

# PROTECTORATION: A Fast and Efficient Multiple-Failure Recovery Technique for Resilient Packet Ring Using Dark Fiber

Martin Maier, *Member, IEEE*, Martin Herzog, Michael Scheutzow, and Martin Reisslein, *Senior Member, IEEE*

**Abstract**—The two protection methods wrapping and steering used in IEEE 802.17 resilient packet ring (RPR) provide fast but very inefficient and limited network failure recovery. Due to the increased length of the backup path, RPR suffers from high traffic loss, a decreased throughput-delay performance, and the lack of resilience against multiple link and/or node failures. To achieve an improved resilience, interconnecting a subset of the ring nodes by means of a dark-fiber single-hop star wavelength division multiplexing (WDM) network is proposed. In doing so, the ring network is divided into separate domains, each being fully recoverable from a single link or node failure without losing full network connectivity. A novel hybrid fault recovery technique, termed protectoration, is proposed and examined by means of probabilistic analysis and simulation in terms of stability, channel utilization, and throughput-delay performance. The proposed protectoration technique 1) combines the fast recovery time of protection and the bandwidth efficiency of restoration, 2) provides full recovery from multiple link and node failures, 3) builds on both wrapping and steering protection methods of RPR and, thus, allows for an evolutionary upgrade of existing RPR networks, and 4) does not require the convergence of routing protocols in response to failures and, thus, improves the routing stability and network availability. Numerical investigations in this paper show that the location of failures has a strong impact on the network performance. For a given failure location, the protectoration technique is able to accommodate multiple ring failures without significant performance loss.

**Index Terms**—Availability, MAC, MAN, protection, reliability, resilience, restoration, RPR, survivability, WDM.

## I. INTRODUCTION

THE NEW IEEE 802.17 resilient packet ring (RPR) standard for optical metropolitan area networks (MANs) aims at improving also, besides throughput efficiency and service differentiation, the resilience of packet-switched bidirectional

dual-fiber ring networks [1]–[3]. To achieve a higher level of resilience, two methods of protection are currently considered in RPR, namely: 1) wrapping; and 2) steering [4]. With wrapping, upon detection of a link or node failure, the two ring nodes adjacent to the failed link or node switch all traffic arriving on the incoming fiber onto the outgoing fiber to reach the destination node going in the opposite direction. Thus, the two ring nodes adjacent to the failure wrap all traffic away from the failed link or node. With steering, after learning that a failure has occurred a given source node injects the traffic in the direction opposite to the link or node failure, i.e., the source node steers the traffic away from the failure. Note that in case of a link and/or node failure, the two protection techniques result in a rather inefficient use of bandwidth resources. Specifically, wrapped traffic travels from the source node to the corresponding wrapping node and then back to the source node, thus, consuming bandwidth without getting closer to the destination. After returning to the source node, the wrapped traffic continues traveling towards the destination node along a secondary path that, in general, is longer than the primary path in terms of hops. (Under normal network operation, the nodes typically deploy shortest-path routing in order to decrease the hop distance between source and destination nodes and thereby to increase the network throughput due to spatial wavelength reuse by nodes downstream of the destination node. As a result, the primary path is shorter in terms of hops than the counterdirectional secondary path.) Note that the bandwidth inefficiency due to wrapping occurs only during a (short) transient period before the failure notification has propagated to the source node. Steering avoids the wasting of bandwidth due to the round trip of traffic between source node and wrapping node, but it does not eliminate the increased bandwidth consumption incurred on the secondary path. For instance, it was shown in [5] that in the event of a failure, the loss of traffic in a 63-node RPR ring network may be as high as 94% due to the increased length of the backup path. It is noted that the traffic scenario in [5] is rather peculiar but realistic, as the authors indicate. It is primarily used to illustrate possible problematic situations of RPR network applications. Furthermore, note that both fault recovery techniques are able to protect traffic only against a single link or node failure. In case of multiple failures, the full connectivity of RPR is lost, i.e., the ring is divided into two or more disjoint subrings.

RPR as a metro ring network will be deployed often in interconnected ring architectures, as typically found in today's

Manuscript received November 20, 2004; revised July 16, 2005. This work was supported in part by the European Commission within the Network of Excellence e-Photon/ONE, the German research funding agency "Deutsche Forschungsgemeinschaft (DFG)" under the graduate program "Graduiertenkolleg 621 (MAGSI/Berlin)," and the DFG Research Center MATHEON "Mathematics for Key Technologies" in Berlin.

M. Maier is with the Institut National de la Recherche Scientifique (INRS), Montréal, QC H5A 1K6, Canada (e-mail: maier@ieee.org).

M. Herzog is with the Telecommunication Networks Group, Technical University Berlin, 10587 Berlin, Germany (e-mail: herzog@tkn.tu-berlin.de).

M. Scheutzow is with the Institute for Mathematics, Technical University Berlin, 10623 Berlin, Germany (e-mail: ms@math.tu-berlin.de).

M. Reisslein is with the Department of Electrical Engineering, Arizona State University, Goldwater Center, MC 5706, Tempe AZ 85287-5706 USA (e-mail: reisslein@asu.edu).

Digital Object Identifier 10.1109/JLT.2005.856165

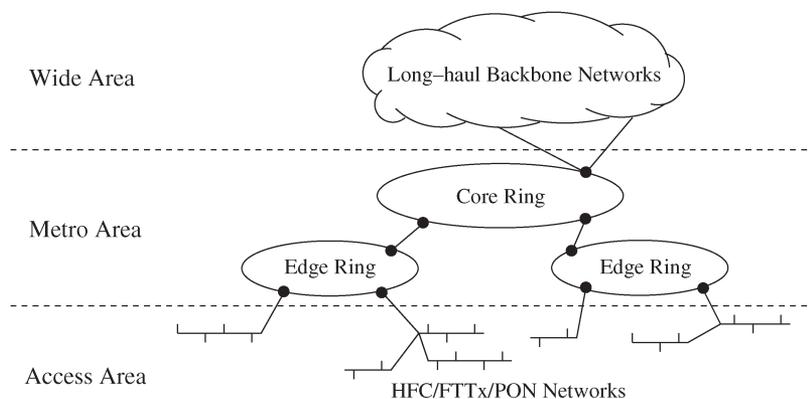


Fig. 1. Metro area networks: Metro core ring interconnects several metro edge rings.

metropolitan areas [6]. Metropolitan interconnected rings are composed of metro core and metro edge rings, where a metro core ring interconnects several metro edge rings, as depicted in Fig. 1. Metro core rings have to meet several requirements. Aside from configurability, reliability, flexibility, scalability, and large capacity, metro core rings have to be extremely survivable [7]. If the metro core ring network fails, all customers are potentially left without service. Thus, survivability in metro core ring networks is crucial. In particular storage networking protocols, being one of the important metro applications without built-in adequate survivability, rely almost entirely on the failure recovery techniques of the optical layer [8].

As mentioned above, RPR with its current protection schemes provides only inefficient and limited survivability. In this paper, a novel fault recovery technique, termed *protection*, is proposed and examined for RPR using dark fibers, which are abundantly available in metropolitan areas. By using already-existing dark fibers, no additional fiber infrastructure needs to be installed, thus, avoiding costly construction work and manpower. However, the available dark fibers need to be lit by deploying additional transceivers at each end. Protectoration aims at combining the benefits of protection and restoration. More precisely, protectoration exploits the fast recovery time of protection mechanisms and the bandwidth efficiency of restoration mechanisms. The proposed resilience technique enables RPR to recover from multiple link and node failures. Moreover, protectoration does not require any major modifications of the basic RPR protocols and mechanisms and makes use of RPR's built-in wrapping and steering protection techniques. Only a subset of nodes needs to be equipped with additional hardware and software, which is used to perform protectoration while all the remaining ring nodes may operate without any modifications. Thus, protectoration allows for an evolutionary upgrade path of existing RPR networks in that a subset of nodes can be upgraded by protectoration in a pay-as-you-grow fashion according to given traffic demands and/or cost constraints. Protectoration operates at the medium access control (MAC) layer. Consequently, it does not require the convergence of routing protocols at the Internet protocol (IP) network layer in response to failures and avoids the complex interworking of protection and restoration schemes of layer 2 and layer 3. (Furthermore, protectoration does not require time-consuming spanning tree

computations.) This results in an improved routing stability and network availability [4]. Note that the improved stability and network availability is not a specific attribute of the protectoration scheme, but is also valid in the original RPR standard. It is noted that transforming the RPR ring into a different network topology has an impact on other major properties of RPR, e.g., packet forwarding, lossless property, in-order packet delivery, and fairness control. This paper focuses on the impact of the proposed protectoration technique on the resilience and throughput performance of RPR. Section IV-C discusses the impact of the modified network topology and protectoration technique on other basic properties of RPR.

It is worthwhile to mention that prestandard RPR solutions exist, which implement advanced resilience solutions in addition to wrapping and steering. These solutions are able to protect traffic against multiple failures on the same fiber ring. For example, in the so-called single ring recovery (SRR) protocol, an extension to the spatial reuse protocol (SRP) of dynamic packet transport (DPT) rings, wrapping and steering take place in the case of a single link/node failure on either ring, similar to RPR. If on one of both rings multiple failures occur, wrapping is not deployed and all ring nodes use only the other failure free ring. In doing so, the failed ring may have multiple failures without losing full network connectivity. This, however, only holds if the other ring is failure free. If there are failures on both rings, full network connectivity is lost and the bidirectional ring is divided into disjoint subrings. Note that SRR affects the entire ring network in that all ring nodes must support SRR. If one or more ring nodes do not support SRR, SRR will have no effect. As opposed to SRR, the proposed protectoration technique requires that only a subset of the ring nodes needs to support the underlying proxy-stripping technique while all the remaining nodes perform only conventional RPR wrapping and steering, giving rise to evolutionary upgrade paths. More importantly, protectoration is able to guarantee full network connectivity also in the presence of multiple failures on both fiber rings, as opposed to SRR.

The remainder of the paper is organized as follows. The following subsection reviews related work. Section II briefly highlights the key properties of RPR. Section III describes the survivable network and node architecture which builds on RPR. Section IV describes the protectoration multiple-failure

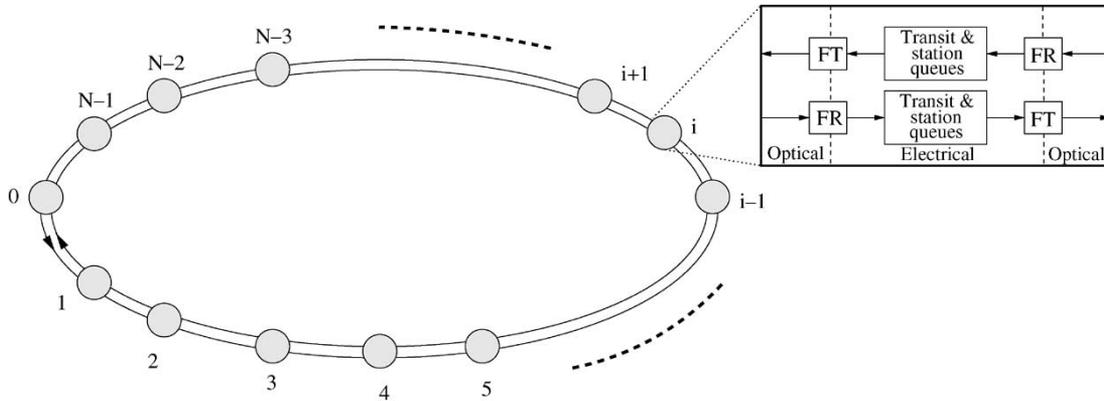


Fig. 2. Generic RPR network and node architecture connecting  $N$  nodes.

recovery technique in greater detail, which is analyzed in Section V. Numerical results in conjunction with verifying simulations are presented in Section VI. Section VII concludes the paper.

#### A. Related Work

The underlying principles and fundamental techniques used for achieving survivability in optical fiber single-channel and wavelength division multiplexing (WDM) networks are discussed in [9] and [10], respectively. An overview of fault management in WDM mesh networks is provided in [11]. The reported fault recovery techniques are categorized into protection or restoration techniques. For more detailed studies of various path and link protection and restoration techniques for WDM networks, the interested reader is referred to [12]–[20] and the references therein. Restoration schemes for IP-over-WDM networks with generalized multiprotocol label switching (GMPLS)-based control signaling were investigated in [21] while protection schemes for these networks were studied in [22]. The provisioning of different levels of fault recovery has recently been studied, see, for instance, [23]–[26] and references therein. Aside from fiber cuts, different equipment failures, e.g., transponder failure, can occur in optical networks. Both fiber and equipment protection schemes and their implementation aspects were examined in [27] and [28], respectively. For more details on interworking aspects between layers 2 and 3 of joint protection/restoration in IP-centric optical WDM networks, the interested reader is referred to [29] and [30]. The graph-theoretical aspects of augmented ring networks that deploy additional shortcut links to the ring have attracted considerable attention [31].

Most previously proposed fault recovery techniques for optical networks are either protection or restoration techniques. In this paper, a hybrid fault recovery technique for optical ring networks, which aims at combining the recovery time of protection (wrapping) and the bandwidth efficiency of restoration by using additional fiber shortcuts, is described and examined. It is noted that the recently reported preconfigured cycles ( $p$ -cycles) [32], [33] and the generalized pre-cross-connected trail (PXT) [34] have the same goal. However, both  $p$ -cycles and PXT are designed for wide area networks (WANs) with a mesh rather than a ring topology.

## II. REVIEW OF RESILIENT PACKET RING

As shown in Fig. 2, RPR is a bidirectional dual-fiber ring network with optical–electrical–optical (OEO) signal conversion at each of the  $N$  nodes [1]–[3]. Every node is equipped with two fixed-tuned transmitters (FTs) and two fixed-tuned receivers (FRs), one for each fiber ring. Broadcasting is achieved by means of source stripping or unidirectional/bidirectional flooding. With source stripping, the source node takes the broadcast packet after one ring round-trip propagation from the ring. With unidirectional and bidirectional flooding, the broadcast packet is taken from the ring by other nodes based on the expiration of the time-to-live (TTL) field in the packet header, which is set by the source node such that all remaining ring nodes are visited. Each node has separate transit and station queues for each ring. Specifically, for each ring, a node has one or two transit queues, one transmission queue termed stage queue, one reception queue, and one add\_MAC queue, which stores control packets generated by the local node.

RPR nodes operate in one of the following two modes: 1) single-queue mode or 2) dual-queue mode. In the single-queue mode, the transit path consists of a single first-in first-out (FIFO) queue termed primary transit queue (PTQ). If the PTQ is not full, highest priority is given to add\_MAC traffic. At the absence of local control traffic, priority is given to in-transit ring traffic over station traffic. In the dual-queue mode, the transit path comprises two queues, one for guaranteed class A traffic (PTQ) and one secondary transit queue (STQ) for class B (committed rate) and class C (best effort) traffic. In the dual-queue mode, if both PTQ and STQ are not full, highest priority is given to add\_MAC traffic (similar to the single-queue mode). If there is no local control traffic, PTQ traffic is always served first. If the PTQ is empty, the local transmission queue (stage queue) is served until the STQ reaches a prescribed queue threshold. When the STQ reaches that threshold, STQ in-transit ring traffic is given priority over station traffic such that in-transit packets are not lost due to buffer overflow. Thus, the transit path is lossless and a packet put on the ring is not dropped at downstream nodes. Furthermore, RPR defines fairness control algorithms that specify how a congested downstream node can throttle the transmission rate of upstream nodes by sending fairness control packets upstream.

To achieve fairness, distributed fairness control algorithms are deployed in RPR according to the so-called ring ingress aggregated with spatial reuse (RIAS) reference model. In RIAS, the level of traffic granularity for fairness determination at a link is defined as an ingress aggregated (IA) flow, i.e., the aggregate of all flows originating from a given ingress node. Moreover, in RIAS, bandwidth can be reclaimed by IA flows when it is unused to ensure maximal spatial reuse. The fairness control in RPR is realized by enabling a backlogged node to send fairness control packets based on its local measurements to upstream nodes in order to throttle their ingress data rates and, thus, alleviate the congestion.

RPR topology discovery provides a reliable and accurate means for all network nodes to discover the topology (number and ordering) of the nodes on the ring network and any changes to that topology due to link/node failures and added/disconnected nodes. This is achieved by collecting information about all nodes and links via the topology discovery protocol. To this end, each node broadcasts topology and protection (TP) control packets on both rings by means of source stripping. TP control packets are sent periodically and when triggered by protection-state changes. Each node uses received TP control packets to build and update its topology database. Among others, the topology database is used by the MAC control entity to compute a complete topology image and determine the shortest path. Note that the topology database enables each node to calculate the number of hops to all the remaining nodes on both rings. For more detailed information on the basic properties of RPR, we refer the interested reader to [1]–[5].

RPR provides a number of advantageous performance features. Among others, the counterrotating rings provide protection against any single link or node failure and the dual-queue operation mode enables service differentiation, e.g., guaranteed quality of service (QoS). Moreover, due to OEO conversion at each node, 3R signal regeneration (reamplifying, reshaping, retiming) can be provided in the electrical domain, which enables optical unamplified transmission between network nodes such that no expensive optical amplifiers are required. Recently, research has begun to propose and investigate fairness performance improvements of RPR [35]–[37].

### III. PROTECTORATION ARCHITECTURE

To guarantee full network connectivity in the presence of multiple link and/or node failures on both fiber rings, additional fiber links are required in the RPR network. Note that for this purpose, dark fibers are preferably used, which are abundantly available in metropolitan areas. The additional (dark) fibers provide alternate physical paths in addition to the fiber rings of RPR. As a result, the network connectivity is increased and, thus, the resultant topology is able to maintain full connectivity in the case of multiple failures, as opposed to RPR, which would be divided into two or more disjoint subrings. Clearly, there are several ways to add fiber links to the bidirectional ring. Among others, individual pairs of ring nodes can be interconnected by means of fiber shortcuts [38] or all ring nodes can be interconnected via a central hub node that consists of multiple working and stand-by wavelength

routers [39]. In this paper, we follow the approach described in [40]. As described in greater detail below, in the protectoration architecture, only a subset of the ring nodes is interconnected by means of a single-hop star WDM network similar to [40]. However, the hub structure differs in that it consists of a single wavelength router in parallel with a single broadcast star coupler, as opposed to a single wavelength router based hub structure proposed in [40]. It was shown in [41] that using a star coupler in parallel with a wavelength router not only protects the wavelength router and, thus, avoids the single point of failure of the star subnetwork but also combines the respective strengths of wavelength routing devices (wavelength router) and wavelength insensitive devices (star coupler) in an efficient manner.

#### A. Building Blocks

Let us first briefly describe the functionality of the underlying building blocks used in the proposed network architecture, which are depicted in Fig. 3(a)–(f).

- 1) *Combiner*: An  $S \times 1$  combiner has  $S$  input ports and one output port, where  $S \geq 1$ . It collects wavelength channels from all  $S$  input ports and combines them onto the common output port. To avoid channel collisions at the output port of the combiner, the collected wavelength channels must be different. Thus, a given wavelength channel can be used only at one of the  $S$  input ports at any time.
- 2) *Splitter*: A  $1 \times S$  splitter has one input port and  $S$  output ports, where  $S \geq 1$ . It equally distributes all incoming wavelength channels to all  $S$  output ports. Hence, a given wavelength channel can be received at all  $S$  output ports.
- 3) *Waveband partitioner*: A waveband partitioner  $\Pi$  has one input port and two output ports. It partitions an incoming set of  $\Lambda$  contiguous wavelength channels into two wavebands (subsets of wavelength channels) of  $\Lambda_A$  and  $\Lambda_B$  contiguous wavelength channels, where  $1 \leq \Lambda_A$ ,  $\Lambda_B \leq \Lambda$  and  $\Lambda = \Lambda_A + \Lambda_B$ . Each waveband is routed to a different output port.
- 4) *Waveband departitioner*: A waveband departitioner  $\Sigma$  has two input ports and one output port. It collects two different wavebands consisting of  $\Lambda_A$  and  $\Lambda_B$  contiguous wavelength channels from the upper and lower input port, respectively. The combined set of  $\Lambda$  wavelength channels is launched onto the common output port, where  $1 \leq \Lambda_A$ ,  $\Lambda_B \leq \Lambda$ , and  $\Lambda = \Lambda_A + \Lambda_B$ .
- 5) *Passive star coupler (PSC)*: A  $D \times D$  PSC has  $D$  input ports and  $D$  output ports, where  $D \geq 1$ . It works similar to a  $D \times 1$  combiner and  $1 \times D$  splitter interconnected in series. Accordingly, it collects wavelength channels from all  $D$  input ports and equally distributes them to all  $D$  output ports. Similar to the splitter, a given wavelength channel can be received at all  $D$  output ports and, similar to the combiner, to avoid channel collisions at the output ports, a given wavelength channel can be used only at one of the  $D$  input ports at any time.

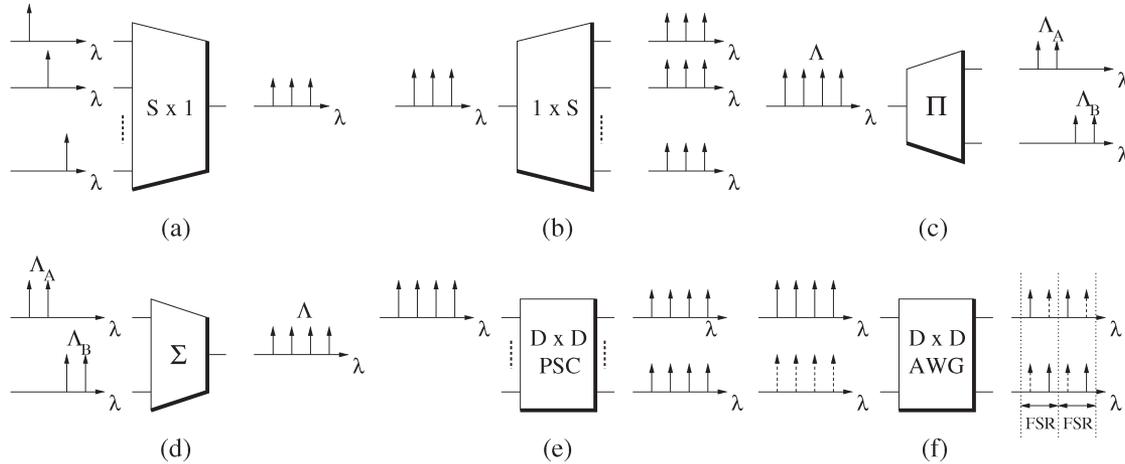


Fig. 3. Architectural building blocks. (a)  $S \times 1$  combiner. (b)  $1 \times S$  splitter. (c) Waveband partitioner. (d) Waveband departitioner. (e)  $D \times D$  PSC. (f)  $D \times D$  AWG with  $D = 2$ .

6) *Arrayed-waveguide grating (AWG)*: A  $D \times D$  AWG has  $D$  input ports and  $D$  output ports, where  $D \geq 1$ . Without loss of generality, we consider a  $2 \times 2$  AWG to explain the properties of an AWG. Fig. 3(f) illustrates a scenario where four wavelengths are fed into both AWG input ports. Let us first consider only the upper input port. The AWG routes every second wavelength to the same output port. This period of the wavelength response is called free spectral range (FSR). In our example, there are two FSRs, each containing two wavelengths. Generally, the FSR of a  $D \times D$  AWG consists of  $D$  contiguous wavelengths, i.e., the physical degree of an AWG is identical to the number of wavelengths per FSR. As depicted in Fig. 3(f), this holds also for the lower AWG input port. Note that the AWG routes wavelengths such that no collisions occur at the AWG output ports, i.e., each wavelength can be applied at all AWG input ports simultaneously. In other words, with a  $D \times D$  AWG, each wavelength channel can be spatially reused  $D$  times, as opposed to the PSC. Furthermore, note that each FSR provides one wavelength channel for communication between a given pair of AWG input and output ports. Hence, using  $R$  FSRs allows for  $R$  simultaneous transmissions between each AWG input-output port pair and the total number of wavelength channels available at each AWG port is given by  $R \cdot D$ , where  $R \geq 1$ .

## B. Network Architecture

As shown in Fig. 4, the proposed network consists of the RPR bidirectional ring subnetwork and a star subnetwork.

1) *Ring Subnetwork*: The RPR ring subnetwork interconnects  $N \geq 1$  nodes that are subdivided into two subgroups of  $N_{rs} = D \cdot S$  ring-and-star homed nodes and  $N_r = N - N_{rs}$  ring homed nodes, with  $D \geq 1$  and  $S \geq 1$ . The  $N_{rs}$  ring-and-star homed nodes are equally spaced among the  $N_r$  ring homed nodes on the ring, as illustrated in Fig. 4 for  $N = 16$  and  $N_{rs} = D \cdot S = 2 \cdot 2 = 4$  (and  $N_r = N - N_{rs} = 12$ ). Unlike the ring homed nodes, the ring-and-star homed nodes are also attached

to the star subnetwork. The ratio  $N/N_{rs}$  denotes the number of ring homed nodes between two adjacent ring-and-star homed nodes, including one of the two ring-and-star homed nodes, and is assumed to be an integer.

2) *Star Subnetwork*: The star subnetwork is based on a central hub that consists of a  $D \times D$  AWG in parallel with a  $D \times D$  PSC, where  $D \geq 1$ . Each ring-and-star homed node  $i$ ,  $i = 1, \dots, N_{rs}$ , has a home channel  $\lambda_i$  on the PSC, i.e., a unique wavelength channel  $\lambda_i$  on which node  $i$  receives data transmitted over the PSC. In addition, there is a control wavelength channel  $\lambda_c$  on the PSC. Consequently, there are  $\Lambda_{PSC} = N_{rs} + 1 = D \cdot S + 1$  wavelength channels on the PSC, which make up the PSC waveband. The AWG waveband consists of  $\Lambda_{AWG} = D \cdot R$  contiguous data wavelength channels, where  $R \geq 1$  denotes the number of used FSRs of the underlying  $D \times D$  AWG. A total of  $\Lambda = \Lambda_{AWG} + \Lambda_{PSC}$  contiguous wavelength channels are operated in the star subnetwork (as detailed further in Section IV).

The signals from  $S$  ring-and-star homed nodes on the  $\Lambda$  wavelength channels are transmitted on  $S$  distinct fibers to an  $S \times 1$  combiner, which combines the signals onto the  $\Lambda$  wavelength channels of one fiber leading to a waveband partitioner. The waveband partitioner partitions the set of  $\Lambda$  wavelengths into the AWG and PSC wavebands, which are fed into an AWG and PSC input port, respectively. The signals from the opposite AWG and PSC output ports are collected by a waveband departitioner and then equally distributed to the  $S$  ring-and-star homed nodes by a  $1 \times S$  splitter. If necessary, optical amplifiers are used between combiner and partitioner as well as splitter and departitioner to compensate for attenuation and insertion losses of the star subnetwork. A total of  $D$  of these arrangements, each consisting of combiner, amplifier, waveband partitioner, waveband departitioner, amplifier, and splitter, are used to connect all  $N_{rs} = D \cdot S$  ring-and-star homed nodes to the central hub.

## C. Node Architecture

Next, let us take a closer look at the structure of both ring homed and ring-and-star homed nodes.

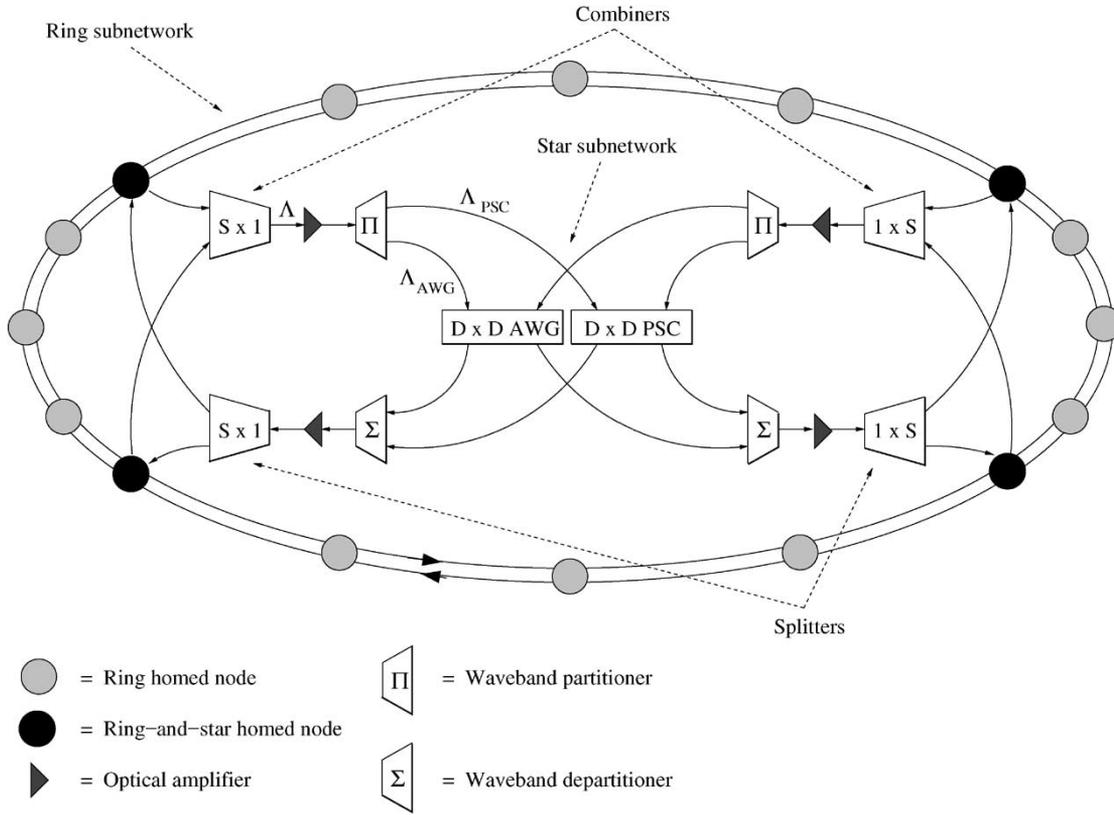


Fig. 4. Network architecture with  $N = 16$  nodes, where  $N_{rs} = D \cdot S = 2 \cdot 2 = 4$  are ring-and-star homed nodes and  $N_r = N - N_{rs} = 12$  are ring homed nodes. There are  $\Lambda_{PSC} = D \cdot S + 1 = 2 \cdot 2 + 1 = 5$  wavelengths on the PSC,  $\Lambda_{AWG} = D \cdot R = 2 \cdot R$  wavelengths on the AWG, for a total of  $\Lambda = \Lambda_{PSC} + \Lambda_{AWG} = 5 + 2 \cdot R$  wavelengths in the star subnetwork.

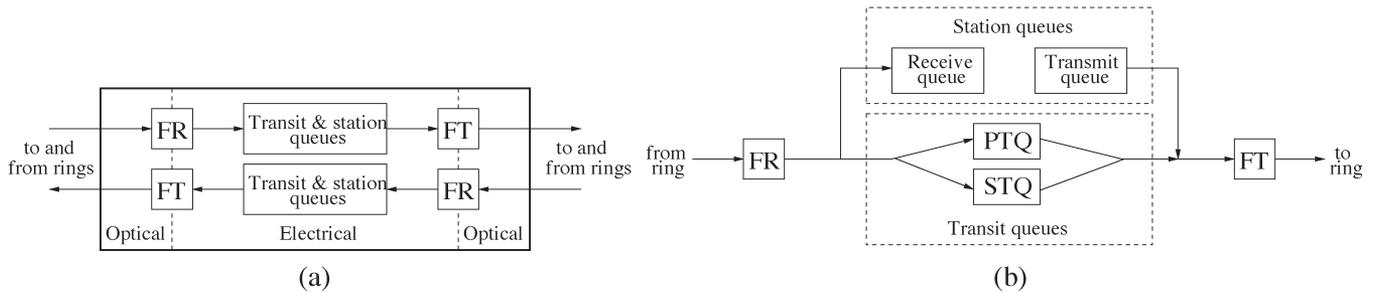


Fig. 5. Ring homed node. (a) Node architecture for both rings. (b) Buffer structure for either ring.

1) *Ring Homed Node*: The architecture of ring homed nodes is identical to that of RPR nodes described in Section II. As shown in Fig. 5(a), every ring homed node is equipped with two FTs and two FRs, one for each ring. Both FT and FR operate at the single wavelength channel of the corresponding ring. Each ring homed node has separate transit and station queues for either ring.

Fig. 5(b) depicts the buffer structure of a ring homed node for either ring in greater detail. For each direction, a ring homed node has one or two transit queues, one transmit queue, and one receive queue (the additional add\_MAC queue for control is not shown). Similar to RPR, ring homed nodes operate in the single-queue mode or dual-queue mode. The service among transmit and transit queues is arbitrated according to the scheduling algorithms reviewed in Section II.

2) *Ring-and-Star Homed Node*: Fig. 6 depicts the architecture of a ring-and-star homed node with PSC data channel  $\lambda_i$ , where  $\lambda_i \in \{1, 2, \dots, D \cdot S\}$ . Each ring-and-star homed node has the same number and type of transceivers and queues as a ring homed node for transmission and reception on both rings. In addition, each ring-and-star homed node has several transceivers that are attached to the star subnetwork by means of a pair of outgoing and incoming fibers. The outgoing fiber is connected to a combiner and the incoming fiber is connected to the splitter that is attached to the opposite AWG-PSC input ports.

As shown in Fig. 6(a), for control transmission on the star subnetwork, each ring-and-star homed node is equipped with a transmitter (FT) fixed tuned to the control wavelength channel  $\lambda_c$  of the PSC waveband, which consists of  $\Lambda_{PSC} = 1 + D \cdot S$

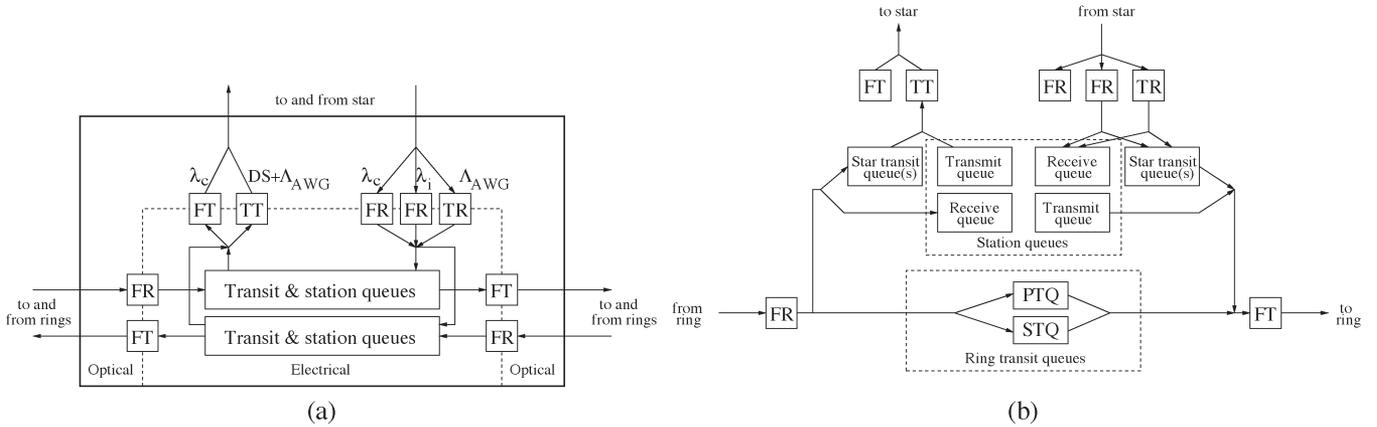


Fig. 6. Ring-and-star homed node with home channel  $\lambda_i \in \{1, 2, \dots, D \cdot S\}$ . (a) Node architecture for both rings. (b) Buffer structure for either ring.

wavelength channels. The remaining  $D \cdot S$  wavelength channels of the PSC waveband and all  $\Lambda_{AWG} = D \cdot R$  wavelength channels of the AWG waveband are accessed for data transmission by a tunable transmitter (TT) whose tuning range equals  $D \cdot S + \Lambda_{AWG} = D(S + R)$ . Similarly, for control reception on the star subnetwork, each ring-and-star homed node is equipped with a receiver (FR) fixed tuned to the control wavelength channel  $\lambda_c$  of the PSC waveband. For data reception on the PSC, each ring-and-star homed node has a separate FR operating at its own dedicated home channel  $\lambda_i \in \{1, 2, \dots, D \cdot S\}$ . Each data wavelength channel of the PSC waveband is dedicated to a different ring-and-star homed node for reception. Data packets transmitted on PSC data wavelength channels do not suffer from receiver collisions (a receiver collision occurs when the receiver of the intended destination node is not tuned to the wavelength channel on which the data packet was sent by the corresponding source node). Moreover, on the wavelength channels of the AWG waveband, data packets are received by a tunable receiver (TR) whose tuning range equals  $\Lambda_{AWG} = D \cdot R$ . All transceivers of the star subnetwork are connected to the station queues. Note that the required tuning range of the TR ( $\Lambda_{AWG}$ ) is smaller than that of the TT ( $D \cdot S + \Lambda_{AWG}$ ). These requirements take into account the current state-of-the-art of tunable transceivers. While fast TTs with a relatively large tuning range have been shown to be feasible [42], [43], TRs are less mature in terms of tuning time and/or tuning range.

Fig. 6(b) depicts the buffer structure of a ring-and-star homed node in greater detail (we show only one ring direction). Each ring-and-star homed node has for each ring direction the following queues: one or two ring transit queues (depending on the operation mode); one ring transmit queue; and one ring receive queue (the additional queue for control is not shown). To send locally generated traffic to and receive traffic destined for itself from the star subnetwork, each ring-and-star homed node has an additional star transmit queue and star receive queue. Furthermore, packets that are pulled from the ring (coming in from both directions of the ring) and forwarded onto the star subnetwork are placed in an additional ring-to-star transit queue (single-queue mode) or one of two additional ring-to-star transit queues according to their priority (dual-queue mode). Traffic that is received from the star subnetwork and needs to be forwarded on either ring is placed in the additional star-to-

ring transit queue (single-queue mode) or one of two additional star-to-ring transit queues according to its priority (dual-queue mode) of the corresponding fiber ring. (The additional queue for sending control on the star subnetwork is not shown.) The service among the transmit and transit queues of the star subnetwork is arbitrated similar to that of their counterparts of the ring subnetwork, as reviewed in Section II. We arbitrate the transmissions from the ring transit queue and the star-to-ring transit queue of a given priority level in a given ring direction for now using a round-robin policy, but note that other more sophisticated arbitration policies are worth examining.

#### IV. PROTECTORATION OPERATION

In this section, we describe and discuss the protection operation technique in greater detail. In Section IV-A, we first describe the normal operation of the network, i.e., without any failures. In Section IV-B, we explain the network operation in the presence of link, node, and other failures.

##### A. Operation Without Failures

1) *Wavelength Channel Allocation Scheme in Star Subnetwork:* Fig. 7 illustrates how the  $\Lambda$  contiguous wavelength channels of the star subnetwork are used for control and data transmission. Time is divided into frames that are repeated periodically. Each frame consists of  $F \geq D \cdot S$  slots, where one slot is equal to the transmission time of a control packet (function and format of a control packet are defined in the following subsection). As shown in Fig. 7, all  $D \cdot S$  home channels of the PSC waveband and all wavelength channels of the AWG waveband are used for data transmission. All these data wavelength channels are not statically assigned to nodes. Instead, access to these wavelength channels is arbitrated by broadcasting control packets on the control wavelength channel  $\lambda_c$  of the PSC prior to transmitting data packets, as explained in greater detail in the following subsection. Control packets are allowed to be sent on  $\lambda_c$  during the first  $D \cdot S$  slots of each frame. More precisely, each of these  $D \cdot S$  slots is dedicated to a different ring-and-star homed node such that channel collisions of control packets are avoided. The remaining  $(F - D \cdot S)$  slots of each frame can be used for data transmission on  $\lambda_c$ . Note that data packets sent during these slots on  $\lambda_c$  are received by all ring-and-star homed

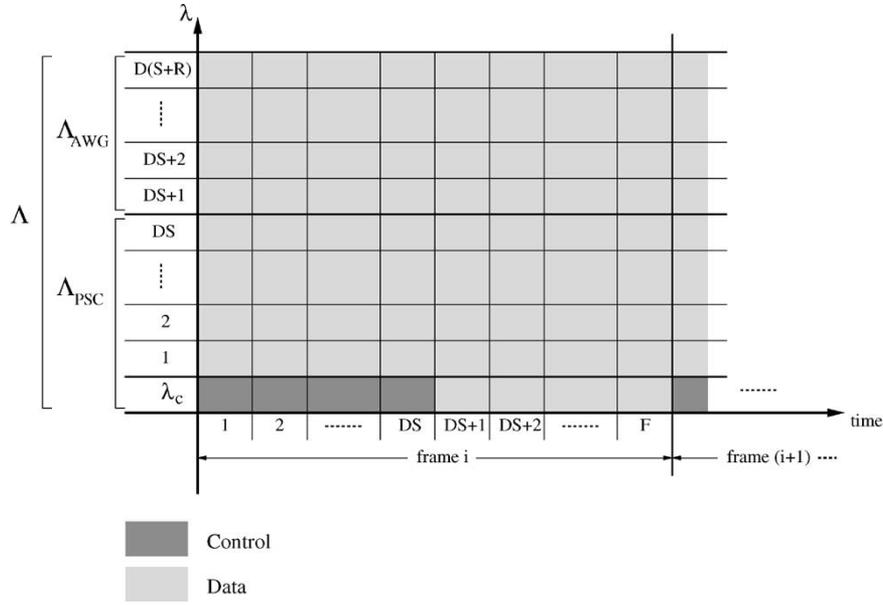


Fig. 7. Wavelength assignment in star subnetwork.

nodes by using their receiver fixed tuned to  $\lambda_c$ . Thus, these slots allow for broadcasting in the star subnetwork.

2) *Wavelength Access*: All  $N$  ring and ring-and-star homed nodes use the single-queue or dual-queue scheduling algorithm to arbitrate service among transit and station queues of the ring subnetwork, as outlined in Section III-C. In the following, we consider unicast (point-to-point) traffic. Each of the  $N$  nodes sends data packets on the shortest path to the corresponding destination node. Next, we specify the shortest-path routing for both ring and ring-and-star homed nodes. Let one hop denote the distance between two adjacent nodes. Adjacent nodes can either be two neighboring nodes on the ring or two nodes interconnected via the single-hop star subnetwork (this holds only for ring-and-star homed nodes). We define the following variables for a given pair of source node  $s$  and destination node  $d$ , where  $s, d \in \{0, 1, \dots, N-1\}$ :

- $h_{s\_rs}$  hop distance between source node  $s$  and its closest ring-and-star homed node;
- $h_{d\_rs}$  hop distance between destination node  $d$  and its closest ring-and-star homed node;
- $h_{\min_{s\_d}}^{\text{ring}}$  minimum hop distance between source node  $s$  and destination node  $d$  on the ring without using the shortcuts of the star subnetwork;
- $h_{\min_{s\_d}}^{\text{star}}$  minimum hop distance between source node  $s$  and destination node  $d$  on the ring with using the shortcuts of the star subnetwork (note that  $h_{\min_{s\_d}}^{\text{star}} = h_{s\_rs} + 1 + h_{d\_rs}$ ).

a) *Ring homed nodes*: Generally speaking, if the hop distance between a given source ring homed node  $s$  and destination node  $d$  is “small enough,” the ring homed node sends the data packet(s) on the ring without using the shortcuts of the star subnetwork. More precisely, if  $h_{\min_{s\_d}}^{\text{ring}} \leq h_{\min_{s\_d}}^{\text{star}}$ , then source node  $s$  sends the data packet(s) to destination node  $d$  along the ring on the shortest path by choosing the appropriate fiber ring. Destination node  $d$  takes the transmitted data packet(s)

from the ring. Note that in this case, intermediate ring-and-star homed nodes forward the data packet(s) on the ring rather than sending them across the star subnetwork. However, if  $h_{\min_{s\_d}}^{\text{ring}} > h_{\min_{s\_d}}^{\text{star}}$ , i.e., if the shortcuts of the star subnetwork form a shorter path between nodes  $s$  and  $d$  than either peripheral fiber ring, the source node  $s$  sends the data packet(s) to its closest ring-and-star homed node. Note that the chosen direction does not necessarily have to be the same as that used in shortest-path routing on the ring. The corresponding ring-and-star homed node pulls the data packet(s) from the ring, as described in greater detail in the following.

b) *Ring-and-star homed nodes*: To pull data packet(s) from the ring, ring-and-star homed nodes perform the so-called *proxy-stripping* technique. Proxy stripping was proposed and examined in [44] in a high-level fashion without specifying any star subnetwork. In the following, we briefly review the main characteristics of proxy stripping and adapt it to our specific star subnetwork.

With proxy stripping, a given ring-and-star homed node pulls only data packets from the ring whose source and destination addresses satisfy the condition  $h_{\min_{s\_d}}^{\text{ring}} > h_{\min_{s\_d}}^{\text{star}}$ . A ring-and-star homed node puts data packet(s) pulled from the ring in one of the two corresponding star transit queues belonging to its TT that is attached to the star subnetwork. The star transit queue of the TT is chosen according to the priority of the pulled data packet(s). The service among these two star transit queues that store in-transit traffic coming from the ring and the transmit queue that stores locally generated traffic is arbitrated by applying the same scheduling algorithms as used on the ring (see Section III-C). That is, ring-to-star in-transit traffic is given priority over star traffic locally generated by the proxy-stripping node. Similar to the transit queues on the ring, the star transit queues, thus, provide a lossless path for in-transit traffic. To guarantee losslessness, the star subnetwork needs to be dimensioned properly, as analyzed in greater detail in Section V-B. Depending on the traffic pattern as well as the

number and location of the proxy-stripping nodes, the amount of proxy-stripped traffic may become large. To provide lossless delivery of proxy-stripped packets, the star subnetwork in general needs to operate at a higher line rate than the ring subnetwork (see Section V-B). Alternatively, each proxy-stripping node may be equipped with more than one star data transceiver, each operating at the same line rate as the ring transceivers. For more details on star WDM networks with multiple transceivers at each node, the interested reader is referred to [45]. Note that the star transit queues (as well as the star station queues) of each ring-and-star homed node need to be added to the RPR MAC layer. Prior to transmitting a data packet, the corresponding ring-and-star homed node broadcasts a control packet on  $\lambda_c$  to all  $N_{rs}$  ring-and-star homed nodes in its assigned slot of the upcoming frame by using its FT. The control packet consists of three fields, namely: 1) destination address of the ring-and-star homed node that is closest to destination node  $d$ ; 2) length of the corresponding data packet; and 3) priority of the corresponding data packet. After announcing the data packet in its assigned control slot, the ring-and-star homed node transmits the corresponding data packet on the home channel  $\lambda_i$  of the addressed ring-and-star homed node in the subsequent  $L$  slots by using its TT, where  $\lambda_i \in \{1, 2, \dots, D \cdot S\}$  and  $L$  denotes the length of the data packet in number of slots. Data packets are sent within the same frame as the corresponding control packet and have a maximum length of  $(F - D \cdot S)$  slots, i.e.,  $1 \leq L \leq F - D \cdot S$ . We note that due to this assumption, a small fraction of each home channel  $\lambda_i$  is not used at the beginning of each frame. However, this could be easily avoided by letting nodes send data packets across the boundary of adjacent frames. After an end-to-end propagation delay of the PSC of the star subnetwork, all ring-and-star homed nodes receive the broadcast control packet by using their FRs fixed tuned to  $\lambda_c$ . The corresponding data packet is successfully received at the addressed ring-and-star homed node by using its FR fixed tuned to  $\lambda_i$ , unless one or more other ring-and-star homed nodes have transmitted data packets on  $\lambda_i$  in at least one of the aforementioned  $L$  slots. In the latter case, all involved data packets are assumed to be corrupted due to (channel) collision and have to be retransmitted. Collided data packets are kept in the queues until the transmission is successful. Note that due to the fact that control packets are sent collision free, all ring-and-star homed nodes are aware of the original order of the corresponding data packets. Consequently, even though collided data packets need to be retransmitted, the receiving ring-and-star homed nodes are able to restore the original sequence of data packets and, thus, maintain in-order packet delivery.

The retransmission of collided data packets works as follows. Due to the dedicated access control of the control channel  $\lambda_c$ , collisions of control packets are prevented. Therefore, for collided data packets, no control packets have to be retransmitted. Instead, each ring-and-star homed node is able to find out which transmitted data packets have experienced channel collision by processing the previously (successfully) transmitted control packets. More precisely, the index  $j$ ,  $1 \leq j \leq D \cdot S$ , of the used control slot and the destination and length fields of the control packet enable each ring-and-star homed node to determine whether the corresponding data packet has collided or not.

Collided data packets are not retransmitted on the home channels of the PSC but across the AWG by using one of the  $\Lambda_{AWG}$  wavelength channels. Given the index  $j$  of the control slot, which uniquely identifies not only the given source ring-and-star homed node but also, more importantly, the input port of the AWG to which it is attached, together with the destination field of the corresponding control packet all ring-and-star homed nodes are able to determine the wavelength in each FSR of the AWG, which provides a single-hop connection between the corresponding pair of source and destination ring-and-star homed nodes. The actual retransmissions on the chosen wavelength channels are scheduled in a distributed fashion by all ring-and-star homed nodes. The scheduling starts at the beginning of frame  $(i + 1)$  upon receiving the control packets in frame  $i$  after one end-to-end propagation delay of the PSC of the star subnetwork. At the end of every frame, each ring-and-star homed node collects all control packets belonging to collided data packets. By using each control packet's priority field, each ring-and-star homed node first processes all high-priority control packets and then all low-priority control packets. Control packets of the same priority class are randomly chosen for scheduling. All ring-and-star homed nodes deploy the same random algorithm and same seed and therefore build the same schedule. Note that randomizing the scheduling counteracts the static control slot assignment, resulting in an improved fairness among the ring-and-star homed nodes. Otherwise, source nodes with a smaller index  $j$  would be more successful in the scheduling than nodes with a larger index. The corresponding data packets of the selected control packets are scheduled on a first-fit basis starting from the lowest possible wavelength channel at the earliest possible time. The data packet is retransmitted on the corresponding AWG wavelength channel at the scheduled time. After the successful retransmission of a given data packet across the AWG, the corresponding ring-and-star homed receiving node puts the data packet in the star subnetwork receive queue belonging to its TR if the data packet is destined for itself. Otherwise, the ring-and-star homed node forwards the data packet on the ring towards the destination node  $d$  on the shortest path by using the appropriate fiber ring and placing the data packet in the corresponding star-to-ring transit queue. Destination node  $d$  finally takes the data packet from the ring. We note that the aggregated length of the collided packets can be larger than  $F - D \cdot S$ . This fact does not pose any problems since the retransmission takes place over the AWG where transmissions are permitted to be scheduled across frame boundaries.

Besides pulling data packets from the ring and forwarding them on the ring, ring-and-star homed nodes also generate traffic. Note that in this case, we have  $h_{s_{rs}} = 0$ . Again, if  $h_{\min_{s,d}}^{\text{ring}} \leq h_{\min_{s,d}}^{\text{star}}$ , then the ring-and-star homed source node  $s$  transmits the data packet on that fiber ring that provides the shortest path to destination node  $d$ . Otherwise, if  $h_{\min_{s,d}}^{\text{ring}} > h_{\min_{s,d}}^{\text{star}}$ , then the ring-and-star homed source node  $s$  sends the data packet across the star subnetwork to the corresponding ring-and-star homed node, which is either the destination itself or forwards the data packet onwards to node  $d$  via the shortest path ring. (Re)transmission and reception of the data packet on the star subnetwork are done in the same way as explained above.

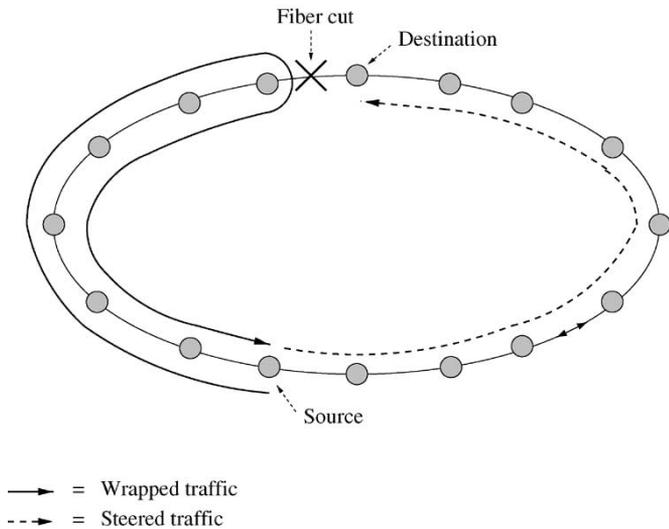


Fig. 8. RPR bidirectional ring with  $N = 16$  nodes using wrapping and steering in the event of a fiber cut.

**B. Operation With Failures**

Aside from link and node failures, other network elements may fail. In the star subnetwork, splitters, amplifiers, combiners, waveband partitioners/departitioners, PSC, or AWG may go down. Note that the various failures affect the network differently. For instance, while a fiber cut between a given ring-and-star homed node and the attached combiner disconnects only a single node from the star subnetwork, the entire star subnetwork goes down if the central hub fails, i.e., if both AWG and PSC fail. In the following, we assume that each node is able to detect any type of failure in both ring and star subnetworks. For a more detailed discussion on available techniques for fault detection in the ring and star subnetwork, we refer the interested reader to [4] and [41], respectively.

Let us first briefly review the wrapping and steering techniques of RPR. Fig. 8 depicts an RPR bidirectional ring with  $N = 16$  nodes, including a pair of source and destination nodes. The source node sends its data packets in clockwise direction since this direction provides the shortest path in terms of hops. For illustration, we assume that a single fiber cut has occurred right before the destination node. Upon detection of the link failure, the node on the left-hand side of the fiber cut wraps the traffic away from the link failure in the counterclockwise direction. In addition, the node on the left-hand side of the fiber cut broadcasts a control packet in the counterclockwise direction in order to inform all other nodes about the link failure. The wrapped traffic travels all the way back to the source node. Upon learning that a link failure has occurred, the source node steers the traffic away from the fiber cut and sends all traffic in the counterclockwise direction. Recall from Section I that the two protection techniques—wrapping and steering—lead to a rather inefficient use of bandwidth resources due to 1) the round trip between source node and wrapping node without getting closer to the destination node and 2) the secondary path (counterclockwise direction in our example), which is longer than the primary path in terms of hops. Furthermore, in case of an additional link or node failure on the secondary path, the

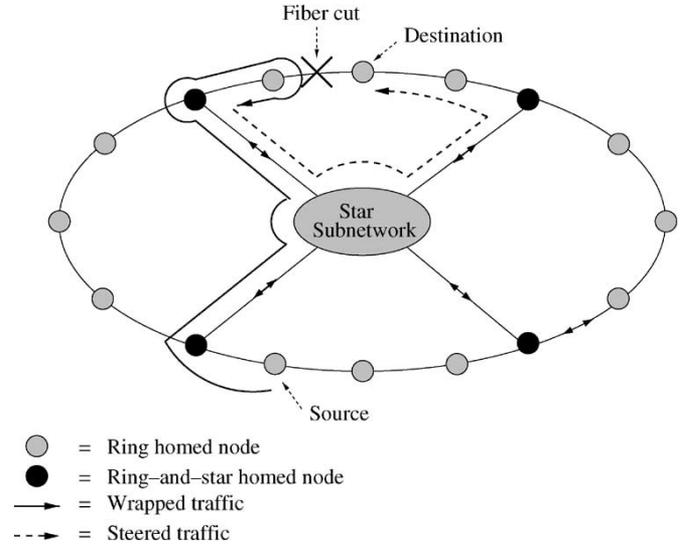


Fig. 9. RPR bidirectional ring with  $N_r = 12$  ring homed nodes and  $N_{rs} = 4$  ring-and-star homed nodes using protectoration in the event of a fiber cut.

source node would be unable to send traffic to the destination node since the ring network would be divided into two disjoint subrings, one containing the source node and the other one the destination node. Thus, the protection techniques of the bidirectional RPR ring network are able to recover only from a single link failure. Likewise, the RPR ring is able to recover only from a single node failure.

Next, we explain the protectoration technique in greater detail. The protectoration technique builds on the wrapping and steering techniques of RPR and, thus, provides an evolutionary upgrade of RPR. Moreover, protectoration makes RPR resilient against multiple link and node failures in an efficient manner, as we shall see shortly. In Section IV-B-1, we first consider link and node failures only in the ring subnetwork while the star subnetwork is assumed to work properly. In Section IV-B-2), we also take failures in the star subnetwork into account.

1) *Failures Only in Ring Subnetwork:* Fig. 9 depicts an RPR bidirectional ring with  $N = 16$  nodes, where  $N_{rs} = 4$  ring-and-star homed nodes are interconnected via the star subnetwork of Section III-B and  $N_r = N - N_{rs} = 12$  are conventional ring homed nodes. Recall from Section IV-A-2)-b) that the ring-and-star homed nodes perform proxy stripping. Again, for illustration, we consider a pair of source and destination nodes and a single fiber cut, as shown in Fig. 9. The source node sends all data packets intended for the destination node to its closest proxy-stripping node, which, in turn, forwards the data packets across the single-hop star subnetwork to the proxy-stripping node that is closest to the destination node. Upon detection of the fiber cut, the data packets are wrapped and return to the proxy-stripping node closest to the destination node. Now, instead of forwarding the wrapped traffic to the source node on the counterclockwise ring (as done in conventional RPR), the corresponding proxy-stripping node sends the wrapped data packets across the single-hop star subnetwork to that proxy-stripping node, which is on the other side of the fiber cut. The latter proxy-stripping node receives the data packets from the star subnetwork and forwards them to the destination node on

the counterclockwise ring. Note that the former proxy-stripping node (on the left-hand side of the link failure) deploys not only proxy stripping but also steering of wrapped traffic. Thus, wrapped traffic does not have to go back all the way to the source node but is sent along a single-hop path to another proxy-stripping node. Compared to Fig. 8, in the event of a single fiber cut, wrapped traffic neither makes a round trip between source and destination nodes nor takes any long secondary path. As a result, wrapped traffic consumes significantly fewer bandwidth resources on the ring network, resulting in a more efficient use of bandwidth. After learning about the fiber cut, the source node—and its closest reachable ring-and-star homed node, in case proxy stripping results in an intact source-to-destination route with a smaller hop count—transmits the data packets along a different path. In our example, after learning that the fiber cut has occurred, the source node sends the data packets intended for the destination node to its closest proxy-stripping node, which, in turn, sends the proxy-stripped packets on the star subnetwork directly to the proxy-stripping node on the right-hand side of the destination node. The latter proxy-stripping node finally forwards the data packets to the destination node on the counterclockwise ring. Hence, protection requires significantly fewer bandwidth resources on the ring than conventional steering, which would use the peripheral counterclockwise ring that is significantly longer than the shortcuts of the star subnetwork in terms of hops.

Note that a ring-and-star homed node is able to determine which packets have to be proxy stripped by using the source and destination addresses available in RPR's MAC address fields of each packet. These MAC addresses, together with the direction a given packet comes from, enable each ring-and-star homed node to determine whether a given data packet has undergone wrapping or not. If wrapping has taken place, the corresponding ring-and-star homed node recomputes the shortest path taking the link failure into account and sends a given wrapped packet along the updated shortest path. Depending on the updated shortest path, the corresponding ring-and-star homed node forwards a given wrapped data packet either on the ring or star subnetwork. Note that each node computes the shortest path by using its topology database. Each node maintains and updates its topology database by means of RPR's built-in topology discovery protocol, as explained above in Section II. The topology discovery protocol is used to disseminate topology information of both ring and star subnetworks.

Clearly, single-failure scenarios also include node failures apart from link failures. While a single failed ring homed node triggers the same procedure as above, special attention has to be paid to a failed ring-and-star homed node. If a ring-and-star homed node goes down, it is no longer available for proxy-stripping traffic from the ring subnetwork and forwarding traffic coming from the star subnetwork. In this case, the two ring homed nodes adjacent to the failed proxy-stripping node detect the failure and inform the remaining nodes by sending control packets. After learning about the failed proxy-stripping node, the remaining nodes do not send traffic to the failed ring-and-star homed node. Instead, the neighboring proxy-stripping nodes of the failed proxy-stripping node take over its role of proxy-stripping regular traffic and steering wrapped traffic.

Next, let us consider multiple failures in the ring subnetwork. If there are multiple link and/or node failures on the ring subnetwork, nodes can use the star subnetwork to bypass the failures. Thus, with an intact star subnetwork, multiple link and/or node failures on the ring subnetwork may occur simultaneously without losing full connectivity. Note, however, that full connectivity in the event of multiple failures is only guaranteed if no more than one link or node failure occurs between each pair of ring-and-star homed nodes. Otherwise, one or more nodes between a given pair of ring-and-star homed nodes are disconnected from the network.

2) *Failures in Both Ring and Star Subnetworks:* Failures in the star subnetwork include fiber cuts and nonfunctional network devices such as failed combiners/splitters, waveband (de)partitioners, AWG, PSC, and amplifiers. Depending on the failure, only one, a subset, or all ring-and-star homed nodes are disconnected from the star subnetwork. More precisely, a fiber cut between a given ring-and-star homed node and the combiner/splitter port to which it is attached disconnects only the ring-and-star homed node from the star subnetwork. If a given combiner/splitter, amplifier, waveband (de)partitioner, or any fiber between these devices goes down, all  $S$  corresponding ring-and-star homed nodes are disconnected from the star subnetwork. If the central hub (AWG and PSC) goes down, the connectivity of the star subnetwork is entirely lost, reducing the network to a conventional bidirectional RPR ring network. If a given ring-and-star homed node detects that it is disconnected from the star subnetwork, it is unable to send and receive traffic to and from the star subnetwork. After detecting disconnection, the affected ring-and-star homed node informs all the remaining nodes by broadcasting a control packet on either ring and acts subsequently as a conventional ring homed node. Failures in the ring subnetwork are handled as described above.

3) *Discussion:* The bidirectional RPR ring network with its two protection techniques wrapping and steering is able to guarantee full connectivity only if no more than one link or node failure occurs. Full connectivity also in the event of multiple link and/or node failures can be achieved by interconnecting several ring nodes via a star subnetwork. In doing so, the ring is divided into several segments, each comprising the nodes between two adjacent ring-and-star homed nodes. Each segment is able to recover from a single link or node failure without losing full connectivity of the network. Thus, the number of fully recoverable link and/or node failures is identical to the number of ring-and-star homed nodes, provided that there is no more than one failure in each segment.

Similar to RPR ring networks, both ring and ring-and-star homed nodes perform wrapping and steering. In addition, ring-and-star homed nodes also perform proxy stripping. By means of proxy stripping, wrapped traffic is sent across single-hop shortcuts to the neighboring ring-and-star homed node, thereby bypassing the link or node failure(s) of the corresponding ring segment(s). As opposed to the RPR bidirectional ring, wrapped data packets do not have to travel back to the corresponding source node. Instead, steering is also done by the ring-and-star homed node that receives wrapped data packets by sending the wrapped data packets across the single-hop shortcuts of star subnetwork rather than along the ring subnetwork. In doing

so, the corresponding ring-and-star homed node restores the network connectivity in a more efficient manner. After learning that a failure has occurred, the source node steers the traffic along the updated shortest path by capitalizing on proxy stripping. Consequently, steered traffic does not have to travel on the longer secondary path along the peripheral ring, requiring fewer bandwidth resources and resulting in an improved bandwidth efficiency.

The proposed multiple-failure recovery technique combines the fast recovery time of protection (wrapping) and the bandwidth efficiency of restoration (steering together with proxy stripping). Accordingly, we call this hybrid approach protectoration.

### C. Impact on RPR

In this section, we discuss the impact of the modified ring-star network topology and the protectoration technique on other important properties of RPR. In particular, we consider their impact on the scheduling algorithms, the lossless property, the service differentiation, the in-order packet delivery, and the fairness control of RPR.

Protectoration does not change the scheduling algorithms of RPR. The service among the station and transit queues of the star subnetwork is arbitrated as that of their counterparts of the ring subnetwork. On the ring, in-transit traffic has priority over station traffic. Similarly, proxy-stripped traffic, which is pulled from the ring, is stored in the ring-to-star transit queue(s) of the corresponding ring-and-star homed node and is sent with higher priority across the star subnetwork than locally generated traffic. Traffic that arrives from the star subnetwork and needs to be forwarded on the ring is put in the star-to-ring transit queue(s) and is forwarded together with ring in-transit traffic in a round-robin fashion with higher priority than traffic locally generated by the corresponding ring-and-star homed node. Hence, in-transit traffic is given priority over station traffic on both ring and star subnetworks. As a result, with proxy stripping, in-transit packets are not lost due to buffer overflow, thus, maintaining the lossless property of RPR. Note that this holds only if the amount of both ring in-transit traffic and star-to-ring in-transit traffic remain below a certain threshold. Since at each ring-and-star homed node, two queues (star-to-ring and ring transit queues) compete for the bandwidth on the outlink to the downstream ring node packets may be lost due to buffer overflow, unless congestion control is applied. We will see shortly how RPR's fairness control algorithm is able to alleviate the congestion and, thus, prevent packet loss in both ring and star-to-ring transit queues.

The protectoration technique supports the service differentiation of the dual-queue mode RPR by storing packets that are sent on or received from the star subnetwork in one of the ring-to-star or star-to-ring transit queues, respectively, according to their priority. Also note that on the star subnetwork reservation control packets carry the priority of the data packets which are subsequently sent across the star. This not only enables the receiving ring-and-star homed node to put an incoming data packet in the corresponding star-to-ring transit queue but also allows for a collided data packet to be retransmitted according

to its priority indicated in the corresponding control packet, which has been sent collision free due to the dedicated time slot assignment on the control channel.

Under stable operational conditions, packets sent between a given pair of source and destination nodes are delivered in order along the shortest path in the hybrid ring-star network by using proxy stripping. Note that proxy stripping does not require any modifications of either of the RPR protection techniques wrapping and steering, which are used in the event of network failures. Like in RPR, packets marked eligible for wrapping are wrapped by the nodes adjacent to a given failure, which are subsequently proxy stripped by the next ring-and-star homed node and sent along the updated shortest path on the star subnetwork, thus, bypassing the failed ring link or node. Steering does not need to be modified either. Proxy stripping can be used for both the strict packet mode, which is the default packet mode in IEEE 802 protocols, and the relaxed packet mode of RPR. These two packet modes affect the steering of traffic by the source node after learning about a failure. In brief, in the strict packet mode, all ring nodes stop adding packets and discard all in-transit packets until their topology database is updated and provides a stable and consistent topology image. In the relaxed mode, ring nodes may steer packets immediately after a failure without waiting for their topology image to become stable and consistent. Note that steering is done by source nodes, whereas proxy stripping is done by intermediate nodes between a given pair of source and destination nodes. To guarantee in-order packet delivery in the hybrid ring-and-star network, source nodes must apply strict-mode steering while intermediate nodes perform proxy stripping of both wrapped traffic and steered traffic, as discussed in detail above.

As for fairness, we note that both the aggressive and conservative mode fairness control algorithms of RPR have been investigated recently. Among others, it was shown in [3] that under unbalanced and constant-rate traffic inputs, the spatial reuse in RPR is significantly decreased due to severe and permanent oscillations that span nearly the entire range of the link capacity, which not only hinder spatial reuse but also decrease throughput as well as increase delay jitter. Due to space limitations, we do not address fairness control in the hybrid ring-star network in this paper. However, we briefly note that the star subnetwork inherently provides fair channel access among the ring-and-star homed nodes due to the randomized scheduling of data packets that need to be retransmitted, as detailed further in [46]. Moreover, in [47], we have examined an extended version of the so-called distributed virtual-time scheduling in rings (DVSR) fairness control algorithm incorporating proxy stripping. DVSR is a novel fairness control algorithm for RPR networks that is able to mitigate the aforementioned oscillations and achieve nearly complete spatial reuse [37]. The modified version of DVSR in [47] enables ring-and-star homed nodes to disseminate their locally measured fair link rates of the star subnetwork to all remaining ring nodes and is able to establish so-called RIAS fair transmission rates in the hybrid ring-star network. (Recall from Section II that RIAS is the fairness reference model used in RPR.) More specifically, one fairness control packet circulates on each ring. Each fairness control packet consists of  $(N + DS/2)$  fields.

The first  $N$  fields contain the fair rates of all ring links and the remaining  $DS/2$  fields contain the fair rates of the star links, where one control packet carries the rates of the even-numbered and the other one the rates of the odd-numbered star links. Each node monitors both fairness control packets and writes its locally computed fair rates in the corresponding fields of the fairness control packets. To calculate the fair link rates, each node measures the number of bytes  $l_k$  arriving from node  $k$ , including the station itself, during the time interval  $T$  between the previous and the current arrival of the control packet. Each node performs separate measurements for either ring using two separate time windows. Proxy-stripping nodes additionally count the number of bytes arriving from the star for each node and use the time window of the fairness control packet that carries the fair rate of the corresponding proxy-stripping node. The fair rate  $F$  of a given link is equal to the max-min fair share among all measured link rates  $l_k/T$  with respect to the link capacity  $C$  currently available for fairness-eligible traffic. Each node limits the data rate of its  $(N - 1)$  ingress flows by using token buckets whose refill rates are set to the current fair rates of the corresponding destinations. Using the same two time windows as in the calculation of the link fair rates above, each node  $i$  counts the bytes  $\rho_{ij}$  sent to destination  $j$  during the time window. Thus, there are two sets of  $(N - 1)$  byte counters, one for each time window. Each time a fairness control packet arrives, a given node calculates the fair rate of each ingress flow as follows. According to RIAS, the total capacity available to a given node on a certain link equals the fair rate  $F$ , which is shared among all its ingress flows crossing that link. Based on the measured ingress rates  $\rho_{ij}/T$  of these flows and the available capacity  $F$ , the max-min fair share  $f$  is calculated for each crossed link. The refill rate of each token bucket is set to the minimum fair share  $f$  of these links.

In the proposed fairness control protocol, only ring-and-star homed nodes need to measure the fair link rates of the star subnetwork and write them into the corresponding fields of each fairness control packet. The remaining ring homed nodes simply gather this information and compute the fair link rates for the ring-star network, as given in their topology database. Hence, the proposed fairness control protocol mainly affects only the subset of ring-and-star homed nodes. Note that the fairness control protocol may also be used by each ring-and-star homed node to control the congestion on the links of both star and ring subnetworks such that packet loss due to buffer overflow is avoided.

In summary, the proposed protection technique used in the hybrid ring-star network does not require any major modifications of the basic RPR protocols and mechanisms and appears to be a viable solution to improve the resilience of RPR dramatically, as we will see shortly.

## V. ANALYSIS

In this section, we develop an analytical model for investigating the protection technique in terms of stability, utilization, and throughput-delay performance. We also address the dimensioning and identify the bottlenecks of the network. We

note that in our analysis, we do not take fairness control into account. The obtained results are intended to give the maximum achievable throughput-delay performance of the protection technique and to provide an upper bound that enables the performance comparison of fairness control mechanisms, e.g., [35]–[37].

### A. Assumptions

In our analysis, we make the following assumptions.

- 1) *Single-queue mode*: We examine the single-queue mode of RPR, i.e., each node is equipped with one PTQ but no STQ on either ring. For sending proxy-stripped traffic across the star subnetwork, each ring-and-star homed node is equipped with an additional star transit queue.
- 2) *Infinite FIFO queues*: All queues are assumed to be FIFO queues of infinite capacity. (This is for analytical simplicity. In our verifying simulations, we use finite-size FIFO queues that provide very good matches between analysis and simulation results.)
- 3) *Propagation delay*: The  $N$  nodes are equally spaced on the ring. The propagation delay between two adjacent ring nodes is given by  $\tau$ . Thus, the round-trip time (RTT) of the RPR ring equals  $N \cdot \tau$ . The propagation delay of the PSC and AWG of the star subnetwork is equal to  $\tau_{PSC}$  and  $\tau_{AWG}$ , respectively. Both  $\tau_{PSC}$  and  $\tau_{AWG}$  are assumed to be the same for all ring-and-star homed nodes. All propagation delays are given in slots.
- 4) *Unicast traffic*: We consider unicast traffic, i.e., all data transmissions are point-to-point.
- 5) *Packet generation process*: At node  $i$ , the average number of locally generated packets that are destined for node  $j$  per frame is equal to  $\sigma(i, j) \geq 0$ . For the stability analysis in Section V-B, the packet generation process is assumed to be stationary and ergodic. For the delay analysis in Section V-D, the packet generation process is assumed to be Poissonian.
- 6) *Packet length distribution*: We consider variable-size data packets with a length of  $L$  slots,  $1 \leq L \leq F - D \cdot S$ . The packet length distribution is independent from source node  $i$  and destination node  $j$ . Let  $T$  be a random variable denoting the packet transmission time (in slots), and let  $E[T]$  denote its mean.

### B. Stability and Dimensioning

Let us introduce the following definitions. For locally generated traffic at node  $i$ , we define the following parameters.

- 1)  $\sigma_i^+(i)$  denotes the mean number of locally generated packets at node  $i$  per frame, which are sent in the clockwise direction of the ring subnetwork.
- 2)  $\sigma_i^-(i)$  denotes the mean number of locally generated packets at node  $i$  per frame, which are sent in the counterclockwise direction of the ring subnetwork.
- 3)  $\sigma_i^*(i)$  denotes the mean number of locally generated packets at node  $i$  per frame, which are sent directly across the star subnetwork (holds only for ring-and-star homed nodes).

Similarly, for in-transit traffic at node  $i$ , we define the following parameters.

- 1)  $\sigma_r^+(i)$  denotes the mean number of packets arriving at node  $i$  per frame, which are forwarded in the clockwise direction of the ring subnetwork.
- 2)  $\sigma_r^-(i)$  denotes the mean number of packets arriving at node  $i$  per frame, which are forwarded in the counterclockwise direction of the ring subnetwork.
- 3)  $\sigma_r^*(i)$  denotes the mean number of packets arriving at node  $i$  per frame, which are forwarded on the star subnetwork (holds only for ring-and-star homed nodes).

We note that if node  $i$  is a ring-and-star homed node, the quantities  $\sigma_r^+(i)$ ,  $\sigma_r^-(i)$ , and  $\sigma_r^*(i)$  account for in-transit traffic that comes from and goes to both the ring subnetwork and the star subnetwork, e.g.,  $\sigma_r^+(i)$  for a ring-and-star homed node  $i$  accounts for both the traffic that arrives from the ring subnetwork and is to be forwarded in the clockwise direction over the ring as well as the traffic that arrives from the star subnetwork and is to be forwarded in the clockwise direction over the ring.

Next, let  $p_{ij}(e)$  denote the probability that a data packet that is generated at node  $i$  and is destined to node  $j$  traverses a given (directed) fiber link  $e$  of the ring subnetwork between two adjacent nodes. Similarly, for ring-and-star homed nodes  $k$  and  $l$ , let  $p_{ij}(k, l)$  denote the probability that a data packet that is generated at node  $i$  and is destined to node  $j$  traverses the star subnetwork from  $k$  to  $l$ . The calculation of the probabilities  $p_{ij}(e)$  and  $p_{ij}(k, l)$  depends on the status of the network and the applied routing. As explained in Section IV, all nodes deploy shortest-path routing, noting that the presence of link and/or node failures may affect the shortest path since failed links and nodes can no longer be traversed. The fault scenarios under consideration are single and multiple link and/or node failures. If for a given pair of source and destination nodes there exist multiple shortest paths, the traffic load is balanced among the multiple shortest paths equally. (Alternatively, the tie could be broken by arbitrating the routing based on the indices of both source and destination nodes. We note that this assumption of load balancing is for analytical simplicity. To guarantee in-order packet delivery a given source node has to send all packets to the corresponding destination node along the same path by choosing one of the multiple shortest paths.) To determine  $p_{ij}(e)$  and  $p_{ij}(k, l)$  for a given scenario, each link that is on the shortest path(s) is weighed by the probability with which it is used by source node  $i$  and destination node  $j$ . Hence, if there is a single shortest path between nodes  $i$  and  $j$ , all links belonging to the shortest path have a weight of 1, whereas the remaining links, which are not part of the shortest path, have a weight of 0. Otherwise, if there are multiple shortest path between a given pair of source and destination nodes, each link belonging to a shortest path is weighed by a factor of 1 over the number of shortest paths. (We do not provide explicit expressions for  $p_{ij}(e)$  and  $p_{ij}(k, l)$  here, but note that for a given scenario either with or without failures, the probabilities can be easily calculated by means of a computer program.)

Now, let  $i^+$  denote the link on the ring subnetwork between node  $i$  and its neighboring node in the clockwise direction

and  $i^-$  denote the link on the ring subnetwork between node  $i$  and its neighboring node in the counterclockwise direction. We then obtain

$$\sigma_t^+(i) = \sum_j p_{ij}(i^+) \cdot \sigma(i, j) \quad (1)$$

$$\sigma_t^-(i) = \sum_j p_{ij}(i^-) \cdot \sigma(i, j) \quad (2)$$

$$\sigma_t^*(i) = \sum_{j,l} p_{ij}(i, l) \cdot \sigma(i, j) \quad (3)$$

$$\sigma_r^+(i) = \sum_{\substack{k,j \\ k \neq i}} p_{kj}(i^+) \cdot \sigma(k, j) \quad (4)$$

$$\sigma_r^-(i) = \sum_{\substack{k,j \\ k \neq i}} p_{kj}(i^-) \cdot \sigma(k, j) \quad (5)$$

$$\sigma_r^*(i) = \sum_{\substack{k,j,l \\ k \neq i}} p_{kj}(i, l) \cdot \sigma(k, j). \quad (6)$$

By using (1), (2), (4), and (5), the traffic loads on the ring subnetwork are calculated as

$$\rho_a^b(i) = \frac{E[T]}{F} \cdot f \cdot \sigma_a^b(i) \quad (7)$$

where  $a \in \{r, t\}$ ,  $b \in \{+, -\}$ , and  $f$  denotes the ratio of the line rate of the star subnetwork and the line rate of the ring subnetwork. Note that, in general, the star subnetwork needs to operate at a higher line rate than the ring subnetwork in order to cope with the proxy-stripped traffic of both fiber rings, i.e.,  $f \geq 1$ .

Given this, we are now able to formulate the stability conditions of the ring subnetwork, the PSC, and the AWG. The stability condition of the ring subnetwork is given by

$$\rho_t^b(i) + \rho_r^b(i) < 1 \quad (8)$$

which needs to hold for every node  $i$ ,  $0 \leq i \leq N - 1$ , and direction  $b \in \{+, -\}$ . Noting that a given ring-and-star homed node  $k$  can send at most one packet per frame over the PSC, the stability condition of the PSC is given by

$$\sigma_t^*(k) + \sigma_r^*(k) < 1 \quad (9)$$

which needs to hold for each ring-and-star homed node  $k$ ,  $k = 1, 2, \dots, D \cdot S$ .

Under the assumption that the stability conditions for the ring subnetwork and PSC hold, we proceed to determine the stability condition of the AWG. Let  $\alpha_{kl}$ ,  $k, l = 1, 2, \dots, D \cdot S$  and  $k \neq l$ , denote the mean number of data packets to be transmitted from (ring-and-star homed) node  $k$  to (ring-and-star homed) node  $l$  on the PSC per frame, which is given by

$$\alpha_{kl} = \sum_{i,j} p_{ij}(k, l) \cdot \sigma(i, j). \quad (10)$$

Note that  $\sum_l \alpha_{kl} = \sigma_t^*(k) + \sigma_r^*(k)$ , which is less than 1 for the assumed stable network by the stability condition of the PSC

[see (9)]. Moreover, note that a given ring-and-star homed node  $k$  can send at most one packet per frame to another ring-and-star homed node  $l$ , hence,  $\alpha_{kl}$  is equivalent to the probability that node  $k$  has a packet to send to  $l$  in a given frame. For simplicity, we assume that all packets sent by two or more nodes to the same receiver across the PSC within the same frame collide and need to be retransmitted (this simplifying assumption still provides quite accurate results, as we will see in Section VI). Assuming independence among the ring-and-star homed nodes, the probability  $P$  that in a given frame a collision occurs at receiver  $l$  is given by

$$P = 1 - \prod_{k=1}^{D \cdot S} (1 - \alpha_{kl}) - \sum_{k=1}^{D \cdot S} \alpha_{kl} \cdot \prod_{\substack{j=1 \\ j \neq k}}^{D \cdot S} (1 - \alpha_{jl}) \quad (11)$$

where  $\alpha_{kk} = 0$  and  $l = 1, 2, \dots, D \cdot S$ . Note that  $\prod_{k=1}^{D \cdot S} (1 - \alpha_{kl})$  is the probability that the (ring-and-star homed) node  $l$  does not receive a data packet. Also note that

$$\sum_{k=1}^{D \cdot S} \alpha_{kl} \cdot \prod_{\substack{j=1 \\ j \neq k}}^{D \cdot S} (1 - \alpha_{jl}) =: r_l \quad (12)$$

is the probability that (ring-and-star homed) node  $l$  receives exactly one data packet, which we denote by  $r_l$ . To see this, note that with probability  $\alpha_{kl}$ , node  $k$  has a packet for node  $l$  in a given frame, and with probability  $\prod_{j=1, j \neq k}^{D \cdot S} (1 - \alpha_{jl})$ , none of the other ring-and-star home nodes  $j$ ,  $j = 1, \dots, D \cdot S$ ,  $j \neq k$ , has a packet for  $l$  in the frame, i.e., the transmission from  $k$  to  $l$  proceeds without collision. Moreover, note that with the approximating assumption that a node can receive at most one packet per frame without collision over the PSC,  $r_l$  is equivalent to the mean number of packets that are transmitted in a given frame without a collision to ring-and-star homed node  $l$ ,  $l = 1, 2, \dots, D \cdot S$ .

Out of the offered load  $\sum_{k=1}^{D \cdot S} \alpha_{kl}$  (in mean number of packets per frame) to ring-and-star homed node  $l$  per frame, the load  $r_l$  is sent across the PSC and the remaining load  $\sum_{k=1}^{D \cdot S} \alpha_{kl} - r_l$  is sent across the AWG. Consequently, we obtain two stability conditions of the AWG. The first stability condition is given by

$$\frac{E[T]}{F} \cdot \left( \sum_{k=1}^{D \cdot S} \alpha_{kl} - r_l \right) < 1 \quad (13)$$

which needs to hold for all  $l$ ,  $l = 1, 2, \dots, D \cdot S$ , and accounts for receiver collisions but does not consider the limited number of available wavelength channels of the AWG. The second stability condition of the AWG takes the limited number of wavelength channels into account and is given by

$$\frac{E[T]}{F} \cdot \sum_{k \in \mathcal{K}_\iota} \sum_{l \in \mathcal{L}_\omega} \left\{ \alpha_{kl} - \alpha_{kl} \cdot \prod_{\substack{j=1 \\ j \neq k}}^{D \cdot S} (1 - \alpha_{jl}) \right\} < R \quad (14)$$

where  $\mathcal{K}_\iota$  and  $\mathcal{L}_\omega$  denote the two subsets of ring-and-star homed nodes that are attached to AWG input port  $\iota$  and

AWG output port  $\omega$ , respectively, with  $\iota, \omega \in \{1, 2, \dots, D\}$ . This condition needs to hold for all AWG input–output port pairs  $\iota, \omega \in \{1, 2, \dots, D\}$ . To understand this second stability condition, note that  $\sum_{k \in \mathcal{K}_\iota} \sum_{l \in \mathcal{L}_\omega} \alpha_{kl}$  is the mean number of packets to be sent by nodes attached to AWG input port  $\iota$  to the nodes attached to AWG output port  $\omega$  per frame, and  $\sum_{k \in \mathcal{K}_\iota} \sum_{l \in \mathcal{L}_\omega} \alpha_{kl} \cdot \prod_{\substack{j=1 \\ j \neq k}}^{D \cdot S} (1 - \alpha_{jl})$  is the mean number of packet that are sent per frame without a collision over the PSC between these considered nodes and, hence, do not require transmission over the AWG.

The network is stable if and only if all four stability conditions (8), (9), (13), and (14) are satisfied. If one or more stability conditions cannot be satisfied, then the network becomes unstable. Thus, for a given traffic load, the network has to be dimensioned such that all four stability conditions are satisfied.

### C. Utilization and Bottleneck

In this section, we briefly describe how in a stable network the channel utilization of the ring and star subnetworks can be found by using (8), (9), (12), and (14), respectively. The utilization of the (data) channel on the ring subnetwork at node  $i$ ,  $0 \leq i \leq N - 1$ , in clockwise and counterclockwise directions is equal to  $\rho_t^+(i) + \rho_r^+(i)$  and  $\rho_t^-(i) + \rho_r^-(i)$ , respectively. In the star subnetwork, the utilization of the control channel  $\lambda_c$  at node  $k$  equals  $\sigma_t^*(k) + \sigma_r^*(k)$ , where  $k = 1, 2, \dots, D \cdot S$ . The utilization of the PSC home data channel  $\lambda_l$  of ring-and-star homed node  $l$  is approximately given by  $E[T] \cdot r_l / F$ , where  $l = 1, 2, \dots, D \cdot S$ . Moreover, the utilization of the  $R$  data channels available between AWG input port  $\iota$  and AWG output port  $\omega$  is approximately given by

$$\frac{E[T] \cdot \sum_{k \in \mathcal{K}_\iota} \sum_{l \in \mathcal{L}_\omega} \alpha_{kl} \cdot \left( 1 - \prod_{\substack{j=1 \\ j \neq k}}^{D \cdot S} (1 - \alpha_{jl}) \right)}{F \cdot R} \quad (15)$$

where  $\mathcal{K}_\iota$  and  $\mathcal{L}_\omega$  denote the two subsets of ring-and-star homed nodes that are attached to AWG input port  $\iota$  and AWG output port  $\omega$ , respectively, with  $\iota, \omega \in \{1, 2, \dots, D\}$ .

Note that the utilization of the various network elements enables the identification of the bottleneck. Clearly, the bottleneck of the network is identical to the network element with the largest utilization.

### D. Delay Analysis

In this section, we analyze the mean delay of the network for Poisson traffic. The mean waiting times in both the transmit queue and transit queue were analyzed in [48] for the case of unidirectional rings. By extending these results to our bidirectional ring subnetwork, we obtain for node  $i$ ,  $0 \leq i \leq N - 1$ , the waiting time (in slots) in the ring transmit queues of both directions approximately as

$$d_t^\pm(i) = \frac{(\rho_r^\pm(i) + \rho_t^\pm(i)) \cdot E[T^2]}{2 \cdot (1 - \rho_r^\pm(i) - \rho_t^\pm(i)) \cdot (1 - \rho_r^\pm(i)) \cdot E[T]} \quad (16)$$

and the waiting time (in slots) in the ring transit queues of both directions approximately as

$$d_r^\pm(i) = \frac{\rho_t^\pm(i) \cdot E[T^2]}{2 \cdot (1 - \rho_r^\pm(i)) \cdot E[T]}. \quad (17)$$

At each ring-and-star homed node, the buffering and sending of data packets (and control packets) across the PSC of the star subnetwork is modelled as an M/D/1 queue, where the service time is equal to one frame. Hence, for ring-and-star homed node  $k$ ,  $k = 1, 2, \dots, D \cdot S$ , the waiting time in the star transmit queue that stores locally generated traffic is approximately given by

$$d_t^*(k) = \frac{\sigma_r^*(k) + \sigma_t^*(k)}{2 \cdot (1 - \sigma_r^*(k) - \sigma_t^*(k)) \cdot (1 - \sigma_r^*(k))} \quad (18)$$

and the waiting time in the star transmit queue that stores proxy-stripped traffic is approximately given by

$$d_r^*(k) = \frac{\sigma_r^*(k) + \sigma_t^*(k)}{2 \cdot (1 - \sigma_r^*(k))}. \quad (19)$$

Note that both  $d_t^*(k)$  and  $d_r^*(k)$  are given in frames.

Data packets that are sent across the PSC of the star subnetwork experience a propagation delay of  $\tau_{\text{PSC}}$ . If a given data packet suffers from a channel collision on the PSC, it will be scheduled for retransmission across the AWG by all ring-and-star homed nodes in a distributed manner. The fraction of traffic  $\beta$  that is sent across the AWG is given by

$$\beta = \frac{\sum_l (\sum_k \alpha_{kl} - r_l)}{\sum_{i,j} \sigma(i,j)}. \quad (20)$$

A given data packet that is sent across the AWG experiences a certain scheduling delay prior to the propagation delay of the AWG  $\tau_{\text{AWG}}$ . We assume that the scheduling delay is significantly smaller than the combined propagation delay  $\tau_{\text{PSC}} + \tau_{\text{AWG}}$  and neglect it in the following. Note that this assumption appears to be reasonable since the very-high-speed star subnetwork operates at a line rate that is by a factor of  $f$  larger than that of the ring subnetwork.

By weighing the different waiting times and propagation delays with the probabilities  $p_{ij}(e)$  and  $p_{ij}(k, l)$ , we obtain the mean delay on the ring subnetwork and the PSC of the star subnetwork for any pair of source node  $i$  and destination node  $j$  as follows. If node  $i$  is a ring homed node, the mean delay  $D_{ij}$  between source node  $i$  and destination node  $j$  in slots is given by

$$\begin{aligned} D_{ij} = & p_{ij}(i^+) \cdot d_t^+(i) + p_{ij}(i^-) \cdot d_t^-(i) + E[T] + \tau \\ & + \sum_{k \neq i, j} [p_{ij}(k^+) \cdot (d_r^+(k) + \tau) + p_{ij}(k^-) \cdot (d_r^-(k) + \tau)] \\ & + \sum_{k, l} p_{ij}(k, l) \cdot (d_r^*(k) \cdot F + \tau_{\text{PSC}}). \end{aligned} \quad (21)$$

If node  $i$  is a ring-and-star homed node, the mean delay  $D_{ij}$  between source node  $i$  and destination node  $j$  in slots is given by

$$\begin{aligned} D_{ij} = & p_{ij}(i^+) \cdot d_t^+(i) + p_{ij}(i^-) \cdot d_t^-(i) + E[T] + \tau \\ & + \sum_{k \neq i, j} [p_{ij}(k^+) \cdot (d_r^+(k) + \tau) + p_{ij}(k^-) \cdot (d_r^-(k) + \tau)] \\ & + \sum_{\substack{k, l \\ k \neq i}} p_{ij}(k, l) \cdot (d_r^*(k) \cdot F + \tau_{\text{PSC}}) \\ & + \sum_l p_{ij}(i, l) \cdot (d_t^*(i) \cdot F + \tau_{\text{PSC}}). \end{aligned} \quad (22)$$

By taking into account the additional delay encountered by traffic sent on the AWG of the star subnetwork, the mean delay of the network  $D$  for a given scenario is given by

$$D = \frac{\sum_{i,j} \sigma(i,j) \cdot D_{ij}}{\sum_{i,j} \sigma(i,j)} + \beta \cdot \tau_{\text{AWG}} \quad (23)$$

where  $\beta$  is given in (20). Note that the mean delay  $D$  in (23) is for a given scenario with certain probabilities  $p_{ij}(e)$  and  $p_{ij}(k, l)$ . The mean delay  $D$  is given in slots.

## VI. RESULTS

In this section, we numerically investigate the performance of the protectoration technique for Poisson and self-similar traffic. Throughout our investigations, we consider uniform unicast traffic that is typically found in metro core networks [6]. (For results of the underlying proxy-stripping technique of protectoration under uniform, hot-spot, symmetric, and asymmetric traffic, the interested reader is referred to [44].) More precisely, at node  $i$ , the average number of locally generated packets that are destined for node  $j$  per frame equals  $0 \leq \sigma \leq 1$ , if  $i \neq j$ , and  $\sigma = 0$ , if  $i = j$ , where  $i, j = 1, 2, \dots, N$ . The parameters are set to the following default values:  $D = 8$ ;  $S = 1$ ;  $R = 1$ ; and  $F = 400$ . The length  $L$  of (data) packets is uniformly distributed over the interval of  $[1, F - D \cdot S] = [1, 392]$  slots, where one slot is 4 B (octets) long (we consider 4 B sufficient for accommodating destination address, length, and priority fields in a control packet). The ring operates at a line rate of 2.5 Gb/s and has a circumference of 100 km. Considering cut-through forwarding on the ring subnetwork, the RTT of the ring subnetwork is given by  $\text{RTT} = 100 \text{ km} / (2 \cdot 10^5 \text{ km/s})$ . For the star subnetwork, we set  $\tau_{\text{AWG}} = \tau_{\text{PSC}} = (100 \text{ km}/\pi) / (2 \cdot 10^5 \text{ km/s})$ . Thus, the RTT of the ring subnetwork is assumed to be  $\pi$  times as large as the one-way end-to-end propagation delay of the star subnetwork.

### A. Poisson Traffic

To verify the accuracy of our analysis, we provide additional simulation results. As opposed to the analysis, in our simulations, we also account for the access delay on the PSC control channel and the scheduling delay on the AWG data channel of the star subnetwork. In each simulation, we have generated

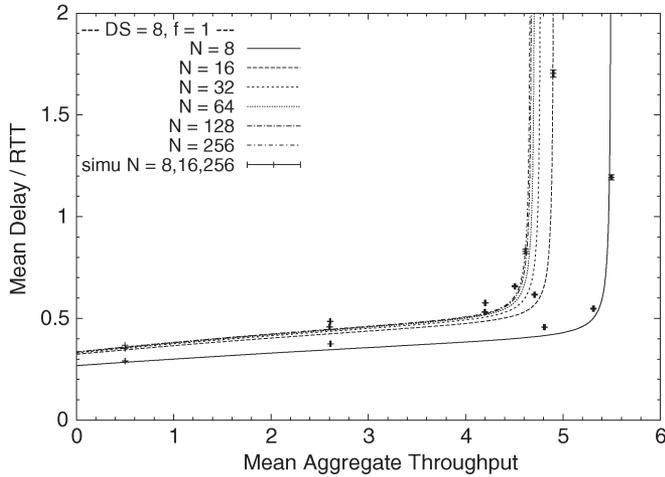


Fig. 10. Mean delay versus mean aggregate throughput with  $N_{rs} = D \cdot S = 8$  ( $D=8, S=1$ ) and  $f=1$  for different  $N \in \{8, 16, 32, 64, 128, 256\}$ .

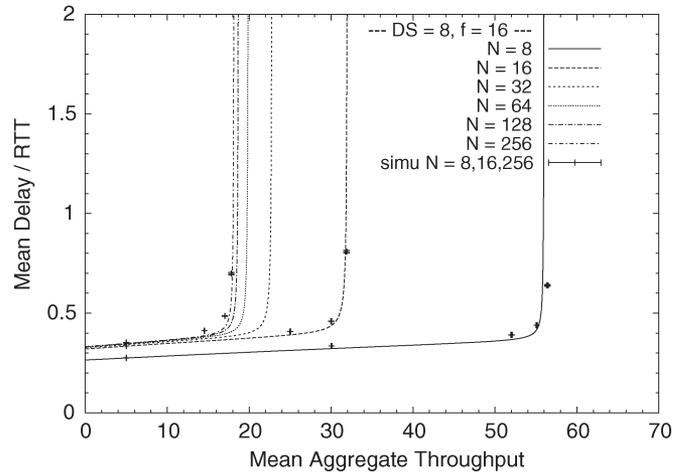


Fig. 12. Mean delay versus mean aggregate throughput with  $N_{rs} = D \cdot S = 8$  ( $D=8, S=1$ ) and  $f=16$  for different  $N \in \{8, 16, 32, 64, 128, 256\}$ .

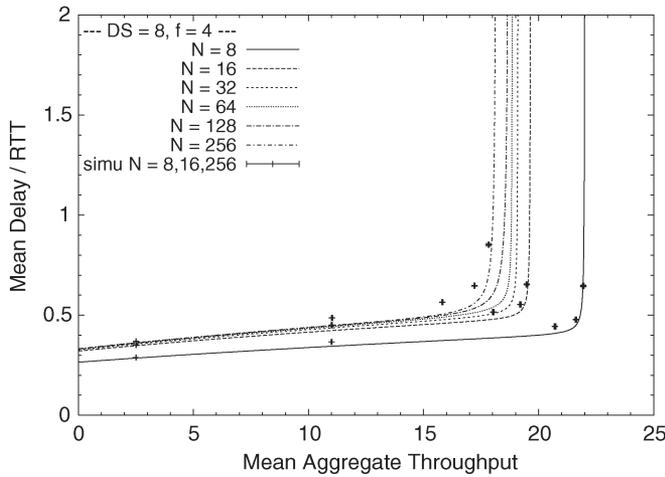


Fig. 11. Mean delay versus mean aggregate throughput with  $N_{rs} = D \cdot S = 8$  ( $D=8, S=1$ ) and  $f=4$  for different  $N \in \{8, 16, 32, 64, 128, 256\}$ .

$10^6$  packets including a warm-up phase of  $10^5$  packets. Using the method of batch means, we calculated the 95% confidence intervals for both mean delay and mean aggregate throughput. The mean delay is given in RTT of the ring subnetwork and the mean aggregate throughput is equal to the mean number of transmitting nodes in steady state.

1) *Operation Without Failures:* Let us first consider the network operating without failures. Figs. 10–12 depict the mean delay versus mean aggregate throughput for a fixed number of  $N_{rs} = D \cdot S = 8$  ring-and-star homed nodes and different numbers of nodes  $N \in \{8, 16, 32, 64, 128, 256\}$ . We consider different speed-up factors  $f$  from  $f = 1$ , i.e., both star and ring subnetworks operate at the same line rate, in Fig. 10 to  $f = 16$ , i.e., the line rate of the star subnetwork is 16 times as large as the line rate of the ring subnetwork and, thus, equals 40 Gb/s, in Fig. 12. The individual curves are obtained by increasing the packet generation probability  $\sigma$  from values close to zero to values that result in very large delays. Focusing for now on Fig. 10, we observe that for  $N = N_{rs} = 8$ , we obtain the lowest delay and largest throughput. For  $N = N_{rs}$ , all nodes are attached to the star subnetwork and can communicate with

each other in a single hop. In particular, each node sends traffic to its two adjacent nodes via the ring subnetwork and to the other nodes via the star subnetwork. At small traffic loads, no significant queuing occurs and the mean delay is mainly dictated by the propagation delay encountered on both the ring and star subnetworks. Due to the fact that the propagation delay between two adjacent ring nodes ( $RTT/8$ ) is smaller than the propagation delay of the star subnetwork ( $RTT/\pi$ ), the mean delay is smaller than  $RTT/\pi$  at small traffic loads. For  $N > 8$ , only a subset of nodes is attached to the star subnetwork and packets need to increasingly traverse multiple nodes on their way from source node to destination node. At small traffic loads, the mean delay for  $N > 8$  is bounded by the propagation delay from the source node to the closest ring-and-star homed node, which does not exceed  $RTT/16$ , plus the propagation delay of the star subnetwork ( $RTT/\pi$ ), plus the propagation delay from the destination node to the closest ring-and-star homed node, which does not exceed  $RTT/16$ . As shown in Fig. 10, for all values of  $N$ , the mean delay and mean aggregate throughput increase with increasing traffic loads (packet generation probabilities  $\sigma$ ). For  $N \in \{8, 16, 256\}$ , we provide verifying simulation results. Analytical and simulation results match quite well. At small traffic loads, the simulation provides a slightly larger mean delay than the analysis. This is because the simulation accounts for the additional access delay on the PSC control channel and the scheduling delay on the AWG data channel of the star subnetwork, as opposed to the analysis. We also observe from Fig. 10 that the maximum mean aggregate throughput slightly decreases with increasing number of nodes  $N$ . This is because with an increasing number of nodes  $N$  and a fixed number  $N_{rs}$  of ring-and-star homed nodes, each ring-and-star homed node collects shortcut traffic from an increasing number of ring-homed nodes. This results in increased loads on the ring segments connecting the ring-homed nodes to the ring-and-star homed nodes and the star subnetwork, which, in turn, makes these ring segments and the star subnetworks the bottlenecks in the network.

Comparing Figs. 10–12, we observe that increasing the speed-up factor of the star subnetwork to  $f = 4$  significantly

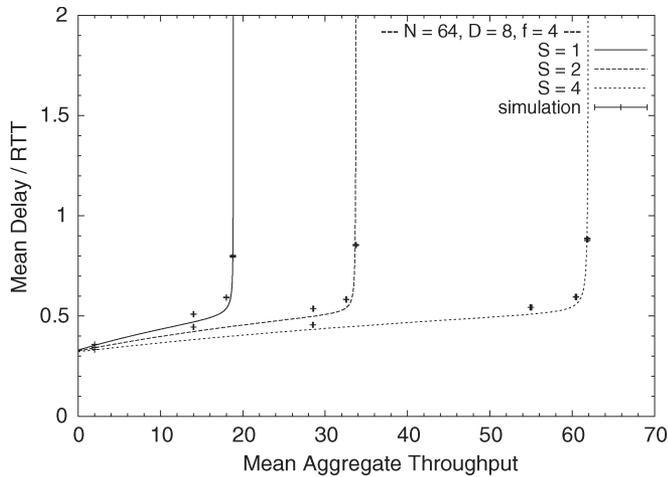


Fig. 13. Mean delay versus mean aggregate throughput with  $N = 64$ ,  $D = 8$ , and  $f = 4$  for different  $S \in \{1, 2, 4\}$ .

increases the maximum mean aggregate throughput for the entire range of number of nodes  $N$ . On the other hand, increasing the speed-up factor further to  $f = 16$  increases the throughput levels significantly for a small number of nodes  $N = 8$  and 16, while for a larger number of nodes  $N \geq 32$ , there is only a minor increase in the throughput. The explanation for these dynamics is as follows. For  $f = 1$ , the star subnetwork is the primary bottleneck in the network, which is relieved by increasing the speed-up factor to  $f = 4$ . As the speed-up factor is further increased to  $f = 16$ , the ring segments connecting the ring homed nodes to the ring-and-star homed nodes become the primary bottleneck, especially for an increasing number of ring-homed nodes between ring-and-star homed nodes. This bottleneck prevents the shortcut traffic from reaching the star subnetwork.

To capitalize on the capacity of the star subnetwork, the bottleneck on the ring subnetwork has to be mitigated. This may be done by reducing the amount of collected shortcut traffic at each ring-and-star homed node. Clearly, this can be achieved by increasing the number of ring-and-star homed nodes  $N_{rs} = D \cdot S$  for a given number of nodes  $N$ . Fig. 13 depicts the mean delay versus mean aggregate throughput with  $N = 64$ ,  $D = 8$ , and  $f = 4$  for different  $S \in \{1, 2, 4\}$ . By connecting more nodes to the star subnetwork, the congestion on the ring subnetwork at each ring-and-star homed node is alleviated and the star subnetwork can be utilized to a larger extent, resulting in a dramatically improved throughput-delay performance at the expense of connecting a larger number of nodes to the star subnetwork.

Note that there exist additional approaches to mitigate the bottlenecks on the star and ring subnetworks and to improve the throughput-delay performance of the network. For instance, the capacity of the PSC control channel could be increased by assigning additional control slots to ring-and-star homed nodes during the last  $(F - D \cdot S)$  slots of each frame on the control channel. The capacity of the ring subnetwork could be increased by operating more than one wavelength in either direction by means of WDM. These modifications are left for future work.

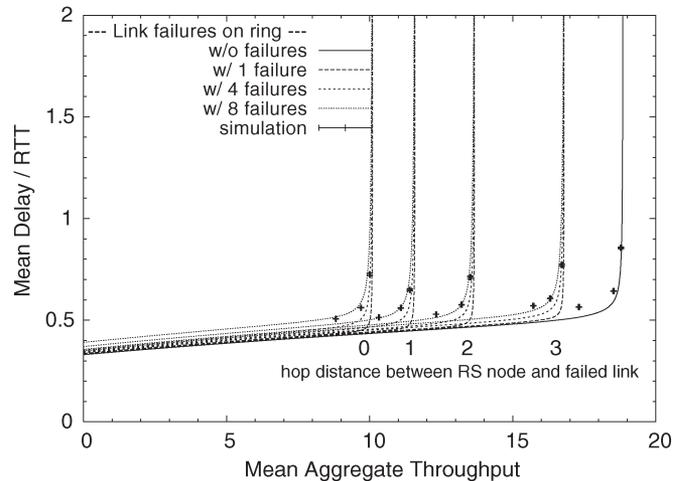


Fig. 14. Mean delay versus mean aggregate throughput for link failures with different locations on the ring subnetwork ( $N = 64$ ,  $D = 8$ ,  $S = 1$ , and  $f = 4$ ).

2) *Operation With Ring Failures:* After gaining some insight into the failure-free operation of the network, we now investigate the protectoration technique in the presence of various link and node failures. Besides a single failure, we consider also multiple failures in both ring and star subnetworks. To assess their impact on the network operation, we compare throughout our investigations the network performance of the various failure scenarios with that of the failure-free scenario. The locations of the failures are chosen as follows. Starting with a single (link or node) failure, the second failure is located at the opposite side of the central hub, i.e., the first failure is mirrored at the hub in order to obtain the location of the second failure. The third failure is placed in the middle of the first and second failures. The location of the fourth failure is found by mirroring the third failure at the hub. This procedure is repeated incrementally until all multiple failures are placed. In our numerical investigations, we focus on failure scenarios that do not split the network into several disjoint subnetworks, i.e., full connectivity is not affected by the various failures. Recall from above, to guarantee full connectivity, no more than one link or one node failure must occur in each ring segment between two neighboring ring-and-star homed nodes. In the following, we set  $N = 64$ ,  $D = 8$ ,  $S = 1$ , and  $f = 4$ , with the other parameters set to their default values.

Let us start with link and node failures on the ring subnetwork while the star subnetwork is completely intact. Fig. 14 depicts the impact of ring link failures and their location on the throughput-delay performance of the network, including verifying simulation results for the two scenarios without failure and with four failures. We consider four different locations of link failures. More precisely, the link failure(s) is (are) zero, one, two, or three hops away from the corresponding next ring-and-star homed node, where one hop denotes the distance between two adjacent nodes on the ring. We observe that the simulation gives a slightly larger mean delay than the analysis at small traffic loads, while at medium to high traffic loads, the results of simulation and analysis match very well. This is due to the fact that the simulation takes the access and scheduling delays

of the star subnetwork into account, as opposed to the analysis. The access and scheduling delays result in a larger delay at small traffic loads, but both become negligible compared to the queueing delays encountered at the various transmit and transit queues as the traffic load increases. Interestingly, Fig. 14 shows that the location of the link failures has a significantly larger impact on the throughput-delay performance of the network than the number of link failures. We observe that for a given failure location, the protection technique is able to make the network very resilient against multiple link failures on the ring subnetwork such that the throughput-delay performance is only slightly deteriorated with an increasing number of link failures. The performance loss, however, strongly depends on the location of the link failure(s). Apparently, link failures that are closer to ring-and-star homed nodes have a significantly more detrimental impact on the throughput-delay performance than link failures that are farther away in terms of hops. This is because of two main effects. First, a link closer to a given ring-and-star homed node carries the traffic of more ring homed nodes that is sent to the ring-and-star homed node for being proxy stripped. If this link fails, more ring homed nodes are affected and need to steer the traffic in the opposite direction towards the neighboring ring-and-star homed node. Second, with a link failure close to a given ring-and-star homed node, the opposite direction towards the neighboring ring-and-star homed node is larger compared to link failures that occur in the middle of two neighboring ring-and-star homed nodes. As a result, with a link failure close to a given ring-and-star homed node, wrapped and steered traffic traverses more intermediate ring homed nodes on the backup path. Both effects—more affected nodes and longer backup paths—lead to an increased congestion on the ring segment and, thus, an increased mean delay and a decreased mean aggregate throughput. Note that this performance loss can be alleviated by increasing the number of ring-and-star homed nodes (see Fig. 13). In doing so, for a given  $N$ , each ring segment between two adjacent ring-and-star homed nodes contains fewer ring homed nodes, resulting in a decreased number of affected nodes and a decreased backup path length.

Next, we consider node failures. To guarantee full connectivity among the remaining functional nodes, no more than one node failure must occur between each pair of adjacent ring-and-star homed nodes. Note that the location of failed ring homed nodes has a similar impact on the throughput-delay performance of the network as the location of failed ring links, which have been examined above. In the following, we concentrate on ring-and-star homed node failures. More precisely, to maintain full connectivity among all nodes, a failed ring-and-star homed node is assumed to be unable to communicate via the ring subnetwork while the transmission and reception via the star subnetwork remains fully intact (below, we will investigate the complementary case where ring-and-star homed nodes are able to transmit and receive only via the ring subnetwork while they are disconnected from the star subnetwork). Fig. 15 depicts the mean delay versus mean aggregate throughput with and without ring-and-star homed node failures on the ring subnetwork. We observe that the maximum achievable mean aggregate throughput decreases significantly

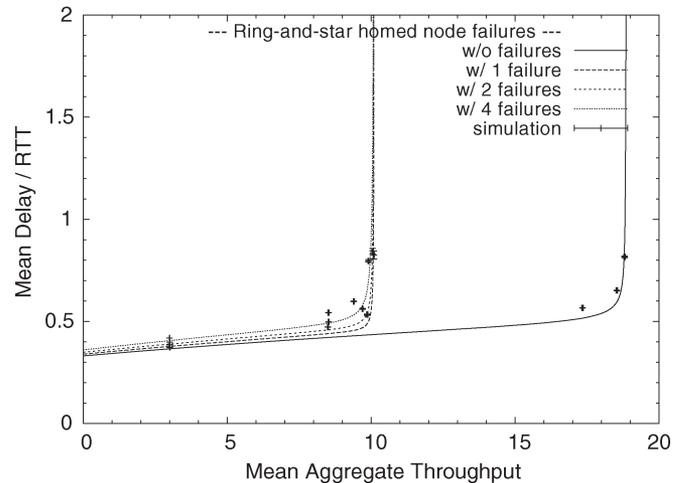


Fig. 15. Mean delay versus mean aggregate throughput for ring-and-star homed node failures on the ring subnetwork ( $N = 64$ ,  $D = 8$ ,  $S = 1$ , and  $f = 4$ ).

if one ring-and-star homed node fails, but it does not decrease any further in the presence of additional ring-and-star homed node failures. Comparing Figs. 14 and 15, we observe that the maximum achievable mean aggregate throughput with failed ring-and-star homed nodes is only very slightly smaller than the maximum mean aggregate throughput with link failures zero hops away from ring-and-star homed nodes. In both cases, ring homed nodes next to a failed node or failed link have to wrap and steer traffic towards the neighboring ring-and-star homed node; with a failed link, this wrapping and steering takes place on one side of the ring-and-star homed node, with a failed ring-and-star homed node on both sides of the node. As discussed above, the protected traffic thereby experiences increased queueing delays at the intermediate ring homed nodes due to congestion. Recall that for a given  $N$ , the congestion could be alleviated by limiting the number of ring homed nodes in each ring segment by increasing the number of ring-and-star homed nodes. Consequently, the detrimental impact of the ring subnetwork on the throughput-delay performance of protected traffic is mitigated, at the expense of more nodes being attached to the star subnetwork.

3) *Operation With Star Failures:* After investigating link and node failures on the ring subnetwork, we now turn our attention to failures on the star subnetwork. We consider failure scenarios where a single, multiple, and all ring-and-star homed nodes are disconnected from the star subnetwork due to link or component (splitter, combiner, amplifier, (de)partitioner, AWG, PSC) failures. In the following, we focus on failed links that connect the ring-and-star homed nodes to the star subnetwork, but note that component failures have a similar impact on the ring-and-star homed nodes in terms of connectivity and throughput-delay performance. As shown in Fig. 16, with one ring-and-star homed node disconnected from the star subnetwork, the maximum achievable mean aggregate throughput decreases significantly. While the throughput-delay performance remains rather unchanged for four link failures, five disconnected ring-and-star homed nodes further decrease the maximum achievable mean aggregate throughput significantly.

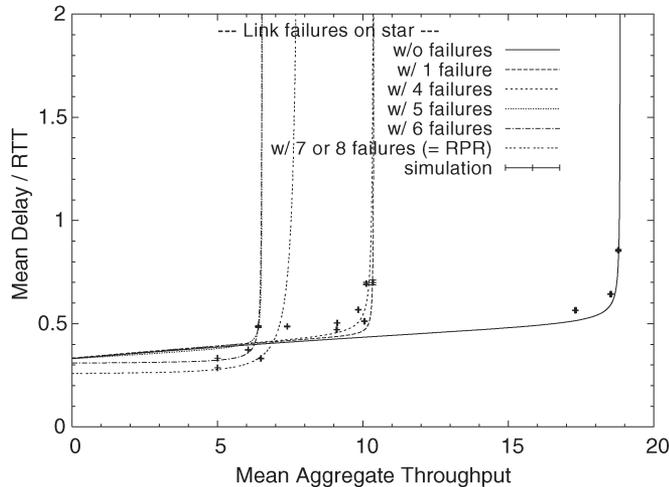


Fig. 16. Mean delay versus mean aggregate throughput for link failures on the star subnetwork ( $N = 64$ ,  $D = 8$ ,  $S = 1$ , and  $f = 4$ ).

Interestingly, with six link failures, i.e., only one pair of ring-and-star homed nodes is interconnected via the star subnetwork, the mean delay at light traffic loads is smaller than in other scenarios with fewer link failures. Note that the mean delay is further decreased with a slightly increased maximum achievable mean aggregate throughput if seven or eight links fail, i.e., no traffic is sent across the star subnetwork at all. Without any star subnetwork, the network is reduced to a conventional bidirectional RPR ring. Thus, we observe that an RPR ring achieves a slightly better throughput-delay performance than our hybrid ring–star network if the number of link failures on the star subnetwork is very large. This is mainly due to the fact that with fewer nodes connected to the star subnetwork, the ring subnetwork gets more congested towards the proxy-stripping ring-and-star homed hot-spot nodes, resulting in an increased delay and a decreased throughput.

Note, however, that our hybrid ring–star network without any failures is able to achieve a dramatically larger maximum mean aggregate throughput (close to 2.5 times larger) than RPR, as depicted in Fig. 16. Furthermore, as shown in Figs. 14 and 15, even in the presence of multiple link and/or node failures on the ring subnetwork, our hybrid ring–star network not only outperforms RPR in terms of maximum achievable mean aggregate throughput but also guarantees full connectivity among all nodes, as opposed to RPR whose connectivity is entirely lost if more than a single link or node fails.

### B. Self-Similar Traffic

In this section, we examine the performance of the protection network for self-similar traffic and compare it with the performance for Poisson traffic using simulations. In addition, we consider the network operation with more realistic finite buffers in this section, in contrast to the infinite buffers considered in the preceding section. More specifically, we consider the same network parameter settings as in the investigation of the failures in the star subnetwork in the preceding section, i.e.,  $N = 64$  nodes,  $D = 8$  AWG and PSC ports,  $S = 1$  combiner input port/splitter output port, star subnetwork speed-up

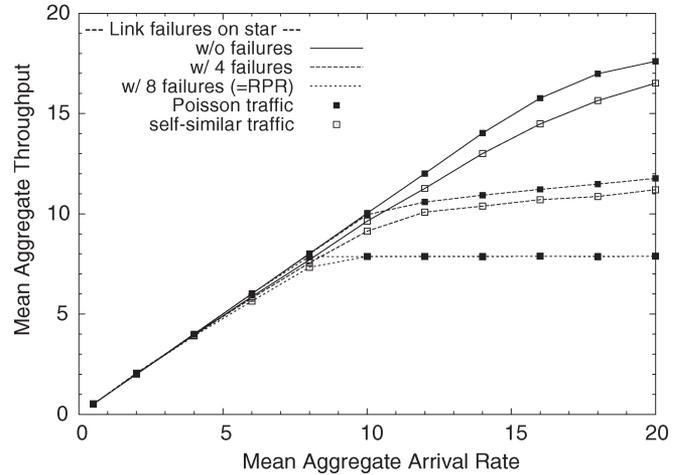


Fig. 17. Mean aggregate throughput versus mean aggregate arrival rate with  $N = 64$ ,  $D = 8$ ,  $S = 1$ , and  $f = 4$  for Poisson and self-similar traffic without and with link failures on the star subnetwork.

factor  $f = 4$ , and all other parameters at their default values. At each node, we generate self-similar packet traffic with Hurst parameter 0.75 for each of the  $N - 1$  destination nodes by aggregating ON/OFF processes with Pareto distributed on-duration and exponentially distributed off-duration [49]. We set the capacity of all buffers to 96 kB. With this choice the transmit buffer for the star subnetwork has approximately the capacity required to hold the packets that can be transmitted within the bandwidth-propagation delay product of the star subnetwork over the PSC. (To see this, note that the bandwidth-propagation delay product is  $4 \cdot 2.5 \text{ Gb/s} \cdot \tau_{\text{PSC}}$ , which corresponds to 124.3 frames of  $400 \cdot 4 \text{ B}$  each. Noting that at most one packet can be transmitted per frame over the PSC and the average packet size is approximately half a frame, we set the buffer capacity to be equivalent to 60 frames, i.e.,  $60 \cdot 400 \cdot 4 = 96 \text{ kB}$ .)

In Figs. 17–19, we plot the mean aggregate throughput, the relative packet loss, and the mean delay as functions of the mean aggregate arrival rate, which is given in the same units as the mean aggregate throughput. The 95% confidence intervals for the Poisson traffic simulations are generally too small to be seen in the plots, except for a few intervals for small loss probabilities in Fig. 18. We observe from Figs. 17 and 18 that for self-similar traffic, the packet loss is generally larger and the throughput smaller than for Poisson traffic, as is to be expected for the more bursty self-similar traffic. These differences become less pronounced as the network saturates, observe, for instance, that for the network with eight failures, the differences disappear when the arrival rate exceeds ten packet generations in steady state or equivalently  $10 \cdot 2.5 \text{ Gb/s}$ . This is because in the saturated network, the buffers tend to be constantly filled to capacity. We observe from Fig. 19 that the self-similar traffic experiences slightly larger mean delays than the Poisson traffic in scenarios where the network is relatively lightly loaded, i.e., in the network with eight (four) failures for a mean arrival rate smaller than approximately five (eleven), and for the network without failures for the entire considered range of arrival rates. On the other

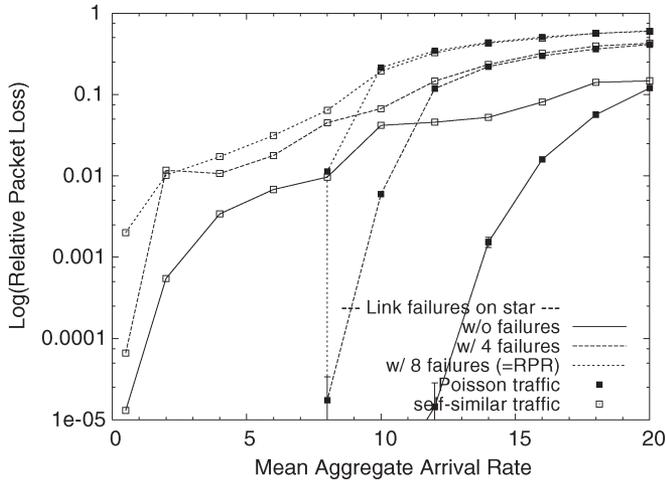


Fig. 18. Relative packet loss versus mean aggregate arrival rate with  $N = 64$ ,  $D = 8$ ,  $S = 1$ , and  $f = 4$  for Poisson and self-similar traffic without and with link failures on the star subnetwork.

hand, the self-similar traffic experiences somewhat smaller mean delays than the Poisson traffic when the network is relatively heavily loaded, i.e., in the network with eight (four) failures for arrival rates larger than approximately five (eleven). This is because with light traffic loads, the buffer occupancy levels tend to be fairly low, thus, bursts of generated packets can typically be held in the buffers and experience larger delays. With heavy loads, on the other hand, the buffers tend to be constantly filled to capacity, especially with the relatively smooth Poisson traffic arrivals. The bursty self-similar traffic arrivals, on the other hand, tend to result in occasional “dips” in the buffer occupancy levels and, consequently, somewhat smaller mean delays. Overall, we observe that the differences in the throughput-delay performance between Poisson and self-similar traffic are relatively small. In addition, from comparison with the analytical and simulation results for the scenario with Poisson traffic and infinite buffers in Fig. 16, we observe that the analysis for the infinite buffer and Poisson traffic scenario predicts the principal behavior of the network for finite buffers and self-similar traffic with reasonable accuracy.

## VII. CONCLUSION

We have developed and evaluated the protection fault recovery technique for resilient packet ring (RPR) networks. The protection technique augments the RPR ring network by a star subnetwork that interconnects a subset of the network nodes using dark fiber and a central hub consisting of an arrayed-waveguide grating (AWG) in parallel with a passive star coupler (PSC). For the fast and efficient recovery from failures, the protection technique uniquely combines the wrapping and steering methods of the conventional RPR ring network to exploit their respective strengths (fast recovery with wrapping, bandwidth efficiency with steering) to achieve a fast and bandwidth-efficient recovery. In contrast to the conventional RPR network, which can recover to full connectivity only after a single failure, our protection technique maintains full network connectivity for multiple failures.

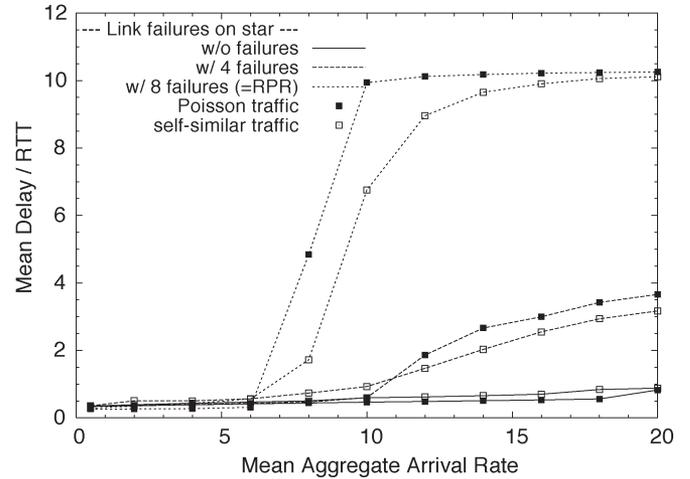


Fig. 19. Mean delay versus mean aggregate arrival rate with  $N = 64$ ,  $D = 8$ ,  $S = 1$ , and  $f = 4$  for Poisson and self-similar traffic without and with link failures on the star subnetwork.

Our analytical and simulation results demonstrate that the protection network without any failures achieves a significantly higher throughput-delay performance than the failure-free RPR network. For a large number of failures on the star subnetwork, the throughput-delay performance of the protection network degenerates to the performance of the conventional RPR network. The impact of the failures on the ring subnetwork on the throughput-delay performance depends largely on the position of the failure (i.e., the distance from the ring-and-star homed node where the ring subnetwork is connected to the star subnetwork) and is largely independent from the number of failures.

We note that one important design goal of RPR was to keep the node architecture and forwarding mechanisms simple. Clearly, the RPR resilience and throughput-delay improvements of our proposed protection technique are achieved at the expense of additional hardware and modified routing. Note, however, that the additional star subnetwork consists only of passive off-the-shelf optical components, except for the optional amplifiers. Also note that only a subset of nodes needs to be WDM upgraded, which can be done in a pay-as-you-grow manner according to given traffic demands and cost constraints. With respect to the increased complexity of protection, we believe that the required modifications of the basic RPR protocols and mechanisms are minor, as outlined in Section IV-C. Protection could be simplified by making the underlying proxy-stripping technique transparent to all ring nodes except the proxy-stripping ring-and-star homed nodes. With transparent proxy stripping, ring-homed nodes are not aware of the presence and location of proxy-stripping nodes. This allows all ring-homed nodes to deploy the same (unmodified) shortest-path routing as in conventional bidirectional RPR networks, at the expense of a somewhat increased mean hop distance and a decreased throughput-delay performance compared to the protection with proxy-stripping-aware ring-homed nodes considered in this paper.

The presented protection technique is not restricted to RPR networks, but it can also be used in other ring networks,

e.g., empty-slot or token rings. There are several exciting avenues for future work on the protectoration technique. One direction is to examine techniques for overcoming the performance limiting bottlenecks, e.g., assigning a larger number of control slots, or operating multiple wavelengths on the ring subnetwork, as identified in this study. Another direction is to examine the transmission of more general traffic patterns, e.g., broadcast and multicast traffic, over the protectoration network.

## REFERENCES

- [1] IEEE, *Resilient Packet Ring*, IEEE Standard 802.17, Sep. 2004.
- [2] F. Davik, M. Yilmaz, S. Gjessing, and N. Uzun, "IEEE 802.17 resilient packet ring tutorial," *IEEE Commun. Mag.*, vol. 42, no. 3, pp. 112–118, Mar. 2004.
- [3] P. Yuan, V. Gambiroza, and E. Knightly, "The IEEE 802.17 media access protocol for high-speed metropolitan-area resilient packet rings," *IEEE Network*, vol. 18, no. 3, pp. 8–15, May/June 2004.
- [4] G. Suwala and G. Swallow, "SONET/SDH-like resilience for IP networks: A survey of traffic protection mechanisms," *IEEE Network*, vol. 18, no. 2, pp. 20–25, Mar./Apr. 2004.
- [5] S. Spadaro, J. Solé-Pareta, D. Careglio, K. Wajda, and A. Szymański, "Positioning of the RPR standard in contemporary operator environments," *IEEE Network*, vol. 18, no. 2, pp. 35–40, Mar./Apr. 2004.
- [6] N. Ghani, J.-Y. Pan, and X. Cheng, "Metropolitan optical networks," in *Optical Fiber Telecommunications*, vol. IVB. San Diego, CA: Academic, 2002, pp. 329–403.
- [7] A. A. M. Saleh and J. M. Simmons, "Architectural principles of optical regional and metropolitan access networks," *J. Lightw. Technol.*, vol. 17, no. 12, pp. 2431–2448, Dec. 1999.
- [8] O. Gerstel and R. Ramaswami, "Optical layer survivability: A post-bubble perspective," *IEEE Commun. Mag.*, vol. 41, no. 9, pp. 51–53, Sep. 2003.
- [9] T.-H. Wu, "Emerging technologies for fiber network survivability," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 58–74, Feb. 1995.
- [10] D. Zhou and S. Subramaniam, "Survivability in optical networks," *IEEE Network*, vol. 14, no. 6, pp. 16–23, Nov./Dec. 2000.
- [11] J. Zhang and B. Mukherjee, "A review of fault management in WDM mesh networks: Basic concepts and research challenges," *IEEE Network*, vol. 18, no. 2, pp. 41–48, Mar./Apr. 2004.
- [12] A. Carena, V. DeFeo, J. Finochietto, R. Gaudio, F. Neri, C. Pignone, and P. Poggiolini, "RingO: An experimental WDM optical packet network for metro applications," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 8, pp. 1561–1571, Oct. 2004.
- [13] P.-H. Ho and H. Mouftah, "Shared protection in mesh WDM networks," *IEEE Commun. Mag.*, vol. 42, no. 1, pp. 70–76, Jan. 2004.
- [14] P.-H. Ho, J. Tapolcai, and H. T. Mouftah, "On achieving optimal survivable routing for shared protection in survivable next-generation Internet," *IEEE Trans. Reliab.*, vol. 53, no. 2, pp. 216–225, Jun. 2004.
- [15] G. Mohan and C. S. R. Murthy, "Lightpath restoration in WDM optical networks," *IEEE Network*, vol. 14, no. 6, pp. 24–32, Nov./Dec. 2000.
- [16] A. Narula-Tam, E. Modiano, and A. Brzezinski, "Physical topology design for survivable routing of logical rings in WDM-based networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 8, pp. 1525–1538, Oct. 2004.
- [17] C. Ou, J. Zhang, H. Zang, L. Sahasrabudde, and B. Mukherjee, "New and improved approaches for shared-path protection in WDM mesh networks," *J. Lightw. Technol.*, vol. 22, no. 5, pp. 1223–1232, May 2004.
- [18] S. Ramamurthy, L. Sahasrabudde, and B. Mukherjee, "Survivable WDM mesh networks," *J. Lightw. Technol.*, vol. 21, no. 4, pp. 870–883, Apr. 2003.
- [19] M. Sridharan, M. V. Salapaka, and A. K. Somani, "A practical approach to operating survivable WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 1, pp. 202–215, Jan. 2002.
- [20] D. Xu, Y. Xiong, C. Qiao, and G. Li, "Failure protection in layered networks with shared risk link groups," *IEEE Network*, vol. 18, no. 3, pp. 36–41, May/June 2004.
- [21] J. Wang, L. Sahasrabudde, and B. Mukherjee, "Path vs. subpath vs. link restoration for fault management in IP-over-WDM networks: Performance comparisons using GMPLS control signaling," *IEEE Commun. Mag.*, vol. 40, no. 11, pp. 80–87, Nov. 2002.
- [22] Q. Zheng and G. Mohan, "Protection approaches for dynamic traffic in IP/MPLS-over-WDM networks," *IEEE Commun. Mag.*, vol. 41, no. 5, pp. S24–S29, May 2003.
- [23] O. Gerstel and G. Sasaki, "Quality of Protection (QoP): A quantitative unifying paradigm to protection service grades," *Opt. Netw. Mag.*, vol. 3, no. 3, pp. 40–50, May/June 2002.
- [24] C. Saradhi, M. Gurusamy, and L. Zhou, "Differentiated QoS for survivable WDM optical networks," *IEEE Commun. Mag.*, vol. 42, no. 5, pp. S8–S14, May 2004.
- [25] M. Tacca, A. Fumagalli, A. Paradisi, F. Unghvary, K. Gadhiraaju, S. Lakshmanan, S. M. Rossi, A. de Campos Sachs, and D. S. Shah, "Differentiated reliability in optical networks: Theoretical and practical results," *J. Lightw. Technol.*, vol. 21, no. 11, pp. 2576–2586, Nov. 2003.
- [26] K. Wu, L. Valcarenghi, and A. Fumagalli, "Restoration schemes with differentiated reliability," in *Proc. IEEE Int. Conf. Communications (ICC)*, Anchorage, AK, 2003, pp. 1968–1972.
- [27] O. Gerstel and R. Ramaswami, "Optical layer survivability: A services perspective," *IEEE Commun. Mag.*, vol. 38, no. 3, pp. 104–113, Mar. 2000.
- [28] —, "Optical layer survivability—An implementation perspective," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 10, pp. 1885–1899, Oct. 2000.
- [29] Y. Ye, S. Dixit, and M. Ali, "On joint protection/restoration in IP-centric DWDM-based optical transport networks," *IEEE Commun. Mag.*, vol. 38, no. 6, pp. 174–183, Jun. 2000.
- [30] A. Fumagalli and L. Valcarenghi, "IP restoration vs. WDM protection: Is there an optimal choice?" *IEEE Network*, vol. 14, no. 6, pp. 34–41, Nov./Dec. 2000.
- [31] W. Aiello, S. N. Bhatt, F. R. K. Chung, A. L. Rosenberg, and R. K. Sitaraman, "Augmented ring networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 12, no. 6, pp. 598–609, Jun. 2001.
- [32] W. D. Grover and D. Stamatelakis, "Cycle-oriented distributed pre-configuration: Ring-like speed with mesh-like capacity for self-planning network restoration," in *Proc. IEEE Int. Conf. Communications (ICC)*, Atlanta, GA, 1998, pp. 537–543.
- [33] G. Shen and W. D. Grover, "Extending the *p*-cycle concept to path segment protection for span and node failure recovery," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 8, pp. 1306–1319, Oct. 2003.
- [34] T. Y. Chow, F. Chudak, and A. M. Ffrench, "Fast optical layer mesh protection using pre-cross-connected trails," *IEEE/ACM Trans. Netw.*, vol. 12, no. 3, pp. 539–548, Jun. 2004.
- [35] P. Yue, Z. Liu, and J. Liu, "High performance fair bandwidth allocation algorithm for resilient packet ring," in *Proc. Int. Conf. Advanced Information Networking and Applications (AINA)*, Xian, China, Mar. 2003, pp. 415–420.
- [36] X. Zhou, G. Shi, H. Fang, and L. Zeng, "Fairness algorithm analysis in resilient packet ring," in *Proc. Int. Conf. Communication Technology (ICCT)*, Beijing, China, Apr. 2003, vol. 1, pp. 622–624.
- [37] V. Gambiroza, P. Yuan, L. Balzano, Y. Liu, S. Sheafor, and E. Knightly, "Design, analysis, and implementation of DVSR: A fair, high performance protocol for packet rings," *IEEE/ACM Trans. Netw.*, vol. 12, no. 1, pp. 85–102, Feb. 2004.
- [38] I. Rubin and J. Ling, "All-optical cross-connect meshed-ring communications networks using a reduced number of wavelengths," in *Proc. IEEE Information Communications (INFOCOM)*, New York, 1999, vol. 2, pp. 924–931.
- [39] A. M. Hill, M. Brierley, R. M. Percival, R. Wyatt, D. Pitcher, K. M. I. Pati, I. Hall, and J.-P. Laude, "Multiple-star wavelength-router network and its protection strategy," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 7, pp. 1134–1145, Sep. 1998.
- [40] M. Herzog, M. Maier, and A. Wolisz, "RINGOSTAR: An evolutionary AWG-based WDM upgrade of optical ring networks," *J. Lightw. Technol.*, vol. 23, no. 4, pp. 1637–1651, Apr. 2005.
- [41] C. Fan, M. Maier, and M. Reisslein, "The AWG||PSC network: A performance-enhanced single-hop WDM network with heterogeneous protection," *J. Lightw. Technol.*, vol. 22, no. 5, pp. 1242–1262, May 2004.
- [42] O. A. Lavrova, G. Rossi, and D. J. Blumenthal, "Rapid tunable transmitter with large number of ITU channels accessible in less than 5 ns," in *Proc. Eur. Conf. Optical Communication (ECOC)*, Munich, Germany, Sep. 2000, pp. 169–170.
- [43] K. Shrikhande, I. M. White, M. S. Rogge, F.-T. An, A. Srivatsa, E. S. Hu, S. S.-H. Yam, and L. G. Kazovsky, "Performance demonstration of a fast-tunable transmitter and burst-mode packet receiver for HORNET," in *Proc. Optical Fiber Communication (OFC)*, Anaheim, CA, 2001, pp. ThG2-1–ThG2-3.
- [44] M. Herzog, S. Adams, and M. Maier, "PROXY STRIPPING: A performance-enhancing technique for optical metropolitan area ring networks," *OSA J. Opt. Netw.*, vol. 4, no. 7, pp. 400–431, Jul. 2005.
- [45] M. Maier and M. Reisslein, "AWG-based metro WDM networking," *IEEE Commun. Mag.*, vol. 42, no. 11, pp. S19–S26, Nov. 2004.
- [46] H.-S. Yang, M. Herzog, M. Maier, and M. Reisslein, "Metro WDM networks: Performance comparison of slotted ring and AWG star

networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 8, pp. 1460–1473, Oct. 2004.

- [47] M. Herzog and M. Maier, *RINGOSTAR: An Evolutionary Performance-Enhancing WDM Upgrade of IEEE 802.17 Resilient Packet Ring (RPR)*, submitted for publication.
- [48] W. Bux and M. Schlatter, "An approximate method for the performance analysis of buffer insertion rings," *IEEE Trans. Commun.*, vol. COM-31, no. 1, pp. 50–55, Jan. 1983.
- [49] *Self-Similar Network Traffic and Performance Evaluation*, K. Park and W. Willinger, Eds. New York: Wiley, 2000.



**Martin Maier** (M'03) received the Dipl.-Ing. and Dr.-Ing. degrees (both with distinctions) in electrical engineering from the Technical University Berlin, Germany, in 1998 and 2003, respectively.

He was a Visiting Researcher at the University of Southern California (USC), Los Angeles, and the Arizona State University (ASU), Tempe. In the summer of 2003, he was a Postdoc Fellow at the Massachusetts Institute of Technology (MIT), Cambridge. He is currently a Professor Researcher at the Institut National de la Recherche Scientifique (INRS), Montreal, QC, Canada. He is the author of the book *Metropolitan Area WDM Networks—An AWG Based Approach* (Norwell, MA: Kluwer, 2003). His research interests include network and node architectures, routing and switching paradigms, protection, restoration, multicasting, and the design, performance evaluation, and optimization of MAC protocols for optical wavelength division multiplexing (WDM) networks, automatically switched optical networks (ASONs), and generalized multiprotocol label switching (GMPLS), with particular focus on metro and access networks. Recently, his research interests have concentrated on evolutionary WDM upgrades of optical metro ring networks and access networks.

Dr. Maier was the recipient of a two-year Deutsche Telekom doctoral scholarship from June 1999 to May 2001. He was also a corecipient of the Best Paper Award presented at The International Society of Optical Engineers (SPIE) Photonics East 2000—Terabit Optical Networking Conference.



**Martin Herzog** received the Dipl.-Ing. degree (with distinction) in computer engineering from the Technical University of Berlin, Germany, in 2002. He is currently working toward the Ph.D. degree at the Telecommunication Networks Group, Technical University Berlin, Germany, and participates in a graduate interdisciplinary engineering research program.

His research interests lie in the area of optical wavelength division multiplexing (WDM) networks with focus on architectures and access protocols for metro networks.



**Michael Scheutzw** received the Diploma in mathematics from the Johann–Wolfgang–Goethe University Frankfurt/Main, Germany, in 1979, and the Ph.D. degree in mathematics (*magna cum laude*) and the Habilitation degree from the University of Kaiserslautern, Germany, in 1983 and 1988, respectively.

From 1988 to 1990, he was a Project Leader at Tecmath GmbH, Kaiserslautern, Germany. Since 1990, he has been a Professor of Stochastics at the Department of Mathematics, Technical University Berlin (TU Berlin). From 1997 to 1999, he was an Associate Chair of the Department. He has visited the Southern Illinois University at Carbondale, Arizona State University, University of California at Irvine, Rutgers University, University of Rochester, Warwick University, and the Mathematical Sciences Research Institute (MSRI) at the University of California (UC) at Berkeley. He is currently a Professor at the Department of Mathematics, TU Berlin, Germany.



**Martin Reisslein** (A'96–S'97–M'98–SM'03) received the Dipl.-Ing. (FH) degree in electrical engineering from Fachhochschule Dieburg, Germany, in 1994, the M.S.E. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1996, and the Ph.D. degree in systems engineering from the University of Pennsylvania in 1998.

During the academic year 1994–1995, he visited the University of Pennsylvania as a Fulbright Scholar. From July 1998 to October 2000, he was a Scientist at the German National Research Center for Information Technology (GMD FOKUS), Berlin, and a Lecturer at the Technical University Berlin. From October 2000 to August 2005, he was an Assistant Professor at the Arizona State University (ASU), Tempe. He is currently an Associate Professor at the Department of Electrical Engineering, ASU. He maintains an extensive library of video traces for network performance evaluation, including frame size traces of Motion Pictures Expert Group (MPEG)-4 and H.263 encoded video, at <http://trace.eas.asu.edu>. His research interests are in the areas of internet quality of service, video traffic characterization, wireless networking, and optical networking.

Dr. Reisslein is the Editor-in-Chief of the *IEEE COMMUNICATIONS SURVEYS AND TUTORIALS* and has served on the Technical Program Committees of IEEE INFOCOM, IEEE GLOBECOM, and the IEEE International Symposium on Computer and Communications. He has organized sessions at the IEEE Computer Communications Workshop (CCW). He is a corecipient of the Best Paper Award of The International Society for Optical Engineers (SPIE) Photonics East 2000—Terabit Optical Networking conference.