New Structures of the Optical Node in Multihop Transparent Optical Networks with Deflection Routing

A. Bononi and P. R. Prucnal
Department of Electrical Engineering
Princeton University
Princeton, NJ, 08544

Abstract

New single-receiver/single-transmitter/single-buffer node structures for two-connected multihop transparent optical packet-switching networks with deflection routing are introduced. A ShuffleNet topology in uniform traffic is adopted to compare various shared optical memory schemes and their control algorithms. These simple structures minimize the number of crossbar switches needed at the node and have moderate control complexity, while still yielding high throughput and low delay. Analytical results are obtained by an extension of the existing theory and verified by simulation.

1 Introduction

Packet switching Transparent Optical Networks (TONs) have recently become the focus of much research towards faster multiuser data communications.

The basic idea behind TONs is to modulate data packets onto a lightwave carrier and let these optical packets propagate through intermediate nodes in the network to the end destination, without conversion to electronic form. Leaving packets in optical form throughout their path allows using much higher data rates than conventional networks and – when the technology matures – possibly a cheaper implementation of the network nodes.

In Wavelength Division Multiplexing (WDM) TONs, optical channels between nodes are provided by dedicated wavelengths within a single fiber, and there may be time-sharing of these channels. This single-medium configuration allows a quite flexible topological rearrangement of the nodes without having to deploy new fiber, and minimizes the required amount of fiber.

In Space Switching (SS) TONs, optical channels are provided by dedicated optical fibers between nodes, and there is time-sharing of a single wavelength within each of these links. Much more fiber has to be deployed, and the topology is not rearrangeable, but 1) the network is much more reliable, an essential requirement for metropolitan and wide area networks, and 2) much higher data rates per channel can be transmitted, having each user the whole fiber bandwidth (actually the optical amplifiers' bandwidth) at its disposal.

As the in/out node degree of multihop SS topologies is increased, the amount of deployed fiber and the node complexity increase, but network reliability and throughput increase as well. For optical implementation, an in/out degree of 2 seems to be a reasonable compromise.

Early work on the structure of all-optical nodes in SS multihop networks concentrated on the optical processing of the packet header and the related control of the crossbar routing switch [1]. Those schemes readily apply to ring networks, where switching is performed by an add/drop switch for reception and transmission of local traffic. Extensions to two-connected mesh networks, in which each node has two optical inputs and two optical outputs, were presented in [2] and [3]. The nodes consisted of an add/drop switch at each optical input and a main routing block at the core of the node to perform the output switching function. Nodes were capable of simultaneously absorbing packets form both links, which was obtained by either providing two receivers per node or a sufficient number of local-reception optical buffers.

2×2 crossbar space switches in $LiNbO_3$ technology are key elements in the implementation of such optical nodes. Minimization of the number of switches at the node is mandatory to maintain low flow-through optical power loss and to limit the cost. To this aim, compact integrated structures are desirable. Amplification can be provided at the output to compensate for this loss. However, optical amplifiers introduce

0743-166X/94 $3.00

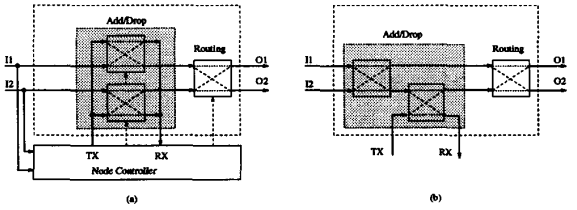

Figure 1: *3Shp node structure with (a) parallel and (b) serial configuration of the add/drop block.*

noise proportionally to their gain. This noise accumulates from node to node in these non-regenerative networks. At extremely high bit rates this might impose severe limits on the physical size of the network [4].

Fast-access optical buffers can be implemented with fiber delay loops. Store-and-forward (S&F) is not feasible in very high bit rate TONs, due to the limited number of optical buffers that can be added at each node to keep a low power loss and low control complexity. Optical amplification in the memory loop – a noisy and costly process – can be avoided with deflection routing [5], [6] if buffered packets are allowed to recirculate in each memory loop only once.

This paper will present new structures of the optical nodes in two-connected multihop Space Switching Transparent Optical Networks in which only one receiver (RX) and one transmitter (TX) are provided at each node. These solutions are meaningful for very high data rates, exceeding 1 Gb/s. Extremely simple nodes, with a few crossbar switches, and using non-priority deflection routing with only one fiber-loop optical memory will be presented and their performance in uniform traffic analyzed. The well-known ShuffleNet topology (SN) [7] will be used to compare the new structures and their control algorithms in terms of throughput and average number of hops. Section 2 introduces the new node schemes and section 3 specifies their control algorithms. Section 4 details the steady state analysis in uniform traffic. Section 5 presents the numerical results.

2 New schemes for the optical node

Fig. 1a shows the node scheme used in [3] adapted for a single TX/RX and no buffers. Solid lines indicate optical fibers and dashed lines electronic controls. Each optical input has an add/drop switch for local traffic, and a routing block performs the output switching. Only one add/drop switch at a time is used for TX/RX operations. If no extra toggles are

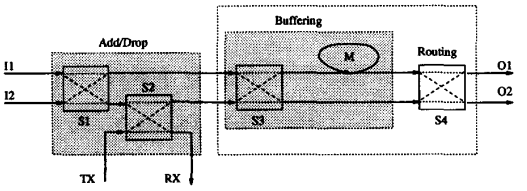

Figure 2: *4SoutM node structure.*

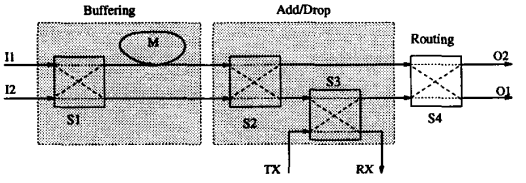

Figure 3: *4SinM node structure.*

used, there is a 3dB power loss for both the TX and RX signal [1]. The advantage of this parallel configuration is that only one switch in the add/drop block is crossed by each flow-through channel. An alternative cascaded configuration is shown in Fig. 1b. Here there is no splitting of the TX/RX signals. However two switches are crossed by one input channel, producing an unbalance of the optical power level at the routing switch. Both figures show that the node consists of a 3x3 optical switch. The minimum number of 2x2 switches to form a non-blocking 3x3 switch is three. This structure will be referred to as *3Shp*, as it implements deflection routing without buffers (hot-potato [5]).

The routing block may instead contain optical buffers. A simple output shared optical memory for this block, making use of fiber delay loops, has been introduced in [8] and a novel control scheme has been proposed and analyzed in [3].

Fig. 2 shows the structure of the node with the above mentioned memory with a one-packet fiber delay used as a buffer. The add/drop block can be either of the two shown in Fig. 1. This output shared memory node with 4 switches will be referred to as *4SoutM*.

If S3 is removed (which logically corresponds to setting it permanently in bar position) a 3-switch node is obtained where switch S1 is shared between TX/RX and buffering operations. It will be referred to as

[1] Also, when receiving a packet and simultaneously transmitting one using the same add/drop switch, part of the TX power loops back to the RX together with the incoming packet. This known interference could in principle be cancelled out, but in practice two extra off-line switches will be added for the TX and RX to toggle between add/drop switches.

3SoutM. Here the add/drop block must be in the cascaded configuration.

Unfortunately in the previous structures the receiver cannot access the optical buffer. If two packets for the node are received in the same slot, one will be missed. Fig. 3 shows that, if the buffering block precedes the add/drop block, the buffer can be ***shared*** between routing and TX/RX operations. Simultaneous reception from both input links is now possible by storing one of the packets. The add/drop block can be either of the two shown in Fig. 1. This structure will be referred to as *4SinM*.

3 Buffer control under non-priority deflection routing

Having defined the structures, the next step is to specify how to control the switches according to the destination of the input packets and of the packet awaiting transmission, given the destination of the possible packet in the buffer. The controller is an electronic processor capable of performing all routing decisions and switch settings within the duration of a packet. If this is too demanding, computations can be broken into sequential steps and pipelined, provided that the processing time of the slowest step be shorter than the packet duration [1].

A slotted network operation will be assumed in the following, as the complexity of the controller is lower than in asynchronous arrivals.

The most general structure of a node with 2 inputs, 2 outputs, single TX/RX and n output shared buffers is shown in Fig. 4. All the proposed 1-buffer structures fit the model with $n = 1$, where only a subset of the possible switch permutations are allowed. Packets arriving at inputs I_1 and I_2 and from the buffers

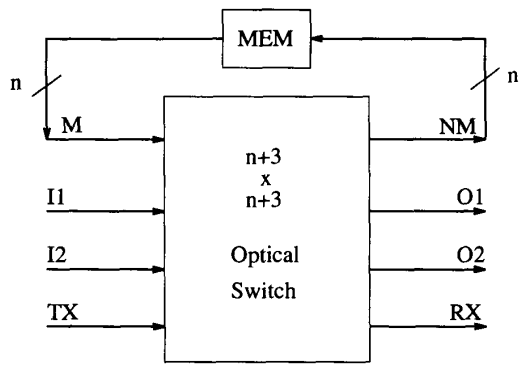

Figure 4: *General node structure.*

$M = M_1, .., M_n$ at each slot can be empty (E), for the node (FM), caring to exit at output 1 (C_1) or at output 2 (C_2) or can be ***don't care*** (DC) when both outputs provide equivalent shortest-paths to their destination. The same holds for packets ready at the TX, except that FM is not a valid option. Deflections occur when two or more packets contend for the same output and there is not enough memory to store the losers. Don't care packets are totally equivalent to empty packets as far as routing is concerned. Their presence helps avoid deflections. However, DCs and Es are not equivalent for transmission, as a new packet can be injected in an empty slot, but not over a flow-through don't care.

The objective of the controller is to maximize the node's throughput T and minimize D, the average number of node-to-node hops for packets to reach their destination. Little's law [9] gives

$$T = \frac{2u}{D} \tag{1}$$

where u is the link utilization. To maximize T, one must minimize the number of deflections and the number of FM misses, which minimizes D.

If all switch permutations are allowed – which implies the buffers are random-access – an efficient controller should do the following to decrease D:

1) Sort all slots at the input of the switch, except TX, into SORT classes: Es, DCs, C_2s, C_1s, FMs. Set TX-possible flag ON if there are Es or FMs.

2) Absorb FM packets first. This will also make room for transmission. A miss will occur only if all sorted slots are FM.

3) If TX-possible, join TX packet to SORT classes.

4) Next assign packets to O_1 and O_2. Priorities are: (C_1s,C_2s),DCs,Es,FMs.

   4.1) Cares are routed first. Unavoidable deflections occur when the two inputs and all buffers care for the same output. It is thus essential to let care packets out as soon as possible. Hence a general rule is to preferably store Es,DCs,FMs, which are equivalent to empty slots for routing purposes.

   4.2) DCs are routed after Cares. This may force DC packets to stay in the node's buffers for many slots. This might not add appreciably to their total delay if the link-propagation delay is much longer than the slot length, as in very high speed networks.

4.3) Route out FMs as a last resort. It is better to deflect a care flow-through packet than to miss a FM packet. This is true even when a miss might add more hops than a specific deflection, since a care packet in memory is likely to cause a deflection at the next slot, while a FM is not since it will be (most probably) absorbed. Note also that in some topologies, like SN, a packet that cannot be absorbed and has to be routed out will have a preferred output, i.e. it becomes a care packet. In some other topologies, like the Manhattan Street Network (MS) [6], a missed packet becomes don't care instead, but this little advantage has a negligible impact on performance in structures with buffers available to the RX, since a miss is a low probability event.

5) All remaining slots are stored in random order.

For the single buffer case, a rearrangeably non-blocking 4x4 optical switch can be built with a minimum of five 2x2 switches [10]. Therefore this complete shared memory structure with the above control will be referred to as $5ShM$.

Next, the controllers of structures $4SoutM$, $3SoutM$ and $4SinM$ will be described. These are the controls that give the best throughput figures we have found so far. Only the setting of the input switch will be detailed, since the TX/RX switch is just for absorption/injection, and the output routing switch implements a simple non-priority hot-potato routing of its input packets.

1) $4SoutM$

RX/TX operations are not coordinated here with buffering operations. Buffering proceeds as in [3]. Care packets at the input of the routing block are stored if in conflict with the memory packet. Else DCs or Es are stored.

2) $3SoutM$

The setting of the memory access switch and memory transitions can be described as follows. Logic symbols $\&, |, !$ mean *and, or, not* respectively and $I_?$ means *one of the two inputs.*

if $(I_1 = I_2) \Rightarrow$ Randomize
elseif $(I_? = FM) \Rightarrow$ receive it
elseif $((I_1, I_2) = (C_1, C_2) \mid (C_2, C_1)) \& (M = E \mid DC) \Rightarrow$ Randomize
elseif $(I_? = E) \& (TX = full) \Rightarrow$ get E on TX branch
elseif $(M, I_?) = (C_2, C_2) \Rightarrow$ store that input
elseif $(M = C_1 \mid FM) \& (I_? = C_1) \Rightarrow$ store that input
else store Es or DCs

Randomization of S1 in line 1 ensures equal treatment of both channels. Absorption of FM packets is the first action, as seen in line 2. Line 3 accounts for the fact that two care non-conflicting packets cannot be routed out directly, as the buffer cannot be bypassed. With no information about next TX packet, either care can be stored. If next TX packet is known, the right care could be stored to avoid a conflict at the next slot. Line 4 routes empty slots to the TX for possible injection. This is called TX-priority. The disadvantage is that this way care packets might get stored, thus increasing the deflection probability at the next slot. However at high loads most empty slots for the TX are provided by absorptions of FMs. Thus TX-priority has effect at low loads and yields higher throughput than a non-TX-priority rule where Es are preferably stored to avoid deflections at the next slot. As shown in lines 5 and 6, conflicts with the memory are resolved by storing the conflicting input.

3) $4SinM$

The main advantage of shifting the buffering block ahead of the add/drop block is that the miss probability gets drastically reduced at almost no expense of deflections, since stored FMs are equivalent to empty slots in most input/TX configurations. Here is a description of the settings of S1 and of memory updates.

if $(I_1 = I_2) \Rightarrow$ Randomize
elseif $((I1, I2) = (FM, C_?) \mid (C_?, FM)) \& (M! = C_?) \Rightarrow$ store FM
elseif $(I_? = FM) \& (M! = FM) \Rightarrow$ receive it
elseif $(I_? = FM) \& (M = FM) \Rightarrow$ store that input
elseif $((I_1, I_2) = (C_1, C_2) \mid (C_2, C_1)) \& (M = E \mid DC \mid FM) \Rightarrow$ store to avoid TX conflicts
elseif $(M, I_?) = (C_2, C_2) \Rightarrow$ store that input
elseif $(M = C_1 \mid FM) \& (I_? = C_1) \Rightarrow$ store that input
else store Es or DCs

The presence of the TX-access switch S2 allows avoiding conflicts with the TX at the present slot, as seen in line 5. The other task of switch S2 is to get Es or FMs to the TX.

The next section will present an analytical procedure to evaluate the performance of these structures.

4 Steady state analysis in uniform traffic

4.1 Definitions and specifications

The steady state behavior in uniform traffic of a two-connected regular mesh network will now be analyzed.

Regular means each node is topologically equivalent to all other nodes. Since in uniform traffic all nodes have identical statistical behavior, it is enough to focus on a single node to get the global network behavior. SN is an example of regular network.

A common clock is distributed to all nodes, so that node operations are performed in fixed length time slots.

New arrivals at each node are collected in an electronic FIFO TX queue, waiting to be injected in the network. Size and average occupancy of this queue will not be treated here, since the main focus is on the optical transport part of the network. Arrivals are assumed to occur at the same rate and independently at each node. It is assumed that at each node the destination of new packets is chosen independently of other nodes and independently of previously admitted packets, and is drawn from a distribution that is uniform on all other nodes. This is the uniform traffic pattern. The assumed regularity of the network and the randomness associated with deflection routing help keep this homogeneous traffic pattern.

As already seen, the node throughput T, i.e. the average number of packets inserted/absorbed per slot by the node at equilibrium, and the number of hops D taken on the average by a packet to reach its destination are related by Little's law (1) to the link utilization u, which is the probability that an input link is occupied by a packet at each clock. Network regularity and uniform traffic pattern ensure that u is the same for both inputs.

The total delay of a packet, once injected in the network, is the sum of the propagation delay, proportional to D, and of the queueing delay D_q at the optical buffers of visited nodes. For deflection routing, D_q is of the order of the number of buffers provided in the routing block. For very high bit rate optical networks, in which each link can contain hundreds of packets in flight at any given time, D_q is small compared to the propagation delay and can thus be neglected. Hence the average network delay is proportional to D.

Let g be the probability that the node's TX buffer has at least one queued packet per slot. It will be referred to as the offered traffic, i.e. the traffic offered to the transport part of the network.

Let r be the probability that an input link contains a packet for the node, given that the link is full. It will be called reach probability.

Let P_{dc} be the probability that an input link contains a flow-through don't care packet, given that the link is full. It will be called don't care probability.

The fundamental assumption of the model is that arrivals at the two input ports are independent white processes [11]. This approximation makes sense in large mesh networks in uniform traffic and with a random routing rule like deflection routing.

At each clock, each node's input link can be E with probability $1-u$, DC with probability uP_{dc}, FM with probability ur and care (C_1 or C_2) otherwise. It is assumed that C_1s and C_2s are equally likely [2].

Also, each packet presented by the TX to the network can be E with probability $1-g$, DC with probability gP_{dc0} and care otherwise, being C_1s and C_2s equally likely in generation by the uniform traffic assumption. P_{dc0} is the fraction of network nodes that can be reached from either output link of the transmitting node in the same minimum number of hops. It just depends on the selected regular topology.

Finally, at steady state, the buffers can be E,DC,C_2,C_1,FM with probabilities depending on u, r, P_{dc} and on the controller algorithm.

4.2 Solution procedure

The offered traffic, g, is the free network parameter and the objective is to express all other quantities as a function of g only. In particular, throughput and delay curves $T(g)$ and $D(g)$ will be found.

The first step is to obtain the steady state memory probabilities. These are obtained as the equilibrium probabilities of a markov chain whose states are the possible memory configurations (for 1 buffer these states are E,DC,C_2,C_1,FM) and whose transitions depend on the controller algorithm and on the input probabilities g, P_{dc0}, u, r, P_{dc} [3].

Then, by conditioning on all possible input triplets of independent random variables $\{I_1, I_2, TX\}$ and on all possible memory configurations $\{M_1, .., M_n\}$ and by averaging out, it is possible, for a specific node structure and control, to find the quantities:

1) a = probability that an input FM packet is absorbed.

2) d_m = probability that an input FM packet is missed and deflected.

3) d = probability that an input care packet is deflected.

4) d_0 = probability that an injected TX packet is deflected.

5) T_{in} = probability per slot of absorbing a packet

[2] Our simulations show that while u is indeed uniform for all links in the network, the frequency of occurence of DCs,C1s,C2s,FMs on the two input links is slightly different. Our model could easily be extended to cope with these different input distributions, at the cost of doubling the dimension of the resolving markov chain described in the appendix. However this gives a negligible improvement on the analytic results.

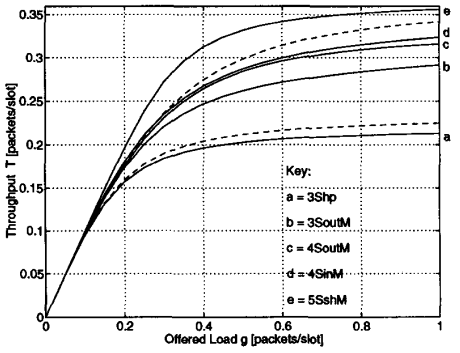

Figure 5: *Throughput in a 64-node Shufflenet vs. offered load.*

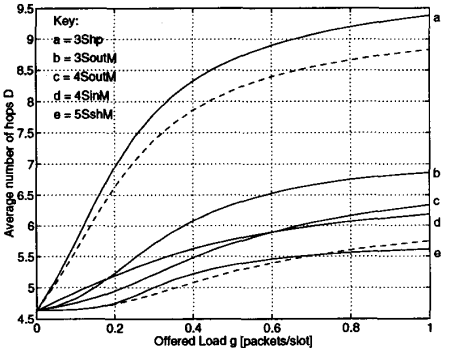

Figure 6: *Delay in a 64-node Shufflenet vs. offered load.*

6) T_{out} = probability per slot of injecting a packet

All of these quantities depend only on g, P_{dc0}, u, r, P_{dc}. At equilibrium, quantities 5) and 6) must be equal. This provides an expression for the link utilization $u = u(g, r, P_{dc}, P_{dc0})$.

These computations are extremely tedious and can be automatized using a symbolic software tool like Mathematica.

It remains now to find r, P_{dc}. As a byproduct, D will also be found. The procedure appeals to the uniform traffic assumption, in which every packet is a "typical" packet. It is thus a matter of following the trajectory of a *typical* or *test* packet hopping towards its final destination in a "uniform gas" of competing packets. The random walk of the test packet can be visualized as an absorbing markov chain whose states coincide with the network nodes [12], [3].

For some topologies like SN it is possible to speed up the computation by drastically reducing the number of states in the chain. This is done by combining together in a single state all nodes with same distance to destination. The test packet thus performs a random walk on the integers $0, 1, ..d_{max}$ where d_{max} is the maximum distance to destination [13].

The procedure yields r, P_{dc}, D as functions of a, d_m, d, d_0, and its improved algorithm is detailed in the appendix.

A 4x4 system of nonlinear equations in the unknowns a, d_m, d, d_0 is thus available, whose solution can be found numerically for every value of g. By back-substitution, the curve $T(g)$ can be found from 5) and 6) and verified by Little's law (1).

5 Results

All the curves presented next have been found for a 64-node ShuffleNet by the previous analytical procedure and verified by simulation. There's a perfect agreement between analysis and simulation on the scale shown. Fig. 5 summarizes throughput results for the proposed structures.

First consider the curves for $3Shp$ and $4SoutM$. Dashed lines refer to the same structures but with two receivers [3]. The gap between solid and dashed lines accounts for the effect of missing FM packets. The throughput degrades more in $4SoutM$, i.e. when buffers are added.

Structure $4SinM$ reduces the miss probability without significantly degrading the deflection probability, so that it has higher throughput than $4SoutM$. This proves the positive effect of shifting buffering at the input for single TX/RX nodes.

Structures with fewer switches present lower throughput since the control is less flexible. However note how well structure $3SoutM$ compares with the 4-switch nodes.

The curve for the non-blocking switch $5SshM$ with a single buffer provides the highest throughput, even higher than $4SoutM$ with 2 RXs, since 1) it better handles two non-conflicting input care packets and 2) it reduces blocking in the TX FIFO queue by storing, when possible, TX packets in the shared optical buffer. However, in a practical optical implementation, buffered packets might need to cross the 4x4 switch many times, each time crossing two to three 2x2 switches. The power loss on such buffered packets could turn out to be unacceptably high. The great advantage of the other structures is to have a number

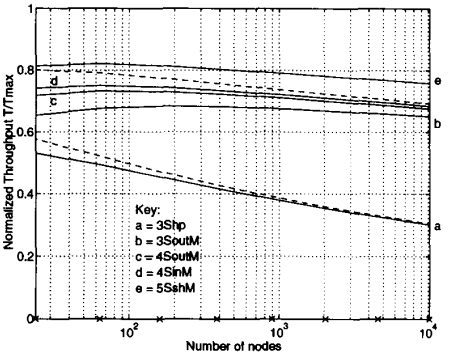

Figure 7: *Normalized Throughput in a 64-node Shuffle net vs network size.*

of 2x2 switch crossings per input channel no higher than 3 (using the parallel structure of the add/drop block when possible). Most importantly, buffered and unbuffered packets will experience the same loss.

Fig. 6 shows the corresponding curves for the average number of hops. Although the $4SoutM$ structure with two receivers (lower dashed line curve) has lower delay than $5SshM$ in the low/mid load g range, the value of u (and the throughput) is higher for $5SshM$ as it uses the optical buffer also to store TX packets that would otherwise be blocked in the electronic TX queue.

The full-load throughput figures should be compared to the highest achievable value $T_{max} = 2/D_{min}$, where D_{min} is the average number of hops when deflections and misses never occur. Fig. 7 shows how the normalized throughput T/T_{max} scales with network size in SN(2,k) topologies, for $k = 3, .., 10$. The relative throughput for $3Shp$ quickly degrades from over 50% to about 30%. Buffered structures show instead a stable behavior, with little degradation with increasing network size. Structure $3SoutM$ shows a maximum for a 160-node SN. The throughput differences among $3SoutM$,$4SoutM$ and $4SinM$ tend to level off for big networks, where their throughput settles around 70% for 10000 nodes. The efficient $5SshM$ degrades less than all other structures, with relative throughput around 80%.

6 Conclusions

New low-loss single-receiver single-buffer optical node structures for deflection routing TONs have been proposed and analyzed in uniform traffic. Results have been presented for a SN topology, although they qualitatively still hold for other regular topologies.

These new structures point to the feasibility of extremely simple, low-loss optical nodes that allow very fast electronic routing control.

The effect on throughput of adding flexibility to the input switching process has been analyzed by comparing nodes with 3, 4 and 5 switches. Important differences in power loss per input channel among the various structures have been pointed out.

It has been shown that throughput results scale well with network size for buffered structures.

Future work will establish how the structures behave in non-uniform traffic with the aid of distance-priority rules to resolve contentions.

Appendix

This procedure can be applied to any regular topology, whether or not a reduced state-space can be used. However, for illustration purposes, a SN topology will be used.

A specific example of the absorbing markov chain describing the random walk of the test packet towards its destination is given in Fig. 8 for a 64-node SN(2,4) topology.

A SN(q,k) topology has $N = k\,q^k$ nodes arranged in k columns of q^k nodes each, and there is a perfect shuffle connection among nodes in adjacent columns [7]. The maximum distance between nodes is $d_{max} = 2k - 1$. Fix a destination node. All nodes reachable in less than $k+1$ hops proceeding backwards are Care with respect to that destination. All the remaining nodes, at distance $k+1, .., 2k-1$ are don't care. A deflection of the test packet flowing towards that destination at a node at distance i brings the packet back to the set of nodes at distance $i+k-1$. A deflection at the destination node brings the packet back at distance $2k-1$, while a miss brings it back at distance $k-1$. Finally, there are q^i nodes at distance $1 \le i \le k-1$, and $q^k - q^{i-k}$ nodes at distance $k \le i \le 2k-1$.

Fig. 8 refers to the initial step of the walk, where the packet is at its injection node. The labels are the transition probabilities, defined in section 4.2. For

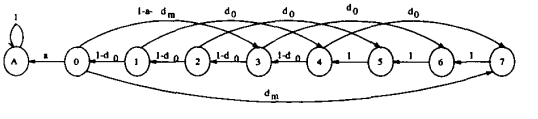

Figure 8: *Markov chain describing the random walk of the test packet in a SN(2,4) topology.*

every step after the first hop, in which the packet is at the TX port of the node, label d_0 changes in d. The nodes represent the distance in hops of the test packet to its destination. A fictitious absorbing state A has been added to take into account the possibility of missing the test packet at its destination.

For this model to hold, it is necessary that the controller's treatment of both input links be the same, so that it is not required to know which link the packet comes from. Also, this reduced-state chain is not directly applicable to topologies like MS in which a set of nodes at the same distance can be partly care and partly don't care.

The transition probabilities can be organized in a transition matrix Π for all steps $t = 1, 2, ...$ Analogously, a matrix Π_0 can be written for the injection step $t = 0$.

Since A is the only absorbing state, if states are ordered as $A, 0, 1, .., 7$ then matrix Π is in its canonical form. Taking off the first row and the first column, matrix Q^T is obtained. From this, the fundamental matrix of the absorbing chain $\mathcal{N} = (I - Q)^{-1}$ is obtained, where I is the 8x8 identity matrix. The entries of $\mathcal{N} = \{n(i,j)\}$ give the mean number of times in each nonabsorbing state j for each possible nonabsorbing state i after the first hop [14]. It is straightforward to derive a closed-form expression of $\mathcal{N}$ for SN.

Consider the set of nonabsorbing states $0, 1, .., 7$. Define the column vectors $fm = [10000000]'$, $dcs = [00000111]'$ (ones in the positions corresponding to don't care states), and $all = [11111111]'$. Let $\mathbf{p}(0)$ be the probability state vector at the injection step. By the uniform traffic assumption and the results on the number of nodes at each distance given before, $\mathbf{p}(0) = [0\,2\,4\,8\,15\,14\,12\,8]/63$. The state after the first hop is $\mathbf{p}(1) = \mathbf{p}(0) * \Pi_0$. Let $\mathbf{p}$ indicate $\mathbf{p}(1)$ with the first component removed.

It is easy to see that :

1) the Expected Number of visits of state 0 before absorption is $EN_{fm} = \mathbf{p} * (\mathcal{N} * fm)$;

2) the Expected Number of visits to don't care nodes at which the test packet is flow-through is $EN_{dc} = \mathbf{p} * (\mathcal{N} * dc)$;

3) the Expected Number of visits to any node before absorption, i.e. the average number of hops before reception, is $D = \mathbf{p} * (\mathcal{N} * all)$.

From these, an estimate of the don't care and reception probability are formed as

$$P_{dc} = \frac{EN_{dc}}{D} \qquad r = \frac{EN_{fm}}{D}. \qquad \text{(A.1)}$$

The first equation estimates P_{dc} as the fraction of time the test packet is don't care flow-through. It is easy to find that $EN_{fm} = 1/a$. Since $r\,a$ is the unconditional probability of absorption, Little's law gives $r\,a = 1/D$, which justifies the second equation in (A.1).

This procedure, making use of the fundamental matrix of the absorbing chain, can be substantially faster than previously reported iterative methods [12], [3] when efficient matrix-inversion algorithms are available.

References

[1] P. R. Prucnal, P. A. Perrier, "Optically-processed routing for fast packet switching," *IEEE LCS Magazine*, vol. 1, pp. 54–67, May 1990.

[2] D. J. Blumenthal, K. Y. Chen, J. Ma, R. J. Feuerstein, and J. R. Sauer, "Demonstration of a deflection routing 2x2 photonic switch for computer interconnects," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 169–173, Feb. 1992.

[3] A. Bononi, F. Forghieri, and P. R. Prucnal, "Analysis of one-buffer deflection routing in ultra-fast optical mesh networks," in *Proc. INFOCOM '93*, San Francisco, CA, March 1993, vol. 1, pp. 303–311.

[4] A. Bononi, F. Forghieri, and P. R. Prucnal, "Design and channel constraint analysis of ultra-fast multihop all-optical networks with deflection routing employing solitons," *IEEE J. Lightwave Technol.*, Dec. 1993.

[5] P. Baran, "On distributed communications networks," *IEEE Trans. Commun. Syst.* vol. 12, pp. 1–9, Mar. 1964.

[6] N. F. Maxemchuk, "Regular mesh topologies in local and metropolitan area networks," *AT&T Tech. J.*, vol. 64, no. 7, pp. 1659–1685, Sep. 1985.

[7] A. S. Acampora, M. J. Karol, and M. G. Hluchyj, "Terabit lightwave networks: the multihop approach," *AT&T Tech. J.*, vol. 66, pp. 21–34, Nov./Dec. 87.

[8] I. Chlamtac and A. Fumagalli, "An all-optical switch architecture for manhattan networks," *IEEE J. Select. Areas in Commun.*, vol. 11, pp. 550–559, May 1993.

[9] D. Bertsekas and R. Gallager, *Data Networks*. Prentice Hall, 1987.

[10] K. Padmanabhan and A. N. Netravali, "Dilated networks for photonic switching," *IEEE Trans. Commun.*, vol. COM-35, pp. 1357–1365, Dec. 1987.

[11] A. G. Greenberg and J. B. Goodman, "Sharp approximate models of deflection routing in mesh networks," *IEEE Trans. Commun.*, vol. 41, pp. 210–223, Jan. 1993.

[12] A. S. Acampora and A. Shah, "Multihop lightwave networks: a comparison of store-and-forward and hot-potato routing," *IEEE Trans. Commun.*, vol. COM-40, pp. 1082–1090, June 1992.

[13] A. Krishna and B. Hajek, "Performance of Shuffle-like switching networks with deflection," in *Proc. INFOCOM '90*, San Francisco, CA, June 1990, vol. 2, pp. 473–480.

[14] J. G. Kemeny, H. Mirkil, J. L. Snell, G. L. Thompson, *Finite Mathematical Structures*. Prentice Hall, 1959, pp. 404–409.