Impacts of Hierarchy in Ethernet Ring Networks on Service Resiliency

Kwang-koog Lee, Jeong-dong Ryoo, and Younglok Kim

In transport networks, a multi-ring architecture is very useful to facilitate network planning and to design and provide more resilient services for customers. Unlike traditional synchronous optical network multi-rings, the service resiliency of Ethernet-based multi-rings is significantly impacted by the ring hierarchy because a link or node failure in a certain level ring triggers filtering database flush actions in all higher level rings as well as in the ring with the failure, and consequently a large amount of duplicated data frames may be flooded. In this paper, we investigate how the ring hierarchy impacts the service resiliency of multi-ring networks. Based on extensive experiments on various single- and multiple-link failures, we suggest two effective inter-ring connection rules to minimize the transient traffic and to ensure more resilient multi-ring networks. In addition, we consider a flush optimization technique called e-ADV, and show that the combination of e-ADV and multi-ring structures satisfying our inter-ring connection rules results in a more attractive survivability performance.

Keywords: Carrier-grade Ethernet, filtering database, flush operation, multi-ring, ring protection, survivability.

I. Introduction

With its simplicity and economical data service delivery, the current Ethernet has emerged as an important player in the next-generation packet-based transport networks [1]. As carrier-grade Ethernet technology continues to make such considerable progress, it is being challenged by service providers who need rapid and reliable recovery capabilities that guarantee the availability of their services. The traditional Ethernet, which focuses on the scope of LAN, relies on a spanning tree protocol (STP) ensuring loop avoidance of data forwarding [2]. However, the STP approach neither acts quickly upon any topology change nor ensures optimal forwarding paths. To resolve these drawbacks, many enhanced STP-based approaches including a rapid STP and multiple STP have been additionally proposed [3], [4]. Yet, their convergence time is still too long to meet the sub-50-ms protection switching time requirement of transport networks.

To replace the conventional STP-based protocols and ensure more resilient Ethernet networks when a ring topology is used, several self-healing Ethernet rings including a resilient packet ring (RPR) [5], Ethernet ring protection (ERP) [6], Ethernet automatic protection switching (EAPS) [7], the resilient Ethernet protocol (REP) [8], and rapid ring protection protocol (RRPP) [9] have been introduced by several standard groups and vendors as the synchronous optical network (SONET) ring has been marked in traditional transport networks [10]. The ring structure provides two paths between any two nodes with an inherent bi-connectivity feature. Hence, the failure of any one path can be simply fixed by the re-selection of another path within the 50-ms protection switching time, which service providers require.

Meanwhile, with the benefit of the ring, the multi-ring

Manuscript received May 30, 2011; revised Sept. 19, 2011; accepted Oct. 19, 2011.

Kwang-koog Lee (phone: +82 42 860 6723, kwangkoog@etri.re.kr) is with the Advanced Communications Research Laboratory, ETRI, Daejeon, Rep. of Korea, and also with the Department of Engineering, University of Science and Technology, Daejeon, Rep. of Korea.

Jeong-dong Ryoo (ryoo@etri.re.kr) is with the Advanced Communications Research Laboratory, ETRI, Daejeon, Rep. of Korea.

Younglok Kim (corresponding author, ylkim@sogang.ac.kr) is with the Department of Electronic Engineering, Sogang University, Seoul, Rep. of Korea.

http://dx.doi.org/10.4218/etrij.12.0111.0341

structure, which interconnects multiple single rings, constitutes a promising architecture for the transport networks [11]. Although a multi-ring structure is more complex than a single-ring structure, it has the following features: A multi-ring structure allows an independent operation within each single ring, such that the recovery speed under a failure condition can be faster than that of a large-scale single-ring structure. In addition, it can survive multiple failures as long as there is not more than one failure on each ring. Therefore, the hierarchical multi-ring architecture provides better survivability than the single ring structure. Focusing on traditional SONET rings, many researchers have considered various inter-ring connection methods [11]-[15].

However, unlike SONET, service resiliency in such Ethernet multi-rings is impacted by their ring hierarchy. ERP technology, which provides sub-50-ms protection switching capability for Ethernet rings based on a typical Ethernet header, basically uses the generic functions of an IEEE 802.1d bridge [2], such as the filtering database (FDB) flush, to remove all learned MAC addresses for correct data forwarding when a link or node fails. However, this behavior allows for a large amount of duplicated data frames inside the ring, which then causes a traffic overshooting problem [16]. In multi-rings, a ring hierarchy is one of the major factors used to determine the volume of such flooded traffic [17], as the occurrence of a single failure in a lower-level ring triggers an FDB flush in both the higher-level ring and lower-level ring. Hence, if a multi-ring network is designed without serious consideration of its ring hierarchy, the volume of such flooded traffic might be far greater than the link capacity, where the majority of frames can be lost or delayed in a node with a finite buffer. In addition to the concern regarding traffic overshooting, when a multi-ring consists of single rings with different capacities, a lack of an appropriate configuration for the ring hierarchy might lead to an undesirable traffic redirection in which normal traffic is redirected to the opposite path with a lower capacity after a failure. Therefore, various multi-ring configuration methods for resilient Ethernet multi-rings should be suggested to minimize the transient traffic and ensure stable protection under a failure condition.

In this paper, we design an Ethernet-based reference network comprising one backbone ring and four access rings, which reflects the current metro network architecture [16], [18]-[20]. Based on the modeled multi-ring, three distinct hierarchical ring structures using ERP are proposed to investigate how their ring hierarchies impact the service resiliency of the multi-ring network. In addition, we consider a non-hierarchical ring structure to understand how valuable the multi-ring structure is over a single ring, where the ERP-based backbone ring is interconnected with the access networks using the Ethernet linear protection (ELP) switching scheme [21]. To evaluate the survivability performance of these ring structures, we implemented both the ERP and ELP protocols in OPNET [22]. We carried out extensive experiments in various single-link or multiple-link failures and compared their performances in terms of recovery time, resource utilization, and data throughput. From the experimental results, we show that the volume of the duplicated frames indeed varies depending on the ring hierarchies, and suggest two inter-ring connection rules for more survivable multi-rings: First, each link within a single ring should have an equivalent link capacity to prevent an unbalanced traffic redirection after a failure. Second, a multi-ring should be designed with as small a depth as possible to avoid consecutive traffic overshooting caused from a failure of a lower-level ring. Meanwhile, we additionally apply a flush optimization technique, called the enhanced selective advertisement (e-ADV) scheme [17], to observe how well the combination of such ring hierarchy design and flush optimization minimizes the transient traffic. The simulation results demonstrate that it fits best for reducing a large amount of duplicated frames and for providing a successful recovery process within 50 ms.

The remainder of this paper is organized as follows. Section II describes the traffic overshooting problem of Ethernet-based rings and introduces several flush optimization techniques and efficient inter-ring connection schemes for multi-rings. Section III illustrates the design of ring-based Ethernet networks, where a reference network model and four hierarchical ring structures are designed. Their survivability performances, including the flush optimization technique, are delineated in section IV. Discussion and arguments for each model are given in the next section. Finally, section VI provides some concluding remarks.

II. Network Stability Issues of Ethernet-Based Self-Healing Rings

This section provides a brief description of flush-based survivable Ethernet rings and introduces the traffic overshooting problem. We then review some flush optimization techniques to resolve a traffic overshoot. In addition, previous works related to an efficient ring interconnection are introduced.

1. Flush-Based Survivable Ethernet Rings and Traffic Overshooting Problem

Ethernet-based self-healing rings, such as ERP [6], RPR [5], EAPS [7], REP [8], and RRPP [9], have been introduced to replace the traditional spanning tree protocol and ensure more resilient carrier-grade Ethernet ring networks. Without a complex computation, provisioning overhead, or excessive information exchange, these methods provide fast and reliable protection switching within 50 ms. All methods except RPR fully comprise a typical Ethernet media access control (MAC) header and have been developed on the principle of utilizing a generic mechanism inherited from the Ethernet bridge functions. In this paper, we focus on the ERP protocol with the multi-ring connection capability and hereafter introduce the ERP technology in detail.

In a normal state, the ERP defines a master node called a ring protection link (RPL) owner on a ring and lets this node block one of two ring ports to create a logical loop-free topology. When a failure occurs in the ring domain, nodes adjacent to the failure (NAFs) detect the failure condition and immediately block the port facing the failure. Next, NAFs generate and multicast a ring automatic protection switching (R-APS) signal fail (SF) message notifying the failure situation along both ring directions. Then, other ring nodes accepting this R-APS (SF) message also recognize that the ring is in a failed state. Among those nodes, the RPL owner releases its logically-blocked port so that the connectivity between any two nodes on the ring is rapidly recovered.

During the fault notification process, every ring node should refresh its FDB table upon receiving the R-APS (SF) message because MAC addresses learned before this failure are no longer valid due to the changed block position. However, this allows a large amount of duplicated frames inside the ring. When the volume of such flooded traffic is far greater than the link capacity, the majority of frames can be lost or delayed in a node with a finite buffer. Further, the transient traffic makes ring nodes extend the address learning period. Thus, the combination of these two impairments can make protection switching and settling time greater than 50 ms. In multi-rings, the traffic overshooting problem is even more critical than that of a single ring. When a single failure occurs in a lower-level ring (sub-ring), it affects forwarding routes of traffic traversing its higher-level ring (major ring) as well as the sub-ring. It makes ring nodes in the higher-level ring also trigger an FDB action. In [17], the authors showed that such consecutive flush operations generate more than ten-times the amount of transient traffic over the normal state. In their experiments, the duplicated traffic was not stabilized even after 300 ms.

2. Flush Optimization Schemes

To minimize the amount of the transient traffic, several flush optimization techniques have been proposed in the previous literature. For a single-ring domain, Rhee and others proposed a method called FDB flipping, which makes the NAFs send the modified R-APS (SF) messages to inform other nodes of the addresses impacted by the protection switching [23]. Since other ring nodes update their FDBs using the addresses in those messages, the transient traffic can be immediately eliminated. Meanwhile, Lee and others introduced an optimization scheme operating on a multi-ring architecture [24]. When a sub-ring topology is changed by a failure, it can avoid unnecessary FDB flush operations in its major ring using the modified ERP protocol. In [25], Lee and others proposed the selective FDB advertisement scheme to let all ring nodes exchange their subnet FDB information with each other after the FDB flush operation. This helps the nodes learn the subnet clients of each node in a short period of time. The e-ADV scheme has also been developed for operation on a multi-ring network [17].

3. Related Works for Hierarchical Inter-ring Connection

Given a multi-ring network, these flush optimization methods can reduce the amount of undesirable transient traffic in their own manner. However, they have no responsibility for guaranteeing the complete recovery of service flows under a failure condition. When a multi-ring is configured with an inappropriate ring hierarchy, it can still experience severe service degradation due to an incorrect traffic redirection where normal traffic is redirected to the opposite path with a lower link capacity after a failure. Therefore, inter-ring connection methods to prevent a service disruption and ensure survivable services under a failure are additionally required.

Several ring interconnection methods have been suggested to achieve cost-effective network planning of traditional SONET rings. Lee and Koh proposed a very simple approach interconnecting multiple access rings into one core ring [14]. Grover and others introduced overlapping multiple rings in which several nodes and links can be shared between different rings [15]. However, this approach is applicable only to the fixed ring topologies. To provide a more flexible inter-ring connection, Shi and Fonseka proposed heuristic algorithms with a hierarchical self-healing ring structure design [11]. The basic idea of these algorithms concentrates on minimizing the traffic routing between different rings while grouping nodes within a small ring. Since this approach makes a multi-ring structure using a single interconnection node, it cannot survive an interconnection node failure. Thus, they proposed an extensive algorithm based on dual interconnection nodes [12].

These SONET-based inter-ring connection methods can be applied for Ethernet-based multi-ring networks. However, they are still not capable of guaranteeing complete recovery of service flows and minimizing a traffic overshoot under a failure condition. To the best of our knowledge, there have not been such inter-ring connection methods for Ethernet multi-rings. For this reason, we attempt to design various inter-ring connections on an Ethernet multi-ring network in the following sections and find effective and survivable inter-ring connection strategies. This paper does not focus on the FDB flush optimization, but we do consider it in our performance evaluations in order to observe how well the combination of the preferable ring hierarchy and flush optimization scheme minimizes the transient traffic. In this paper, we use the e-ADV scheme as it guarantees the most reliable and feasible protection switching among these flush optimizations.

III. Design of Hierarchical Ethernet Rings

Designing an optimal hierarchical Ethernet ring minimizing the transient traffic and guaranteeing survivable services under a failure is a complicated optimization problem. With the construction of an optimal multi-ring, there can be many decision variables, such as the number of nodes on individual rings, the number of rings within the multi-ring network, the ring interconnection rules, and the formation of a hierarchy in a multi-ring. In addition, depending on the positions of blocks under normal and failure conditions, the amount of flooded traffic during the switchover period varies.

For these reasons, we assume that an Ethernet-based multi-ring network is pre-established as a reference network. This reflects the current metropolitan area network environment in a simple manner and is also quite often cited in several papers [11]-[13]. On the basis of this multi-ring architecture, we designed three distinct hierarchical ring structures using the ERP scheme. In addition, we present one non-hierarchical single-ring approach.

1. Reference Network Model for Transport Networks

In general, a transport network is designed as a two-tier structure of a core (backbone) network and an access network [18]-[20] where several access networks are connected to a single core network. Even though a mesh is a general topology for numerous networks, carrier-grade networks normally apply a ring topology because the management in a ring-based network is lighter, and certain behaviors such as the recovery process are more predictable than in a mesh network.

For the connection of two single rings, two types of inter-ring connection mechanism can be used: *single-homing* [11] and *dual-homing* [12]. Single-homing makes a higher-level ring connect to lower-level rings by choosing one node from each lower-level ring. On the other hand, the dual-homing makes lower rings share two interconnecting nodes with a higher level ring. The dual-homing architecture is considered to be more survivable because it can survive all single-point failures including single-node failures.

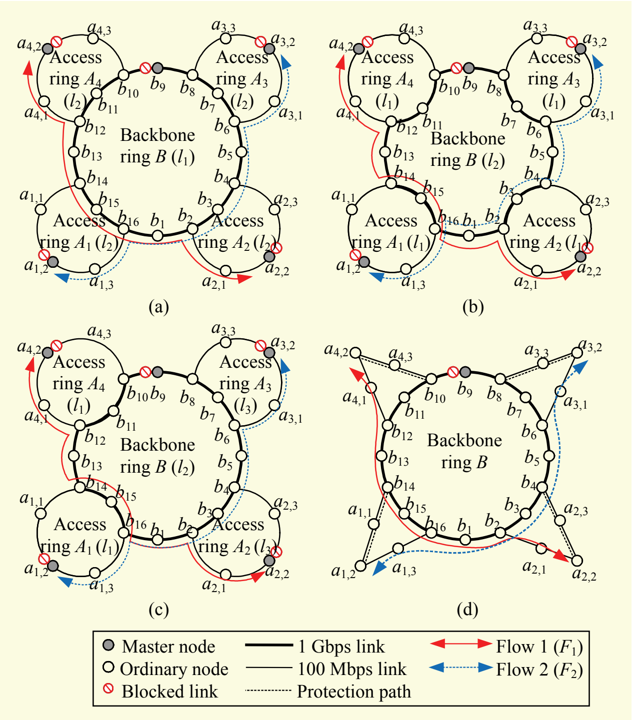


Fig. 1. Hierarchical multi-ring architectures: (a) one major ring at first level and four sub-rings at second level, (b) four major rings at first level and one sub-ring at second level, (c) two major rings at first level, one sub-ring at second level, and two sub-rings at third level, and (d) single-ring architecture with linear protection scheme.

From these observations, in this paper, we determine a multi-ring network with four access rings and a large-scale backbone ring as a reference network model. Here, all access rings are interconnected to the backbone using the dual-homing scheme. The suggested reference network is presented in Fig. 1, which will be used in our performance evaluations later.

2. Inter-ring Connections of the Ethernet Multi-ring

In Figs. 1(a) through 1(c), the backbone ring B includes sixteen Ethernet nodes (b_i, $1 \leq i \leq 16$), each of which is connected to two adjacent ring nodes. On the other hand, each of the four access rings (A_i, $1 \leq i \leq 4$) has three Ethernet nodes ($a_{i,j}$, $1 \leq i \leq 4$, $1 \leq j \leq 3$) and is interconnected with two nodes on the backbone ring. On the backbone ring, ring nodes other than interconnection nodes (b_i, $i = 2n$, $1 \leq n \leq 8$) individually have one subnet network where numerous clients reside. We define nodes b_9, $a_{1,2}$, $a_{2,2}$, $a_{3,2}$, and $a_{4,2}$ as a master node on each ring. They block one ring port to guarantee a loop-free topology under normal conditions.

With the reference network model, we now discuss the design of three different hierarchical ring structures. To indicate multi-rings with different ring hierarchies, we named them

l_1+4l_2, $4l_1+l_2$, and $2l_1+l_2+2l_3$, where l_i and '+' denote the level i of a ring and the interconnection between rings with distinct levels, respectively. A lower-level ring is interconnected to two nodes in a higher-level ring in a hierarchical manner. Therefore, the level increases from top to bottom in ascending order. The level of a ring on top is always one.

The first model in Fig. 1(a), l_1+4l_2, has two levels of the hierarchical ring structure. The backbone ring is at the top of the multi-ring and four access rings are in the second level. Thus, all backbone ring nodes are associated with the first level. However, the levels between the backbone and access rings can also be swapped. In this way, the second structure in Fig. 1(b), $4l_1+l_2$, has four access rings at the first level and one backbone ring at the second level. Since the four access rings are on top of the multi-ring, four sets of backbone nodes, $\{b_{14}, b_{15}, b_{16}\}$, $\{b_2, b_3, b_4\}$, $\{b_6, b_7, b_8\}$, and $\{b_{10}, b_{11}, b_{12}\}$, are included in access rings, A_1, A_2, A_3, and A_4, respectively. Finally, the third model in Fig. 1(c), $2l_1+l_2+2l_3$, is designed from a combination of the first model and second model. This results in a three-level ring hierarchy, where access rings A_1 and A_4 are on top of the multi-ring, the backbone ring is in the second level, and access rings A_2 and A_3 are in the third level.

As shown in Figs. 1(a) through 1(c), we suppose that two end-to-end flows F_1 and F_2 exist where flow F_1 is used between node $a_{2,2}$ and node $a_{4,2}$, and flow F_2 is used between $a_{1,2}$ and node $a_{3,2}$. Although the designed multi-ring structures have different hierarchy levels, the two end-to-end flows are on the same paths. This is because their link blocks are configured in the same positions. However, despite the same physical topology, their survivability performance can be different according to their hierarchies. In the next section, we evaluate these topologies under a variety of single- and multiple-link failures through extensive simulations.

3. Single Ring with End-to-end Survivability Schemes

Along with the multi-ring models above, in this subsection, we present a single-ring approach. Even though the multi-ring architecture is regarded as a promising structure for survivable networks, such networks could experience a severe traffic overshooting problem from single failures on a lower-level ring, as an FDB flush operation is needed at all higher-level rings as well as at the lower-level ring that suffers the failure. Therefore, it is necessary to compare a multi-ring structure with a non-hierarchical single ring structure where access networks use a different survivability scheme.

As depicted in Fig. 1(d), the single-ring approach is also applied to the same physical topology as in the previous multi-ring structures. The backbone network employs ring-based survivable Ethernet technologies, but in the four access networks, the end-to-end flows F_1 and F_2 of our interest are protected by a 1:1 Ethernet linear protection (ELP) switching method [21], in which a disjointed backup path is dedicated to the working path, and, under a failure condition, normal traffic on the working path is immediately switched to the backup path. In Fig. 1(d), the working path of flows F_1 and F_2 are the same as the paths in the multi-ring approach. Although their backup paths are not defined in the backbone ring domain, they are still supported by a ring-based survivability scheme activated in the backbone.

Meanwhile, nodes $a_{2,2}$, $a_{4,2}$, $a_{1,2}$, and $a_{3,2}$, which are located at ends of each linear protection domain, should have their own hold-off timers to prevent a racing condition in which two distinct protection schemes try to recover a failure condition at the same time. This timer allows an inner protection group, that is, ERP, to attempt to restore the traffic before switching at the outer protection group, that is, ELP. In other words, the outermost node does not immediately operate protection switching even when detecting a new defect. It waits until the hold-off timer expires and decides to start its own recovery procedure to determine whether the defect still exists. In practice, the range of the hold-off timer varies according to different standards. The period usually ranges from 0 to 10 seconds in steps of x ms; for example, the x value is 100 ms for Ethernet and SDH and 500 ms for ATM. However, the suggested range in an optical transport network (OTN) is from 0 ms, 20 ms, or 100 ms to 10 seconds in steps of 100 ms. Note that the previous three multi-ring approaches applying a unified ring protection technique do not need this timer.

IV. Performance Evaluations

The survivability performances of the four ring structures, l_1+4l_2, $4l_1+l_2$, $2l_1+l_2+2l_3$, and a *single* ring, were evaluated using the OPNET simulator [22]. To evaluate their protection performances and make the simulations as practical as possible, we implemented both the ERP and ELP protocols in the commercial carrier Ethernet devices provided in OPNET. We carried out numerous experiments in various single-link or multiple-link failures and measured their performances by focusing on the recovery time, resource utilization, and data throughput.

Each ring node in the backbone is connected to two adjacent ring nodes with a 20-km 1-Gbps full-duplex link (propagation delay of 0.1 ms). In contrast, the nodes in each access area are connected in series using a 10-km 100-Mbps full-duplex link, and two end-nodes are then interconnected with two nodes in the backbone ring using the 100-Mbps link. Each subnet of the non-interconnection nodes has 1,000 clients, each of which exponentially transmits an average of 80-kbps of traffic toward their destinations equally distributed among all other clients in the remote subnets as well as in the same subnet. All of the generated traffic is symmetric, bidirectional, and co-routed, and thus the same amounts of traffic between source and destination pairs are delivered through the same set of links and nodes in each direction. Regarding the traffic demands of the backbone, two bi-directional end-to-end flows via the backbone ring exist. As mentioned in the previous section, flow F_1 is for between node $a_{2,2}$ and node $a_{4,2}$, while flow F_2 is for between $a_{1,2}$ and node $a_{3,2}$. Each client located at the ends of these flows transmits 40 Mbps of exponential traffic on average toward its own destination in a remote access ring. Finally, the service rates of the data and R-APS protocol message were set to 1.0 million packets per second (mpps) and 0.01 mpps, respectively.

Our experiments were divided into two aspects: a variety of single-link failures and multiple-link failures. The former were further divided into two cases: single failures in both the backbone network and the access networks. Experiments on a flush optimization technique were also carried out. In the following subsections, we compare the experimental results in detail.

1. Single-Link Failure Scenarios in a Backbone Network

For single-link failure conditions in a backbone ring, each backbone link failed at 2.0 s. For each experiment, we measured the link utilization of both the backbone ring and four access networks every 40 ms. We also gauged the end-to-end delay and number of received frames in 40 ms intervals for flows F_1 and F_2. The averaged results from all the experiments are shown in Fig. 2.

In Fig. 2(a), the transient traffic of both l_1+4l_2 and *single* structures follows the same curve and is the heaviest, as the ring nodes on the large backbone network perform an FDB flush operation. At the peak, their average link utilizations are about 31%, which is more than double that in a normal state. The flooded traffic was stabilized after about 600 ms. The $4l_1+l_2$ structure, on the other hand, experiences the smallest traffic overshooting because half of the backbone nodes grouped into small access rings can help prevent the spreading of transient traffic generated in the backbone area. However, $4l_1+l_2$ shows the heaviest link utilization after 2.3 s because a failure on the backbone links associated with the level-1 access rings allows its normal traffic to be redirected to the access links with a capacity of 100 Mbps. Even though the utilization of each link is not displayed in this paper, the failure condition on each of the bidirectional links, b_{14}-b_{15} and b_{15}-b_{16} (or b_2-b_3 and b_3-b_4), made the access network A_1 (or A_2) experience full link utilization.

The average end-to-end delay and received data frames of flows F_1 and F_2 are depicted in Figs. 2(b) and 2(c), respectively. Of all the ring structures under a failure condition, flows F_1 and F_2 are abnormally delayed for about 200 ms. During this time, both l_1+4l_2 and *single* structure show the longest end-to-end delay and lose 70% of the end-to-end flows. However, their traffic was stabilized after 2.2 s. Note that their end-to-end delay slightly increases after the failure because the length of the protection paths in most of the failure conditions is longer than that of the working paths. Meanwhile, $4l_1+l_2$ and $2l_1+l_2+2l_3$ reveal a shorter end-to-end delay and smaller loss of

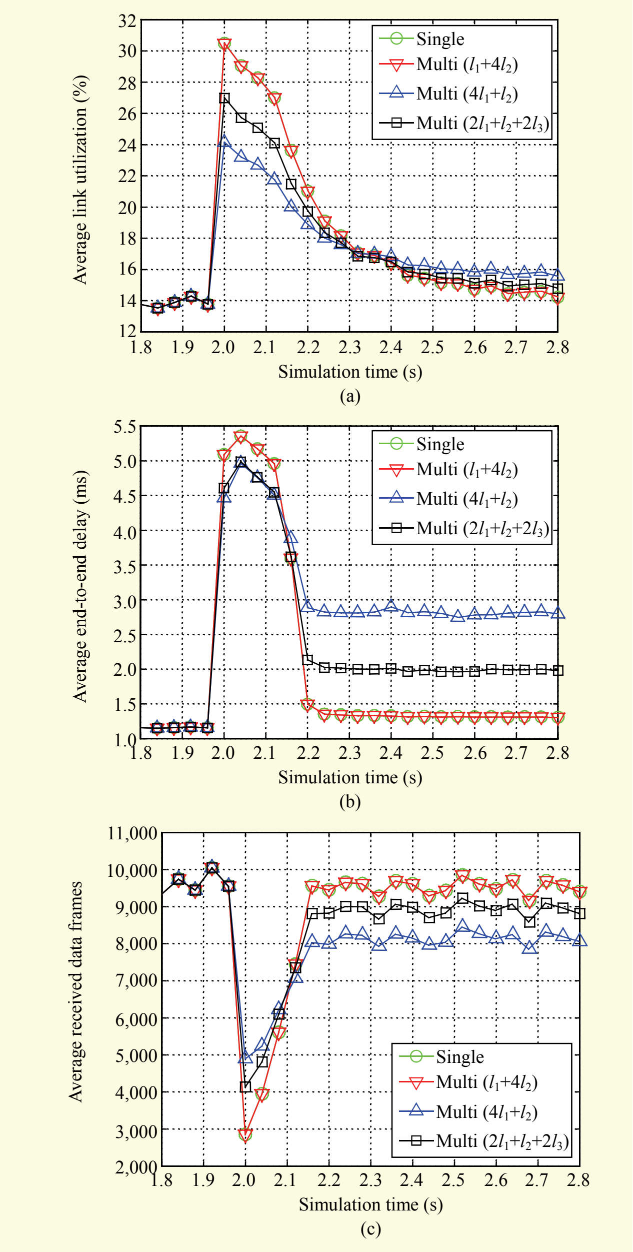


Fig. 2. Single-link failures in backbone network: (a) average link utilization in metro area, (b) average end-to-end delay of flows F_1 and F_2, and (c) average number of received data frames of flows F_1 and F_2.

data traffic than both l_1+4l_2 and *single* structure from 2.0 s to 2.2 s. However, flows F_1 and F_2 in their ring structures are delivered across an access ring suffering from full utilization in their failure scenarios so that the average end-to-end delay of $4l_1+l_2$ and $2l_1+l_2+2l_3$ was increased to about 2.5-times and 1.7-times, respectively. Moreover, they lost about 20% and 10% of their end-to-end flows after 2.2 s.

2. Single-Link Failure Scenarios in Access Networks

We evaluated the survivability performances of single-link failures in the four access domains. Each of the four access links on the working paths for flows F_1 and F_2, b_{12}-$a_{4,1}$, b_6-$a_{3,1}$, b_{16}-$a_{1,3}$, and b_2-$a_{2,1}$, failed at 2.0 s in one experiment, and we repeated the same experiments for other access link failures. We monitored the same performance parameters as in the previous metro backbone failure.

The average link utilization in the metro area is drawn in Fig. 3(a). First, the *single* structure does not produce transient traffic after a failure, but its normal traffic decreases about 15% for 100 ms. As aforementioned, the end nodes of each linear protection domain have hold-off timers to allow the protection in the metro area to repair the failure first. In our experiments, the hold-off timer value is set to 100 ms, which is the minimum value recommended for Ethernet linear protection. Thus, normal traffic due to the failure on an access link is switched to its backup path after 100 ms. As shown in Fig. 3(c), the amount of received data frames drops to one-half of the normal state during that period. However, since the linear protection does not invoke any FDB flush operation, no transient traffic is produced. Therefore, the average end-to-end delay of the *single* structure depicted in Fig. 3(b) is slightly increased after 2.1 s. Also, the amount of received data frames is perfectly recovered at that time.

Meanwhile, the l_1+4l_2 design shows the heaviest transient traffic for about 400 ms because single failures of the access rings make all ring nodes in both the backbone and failed access ring flush their FDB tables. The link utilization in Fig. 3(a) also supports the results of Figs. 3(b) and 3(c). Although each access ring completes the ring protection switching process for its fault condition within 50 ms, the average received data frames are less than in the single-ring model with a hold-off timer.

In contrast, the $4l_1+l_2$ structure produces barely any transient traffic, as all the access rings are major rings. Subsequently, traffic loss of the two flows is very little and the end-to-end delay is quickly stabilized. Finally, the performance measures of $2l_1+l_2+2l_3$ are in between the l_1+4l_2 and $4l_1+l_2$ structures.

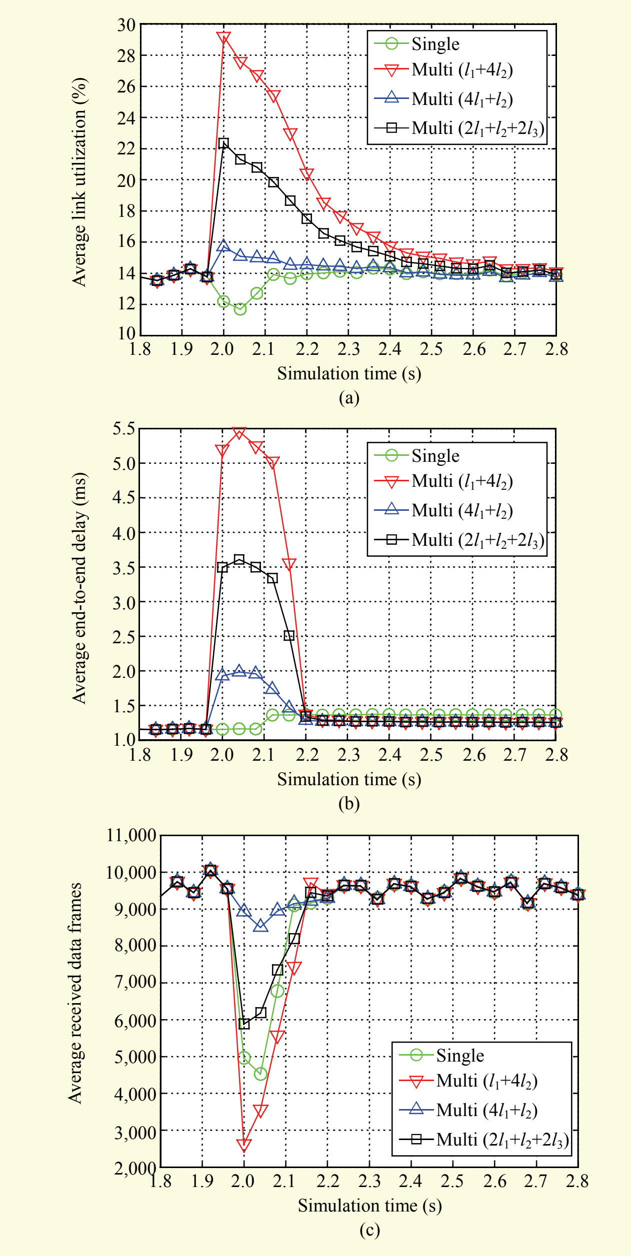


Fig. 3. Survivability performances under single-link failures on access networks: (a) average link utilization in metro area, (b) average end-to-end delay, and (c) average number of received data frames of two end-to-end connections.

3. Multiple-Link Failure Scenarios

We carried out additional experiments to evaluate how each ring structure protects the end-to-end service flows even in multiple-link failures. The failures occur in the following links in sequence: link b_1-b_2 in the backbone at 2.0 s, link $a_{4,1}$-b_{12} in the access network A_4 at 3.0 s, and link $a_{2,2}$-$a_{2,3}$ in the access network A_2 at 4.0 s. The average end-to-end delay and the average received data frames are measured for each flow every 100 ms.

When the first failure occurs at 2.0 s, as shown in Fig. 4, the survivability performances depicted in Fig. 6 show similar

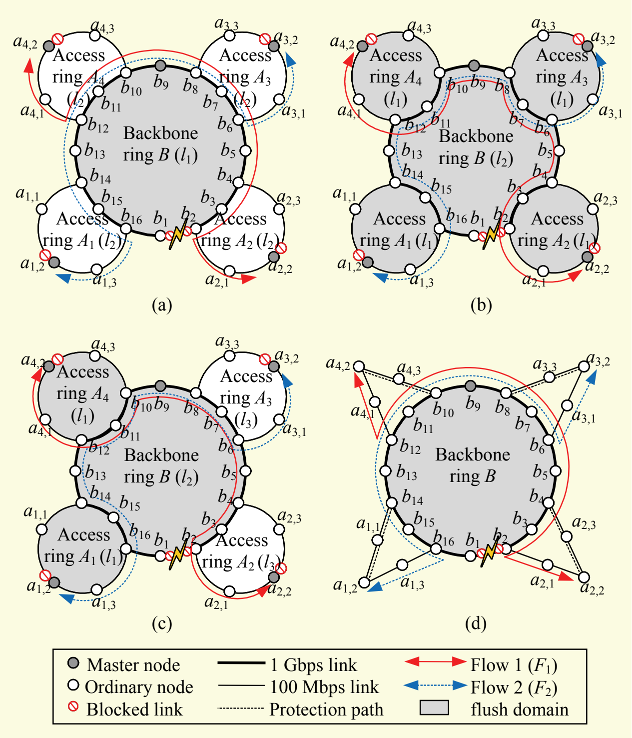


Fig. 4. First failure case at 2.0 s and flush domain of ring structures.

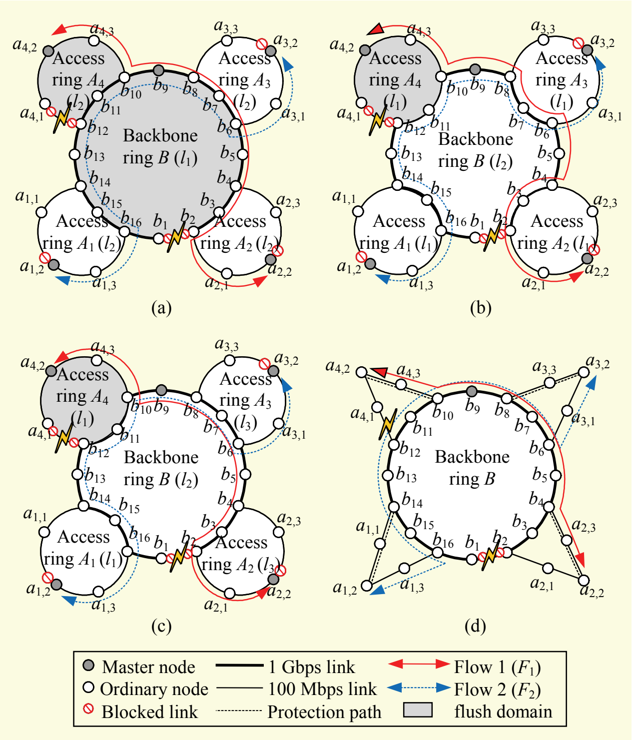


Fig. 5. Second failure case at 3.0 s and flush domain of ring structures.

aspects for each model. Regardless of the hierarchical structure of each model, the block location change by this failure allows most of the backbone ring nodes to trigger an FDB flush action. Thereby, flows F_1 and F_2, passing by the backbone, experience significant delay and loss for about 200 ms. After the failure occurs, the end-to-end paths of flows F_1 and F_2 are still the same in all cases. However, their paths are increased by four hops; therefore, the end-to-end delays of flows F_1 and F_2 in Figs. 6(a) and 6(c) are slightly prolonged after 2.0 s.

Meanwhile, referring to Fig. 5, the results from the second failure invoked at 3.0 s are quite different from the first one. In the case of all three multi-ring models, flow F_1 suffers from traffic overshooting in access ring A_4. Since the l_1+4l_2 model allows backbone ring nodes on the top level to perform a flush operation, flow F_2 is also delayed and lost. In contrast to the multi-ring models, the single-ring approach sets up a hold-off timer in the end nodes of the linear protection domains. As depicted in Figs. 6(a) and 6(b), flow F_1 is cut off and no results are found during 100 ms, after which, its normal traffic is redirected to a backup path. Despite such a distortion, the 1:1 Ethernet linear protection scheme does not require any flush actions in the access area. Its normal traffic is therefore restored faster than in the other multi-ring models. Moreover, as depicted in Fig. 5(d), under the first failure condition, the backup path of flow F_1 is shorter than the working path. This helps reduce the link utilization in the backbone ring, and thus the end-to-end delay of flow F_2 slightly decreases after 3.1 s.

Finally, when the last failure occurs at 4.0 s, flow F_1 in the *single* model is no longer protected, but in the multi-ring models it is still survivable. In the multi-ring cases, the failure condition of the blocked link does not change the ring topology. Flows F_1 and F_2 in the 4l_1+l_2 model are barely affected since that failure occurs in the first-level ring. However, the other two models let the ring nodes on the higher-level ring trigger the flush action. As a result, their flows experience a traffic overshoot for about 200 ms.

4. Experiment Results of Flush Optimization

As evaluated in the previous subsections, a duplication of frames leads to a significant loss and delay of the data frames. To minimize this transient traffic more effectively, we plugged a flush optimization scheme, e-ADV, into the l_1+4l_2 model, and this produced the heaviest transient traffic but guaranteed the complete recovery of service flows after a failure event. The averaged link utilization from all single-link failures is depicted in Fig. 7. We note that the l_1+4l_2 model with the e-ADV scheme is named l_1+4l_2(opt) in the figure.

Similar to the previous experiment results, the l_1+4l_2 model still generates the heaviest transient traffic. It increased to more than double the normal state. Meanwhile, the *single* model shows less traffic overshooting than l_1+4l_2 owing to its linear

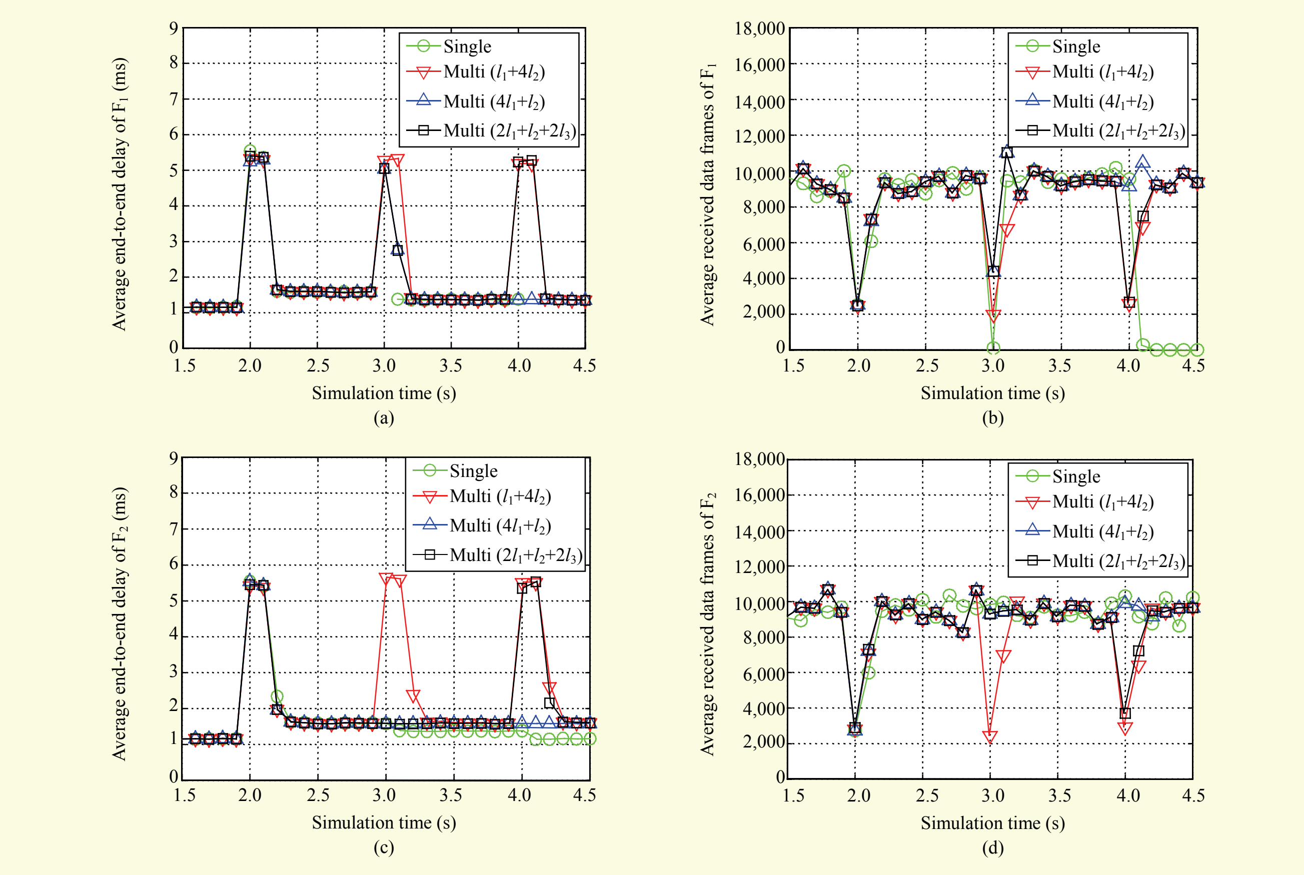


Fig. 6. Survivability performances under multiple link failures: (a) average end-to-end delay of flow F_1, (b) average number of received data frames of flow F_1, (c) average end-to-end delay of flow F_2, and (d) average number of received data frames of flow F_2.

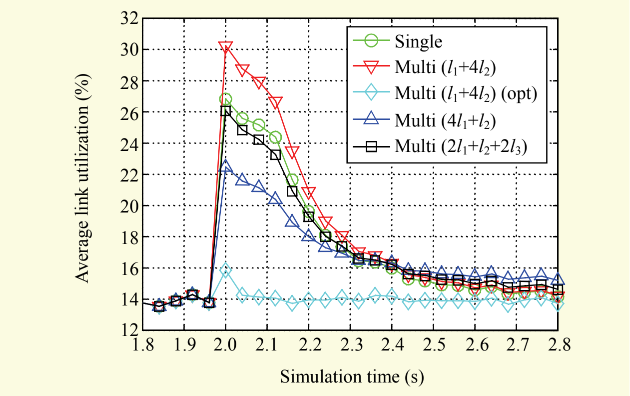


Fig. 7. Averaged link utilization under single link failures.

protection scheme. Then, $4l_1+l_2$ and $2l_1+l_2+2l_3$ recorded link utilizations of about 26% and 22%, respectively, at the peak. The transient traffic is almost relieved after 2.4 s in all cases except l_1+4l_2(opt). However, the link utilizations of $4l_1+l_2$ and $2l_1+l_2+2l_3$ slightly increase because of their unbalanced traffic redirection. In opposition to the former cases, the l_1+4l_2(opt) model significantly reduces the transient traffic. It generates one spike for only a few milliseconds. This spike increased to only 1.14-times the normal state, but it was immediately suppressed by the indirect MAC learning process of e-ADV.

Although we do not display the average end-to-end delay or average number of received data frames for flows F_1 and F_2 here, except for l_1+4l_2(opt), the results were similar to those shown in Fig. 2(b) and 2(c). The l_1+4l_2(opt) approach showed the shortest end-to-end delay and most stable frame reception.

V. Summary and Arguments

The single-ring approach does not deal with an inter-ring connection. However, a large-scale single ring is difficult to realize in a practical environment. Moreover, the number of nodes on a ring is usually limited by the capacity requirement and the latency of the control message circling around the ring. Moreover, the single ring cannot survive under multiple failures. When multiple protection domains are nested, the use of a hold-off timer is essential to avoid any possible contention from multilayered protection schemes under a failure condition. Therefore, normal traffic on a failed working path can suffer from service suspension until that timer expires. However, the single ring has an advantage of avoiding the consecutive traffic overshooting problem invoked by a failure event of a lower-level ring. We can also conclude that instead of a ring-based survivability scheme, the application of other methods such as the linear protection in a small local network is more preferable because it can lower the total level of the multi-ring network.

In the meantime, the hierarchical ring structures require no hold-off timer due to the unified protection behavior among the individual rings. Thus, under any single failure, they are indeed capable of completing the protection switching process within 50 ms. However, they can experience severe transient traffic from FDB flush actions triggered by the failure of a lower level ring. Our experiments demonstrated that the end-to-end service flows are not properly transmitted for hundreds of milliseconds. We also observed that the amount of transient traffic varies depending on the structure of the ring hierarchy. Moreover, an arbitrary ring hierarchy can lead to undesirable protection behavior where disrupted normal traffic is redirected to a path with a lower link capacity. Consequently, it is important to design a hierarchical ring structure of a multi-ring network while preserving the survivability performance. From the experiment results, we recommend the following two inter-ring connection rules for more survivable multi-rings. First, each link within a single ring should have equivalent link capacity to prevent an unbalanced traffic redirection after a failure. Second, a multi-ring should be designed with as small a depth as possible to avoid consecutive traffic overshooting caused by the failure on a lower-level ring.

Finally, we considered a flush optimization technique called e-ADV to investigate how it properly minimizes the transient traffic in the proposed hierarchical ring structures. This method is applied to the l_1+4l_2 structure satisfying our inter-ring connection rules. Even though this ring structure exposed the worst survivability performance in the first evaluation, the results with the e-ADV showed the fastest and best survivability performance. Of course, the flush optimization scheme plays a key role in minimizing the traffic overshooting. However, unless the ring hierarchy is considered, such optimization is not meaningful. Therefore, it is quite valuable to investigate how the hierarchical ring structures impact the survivability performances of Ethernet-based multi-rings.

VI. Conclusion

This paper discussed the hierarchical ring structures of Ethernet-based multi-ring networks. Based on our reference network model, three distinct hierarchical ring structures and one non-hierarchical single ring structure were presented. From extensive simulations, we showed that a lack of consideration for a hierarchical ring structure may lead to an unbalanced protection switching process and a serious broadcast storm problem. For more survivable multi-rings, we recommended two inter-ring connection rules. The results from the combination of a flush optimization scheme, e-ADV, and a hierarchical ring structure satisfying the two recommendations showed that the combination fits best for providing a successful recovery process within 50 ms.

References

[1] K. Fouli and M. Maier, “The Road to Carrier-grade Ethernet,” *IEEE Commun. Mag.*, vol. 47, no. 3, Mar. 2009, pp. S30-S38.

[2] IEEE Std. 802.1D, “IEEE Standard for Local and Metropolitan Area Networks – Media Access Control (MAC) Bridge,” June 2004.

[3] IEEE Std. 802.1w, “IEEE Standard for Local and Metropolitan Area Networks, Amendment 2: Rapid Reconfiguration for Spanning Trees,” June 2001.

[4] IEEE Std. 802.1s, “IEEE Standard for Local and Metropolitan Area Networks – Virtual Bridged Local Area Networks, Amendment 3: Multiple Spanning Trees,” Dec. 2002.

[5] IEEE Std. 802.17, “Part 17: Resilient Packet Ring (RPR) Access Method and Physical Specifications,” 2004.

[6] ITU-T Recommendation G.8032/Y.1344, “Ethernet Ring Protection Switching,” 2008.

[7] IETF RFC 3619, “Extreme Networks’ Ethernet Automatic Protection Switching (EAPS) Version 1,” 2003.

[8] Cisco Systems, “White Paper for Resilient Ethernet Protocol (REP),” *Cisco Systems*, Technical Report, C11-427224-00, 2007.

[9] Huawei Corp., “Technical White Paper for Rapid Ring Protection Protocol (RRPP),” *Huawei Corp.*, Technical Report, 2007.

[10] T.H. Wu., *Fiber Network Service Survivability*, Massachusetts: Artech House, 1992.

[11] J. Shi and J.P. Fonseka, “Hierarchical Self-Healing Rings,” *IEEE/ACM Trans. Netw.*, vol. 3, no. 6, Dec. 1995, pp. 690-697.

[12] J. Shi and J.P. Fonseka, “Interconnection of Self-Healing Rings,” *Proc. IEEE Int. Conf. Commun*, 1996, pp. 1563-1567.

[13] A. Proestaki and M.C. Sinclair, “Interconnection Stratiges for Dual-homing Multi-ring Networks,” *Proc. 16th Int. Teletraffic Congress*, June 1999.

[14] C.Y. Lee and S.J. Koh, “A Design of the Minimum Cost Ring-chain Network with Dual-Homing Survivability: A Tabu Search Approach,” *Comput. Operations Research*, vol. 24, vol. 5, 1997, pp. 883-897.

[15] W.D. Grover, J.B. Slevinsky, and M.H. MacGregor, “Optimized Design of Ring Based Survivable Networks,” *Canadian J. Electr. Comput. Eng.*, vol. 20, no. 3, 1995. pp. 139-149.

[16] J. Ryoo et al., “Ethernet Ring Protection for Carrier Ethernet Networks,” *IEEE Commun. Mag.*, vol. 46, no. 9, Sept. 2008, pp. 136-143.

[17] K. Lee and J. Ryoo, “Flush Optimizations to Guarantee Less

Transient Traffic in Ethernet Ring Protection," *ETRI J.*, vol. 32, no. 2, Apr. 2010, pp. 184-194.

[18] M. Huynh and P. Mohapatra, "Metropolitan Ethernet Network: A Move from LAN to MAN," *Elsevier Comput. Netw.*, vol. 51, no. 17, 2007, pp. 4867-4894.

[19] T. Atmaca et al., "A Solution for Synchronization Problem of Interconnected Metro Access and Metro Core Ring Networks," *Euroview*, Aug. 2011.

[20] T. Atmaca and D.T. Nguyen, "End-to-End Performance Evaluation of Interconnected Optical Multi-ring Metropolitan Networks," *Springer Wireless Developing Countries Netw. Future*, 2010, pp. 206-216.

[21] ITU-T Recommendation G.8031/Y.1342, "Ethernet Linear Protection Switching," 2011.

[22] OPNET Technologies Inc. Available: http://www.opnet.com

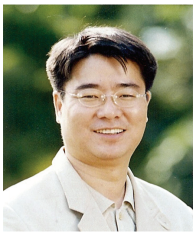
[23] J.K. Rhee, J. Im, and J. Ryoo, "Ethernet Ring Protection Using Filtering Database Flip Scheme for Minimum Capacity Requirement," *ETRI J.*, vol. 30, no. 6, Dec. 2008, pp. 874-876.

[24] D. Lee et al., "Efficient Ethernet Multi-ring Protection System ," *Proc. IEEE Int. Workshop Design Reliable Commun. Netw.*, 2009.

[25] K. Lee, J. Ryoo, and S. Min, "An Ethernet Ring Protection Method to Minimize Transient Traffic by Selective FDB Advertisement," *ETRI J.*, vol. 31, no. 5, Oct. 2009, pp. 631-633.

Kwang-koog Lee received his BS and MS in electronic communication engineering from Kangwon National University, Chuncheon, Rep. of Korea, in 2006 and 2008, respectively. He is currently working toward a PhD with broadband network engineering in the University of Science and Technology, Daejeon, Rep. of Korea. Since 2008, he has also been with ETRI, Daejeon, Rep. of Korea, where he has worked on switching and management technologies for high-speed optical transmission systems. His research interests are in packet-based transport networks, such as carrier Ethernet and MPLS-TP, and optimized protocol design and performance analysis in wired/wireless networks.

Jeong-dong Ryoo is a principal member of the research staff in ETRI, Daejeon, Rep. of Korea. He holds a master's degree and a doctorate in electronic engineering from the Polytechnic Institute of NYU, NY, USA, and a bachelor's degree in electronic engineering from Kyungpook National University, Rep. of Korea. After completing his PhD in the area of telecommunication networks and optimization, he started working for Bell Labs, Lucent Technologies, NJ, in 1999, where he was mainly involved with performance analysis, evaluation, and enhancement study for various wireless and wired network systems. Since joining ETRI in 2004, his work has focused on next generation network, carrier class Ethernet and MPLS-TP technology research, especially participating in OAM and protection standardization activities in ITU-T. He co-authored TCP/IP Essentials: A Lab-Based Approach (Cambridge University Press, 2004). He is a member of Eta Kappa Nu association.



Younglok Kim received his BS in electronic engineering from Sogang University, Seoul, Rep. of Korea, in 1991, and his MS and PhD in electrical engineering from the Polytechnic Institute of NYU in 1993 and 1998, respectively. From 1999 to 2003, he joined InterDigital Communication Corporation as a senior system engineer at in Melville, NY, working on signal processing for wireless communication systems. Since 2003, he has been with the Department of Electronic Engineering, Sogang University, where he is now an associate professor. His research interests are algorithm and hardware implementation for communication and radar signal processing.