Queueing Analysis of Scheduling Policies in Copy Networks of Space Based Multicast ATM Switches

Biplab Sikdar
Dept. of ECSE
Rensselaer Polytechnic Institute
Troy, NY 12180 USA
Ph: +1-518-276-8289
bsikdar@networks.ecse.rpi.edu

D. Manjunath
Dept. of Electrical Engg.
Indian Institute of Technology
Powai, Mumbai 400 076 INDIA
Ph: +91-22-5767427
dmanju@ee.iitb.ernet.in

Abstract— **Space based multicast switches use copy networks to make copies of the inputs before forwarding them to the multiple destinations. In a copy network a copy scheduling policy determines the order in which the copies of the input packets are made and this policy affects the performance of the multicast switch. In this paper we present queueing models for the copy network for three scheduling policies – acyclic service without fanout splitting and cyclic service with and without fanout splitting. From these queueing models we obtain the average copy delay and the** ***sustainable throughput*****, the maximum load that can be applied to all the input ports without resulting in an unstable queue at any of the inputs, for the above scheduling policies.**

I. Introduction

Two basic design paradigms are known for multicast packet switches, - *space based* and *time based* [1]. In a space-based space division multicast switch a copy network is followed by a routing stage. In the copy network, a copy scheduling algorithm is used to determine the order in which the input ports are served. Also, since the total number of copies requested in a time slot can exceed its capacity some of the copy requests may need to be buffered leading to an additional delay stage in the switch. Most literature on performance analysis of multicast switches, for example [3], [4], [5], do not account for the copy network structure. Lee [2] and Liew [6] analyse the copy generation process in terms of overflow and loss probabilities but queueing processes are not analysed. In this paper we do a queueing analysis of the copy network under three copy scheduling policies - acyclic service without fanout splitting and cyclic service with and without fanout splitting. We also introduce and evaluate *sustainable throughput* for a copy network, defined as the maximum load that can be applied to all inputs without causing an unstable queue at any one, for the above scheduling policies. In Section II queueing models for the different scheduling policies are presented and in Section III numerical results from these models are compared with those from simulation.

II. Performance Analysis of the Copy Network

We consider a copy network proposed by Lee [2] and based on a broadcast banyan network. We call this the Lee Copy Network (LCN). Time is slotted and all the inputs are synchronised to have packet arrivals at the beginning of a slot. Packet lengths are fixed and equal to the slot length. The copy network is a cascaded combination of a packet header encoder and a decoder. The switching elements in the broadcast banyan network are capable of routing an input packet to either or both its outputs and the routing algorithm used is the boolean interval splitting algorithm described in [2]. (See Figure 1.) Consider an $M \times N$ copy network that can make a total of at most N copies in a slot. At the beginning of a slot let input i request c_i copies. In the simplest copy scheduling algorithm that we consider, acyclic service, the running adder begins with input 1 and obtains the running sum of c_is starting from port 1. In this service policy port i is serviced (all the copies requested by the input packet are made) if $\sum_{j=1}^{i} c_j \leq N$ in that slot. Other scheduling policies are discussed later.

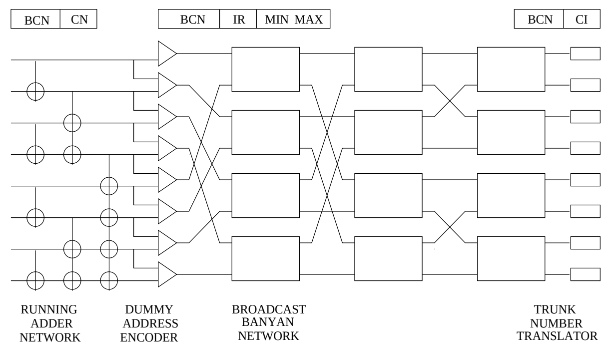
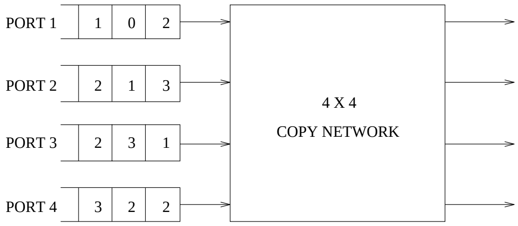


Fig. 1. Lee's Copy Network for a Multicast Packet Switch [2]

The copy network can produce only N copies in a time slot. In any slot if the sum of copy requests at the heads of input port queues exceeds N, some requests are not served in the slot. For the case when the requests that are not served in the slot are dropped, an exact analytical model for the loss probability is reported in [7]. In this paper, we consider queueing of the packets that are not served in the slot. We assume that within each input queue the packets are served in FIFO order and only the packet at the head of the queue is considered for service in a slot. Thus, HOL blocking limits the maximum achievable throughput from each queue. Also, "predecessor port blocking" further limits the achievable throughput. To explain this latter phenomenon, consider the following example. Let ports



In a slot, assume that service begins with port 1. The head of the line packets at the other ports is not served because the running adder sum exceeds 4 at port 2. The second packet at port 2 could have been served but was blocked due to the packet at the head of the line. Also, the packet at port 3 could have been served along with the packet at port 1 if the running adder had skipped port 2 but is blocked because of "predecessor port blocking".

Fig. 2. Blocking in a Copy Network

$i-1$, i and $i+1$ of request c_{i-1}, c_i and c_{i+1} copies respectively. Ports i and $i+1$ will not be served if $\sum_{j=1}^{i} c_j > N$ even if $\sum_{j=1}^{i-1} c_j + c_{i+1} \leq N$, i.e., port $i+1$ is blocked because the running sum exceeds N at i although $i+1$ could have been served by not serving i. Figure 2 discusses both these phenomena. It is known that in input queued unicast switches, HOL blocking limits maximum throughput to 0.586 per port. *Sustainable throughput* is used to obtain a similar characterisation for the copy network.

Traffic into the copy network is assumed to be independent and identical Bernoulli processes at each port and the copy request of the packets, c_is, at the head of the line at the inputs in a given slot are independent of those from the previous slots. This assumption is necessary to keep the analysis simple and manageable. Finally, we assume that the input buffers have infinite capacity. In the analysis technique that we develop here, most of these assumptions can be relaxed at the cost of computational complexity.

A. Acyclic Service without Fanout Splitting

Under this policy, in each slot, input ports are served serially, starting from port 1. All ports for which the running adder sum is less than or equal to N are served in the slot. Since service always starts from port 1, the policy is unfair to ports with higher addresses and the service rate decreases as port address increases. Thus the sustainable throughput is the throughput achievable by port M. Also, since fanout splitting is not used, in cases where a part of a copy request could be served, resources remain unutilised.

Port 1 is served in every slot. Input ports $2 \cdots M$ are modeled as discrete time M/M/1 queues. The rate of the Bernoulli arrival process at each of the ports is ρ and active inputs request k copies with probability $q(k)$. Let X_i denote the number of copies requested by port i. Let $f_{H,i}(k)$ be the probability that k copies are requested by packet at head of input queue i. Let $P_{0,i}$ denote the probability that the head of queue i is non empty. If $Q(z)$ is the probability generating function of $q(k)$, $f_{H,i}(k)$ and its probability generating function will be,

$$f_{H,i}(x_i) = \begin{cases} P_{0,i}, & x_i = 0 \\ (1-P_{0,i})\, q(x_i), & x_i = 1, 2, \cdots \end{cases}$$

$$\mathcal{F}_{H,i}(z) \equiv \sum_{x_i} f_{H,i}(x_i) z^{x_i} = P_{0,i} + (1-P_{0,i})\, \mathcal{Q}(z)$$

Effective service rate at input i, $\mu(i)$, depends on the arrival processes and effective service rates at ports $j = 1, \cdots i-1$. We are approximating the service times by a geometric distribution of rate $\mu(i)$. Let $W(i)$ denote the mean waiting time of the packets arriving to input i. Since port 1 is always served in each slot irrespective of the copy requests of the other ports,

$$\mu(1) = 1.0 \quad W(1) = 1.0 \quad P_{0,1} = 1-\rho \tag{1}$$

$\mu(i)$, $W(i)$ and $P_{0,i}$ for $i = 2, \cdots M$, are obtained as follows. Consider a slot in which port i is active and requesting k copies. This request will be served in the slot if $\sum_{j=1}^{i-1} x_j \leq N-k$. Unconditioning on k, $\mu(i)$, the probability that a copy request of port i is served in a slot, is

$$\begin{aligned}\mu(i+1) &= \sum_{k=1}^{N} q(k) \text{ Prob} \left\{ \sum_{j=1}^{i} X_j \leq N-k \right\} \\ &= \sum_{k=1}^{N} q(k) \left[\sum_{\sum_{j=1}^{i} x_j \leq (N-k)} \prod_{j=1}^{i} f_{H,j}(x_j) \right] \end{aligned} \tag{2}$$

The second summation on the RHS of Eqn 2 is over all possible combinations of copy requests from ports 1 to i that sum to less than or equal to $N-k$. This summation is similar to obtaining the normalisation constant in a stochastic knapsack. Therefore, following [8], we can obtain $mu(i+1)$ as follows. Define

$$\delta(k) \equiv \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases} \qquad \Phi_N(k) \equiv \begin{cases} 1 & \text{for } k \leq N \\ 0 & \text{for } k > N \end{cases}$$

Contour integral representations of $\delta(k)$ and $\Phi_N(k)$ are

$$\delta(k) = \oint z^{(k-1)} dz \quad \Phi_N(k) = \oint \left[\frac{z^{(N+1)} - 1}{z-1} \right] \left[\frac{z^k}{z^{(N+1)}} \right] dz$$

with the unit circle in the complex plane as the contour of integration. We can use $\Phi_N(k)$ to represent the summation in Eqn 2 as follows.

$$\begin{aligned}\mu(i+1) &= \sum_{k=1}^{N} q(k) \sum_{x_1=0}^{N} \cdots \sum_{x_i=0}^{N} \prod_{j=1}^{i} f_{H,j}(x_j) \Phi_{N-k}(x_1 + \cdots + x_i) \\ &= \sum_{k=1}^{N} q(k) \oint \sum_{x_1=0}^{N} f_{H,1}(x_1) z^{x_1} \cdots \sum_{x_i=0}^{N} f_{H,i}(x_i) z^{x_i} \\ &\qquad \left[\frac{z^{(N-k+1)} - 1}{z-1} \right] \left[\frac{1}{z^{(N-k+1)}} \right] dz \\ &= \sum_{k=1}^{N} q(k) \oint \left[\frac{z^{(N-k+1)} - 1}{z-1} \right] \left[\frac{1}{z^{(N-k+1)}} \right] \prod_{j=1}^{i} \mathcal{F}_{H,j}(z) dz \end{aligned} \tag{3}$$

It is easy to see that the only poles of the integrand in Eqn 3 inside the unit circle are those at $z = 0$. Hence, by the residue theorem, the integral is equal to the sum of the residues at $z = 0$ and $\mu(i)$ are obtained. From the theory of discrete time M/M/1 queues [9], $W(i)$ and $P_{0,i}$ will be

$$W(i) = \frac{1-\rho}{\mu(i)-\rho} \qquad P_{0,i} = \frac{(\mu(i)-\rho)}{\mu(i)} \tag{4}$$

Eqns 1, 3 and 4 are used to recursively obtain the $W(i)$.

Clearly, $\mu(i)$ decreases with increasing i. From its definition, the sustainable throughput of the copy network under the acyclic scheduling policy without fanout splitting, $\lambda_{A:NFS}$, is determined by the effective service rate at port M. Thus

$$\lambda_{A:NFS} = Q'(1) \max_{\rho}\{\mu(M)\}$$

where $Q'(1)$ is the average number of copies requested by a packet and $\mu(M)$ is obtained as above.

B. Cyclic Service without Fanout Splitting

In this scheduling policy, the inputs are served cyclically. In a slot, the running adder sums the copy requests at the head of the queues beginning at the first input port that was not served in the previous slot and continues sequentially till the sum exceeds N or it has summed request from M ports. The ports where the sum exceeds N are not served in the slot. Fanout splitting is not allowed. Cyclic service provides fairness. If we assume the traffic to be independent and identically distributed at all the ports, the performance metrics will be identical at each input port.

In a slot, the probability that a tagged input port having a packet requesting k copies gets served depends on how many ports are to be served before it in the slot (uniform in $[0, M-1]$), and on the copy request distribution at these ports. The probability of service for the tagged packet can then be obtained by summing the probability that the sum of the copies requested by the head of the line packets of the preceding i ports is at most $N-k$, for all values of i. For $i = 0$, i.e. tagged port is the first one to be served, the probability of service is 1. Thus,

$$\mu(k) = \frac{1}{M} + \sum_{i=1}^{M-1} \frac{1}{M} \left[\sum_{\sum_{j=1}^{i} x_j \leq (N-k)} \prod_{j=1}^{i} f_{H,j}(x_j) \right] \tag{5}$$

Probability of a copy request at the tagged port being served, μ, is obtained by unconditioning Eqn 5 on k.

$$\mu = \frac{1}{M} \sum_{k=1}^{N} q(k) \left[1 + \sum_{i=1}^{M-1} \left[\sum_{\sum_{j=1}^{i} x_j \leq (N-k)} \prod_{j=1}^{i} f_{H,j}(x_j) \right] \right]$$

Following the approach for Eqn 3, we get,

$$\begin{aligned} \mu \quad = \quad & \frac{1}{M} \sum_{k=1}^{N} q(k) \left[1 + \sum_{i=1}^{M-1} \oint \left[\frac{z^{(N-k+1)}-1}{z-1} \right] \right. \\ & \left. \left[\frac{1}{z^{(N-k+1)}} \right] \prod_{j=1}^{i} \mathcal{F}_{H,j}(z) dz \right] \end{aligned} \tag{6}$$

We use the μ obtained above in the discrete time M/M/1 queue model that we use for each input port and use results from [9] to obtain the mean waiting time and the P_0.

$$W = \frac{1-\rho}{\mu-\rho} \qquad P_0 = \frac{(\mu-\rho)}{\mu} \tag{7}$$

μ is evaluated iteratively from Eqns 6 and 7. Iterations are continued till W and P_0 converge for a given arrival rate. We have not investigated the proof of convergence but for all the examples considered, encompassing diverse types of input traffic, convergence to the fifth place after decimal occurs within fourteen iterations.

Cyclic service ensures fairness and thus the effective service rate is the same at all the ports. Thus, the sustainable throughput for cyclic service without fanout splitting policy, $\lambda_{C:NFS}$, corresponds to the maximum input load that can be supported by any input port. Thus,

$$\lambda_{C:NFS} = Q'(1) \max_{\rho}\{\mu)\}$$

where $Q'(1)$ is the average number of copies requested by a packet and μ is given by Eqn 6.

C. Cyclic Service with Fanout Splitting

Fanout splitting of copy requests whose requirements can be partially served increases the throughput of the copy network. We have seen that with identical input traffic at each port, cyclic service will result in identical service time distributions at each port and the sustainable throughput will be the maximum achievable throughput at any port. To analyse this scheduling policy, we approximate the copy network as a single, discrete time GI/D/N queue. The N servers account for the fact that under this policy, if the sum of the copy requests at the heads of the input queues is greater than or equal to N, a total of N copies is generated. If this sum is less than N, all of the requests are served. Cyclic service ensures that the service time distribution at each port is identical and a copy from any port will experience the same amount of average delay. This insensitivity of the service time distribution on the port number along with the fact that if there are N or more requests, N will be served, and if there are less than N, all will be served, allows us to model the copy network as a single queue. The arrival process into the GI/D/N queue is the arrival of copy requests into the copy network in each slot. Let $A(k)$ denote the number of new copies request arrivals in slot k. The distribution of $A(k)$ is the M-fold convolution of the distribution of copy request arrivals at the M input ports. The probability generating function of $A(k)$, $\mathcal{A}(z)$, is given by,

$$\mathcal{A}(z) = \prod_{i=1}^{M} \mathcal{F}_i(z) = ((1-\rho) + \rho\mathcal{Q}(z))^M$$

We can now use results from the queueing analysis of discrete time GI/D/N queues [10], with the above arrival process description and determine the average waiting time, W, for a copy in the network.

$$W = \frac{1}{A'(1)} \left[\frac{1}{2} \sum_{j=1}^{N-1} \frac{(1+z_j)}{(1-z_j)} + A'(1) + \frac{A''(1)-(N-1)\,A'(1)}{2\,(N-A'(1))} \right]$$

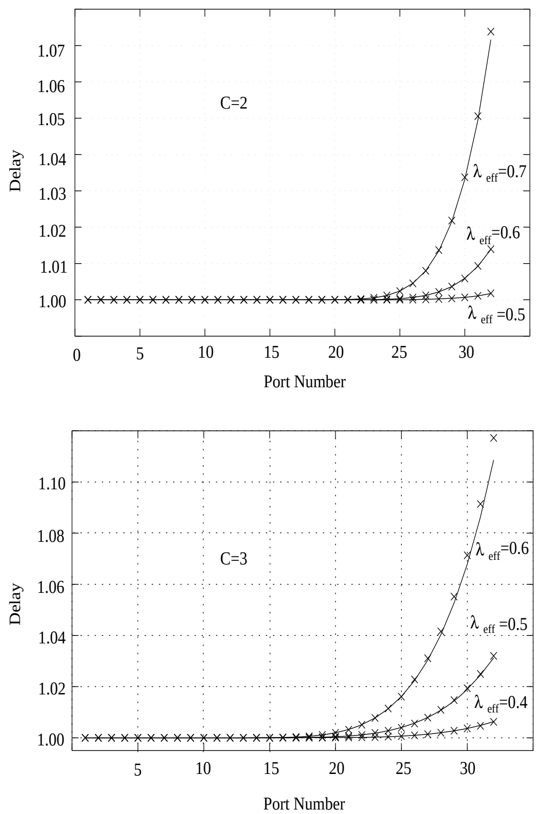


Fig. 3. Delay characteristics of a 32×32 copy network under acyclic service without fanout splitting scheduling policy with deterministic copy requests.

If all the ports are identically loaded, the sustainable throughput under this discipline is obviously 1 for copy request distributions.

It is easy to see that the copy network is not exactly a GI/D/N queue. When the head of the line copy requests sum to less than N, the copy network serves the requests of only the head of the line packets even though there may be packets in the inputs which could have been served. Thus work conservation is not preserved in the copy network whereas in a GI/D/N queue model these packets would have been served and work conservation would have been preserved. This approximation leads to a slight underestimation of the waiting time for the copies. Also, unlike the GI/D/N queue, the service discipline of the copy network is not FCFS.

III. Numerical Results

We now present the numerical results for the delay and sustainable throughputs from the queueing models of the previous section and compare them with those obtained from simulation. We present results for geometric and deterministic distributions for c_i. The effective load, λ_{eff}, at each input port will be $\lambda_{eff} = \rho\bar{Q}$ where $\bar{Q}$ is the average copy request number at each port.

With the service beginning with port 1 in each slot, the acyclic scheduling policy is unfair to the higher numbered ports. A lower service rate and consequently, a higher delay is thus expected at ports with higher addresses. In

Acyclic service without fanout splitting; $\bar{Q} = 2$

No. of Ports	Geometric		Constant	
	Anal.	Simul.	Anal.	Simul.
4	0.61	0.54	0.69	0.69
8	0.67	0.58	0.74	0.73
16	0.73	0.67	0.78	0.77
32	0.77	0.76	0.82	0.82

Cyclic service without fanout splitting; $\bar{Q} = 2$

No. of Ports	Geometric		Constant	
	Anal.	Simul.	Anal.	Simul.
4	0.72	0.66	0.75	0.79
8	0.80	0.72	0.86	0.85
16	0.86	0.81	0.90	0.89
32	0.90	0.87	0.93	0.92

TABLE I
Sustainable throughputs for policies without fanout splitting

Figure 3, this delay is shown against the port number for a 32×32 copy network. The results for the case of deterministic copy requests are more accurate than those of the truncated geometric distribution. This is because the copy requests of the head of the line packets in successive slots are indeed independent in the case of deterministic copy requests. In cases where the copy request is random, packets with larger fanout would be more likely to be blocked and stay at the head of the queue. The independence assumption fails to take this into account thereby resulting in inaccuracies in the results. Nevertheless, the isolated, discrete time M/M/1 queue model of each input port can be seen to be a very good approximation as the worst case difference between the simulation and analytical results in the delay is about 12%. Table 1 lists the sustainable throughput of the copy network under an acyclic service without fanout splitting scheduling policy. The simulation results for the sustainable throughputs were obtained by successively increasing the input load and marking the point where the service rate becomes less than the arrival rate.

The scheduling policies with cyclic service have the same service rate at all the input ports. The performance metrics at the input ports are thus identical. Hence, the metrics of a port reflect the performance characteristics of the copy network as a whole. Figure 4 shows the delay versus normalised load characteristics for a 32×32 copy network under acyclic service without fanout splitting scheduling policy. The results obtained from the proposed model match very closely with the simulation results for the case where all incoming packets have a deterministic copy request number. For the truncated geometric distribution, the worst case error in the delay is about 20%. The sustainable throughputs for the copy network are tabulated in Table 1 for the two distributions under consideration.

The delay characteristics for a 32×32 copy network under cyclic service with fanout splitting scheduling policy as

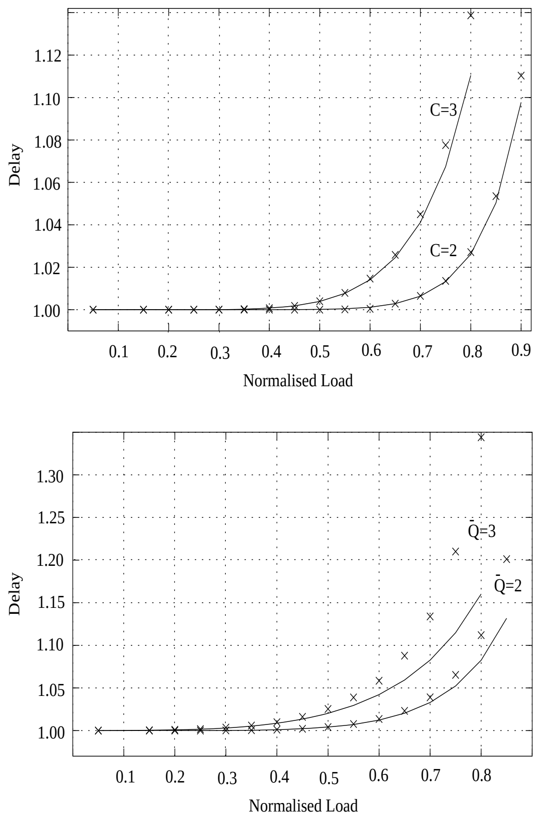


Fig. 4. Delay characteristics of a 32×32 copy network under cyclic service without fanout splitting scheduling policy with deterministic and geometric copy requests.

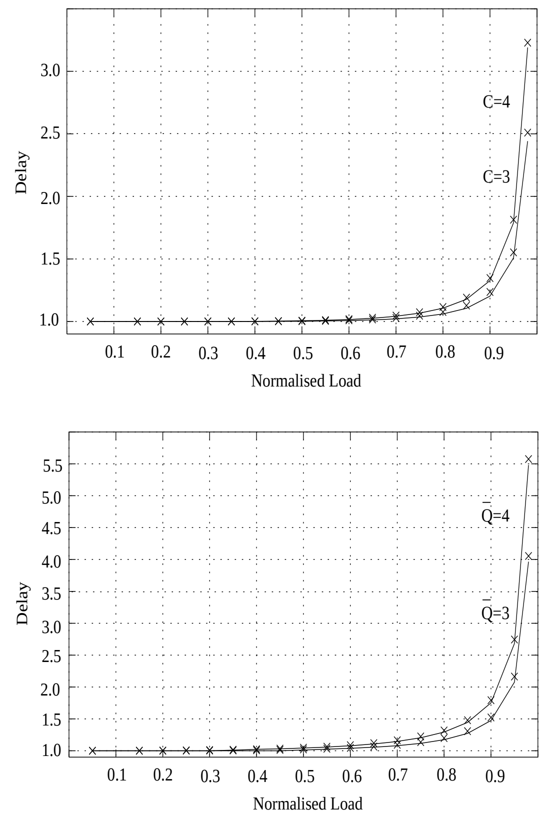


Fig. 5. Delay characteristics of a 32×32 copy network under cyclic service with fanout splitting scheduling policy with deterministic and geometric copy requests.

given by the approximate discrete time GI/D/N model is shown in Figure 5. As evident from the figures, the model is a very good approximation at all loads and the worst case error in the delay is less than 3%. The sustainable throughput of the copy network under this scheduling policy is 1. Hence, this scheduling policy has the best performance characteristics amongst the scheduling policies we have discussed.

IV. Discussions and Concluding Remarks

For the various copy scheduling policies, our interest was primarily in obtaining the delay in copy generation and the sustainable throughputs for each of the scheduling policies. In all the scheduling policies that we model, the analytical results for the deterministic copy request match very well with simulation results. This has been explained earlier.

The fanout splitting scheduling policies have a lower delay than the non fanout splitting policy for the same input load. This is because the fanout splitting policies have a better system usage. It can also be seen that there is an improvement in the sustainable throughput for cyclic service with no fanout splitting compared to acyclic with no fanout splitting policy. The increase is as much as 17% for a 32×32 copy network with copy requests having a truncated geometric distribution with a mean request size of 2 copies.

An attractive feature of cyclic service policies is their inherent fairness to all ports, although it adds to the implementation complexity. Cyclic service, combined with fanout splitting gives the lowest delay for the sechduling policies analysed. Also, the sustainable throughput in this copy scheduling policy is 1, the best amongst all the scheduling policies considered.

References

[1] Chin-Tau Lea, "A Multicast Broadband Packet Switch," *IEEE Trans on Commun,* vol. 41, no. 4, pp. 621-630, Apr 1993.

[2] T. T. Lee, "Nonblocking Copy Networks for Multicast Packet Switching," *IEEE Jour on Sel Areas of Commun,* vol. 6, no. 9, pp. 1455-1467, Dec 1988.

[3] R. Ahuja, B. Prabhakar and N. McKeown, "Multicast Scheduling Algorithms for Input-Queued Switches," *IEEE Jour on Sel Areas of Commun,* vol. 15, no. 5, pp. 855-866, Jun 1997.

[4] J. F. Hayes, R. Breault and M. K. Mehmet-Ali, " Performance Analysis of a Multicast Switch," *IEEE Trans on Commun,* vol. 39, no. 4, pp. 581-587, Apr 1991.

[5] J. F. Hui and T. Renner, "Queueing Analysis for Multicast Packet Switching," *IEEE Trans on Commun,* vol. 42, pp. 723-731, 1994.

[6] S. C. Liew, "A General Packet Replication Scheme for Multicasting With Application to Shuffle-Exchange Networks," *IEEE Trans on Commun,* vol. 44, no. 8, pp. 1021-1033, Aug 1996.

[7] Biplab Sikdar and D. Manjunath, "Evaluation of the Exact Overflow Probabilities in a Space Based Multicast Switch Copy Network," *Proc of the Conf on Inform Sciences and Sys (CISS),* Baltimore MD, Mar 1999.

[8] D. Manjunath, "Queueing Networks with Multiple Constraints," unpublished manuscript, 1992.

[9] M. E. Woodward, "Communication and Computer Networks : Modeling with Discrete-Time Queues," Pentech Press, 1993.

[10] H. Bruneel and B. G. Kim, "Discrete-Time Models for Communication Systems Including ATM," Kluwer Publishers, 1993.