Fault Tolerance in Optical Packet-Switched Shuffle Networks with Deflection Routing

Ben Y. Yu, Ivan Glesk, and Paul R. Prucnal

Department of Electrical Engineering
Princeton University, Princeton, NJ 08544, USA
Tel: (609) 258-2041; Fax: (609) 258-2876

ABSTRACT

In this paper, we propose a distributed fault-tolerance algorithm using store-and-forward and deflection routing schemes suitable for ultrafast packet switching in self-routed optical shuffle networks. An OTDM multihop prototype shuffle network was built at Princeton University to demonstrate optical packet switching at ultra-fast bit rates [1,2] based on a bit-level self-routing scheme [3] we recently developed. We present in this paper the results on the performance and survivability of the network using the fault-tolerance algorithm under single/multiple node/link failures.

Fault Tolerance in Optical Packet-Switched Shuffle Networks with Deflection Routing

Ben Y. Yu, Ivan Glesk and Paul R. Prucnal

Department of Electrical Engineering

Princeton University, Princeton, NJ 08544, USA

Tel: (609) 258-2041; Fax: (609) 258-2876

Recent research works on optical switching indicate that bit rates exceeding tens of gigabits per second can be achieved in transparent optical networks using optical time-division multiplexing and wavelength-division multiplexing techniques. We developed a testbed to demonstrate ultrafast optical packet switching in an 8-node shuffle network at 100 Gbps using a novel self-routing scheme. An one-packet shared medium optical buffer is provided at each node, and deflection routing is employed as the principle contention resolution in the network. High throughput and short processing delays have been achieved under normal steady-state network operations. This paper discusses the survivability of optical shuffle networks with deflection routing when one or multiple link/node failures take place. The structure of our network node is shown in Fig 1b. Each node has a main nonblocking routing switch which consists of $LiNbO_3$ crossbar switches connecting two input data links to two output data links and a transmitter and a receiver which are capable of injecting and absorbing packets from both links. The one-packet buffer is implemented by inserting a recirculating fiber loop between a pair of input and output ports. Each packet carries a header and a payload. The routing tag in the header contains $(k+1)$ 3-bit groups, where k is the distance from the source to destination in hops. Only one group is needed for routing at each hop, and the last group is used for packet dropping at the destination. A clock bit is included in the header for selecting the appropriate groups for routing at every node, and is extracted by a thresholding device at the network nodes. The clock position in the header is controlled by the clock shifter. Deflecting routing is also implemented using this clock control. This design lead to simple node structures and very short processing delays. In each time slot, every node in the network also sends two-bit signals indicating the status of the node to each of its two upstream nodes in the opposite direction to the flow of data packets. These two bits are generated at twice the packet rate in the polarization orthogonal to the data streams, with "11" indicating normal operations, and "10" indicating that both output data links are broken and the node is only accepting packets destined for itself. If a node does not receive the status signals from one of its downstream nodes in two time slots, it assumes that the output data link to that particular node is down. Since node failures can be interpreted as a subset of link failures, the routing algorithm focuses on the adaptation of link failures. When one of its output links fails, a node has two options in routing transiting packets to reduce packet loss: store and forward by using electronic buffers and packet regeneration after the link is fixed; or use deflecting routing so that the packets can still arrive at their destinations by taking alternative routes. When a large number of link failures occur, store-and-forward may be the only realistic solution and the operation of the network may halt temporarily. However, if only a few links fail, the network can still survive if the deflecting routing method is used. The effect of a link failure on the performance of the network is dependent on the size of the network and the traffic load in the network. The simulation results in Fig. 2 compare the network throughput performance under normal operation and a single link failure in an 8-node shuffle network employing self-routing and deflecting routing algorithms.

References

1. P.R. Prucnal, D.J. Blumenthal, and M.A. Santoro, "A 12.5 Gbps fiber optical network using all-optical processing," *Elec. Lett.*, 23, pp. 629-630, 1987.
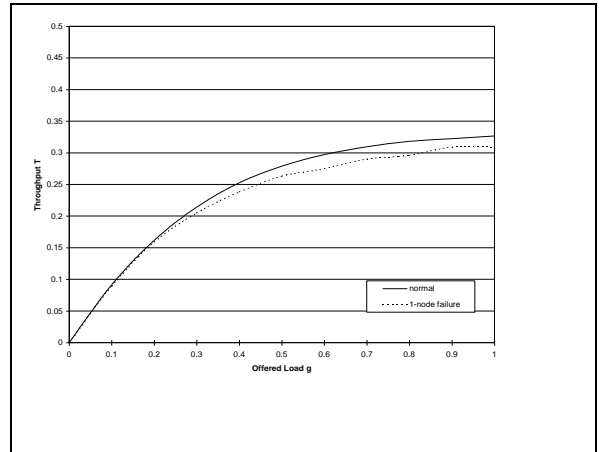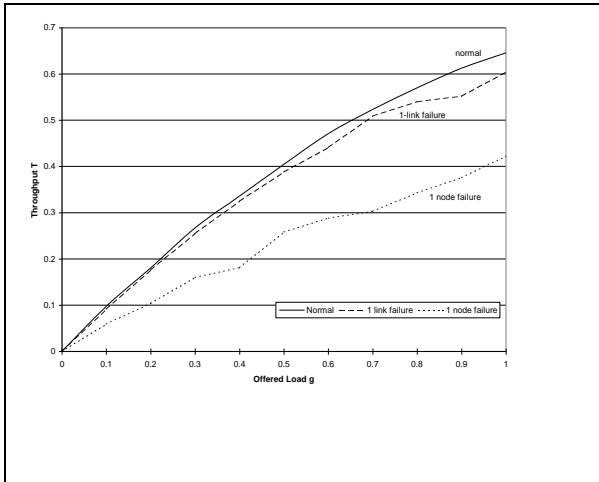
Figure 1: Network Node Structure

Figure 2: a) Throughput in 8-node shufflenet under normal operation, single link failure, and single node failure; b) 64-node shufflenet under normal operation and single node failure