Variable Length Packet Switches: Input Queued Fabrics with Finite Buffers, Speedup and Parallelism

D. Manjunath[1] and Biplab Sikdar[2]

[1] Department of Electrical Engineering, Indian Institute of Technology, Powai Mumbai 400 076 INDIA
[2] Department of ECSE, Rensselaer Polytechnic Institute, Troy NY 12180 USA

Abstract. We investigate non blocking, variable length packet switches by focusing on performance evaluation and architectures to increase the throughput of such switches. With TCP/IP becoming the dominant protocol suite in the Internet, the analysis of variable length packet switches is necessary to understand the performance of the core routers and switches. We first present analytic models for delays and overflow probabilities in a variable length packet switch with finite buffers for both Poisson and self-similar packet arrival processes. The second part of the paper investigates various means to increase the throughput of these switches. As an alternative to VOQ-CIOQ switches that are known to be necessary to provide practical 100% throughput and QoS we consider a FIFO-CIOQ switch with speedup and multiple parallel planes of switches to minimise the delay in the input queue. We present analytic models for evaluating the impact of speedup and parallelism on increasing the throughput of the switch and show that with a parallelism of 4, it is possible to achieve 99.9% throughput.

1 Introduction

Fixed length packet switches have been studied extensively in the context of ATM switching fabrics. However, the Internet is primarily TCP/IP with variable length packets and arrivals not slotted in time and it is become important to analyse switching architectures for variable length packets. Also, while output queued (OQ) switches can provide 100% throughput and arbitrary QoS provisions efficiently, they are deemed infeasible to implement at high speeds and high port densities due to the switch and memory speed requirements. Thus there is considerable interest in architecture and performance of input queued (IQ) switches. Much of switching literature, be it on performance analysis, design and architectures or QoS issues, assume fixed length packets. Variable length packets can be switched using by breaking them up into fixed length cells, switch the cells and then reassemble the cells at the output but they have overheads from cell padding and headers and have additional circuit and control complexities for fragmentation and reassembly. This leads us to believe that for variable length packet switches with non slotted arrivals, a FIFO-CIOQ switch with speedup will be architecturally simpler than a VOQ-CIOQ switch and will achieve low latencies in the input queue, which in turn will enable it to use output port QoS schedulers. Alternative to speedup, FIFO-CIOQ switch could have multiple parallel switching planes, such that more than one packet is switched to an output queue *simultaneously.* This latter feature will be called parallelism. In this paper we develop analytical models for the throughput, delay and loss performance of a variable length FIFO-CIOQ switch with speedup and parallelism.

There is surprisingly little study of non time-slotted variable length packet switches. The only architecture that does not breakup variable length packets into smaller cells is reported by Yoshigoe and Christensen [19]. Performance models for variable length packet switches are by Fuhrman [7] and Manjunath and Sikdar [11, 12]. In [7], the packet delay in a non blocking $M \times N$ input queued packet is analysed with variable length packets arriving according to an i.i.d. Poisson processes at all the inputs and having uniform routing probabilities. In [11, 12] we analysed these switches for arbitrary Poisson and self similar arrival and routing probabilities and infinite buffers. In this paper we continue that work with finite buffer analysis for Poisson and self-similar packet arrivals. We also analyse speedup and parallelism. The rest of the paper is organised as follows. In the next section we describe the model and the analysis method. In Section 3 we present the analysis and results from the analysis for a finite buffered input queued switch with Poisson and self-similar arrivals. In Section 4 we discuss the comparative merits of speedup and parallelism and present some results from our analysis for CIOQ switches.

2 Analytic Models: Solution Method

We consider a single stage unslotted, internally nonblocking $M \times N$ input queued packet switch with packet arrivals of rate λ_i to input port i. Packet lengths are exponential with unit mean and a packet input to port i chooses a destination j with probability p_{ij}. The line rate on output port j is μ_j. Input queue is FIFO. When a packet moves to the head of its input queue, if its destination is busy, the packet will wait at the head of the input queue till the destination output port is free and chooses to evacuate it. When an output port finishes service, the first blocked HOL packet for that output is served. From above, arrival rate to output port j, Λ_j, and its utilisation, η_j, are

$$\Lambda_j = \sum_{i=1}^{M} \lambda_i p_{ij} \qquad \eta_j = \frac{\Lambda_j}{\mu_j} \tag{1}$$

There are two components to the sojourn time of an input packet in the switch - waiting time till it moves to the head of the line (HOL) and the time spent at the HOL till the HOL packets blocked earlier at other input queues finish their service and the packet is evacuated. The time spent at the HOL of the input queue corresponds to the “service time” in the input queue. For Poisson packet arrivals, each input queue is a M/G/1/K queue with service time equal to the time spent by a packet at its HOL. Thus, we now need the distribution of the time spent at the HOL of the queue, which is obtained using techniques similar to the analysis of blocking queueing networks [18].

The analysis is similar to the technique outlined in [12]. We summarise that technique here. Consider output port j. Output port j evacuates a packet from the HOL of the inputs at rate μ_j. The throughput from output j will be $\Lambda_j(1-PB_j)$, (PB_j is the blocking probability of packets destined for output j) because of finite input buffers. Now, we approximate the arrival process to the virtual queue of output j by a Poisson process of throughput $\Lambda_j(1-PB_j)$ and model the output queue as a M/M/1/M queue. In fact, we can show that as $M \to \infty$, the arrival process to the output queue is indeed Poisson following the technique in [9]. Therefore the "arrival rate" corresponding to this throughput, let us call this the effective arrival rate Λ'_j, will be obtained by solving for Λ'_j in the equation

$$\Lambda_j(1-PB_j) = \Lambda'_j \left[\frac{1-{\eta'}_j^M}{1-{\eta'}_j^{M+1}}\right] \tag{2}$$

where $\eta'_j = \Lambda'_j/\mu_j$. The probability that a packet arriving to the head of an input queue wanting to go to output j sees k packets ahead of it, $\pi_j(k)$, using M/M/1/M queueing theory is given by

$$\pi_j(k) = \left[\frac{(1-\eta'_j)(\eta'_j)^k}{1-(\eta'_j)^M}\right] \qquad \text{for } k = 0, 1, \cdots M-1 \tag{3}$$

In the virtual queue of output port j if there are k packets ahead of it, the packet has to wait for the evacuation of these packets before it can begin its service and its waiting time is a k stage Erlangian distribution (sum of the k independent, exponentially distributed evacuation times). In addition to the blocking delay there is the evacuation time that has an exponential distribution of mean $1/\mu_j$. The Laplace-Stieltjes Transform (LST) of the unconditional distribution of the sojourn time at the head of input i, $\mathcal{X}_i(s)$, can be seen to be

$$\mathcal{X}_i(s) = \sum_{j=1}^{N} p_{ij} \left[\sum_{k=1}^{M-1} \pi_i(k) \left(\frac{\mu_j}{\mu_j + s}\right)^k\right] \left[\frac{\mu_j}{\mu_j + s}\right] \tag{4}$$

Here the term in the first square brackets corresponds to the blocking delay and that in the second corresponds to the evacuation time given that the packet wants to go to output j. The first three moments of the blocking delay at input queue i, $\overline{B_i}$, $\overline{B_i^2}$ and $\overline{B_i^3}$ respectively, are

$$\overline{B_i} = \sum_{j=1}^{N} p_{ij} \sum_{k=1}^{M-1} \pi_j(k) \frac{k}{\mu_j} \qquad \overline{B_i^2} = \sum_{j=1}^{N} p_{ij} \sum_{k=1}^{M-1} \pi_j(k) \frac{k(k+1)}{\mu_j^2}$$
$$\overline{B_i^3} = \sum_{j=1}^{N} p_{ij} \sum_{k=1}^{M-1} \pi_j(k) \frac{k(k+1)(k+2)}{\mu_j^3} \tag{5}$$

Likewise, the first three moments of the service time for the input queue, $\overline{X_i}$, $\overline{X_i^2}$, $\overline{X_i^3}$ respectively, are

$$\overline{X_i} = \overline{B_i} + \sum_{j=1}^{N} \frac{p_{ij}}{\mu_j} \qquad \overline{X_i^2} = \overline{B_i^2} + 2\overline{B_i} \sum_{j=1}^{N} \frac{p_{ij}}{\mu_j} + 2 \sum_{j=1}^{N} \frac{p_{ij}}{\mu_j^2}$$
$$\overline{X_i^3} = \overline{B_i^3} + 3\overline{B_i^2} \sum_{j=1}^{N} \frac{p_{ij}}{\mu_j} + 6\overline{B_i} \sum_{j=1}^{N} \frac{p_{ij}}{\mu_j^2} + 6 \sum_{j=1}^{N} \frac{p_{ij}}{\mu_j^3} \tag{6}$$

3 Finite Buffer Analysis for Poisson and Self-Similar Arrivals

Along with the approximate distribution of the input queue service time, we use well known results from M/G/1/K and MMPP/G/1/K queueing theory to analyse the input queued switch for these packet arrival models. We first derive the expressions for the Poisson arrivals and then for self-similar arrivals.

3.1 Poisson Arrivals

The queue at input port i is modeled as a M/G/1/K_i queue with Poisson arrivals of rate λ_i and service time distribution given by Eqn. 4. We analyse this queue using the diffusion approximation method of [8]. From [8], the probability that there are n packets in input queue i, $\zeta_i(n)$, is given by

$$\zeta_i(n) = \begin{cases} c_i \hat{\zeta}_i(n), & 0 \le n < K_i \\ 1 - \frac{1 - c_i(1-\rho_i)}{\rho_i}, & n = K_i \end{cases} \tag{7}$$

$$\text{where} \quad c_i = \left\{ 1 - \rho_i \left[1 - \sum_{n=0}^{K_i - 1} \hat{\zeta}_i(n) \right] \right\}^{-1} \tag{8}$$

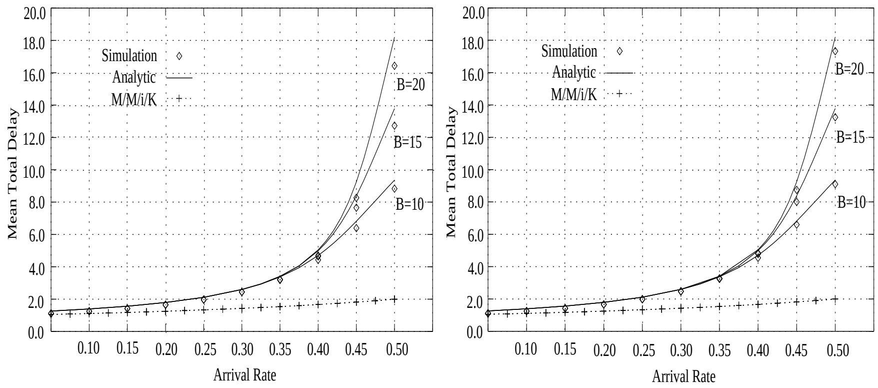
Here $\rho_i = \lambda_i \overline{X_i}$, $\hat{\zeta}_i(n)$ is the probability that there are n packets in an M/G/1/∞ with arrival rate λ_i and service time distribution of Eqn. 4. Using the diffusion approximation, the probabilities $\hat{\zeta}_i(n)$ can be written as

$$\hat{\zeta}_i(n) = \begin{cases} 1 - \rho_i, & n = 0 \\ \rho_i (1 - \hat{\rho}_i)(\hat{\rho}_i)^{n-1} & n \ge 1 \end{cases} \tag{9}$$

$$\text{where} \quad \hat{\rho}_i = \exp\left(\frac{2\overline{X_i}^2 (\lambda_i \overline{X_i} - 1)}{\lambda_i \overline{X_i}^3 + \overline{X_i^2} - \overline{X_i}^2} \right) \tag{10}$$

The mean packet delay from input i, D_i, is obtained from Little's theorem.

$$D_i = \frac{1}{\lambda_i (1 - \zeta_i(K_i))} \sum_{n=0}^{K_i} n \zeta_i n \tag{11}$$



(a) 32 × 32 switch (b) 64 × 64 switch

Fig. 1. Mean delay for a 32 × 32 and 64 × 64 switch from analysis and simulation

	32 × 32, B=15			32 × 32, B=20			64 × 64, B=15			64 × 64, B=20		
λ	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q
0.4	0.0006	0.0008	6.4E-7	0.0001	0.0001	6.6E-9	0.0007	0.0008	6.4E-7	0.0001	0.0001	6.6E-9
0.5	0.0363	0.0443	1.5E-5	0.0256	0.0335	4.7E-7	0.0401	0.0443	1.5E-5	0.0295	0.0336	4.8E-7
0.6	0.1595	0.1718	0.0002	0.1560	0.1687	1.5E-5	0.1657	0.1718	0.0002	0.1622	0.1687	1.5E-5

Table 1. Loss probabilities in a 32 × 32 and 64 × 64 switch

Since arrivals are Poisson, blocking probability at input i will be $\zeta_i(K_i)$. Note that the effective arrival rate for a virtual queue of an output was derived using the blocking probability for packets destined for output port i and

$$PB_j = \sum_{i=1}^{M} \zeta_i(K_i) \frac{p_{ij}\lambda_i}{\Lambda_j} \tag{12}$$

The analytical model is solved by iterating on Eqns 1-12.

Numerical results are obtained for 32 × 32 and 64 × 64 switches with i.i.d. input arrivals and uniform routing. Fig. 1 shows the mean queuing delay. The good agreement between the simulation and the analytical results justifies our approximations. The significant "delay-penalty" for input queued switch can be seen by comparing with that output queued switch, modeled as a M/M/1/K queue. Table 1 shows the loss probabilities. Once again, the close match between analytical and simulation results indicate the goodness of our approximations. Also observe that the "blocking-penalty" of input queueing is significant.

3.2 Self-Similar Arrivals

Having modeled switch behavior for a "simplistic" Poisson arrivals, we now analyse under a more realistic model of self similar packet arrival processes. Among the many models for simulation of exact and approximate self similar processes, we will choose the one by Andersen and Nielsen [1]. Here a Markovian approach in which a self-similar process is obtained by superposing a number of two state Interrupted Poisson Processes (IPPs) with the resultant arrival process being an MMPP. This method, in addition to allowing the burstiness over a number of time scales with the desired covariance structure, allows us to use the well developed MMPP queuing theory for our analysis.

The distribution of the time spent by a packet at head of the input queue, its "service time" distribution, is approximated by a phase type distribution whose LST is given by Eqn. 4. Let the generator and the rate matrix of the MMPP process corresponding to the arrivals at input port i be denoted by $\mathbf{Q_i}$ and $\mathbf{R_i}$ respectively. We only report the loss analysis for the MMPP arrival process. Having obtained the service time distribution and described the arrival process we use the MMPP/G/1/K analysis of [2] and the efficient evaluation techniques for evaluating the loss probabilities from [3] to solve our model. The following notation will be used for each input i and to simplify the notation we will omit the subscript corresponding to the input port.

$\mathbf{U}$: $m \times m$ matrix given by $(\mathbf{R}-\mathbf{Q})^{-1}\mathbf{R}$

$\mathbf{P}(n,t)$: $m \times m$ matrix whose (p,q)th element denotes the conditional probability of reaching phase q of the MMPP and having n arrivals during a time interval of length t, given that we start with phase p at time $t=0$.

$\mathbf{A}_n$: $m \times m$ matrix given by $\int_0^\infty \mathbf{P}(n,t)dX(t)$, $n \geq 0$ where $X(t)$ is the distribution of the service time as obtained previously.

$\mathbf{A}$: $m \times m$ matrix given by $\sum_{n=0}^\infty \mathbf{A}_n = \int_0^\infty e^{\mathbf{R}t}dX(t)$

$\mathbf{\Pi}(i)$: m dimensional vector whose p^{th} element is the limiting probability at the embedded epochs of having i packets in the queue and being in phase p of the MMPP, $i = 0, 1, \cdots, K-1$.

$\mathbf{I}$: $m \times m$ identity matrix

The steady-state probability distribution of the queue length of the embedded Markov chain at the departure instants can be calculated using the following approach. The matrices $\mathbf{A}_n$ are first calculated using the technique described in [6]. From $\mathbf{P}$, we then find the matrix sequence $\{\mathbf{C}_i\}$, independent of the buffer size K, such that $\mathbf{\Pi}(i) = \mathbf{\Pi}(0)\mathbf{C}_i$ for $i = 0, 1, \cdots, K-1$. The matrices $\mathbf{C}_i$ are calculated using the following equation

$$\mathbf{C}_{i+1} = \left[\mathbf{C}_i - \mathbf{U}\mathbf{A}_i - \sum_{v=1}^{i} \mathbf{C}_v \mathbf{A}_{i-v+1}\right] \mathbf{A}_0^{-1} \qquad i = 1, \cdots, K-2 \tag{13}$$

beginning with $\mathbf{C}_0 = \mathbf{I}$. The vector $\mathbf{\Pi}(0)$ is then determined using

$$\mathbf{\Pi}(0)\left[\sum_{v=0}^{K-1} \mathbf{C}_v + (\mathbf{I}-\mathbf{U})\mathbf{A}(\mathbf{I}-\mathbf{A}+\mathbf{e}\mathbf{\Phi})^{-1}\right] = \mathbf{\Phi} \tag{14}$$

The loss probability can then be found using the following expression

$$\mathrm{P}_{\text{loss}} = 1 - (\overline{X}\mathbf{\Phi}\mathbf{R}\mathbf{e})^{-1}\left[1 + \mathbf{\Pi}(0)(\mathbf{R}-\mathbf{Q})^{-1}\overline{X}^{-1}\mathbf{e}\right]^{-1} \tag{15}$$

	`pAug.TL` B=500			`pAug.TL`, B=1000			`pOct.TL`, B=500			`pOct.TL`, B=1000		
λ	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q
0.1	0.057	0.065	0.049	0.054	0.061	0.045	0.022	0.030	0.0003	0.014	0.020	6.9E-5
0.2	0.076	0.099	0.057	0.075	0.094	0.054	0.086	0.110	0.027	0.074	0.095	0.018

Table 2. Loss probabilities in 8×8 switch for Bellcore traces `pAug.TL` and `pOct.TL`.

Numerical results are obtained as follows. The Hurst parameter, correlation at lag 1 (ρ), arrival rate λ and the time scales over which burstiness occurs in the Bellcore traces `pAug.TL` and `pOCT.TL` [10] are used to fit the parameters c_i^{1j}, c_i^{2j} and r_i for $j = 1, \cdots, 4$ of the MMPP model described in [1]. The analytical results are obtained for the MMPP/G/1 queue as described earlier. We also develop a simulation model in which the arrivals are MMPP with parameters derived above. The arrival process generator is validated by simulating a single server queue and comparing with the results given in [5]. A separate and independent MMPP arrival process generator is used for each input port with the traces generated by each of the sources having identical statistical properties. Thus, statistically identical self-similar traces but with different sample paths are used as the input processes to the simulation model. We use the Bellcore traces `pAug.TL` ($H = 0.82$ and $\rho = 0.582$) and `pOct.TL` ($H = 0.92$ and $\rho = 0.356$) with burstiness over 4 time scales.

Table 2 shows the overflow probabilities for a 8×8 switch fed with Bellcore traces `pAug.TL` and `pOct.TL`. Packet loss is significant for low loads and the drop in loss rates does not scale with increase buffer sizes. Such behavior of queues for long-range dependent arrivals are well known [5]. A large number of buffers are required with long-range dependent traffic to support even moderate loss probabilities. The tables also show the loss probabilities for the MMPP/M/1/K queue for different values of K. We only use it as an approximation to get an indication of the "input queueing penalty" for the more realistic model of self-similar packet arrivals. (Bernoulli splitting of packet arrivals into N streams corresponding to splitting the packet arrivals to the N outputs and combining of M such streams at output queue will make the characteristics of the arrival process to the output queue different from that to the input queue.)

4 Increasing Throughput in an Input Queued Switch

For Bernoulli packet arrivals and uniform routing, it is well known that the maximum throughput of an input queued switch is 0.586 [9] and it is 0.5 for variable length packet switches with Poisson arrivals and exponential packet lengths [12]. Output queued switches can provide 100% throughput, but with higher implementation complexity, typically of $\mathcal{O}(N)$ more than the input queued switch. The following two mechanisms can approximate the output queued switch but with reduced complexity in fixed length packet switches. The first method is a virtual output queued (VOQ) switch was shown to be able to achieve 100%

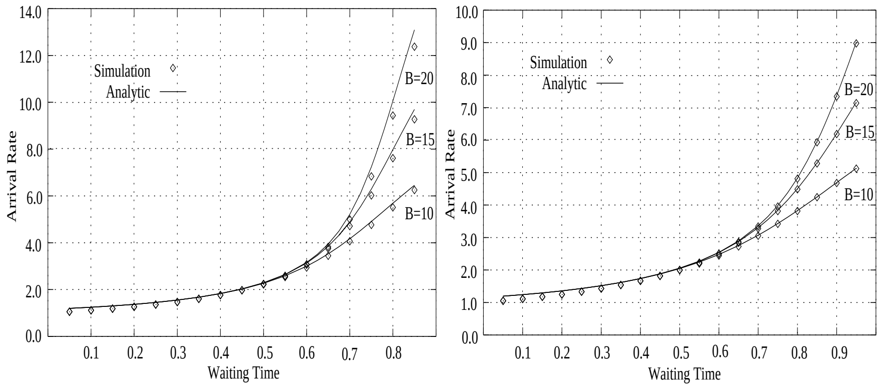
throughput in [13]. Switch complexity is traded for scheduling complexity. The scheduling algorithms of [13] cannot be used in practical switches and many "practical" algorithms that have been reported in literature [14–16]. The algorithms of [14] impractical to be implemented in hardware [15] and the "practical algorithm" of [15] has a complexity of $\mathcal{O}(N^{2.5})$. Thus, although VOQ reduces the effect of HOL blocking, they are complex and do not scale well enough to offset the throughput disadvantage of input queued switches for large N. Further, note that the result of [13] is for iid arrivals and uniform destinations. In [4] it has been shown that if in each slot there are two scheduling cycles, then 100% throughput is achieved under any arrival process and destination distributions. However, this will require the use of queues at the output! Further, for variable length packets, an obvious implementation will be to break the packets into fixed length cells, switch the cells and reassemble them at the output. This method, in addition to requiring disassembly and reassembly circuits, will also mean that the switching hardware has to operate at a higher rate than the line rate to account for padding and cell header overheads.

The second method is the combined input-output queued (CIOQ) switch in which either the switch fabric operates at a higher rate than the line rate on the output ports or there are multiple parallel switching fabrics or both. Given that to achieve high throughput in a practical VOQ switch we need to operate the switch at a rate higher than the line rate, a CIOQ switch with speedup or parallel fabrics seems appealing. At this time we would like to separate the concepts of parallelism and speedup that have been used interchangeably in switching literature. In the former multiple packets can be evacuated to the same output and/or from the same input *simultaneously* or in parallel possibly through multiple parallel switching fabrics. In the case of speedup, the evacuation rate from the input queue HOL is higher than the line rate. Oie et al [17] use the term speedup to connote parallelism. In our analysis speedup is modeled by using a higher μ for the evacuation rate, or equivalently a lower λ for the arrival rate. In our analysis speedup is modeled by using a higher μ for the evacuation rate. We present only the analysis for the case of parallelism.

4.1 Analysis for Parallelism Factor m, $m > 1$

The delay analysis is similar to the previous analyses. First, consider the infinite buffer case. If there are more than m HOL packets at the inputs destined for a particular output port, m of them are served simultaneously while the others are blocked. We approximate the input process to the queue to be Poisson. Thus the virtual queue of each output port will be modeled as an M/M/m/M queue and the effective arrival rate to output port j corresponding to a throughput of Λ_j is obtained by solving for Λ'_j in

$$\Lambda_j(i - PB_j) = \Lambda'_j \left[1 - \frac{\left[\frac{(\eta'_j)^M m^m}{m!}\right]}{\left[\sum_{k=0}^{m-1} \frac{(m\eta'_j)^k}{k!} + \sum_{k=m}^{M} \frac{(\eta'_j)^k m^m}{m!}\right]} \right] \tag{16}$$



(a) Parallelism of 2 (b) Parallelism of 4

Fig. 2. Mean delay for a 32×32 and 64×64 switch from analysis and simulation

where $\eta'_j = \frac{\Lambda'_j}{m\mu_j}$. $\theta_j(k)$, $\pi_j(k)$ and $\mathcal{X}_i(s)$ are obtained as before by considering an M/M/m/M queue at the outputs. Similarly the blocking and total delay moments are also obtained like before and are given by,

$$\begin{aligned}\overline{B_i} &= \sum_{j=1}^{N} p_{ij} \sum_{k=m}^{M-1} \pi_j(k) \frac{k-m+1}{\mu_j} \\ \overline{B_i^2} &= \sum_{j=1}^{N} p_{ij} \sum_{k=m}^{M-1} \pi_j(k) \frac{(k-m+1)(k-m+2)}{\mu_j^2} \\ \overline{B_i^3} &= \sum_{j=1}^{N} p_{ij} \sum_{k=m}^{M-1} \pi_j(k) \frac{(k-m+1)(k-m+2)(k-m+3)}{\mu_j^3}\end{aligned} \tag{17}$$

We now use these moments of the blocking delay in Equation 6 to obtain the first three moments of the total delay. The summations over the index k for the blocking delay is from $k = m$ to $k = M - 1$ because only when there are $\geq m$ waiting in the virtual queue will the packet at the HOL of an input queue have to wait. Mean delay and loss probabilities at the input ports are obtained as in the previous section by modeling each input port as a M/G/1/K and MMPP/G/1/K queue for Poisson and self-similar packet arrivals respectively.

Fig. 2 plots the throughput-mean delay characteristics for a CIOQ switch with Poisson arrivals, and parallelism of 2 and 4 for various buffer sizes and arrival rates normalised to the switch rate. Note the excellent agreement between analytic and simulation results. Tables 3 and 4 show the blocking probabilities for parallelism of 2 and 4 for a 64×64 switch with Poisson arrivals and an 8×8 switch self similar arrivals respectively. The maximum throughput for a given parallelism is obtained by solving for λ in $\lambda\overline{X} = 1.0$. We find that for a parallelism of 4 the achievable throughput is 99.3% while it is 82.8% for a parallelism of 2.

	Parallelism=2								
	K=10			K=15			K=20		
λ	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q
0.6	0.0044	0.0048	0.0024	0.0005	0.0006	0.0002	0.0001	0.0001	1.46E-5
0.7	0.0196	0.0202	0.0086	0.0057	0.0060	0.0014	0.0017	0.0018	0.0002
0.8	0.0571	0.0590	0.0235	0.0333	0.0352	0.0072	0.0214	0.0233	0.0023
	Parallelism=4								
0.60	0.0024	0.0026	0.0024	0.0002	0.0002	0.0002	1.40E-5	1.80E-5	1.46E-5
0.70	0.0087	0.0090	0.0086	0.0015	0.0015	0.0014	0.0003	0.0003	0.0002
0.80	0.0238	0.0240	0.0235	0.0074	0.0075	0.0072	0.0024	0.0024	0.0023

Table 3. Loss rates in a 64×64 switch with parallelism of 2 and 4 for Poisson arrivals.

	Parallelism=2								
	K=250			K=500			K=1000		
λ	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q	Sim.	Ana.	o/p Q
0.10	0.0573	0.0505	0.0506	0.0516	0.0487	0.0488	0.0444	0.0453	0.0454
0.15	0.0643	0.0549	0.0549	0.0523	0.0531	0.0531	0.0513	0.0497	0.0498
0.20	0.0686	0.0596	0.0591	0.0595	0.0579	0.0573	0.0559	0.0545	0.0540
	Parallelism=4								
0.10	0.0544	0.0502	0.0506	0.0504	0.0484	0.0488	0.0439	0.0450	0.0454
0.15	0.0631	0.0541	0.0549	0.0513	0.0523	0.0531	0.0510	0.0489	0.0498
0.20	0.0681	0.0581	0.0591	0.0603	0.0563	0.0573	0.0546	0.0530	0.0540

Table 4. Loss rates in a 8×8 switch with parallelism of 2 and 4 for self-similar arrivals corresponding to the Bellcore trace `pAug.TL`

.

5 Conclusion

We presented a general analytical model for finite buffer analysis of input queued, variable length packet switches. This analytical model can also be used with other arrival process and for random order of service and priority queues at the input.

Numerical results show that the buffer requirements to provide acceptable blocking probabilities for self-similar arrivals are considerably higher than those for Poisson arrivals, even for low utilisations. These results emphasise the fact that realistic assumptions about the packet arrival process needs to be made to resonably estimate the performance characteristics.

We also discussed two alternatives to increase the throughput of input queued switches, VOQ and FIFO-CIOQ, and analysed FIFO-CIOQ switches based on parallelism. Our analysis suggests that FIFO-CIOQ can be a viable alternative to time slotted, fixed packet length VOQ-CIOQ switches especially if the latter have speedup ≥ 1. A complex centralised scheduler that has to collect information about N^2 queues and fragmentation and reassembly overhead is traded for multiple switching planes. Although a VOQ-CIOQ switch can suppor arbitrary

QoS, technology limitations will make the centralised QoS scheduler infeasible in the Tbps region. Instead, a FIFO-CIOQ switch with speedup and parallelism can minimise input queueing delay and an output QoS scheduler can be used.

References

1. A. T. Andersen and B. F. Nielsen, "A Markovian approach for modeling packet traffic with long-range dependence," *IEEE Jour on Sel Areas in Commun,* vol. 16, no. 5, pp. 719-732, June 1998.
2. A. Baiocchi and N. Bléfari-Melazzi, "Steady-State Analysis of the MMPP/G/1/K Queue," *IEEE Trans on Commun,* vol. 41, no. 4, pp. 531-534, Apr 1993.
3. A. Baiocchi and N. Bléfari-Melazzi, "Analysis of the Loss Probability of the MAP/G/1/K Queue Part II: Approximations and Numerical Results," *Stochastic Models,* vol. 10, no. 4, pp. 895-925, 1994.
4. J.G. Dai and Balaji Prabhakar, "The Throughput of Data Switches with and without Speedup," *Proc of IEEE Infocom-2000,* Mar 2000.
5. A. Erramilli, O. Narayan and W. Wilinger, "Experimental Queuing Analysis with LRD Packet Traffic," *IEEE/ACM Trans on Networking,* vol 4, no 2, Apr 1996.
6. W. Fischer and K. Meier-Hellstern, "The Markov-modulated Poisson process (MMPP) cookbook," *Performance Evaluation,* vol. 18, no. 2, pp. 149-171, 1993.
7. S. W. Fuhrman, "Performance of a Packet Switch with a Crossbar Architecture," *IEEE Trans. on Commun.,* vol COM-41, pp. 486-491, 1993.
8. E. Glenbe and G. Pujolle, "Introduction to Queueing Networks," *John Wiley and Sons,* 1987.
9. M. J. Karol, M. G. Hluchyj and S. P. Morgan, "Input Versus Output Queuing on a Space-Division Packet Switch," *IEEE Trans on Commun,* vol. COM-35, no. 12, pp. 1347-1356, Dec 1987.
10. W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, "On the self-similar nature of Ethernet traffic (Extended Version)," *IEEE/ACM Transactions on Networking,* vol. 2, no. 1, pp. 1-15, Feb 1994.
11. D. Manjunath and B. Sikdar, "Input Queued Packet Switches for Variable Length Packets : A Continuous Time Analysis," *Proc of IEEE Broadband Switching Systems,* Kingston, ON, pp. 65-69, June 1999.
12. D. Manjunath and B. Sikdar, "Variable Length Packet Switches: Delay Analysis of Crossbar Switching under Poisson and Self Similar Traffic," *Proc of IEEE Infocom-2000,* Tel Aviv, Israel, Mar 2000.
13. N. McKeown, V. Anantharam and J. Walrand, "Achieving 100% throughput in an Input Queued Switch," *Proc of IEEE Infocom-96,* pp.296–302.
14. N. Mckeown and A. Mekkittikul, "A Starvation Free Algorithm for Achieving 100% Throughput in an Input Queued Switch ", *Proc of ICCCN 96,* Oct.1996.
15. N. McKeown and A. Mekkittikul, "A Pratical Scheduling Algorithm to Achieve 100% Throughput in Input-Queued Switches", *Proc of IEEE INFOCOM-98,* Apr.1998.
16. N. McKeown, "iSLIP: A Scheduling Algorithm for Input Queued Switches," *IEEE/ACM Trans on Networking,* vol 7, pp 188-201, 1999.
17. Y. Oie, M. Murata, K. Kubota and H. Miyahara, "Effect of Speedup in Nonblocking Packet Switch," *Proc of IEEE ICC, 1989,* pp 410-414.
18. H. Perros, "Queuing Networks with Blocking," *Oxford University Press,* 1994.
19. K. Yoshigoe and K. Christensen, "A Parallel Polled Virtual Output Queued Switch with a Buffered Crossbar," *Proc of IEEE HPSR,* pp. 271-275, May 2001.