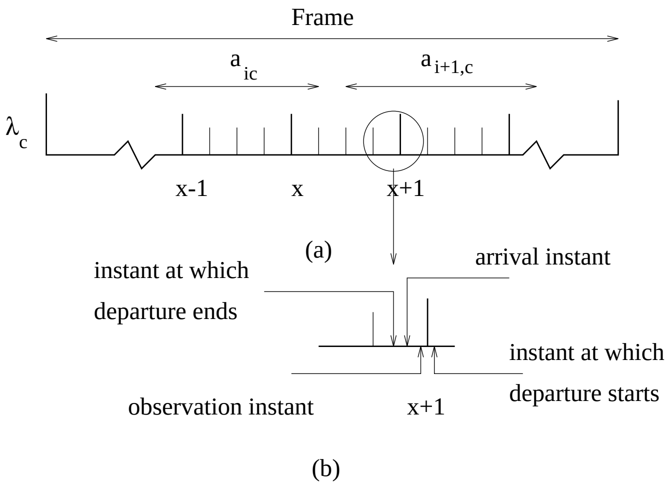

Figure 5: (a) Arrivals to output port j from input ports i and $i+1$, and (b) detail showing the relationship of departure, arrival, and observation instants

Table 2: Transition probabilities out of state (x, w) of the Markov chain

Current State	**Next State**	**Transition Probability**
(x, w)	$(x \oplus 1, \min\{B_j^{(out)}, \max\{0, w-1\} + s\})$ $0 \leq s \leq u_j(x)$	$\sum_{s_1+\cdots+s_N=s} \prod_{i=1}^{N} L_i(s_i, x)$

In Table 2, $L_i(s_i, x)$ is the probability that input port i transmits s_i cells to output port j in arrival slot x. To compute $L_i(s_i, x)$, define r'_{ij} as the conditional probability that a cell is destined to output port j, given that the cell is destined to be transmitted on λ_c, the receive wavelength of output port j:

$$r'_{ij} = \frac{r_{ij}}{\sum_{k \in \mathcal{R}_c} r_{ik}} = \frac{r_{ij}}{r_{ic}} \tag{10}$$

Also, define $\pi_{ic}(y \mid x)$ as the conditional probability of having y cells at the i-th input queue given that we are at the beginning of slot x:

$$\pi_{ic}(y \mid x) = \frac{\pi_{ic}(x, y)}{\sum_{k=0}^{B_{ic}^{(in)}} \pi_{ic}(x, k)} \tag{11}$$

Then, for $r'_{ij} < 1$, the probability $L_i(s_i, x)$ is given by

$$L_i(s_i, x) = \begin{cases} 0, & s_i > v_{ic}(x) \\ \sum_{y=s_i}^{B_{ic}^{(in)}} \pi_{ic}(y \mid x) \begin{pmatrix} \min\{y, v_{ic}(x)\} \\ s_i \end{pmatrix} (r'_{ij})^{s_i}(1 - r'_{ij})^{\min\{y, v_{ic}(x)\}-s_i}, & s_i \leq v_{ic}(x) \end{cases} \tag{12}$$

Expression (12) can be explained by noting that input port i will transmit s_i cells to output port j during arrival slot x if (a) $v_{ic}(x) \geq s_i$, (b) input port i has $y \geq s_i$ cells in its queue for λ_c at the beginning of the slot, and (c) exactly s_i of $\min\{y, v_{ic}(x)\}$ cells that will be transmitted by this queue in this arrival slot are for output j.

If $r'_{ij} = 1$, in which case j is the only port listening on wavelength λ_c, the expression for $L_i(s_i, x)$ must be modified as follows:

$$L_i(s_i, x) = \begin{cases} 0, & s_i > v_{ic}(x) \\ \pi_{ic}(s_i \mid x), & s_i < v_{ic}(x) \\ \sum_{y=s_i}^{B_{ic}^{(in)}} \pi_{ic}(y \mid x), & s_i = v_{ic}(x) \end{cases} \tag{13}$$

Expressions (12) and (13) are based on the assumption that $v_{ic} < B_{ic}^{(in)}$ and $v_{ic} < B_j^{(out)}$, which we believe is a reasonable one. In the general case, quantity v_{ic} in both expressions must be replaced by $\min\{v_{ic}(x), B_{ic}^{(in)}, B_j^{(out)}\}$.

The transition matrix $\mathbf{T}_j$ of the Markov chain defined by the evolution of the state (x, w) of output queue j has the following form, which is similar to that of matrix $\mathbf{S}_{ic}$ given by (5):

$$\mathbf{T}_j = \begin{array}{c} \left[\begin{array}{cccccc} 0 & \mathbf{U}_j(0) & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{U}_j(1) & 0 & \cdots & 0 \\ 0 & 0 & 0 & \mathbf{U}_j(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \mathbf{U}_j(M-2) \\ \mathbf{U}_j(M-1) & 0 & 0 & 0 & \cdots & 0 \end{array} \right] \end{array} \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ M-2 \\ M-1 \end{array} \tag{14}$$

$\mathbf{U}_j(x)$ is a $(B_j^{(out)} + 1) \times (B_j^{(out)} + 1)$ matrix that governs changes in random variable w of the state of the output queue. The elements of this matrix can be determined using Table 2 and expressions (12) and (13). Since they depend on $v_{ic}(x)$ and $u_j(x)$, $\mathbf{U}_j(x)$ also depends on x, the slot number within the frame.

We observe that $\mathbf{T}_j$ also defines a *p-cyclic* Markov chain. We have used the power method in [19, ch. 7] to obtain $\pi_j(x, w)$, the steady-state probability that output queue j has w cells at the beginning of slot x.

3.3 Summary of the Decomposition Algorithm

Below we summarize our approach to analyzing the sub-network of Figure 3 corresponding to wavelength λ_c. We assume that quantities $\{a_{ic}\}$ and the corresponding schedule (see [9]) are given.

1. For each arrival slot x, use the schedule and expressions (4) and (9) to compute the quantities $v_{ic}(x)$ and $u_j(x)$, $i = 1, \cdots, N,\ j : \lambda(j) = \lambda_c$.

2. For each input queue i, construct the transition probability matrix $\mathbf{S}_{ic}$ from (2), (3), (5), (6), and (7). Solve this matrix using any of the techniques in [19, ch. 7], and use (8) to obtain the steady-state probability $\pi_{ic}(x, y)$ that input queue i has y cells at the beginning of the x-th slot of the frame.

3. For each output port j with λ_c as its receive wavelength, use $\pi_{ic}(x, y)$ derived in Step 2, and (12)-(13) to construct the transition matrix $\mathbf{T}_j$ given by (14). Solve the matrix as in Step 2 to obtain $\pi_j(x, w)$, the steady-state probability that port j has w cells in its queue at the beginning of slot x.

Note that the complexity of this approach is dominated by Step 2. For each of the N input queues we have to solve a matrix of dimensions $\left[2M(B_{ic}^{(in)} + 1)\right] \times \left[2M(B_{ic}^{(in)} + 1)\right]$, where M is the length of the schedule (in arrival slots) and $B_{ic}^{(in)}$ is the capacity of the respective queue. Inverting a $K \times K$ matrix takes time $O(K^3)$, although some of the techniques in [19, ch. 7] can take advantage of the fact that the matrix is sparse to solve for the queue-length distributions at a significantly faster rate. Thus, in the worst case, the overall complexity of our algorithm is $O(NM^3B^3)$, where $B = \max_i\{B_{ic}^{(in)}\}$.

4 Cell-Loss Probability and Delay Distribution

We now use the queue-length distributions for the input and output ports, $\pi_{ic}(x, y)$ and $\pi_j(x, w)$, respectively, derived in the previous section, to obtain the cell-loss probability and the delay distribution to traverse the switch.

4.1 The Cell-Loss Probability at an Input Port

Let $\Omega_{ic}(x)$ be the probability that a cell arriving to the c-th queue of input port i during arrival slot x of the schedule will be lost. Let us also refer to Figure 4(b) which shows the service completion, arrival, and observation instants within slot x. We observe that, if the number $v_{ic}(x)$ of slots during which this queue is serviced within arrival slot x is not zero (i.e., $v_{ic}(x) > 0$), no arriving cell will be lost. This is because at most one cell may arrive in slot x. Even if the c-th queue at input port i is full at the beginning of slot x, $v_{ic}(x) \geq 1$ cells will be serviced during this slot, and the order of service completion and arrival instants in Figure 4(b) guarantees that an arriving cell will be accepted. On the other hand, if $v_{ic}(x) = 0$ for slot x, then an arriving cell will be discarded if and only if the queue is full at the beginning of x. Thus,

$$\Omega_{ic}(x) = \begin{cases} 0, & v_{ic}(x) > 0 \\ \pi_{ic}(x \ominus 1, B_{ic}^{(in)}), & v_{ic}(x) = 0 \end{cases} \tag{15}$$

where $\ominus$ denotes regular subtraction with the exception that, if $x = 0$, then $x \ominus 1 = M - 1$.

The probability Ω_{ic} that a cell arriving to the c-th queue of input port i is lost, regardless of the arrival slot x can now be computed from (15) as follows:

$$\begin{aligned} \Omega_{ic} &= \sum_{x=0}^{M-1} \Omega_{ic}(x) \; Pr[\text{a cell arrives at slot } x] \\ &= \sum_{x: v_{ic}(x)=0} \pi_{ic}(x \ominus 1, B_{ic}^{(in)}) \; Pr[\text{a cell arrives at slot } x] \end{aligned} \tag{16}$$

where the summation runs over all x for which $v_{ic}(x) = 0$.

Note that the state of the arrival process to the i-th input port is independent of the slot number. Thus, the probability that a cell arrives to the c-th queue of input port i, at some slot x, is equal to the steady-state arrival probability for the 2-MMBP times the probability r_{ic} that this cell is destined for an output port listening to wavelength λ_c. From (2) and [17], the steady state arrival probability for this 2-MMBP is

$$\gamma_i = \frac{q_i^{(10)} a_i^{(0)} + q_i^{(01)} a_i^{(1)}}{q_i^{(01)} + q_i^{(10)}} \tag{17}$$

Thus, we can rewrite (16) as:

$$\Omega_{ic} \quad = \sum_{x: v_{ic}(x)=0} \pi_{ic}(x \ominus 1, B_{ic}^{(in)}) \; \gamma_i \; r_{ic} \tag{18}$$

4.2 The Cell-Loss Probability at an Output Port

The cell-loss probability at an output port is more complicated to calculate, since we may have multiple cell arrivals to the given output port within a single arrival slot (refer to Figure 5(a)). Let us define $\Omega_j(x, n)$ as the probability that n cells will be lost at output queue j within arrival slot x. An output port will lose n cells in slot x if (a) the port had $w, 0 \leq w \leq B_j^{(out)}$ cells at the beginning of slot x, (b) one cell departed during the slot (if $w > 0$), and (c) exactly $B_j^{(out)} - \max\{0, w-1\} + n$ cells arrived during slot x (the max operation in this expression takes care of the case $w = 0$ when no cell may depart). We can then write:

$$\Omega_j(x, n) \quad = \sum_{w=0}^{B_j^{(out)}} \pi_j(x \ominus 1, w) \; Pr[B_j^{(out)} - \max\{0, w-1\} + n \text{ cells arrive in slot } x] \tag{19}$$

The last probability in (19) can be obtained using (12) and (13):

$$Pr[s \text{ cells arrive to output port } j \text{ in slot } x] \quad = \sum_{s_1 + \cdots + s_N = s} \prod_{i=1}^{N} L_i(s_i, x) \tag{20}$$

Note that at most $u_j(x)$ cells may arrive (and get lost) in arrival slot x. We can then get the probability $\Omega_j(x)$ that a cell is lost given that it arrives to output port j in slot x:

$$\Omega_j(x) \quad = \quad \frac{\sum_{n=1}^{u_j(x)} n \Omega_j(x, n)}{E[\text{number of arrivals to output port } j \text{ in slot } x]} \tag{21}$$

The expected number of arrivals in slot x can be computed using (20):

$$E[\text{number of arrivals in slot } x] \quad = \sum_{s=1}^{u_j(x)} s \; Pr[s \text{ cells arrive in slot } x] \tag{22}$$

Finally, the probability Ω_j that an arriving cell to output port j will be lost regardless of the arrival slot x can be found by unconditioning (21) on the arrival slot. We have:

$$\Omega_j \quad = \sum_{x=0}^{M-1} \Omega_j(x) \; Pr[\text{a cell arrives to port } j \text{ in slot } x] \tag{23}$$

This last probability is equal to

$$\frac{E[\text{number of arrivals to output port } j \text{ in slot } x]}{\sum_{x'=0}^{M-1} E[\text{number of arrivals to output port } j \text{ in slot } x']} \tag{24}$$

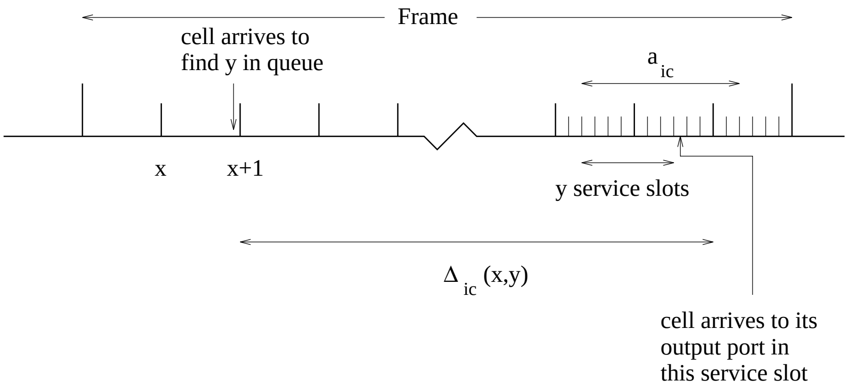

Figure 6: Definition of $\Delta_{ic}(x,y)$ for $y < a_{ic}$

Thus, we obtain

$$\Omega_j = \frac{\sum_{x=0}^{M-1} \Omega_j(x) \; E[\text{number of arrivals to output port } j \text{ in slot } x]}{\sum_{x'=0}^{M-1} \; E[\text{number of arrivals to output port } j \text{ in slot } x']} \tag{25}$$

4.3 The Delay Distribution

In this section we calculate the distribution of the number of arrival slots that elapse from the instant that a cell arrives to an input queue to the instant that the cell departs from an output queue. We note that the cell arrives at, and departs from, the switch on arrival slot boundaries.

Let us tag a cell arriving to the c-th queue of input port i in arrival slot x. Let j be the destination output port of this cell, and let λ_c be the wavelength of j. We assume that the tagged cell sees y cells in the input queue, where $y < B_{ic}^{(in)}$. Define $\Delta_{ic}(x,y)$ to be the number of arrival slots between the end of slot x and the end of the arrival slot during which the tagged cell is transmitted to the output queue j on wavelength λ_c. If $y < a_{ic}$, then $\Delta_{ic}(x,y)$ can be calculated very easily, as shown in Figure 6. (Since x and y are given, we can calculate how many arrival slots will elapse until the input queue is served by wavelength λ_c, and subsequently we can calculate the number of arrival slots required to transmit y cells on this wavelength.) We note that for $y < a_{ic}$ we have $\Delta_{ic}(x,y) < M$, where M is the length of the schedule in arrival slots.

Now, let us assume that $y \geq a_{ic}$. Then, y can be written as $y = ka_{ic} + y'$, where k is an integer $k \geq 1$, and $y' < a_{ic}$. In this case, we have $\Delta_{ic}(x,y) = kM + \Delta_{ic}(x,y')$.

Finally, let w, where $w < B_j^{(out)}$, be the number of cells that the tagged cell will find in the output queue j upon arrival to this queue. Then, it will wait exactly w arrival slots before it is transmitted out of the switch, for a total of $w+1$ slots.

We can now compute $\Phi_{ij}(m)$, the probability that a cell with destination j arriving to port i will spend m arrival slots in the switch as the product of (a) the probability that the cell will spend $l < m$ slots in its input queue, and (b) the probability that the cell will spend exactly $m - l$ slots in the output queue:

$$\Phi_{ij}(m) = \sum_{l=1}^{m-1} Pr[l \text{ slots in input queue } i] \; Pr[m - l \text{ slots in output queue } j] \qquad (26)$$

Since the state of the arrival process to input port i is independent of the arrival slot and the number in the queue, using the same reasoning as in (18) we can write:

$$Pr[l \text{ slots in input queue } i] = \sum_{x,y:\Delta_{ic}(x,y)=l} \pi_{ic}(x,y)\, \gamma_i \; r_{ij} \qquad (27)$$

where the sum is over all states (x, y) of the input queue such that the cell will spend exactly l arrival slots in the queue.

In order to compute $Pr[m - l$ slots in output queue $j]$, let us return to the tagged cell arriving to the c-th queue of input port i in slot x. Suppose that, at the time of its arrival to the switch, its input and output queues have y and w cells, respectively ($y < B_{ic}^{(in)}$). Then, the amount of time that the tagged cell spends in the switch is a function of x, y, and w. Note that the cell will arrive to its output port in arrival slot $x' = x \oplus \Delta_{ic}(x, y)$, where $\oplus$ denotes addition modulo-M (refer also to Figure 6). If it finds $w' < B_j^{(out)}$ cells in its output port at that time, it will spend another $w' + 1$ slots in the switch, for a total of $\Delta_{ic}(x, y) + w' + 1$ slots. To obtain an exact expression for $Pr[m - l$ slots in output queue $j]$, we must compute the conditional probability that the cell will find w' cells in its output queue in slot x', given that there were w cells in that same queue in slot x. This conditional probability, however, is difficult to calculate, as it depends on (a) the schedule, (b) the occupancy of the c-th queue at all other input ports, and (c) the routing probabilities.

Alternatively, we can make the simplifying assumption that, when a cell is transmitted to its output queue, the probability that it will find w cells in this queue is equal to the steady-state probability of having w cells in the queue $\pi_j(w) = \sum_{x=0}^{M-1} \pi_j(x, w)$. This is a reasonable approximation when (a) there is a relatively large number of input ports, and (b) the destination of one cell does not affect the destination of the cell behind it (as we assumed in this paper). Then, we can write:

$$Pr[m - l \text{ slots in output queue } j] = \pi_j(m - l - 1) \qquad (28)$$

Finally, we can use (27) and (28) to rewrite the probability (26) that a cell will spend m arrival slots in the switch as

$$\Phi_{ij}(m) \quad = \quad \sum_{l=1}^{m-1} \left\{ \sum_{x,y:\Delta_{ic}(x,y)=l} \pi_{ic}(x,y)\,\gamma_i\; r_{ij}\; \pi_j(m-l-1) \right\} \tag{29}$$

5 Concluding Remarks

We have studied the performance of a photonic ATM switch based on the single-hop WDM architecture. We have developed an exact decomposition algorithm to obtained the queue-length distributions at the input and output ports, and we have presented analytic expressions for the cell-loss probability and delay distribution to traverse the switch. Our current research focuses on extending the queueing model presented here in three directions. First, we will consider the scenario where the C queues at each input port are not independent but share a common buffer space. Second, we will introduce backpressure mechanisms to prevent an input port from transmitting a cell to an output port that is full. Finally, we plan to investigate several approaches for handling multicast traffic.

References

[1] F. A. Tobagi. Fast packet switch architectures for broadband integrated services digital networks. *Proceedings of IEEE*, pages 133–167, January 1990.

[2] H. Ahmadi and W. E. Denzel. A survey of modern high-performance switching techniques. *IEEE JSAC*, pages 1227–1237, September 1989.

[3] P. E. Green. *Fiber Optic Networks*. Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[4] A. Sneh and K. M. Johnson. High-speed tunable liquid crystal optical filter for WDM systems. In *Proceedings of IEEE/LEOS Summer Topical Meetings on Optical Networks and their Enabling Technologies*, pages 59–60, Lake Tahoe, CA, July 1994.

[5] K. Nakagawa, S. Nishi, K. Aida, and E. Yoneda. Trunk and distribution network application of erbium-doped fiber amplifier. *Journal of Lightwave Technology*, LT-9, February 1991.

[6] L. Thylen, G. Karlsson, and O. Nilsson. Switching technologies for future guided wave optical networks: Potentials and limitations of photonics and electronics. *IEEE Communications Magazine*, 34(2):106–113, February 1996.

[7] A. Jaszczyk and H. T. Mouftah. Photonic fast packet switching. *IEEE Communications Magazine*, pages 58–65, February 1993.

[8] G. N. Rouskas and M. H. Ammar. Analysis and optimization of transmission schedules for single-hop WDM networks. *IEEE/ACM Transactions on Networking*, 3(2):211–221, April 1995.

[9] G. N. Rouskas and V. Sivaraman. On the design of optimal TDM schedules for broadcast WDM networks with arbitrary transceiver tuning latencies. In *Proceedings of INFOCOM '96*, pages 1217–1224. IEEE, March 1996.

[10] D. A. Levine and I. F. Akyildiz. PROTON: A media access control protocol for optical networks with star topology. *IEEE/ACM Transactions on Networking*, 3(2):158–168, April 1995.

[11] T. T. Lee, M. S. Goodman, and E. Arthurs. A broadband optical multicast switch. In *Proceedings of ISS '90*, 1990.

[12] P. A. Humblet, R. Ramaswami, and K. N. Sivarajan. An efficient communication protocol for high-speed packet-switched multichannel networks. *IEEE Journal on Selected Areas in Communications*, 11(4):568–578, May 1993.

[13] I. M. I. Habbab, M. Kavehrad, and C.-E. W. Sundberg. Protocols for very high-speed optical fiber local area networks using a passive star topology. *Journal of Lightwave Technology*, LT-5(12):1782–1793, December 1987.

[14] J. Jue, M. Borella, and B. Mukherjee. Performance analysis of the Rainbow WDM optical network prototype. In *Proceedings of ICC '95*, pages 282–286. IEEE, 1995.

[15] G. Pujolle and H. G. Perros. Queueing systems for modelling ATM networks. In *Int'l Conf. on the Performance of Distributed Systems and Integrated Comm. Networks*, pages 10–12, Kyoto, Japan, September 1991.

[16] B. Mukherjee. WDM-Based local lightwave networks Part I: Single-hop systems. *IEEE Network Magazine*, pages 12–27, May 1992.

[17] D. Park, H. G. Perros, and H. Yamashita. Approximate analysis of discrete-time tandem queueing networks with bursty and correleated input traffic and customer loss. *Operations Research Letters*, 15:95–104, 1994.

[18] A. O. Zaghloul and H. G. Perros. Approximate analysis of a discrete-time polling system with bursty arrivals. *Modelling and Performance Evaluation of ATM Technology (Perros, Pujolle, Takahashi Eds.)*, 1993.

[19] W. Stewart. *Numerical Solutions of Markov Chains.* Princeton University Press, Princeton, New Jersey, 1994.