Efficient Ethernet Multi-Ring Protection System

Doojeong Lee and J.-K. Kevin Rhee
Dept. of Elec. Engr. and Dept. of Info. and Comm.
KAIST
Daejeon, S. Korea
{dj.lee | rhee.jk}@kaist.ac.kr

Kyusang Lee and Peter Cho
Actus Networks, Inc.
Seoul, S. Korea
{sangi | choho}@actusnetworks.com

Abstract—Ethernet ring protection (ERP) introduced by ITU-T G.8032 Recommendation is a new class of packet switched network protection technology that can provide ring automatic protection switch for nested multiple Ethernet rings. An ERP system is required to provide network protection against link and node failures within 50ms. Whenever the protection scheme switches the ring to have a new topology, the standard requires all nodes to flush their filtering data base (FDB) resulting in congestion in the access networks. In this paper we evaluate the protection performance of the current G.8032 scheme and introduce additional novel schemes for performance enhancements in nested multi ring networks.

Keywords—Ethernet ring protection, ERP, APS, G.8032, traffic overshoot phenomenon

I. INTRODUCTION

Carrier Ethernet technologies are considered as a prevailing solution for carrier networks due to its simplicity and cost-effectiveness. Currently emerging IEEE standards for PBB and PBB-TE are notable examples of carrier Ethernet technologies well suited for this role [1][2]. However, Ethernet technologies still have a critical challenge of reliability achieved by SONET/SDH [3][4]. The IEEE has standardized 802.17 resilient packet ring protection switching for future replacement of SONET/SDH ring networks, and the ITU-T is currently standardizing Ethernet ring protection (ERP) in G.8032 for carrier grade Ethernet services [4][5][6]. Ring protection is an essentially important technology because the existing dark fiber cables mostly form ring topologies. The G.8032 Recommendation introduces a cost-efficient but highly scalable and reliable ring protection schemes, as the scheme reuses the generic Ethernet functions that are standardized at IEEE802.1 and IEEE802.3 Standards Associations.

The new ERP function can be implemented in most applications as a software addition with no or minimal hardware modifications of standard Ethernet switches. In the G.8032 Version 1 Amendment, the nested multiple ring protection scheme is added for practical market applications and the next version will include feature-rich protection switching including operator command sets and the full support for more complicated multiple ring topology. This new technology, however, is just introduced to the field, requiring better understandings of the protection switching behavior and performance. As well, there are rooms for improvements for better performance such as a network congestion problem after flushing FDBs. This paper investigates the performance of multi-ring Ethernet protection switching behavior for access network applications and discusses possibilities to improve the multi-ring protection performance of currently considered ERP schemes. The rest of this paper is organized as follows. In Section II, we briefly discuss an overview of protection mechanisms of current G.8032 recommendation. In Section III, we evaluate the performance of G.8032 by OPNET simulation with our own ERP modules. Section IV proposes novel protection mechanisms that demonstrate less protection lead times and less link capacity requirements, and finally we conclude our results in Section V.

II. OVERVIEW OF G.8032

As an application of the G.8032 ERP, we consider an Ethernet ring network consisting of four components: Ethernet Ring Node (Node), Ring Protection Link (RPL), RPL owner and RPL node. Nodes have at least two ports, one for the east and the other one for the west. Nodes can interconnect several rings to extend their coverage and form interconnected ring networks. The special node for ring interconnection is called an interconnection node, and links between them are called shared links. The interconnection nodes can have multiple shared links in-between. Ethernet Rings connected to another ring or network through the use of interconnection nodes are called sub rings and the ring to which a sub ring is connected is called the upper ring. The Ethernet forwarding rule requires a loop-free network topology. In G.8032 ERP scheme, loop freeness is guaranteed by placing a logical block in each ring in the idle state. This block forms an RPL and the node of this block is called RPL owner. The node at the other end of the RPL is referred to as an RPL node. The RPL node port connected to the RPL may or may not be blocked. When a link or node failure is detected, the ERP state changes to a protection state...
with unblocking RPL ports and blocking ports associated with the failed link. ERP functions of ring nodes exchange messages with each other through a special channel defined as a ring automatic protection switch (R-APS) channel. An R-APS channel is rendered as an in-band multicast channel by use of a generic Ethernet virtual local area network (VLAN) provided by IEEE 802.1Q Standard. For sub-rings, the two interconnection nodes may exchange R-APS messages via a special virtual R-APS channel on the shared links. For certain networks, this may not be used as it is an optional feature. Fig. 1 shows an example of a G.8032 Ethernet ring network model consisting of three rings. Each ring has its own RPL, preventing loop formation. Logical blocks formed by RPLs make network have a tree topology.

A. Operation Principle of Ethernet Ring in Idle State

In an idle state, ports at both ends of an RPL are blocked by the RPL owner and the RPL node. According to the current recommendation, RPL nodes are optional. In the case where an RPL node is not used, only one end of the RPL is blocked by the RPL owner. The RPL owners, then, periodically send R-APS No Request RPL Blocked (R-APS(NR,RB)) messages through both the east and the west ports. Upon receiving R-APS(NR,RB) messages, nodes unblock non-failed ports and forward messages to the next ring node. Fig. 2 shows you an Ethernet ring network in the idle state. In this example, a network consists of three rings interconnected by interconnection nodes 0, 3, 5, and 8. Each ring forms logical bus by using RPL; and R-APS(NR,RB) messages generated by each RPL owner circulates in each ring in both directions indicating that the network is in the idle state and both or either RPL ports are blocked.

B. Operation Principle of Ethernet Ring in Protection State

Link failure and protection switching result in a topology change. According to the current recommendation, in order to mitigate inconsistency between network topology and forwarding information in the case of network failure, nodes flush FDBs and reorganize FDBs by using flooding base MAC source address learning.

If a link failure is detected by nodes, local Signal Fail (SF) signals are enabled at the nodes detecting the failure, so that these immediate nodes can block their ports associated with the failed link. Immediate nodes, then, start to periodically issue R-APS(SF) messages indicating a Signal Fail condition happened in the ring. If the failed ports are not logically blocked, they are blocked immediately. In the subsequent ERP procedure, these immediate nodes disseminate a Signal Failure message R-APS(SF) to inform (i) the RPL owner and RPL node to unblock PRL ports, and (ii) the ring nodes to switch its ERP state to a protection state. Once ring nodes receive R-APS(SF) message, they flush their FDBs, meaning that the all entries of the FDB are cleared. If failed ports are already blocked, e.g. an RPL detects a failure, a link failure does not cause a topology change, meaning that FDBs need not be flushed. In this case, Do Not Flush (DNF) flag in an R-APS(SF) message is set. After protection process, the ring forms a new logical tree.

In the case of a link failure of an upper ring, the topology change does not require FDB flush in the sub rings. However, if a link failure happens in a sub ring, not only a sub ring but also upper rings require FDB flush. Fig. 3 shows you the topology changes of two failure scenarios. In the first scenario, a link failure happens in the upper ring, ERP 1, between nodes 8 and 9. Due to a change of link blocking position in the upper ring, FDB information of nodes in the upper ring nodes becomes stale. However, the link failure of upper ring does not affect forwarding information of sub ring nodes because the forwarding information associated with interconnection nodes is not changed. Thus the only nodes in the upper ring need to flush FDBs. In the second scenario, however, link failure happened in the sub ring, ERP 2, between nodes 10 and 11; and this failure invalidates FDB information of nodes 0, 8, 10, 11, 12, 13 in the sub ring and node 9 in the upper ring. Nodes 10 and 11 detect a link failure, and generate R-APS SF messages requesting FDB flush in the sub ring. However, unfortunately node 9, in the upper ring, cannot be informed of the link failure of the sub ring. In order to make those upper ring nodes to flush their FDBs, the current recommendation requires interconnection nodes to inject R-APS (EVENT, Flush Request) messages into their upper ring if a topology change of sub ring is detected. Upper ring nodes receiving
R-APS(EVENT, Flush Request) messages flush their FDBs; and finally, all nodes in ERP 1 and ERP 2 flush their FDBs. This process is referred to as flush propagation. In this case, even if only nodes 0, 8, 9, 10, 11, 12, and 13 need to flush their FDBs, all nodes in the sub and upper rings flush FDBs. When FDBs are flushed, Ethernet nodes start to broadcast the packets with unknown destinations. As a result, the network experiences traffic congestion resulting in long end-to-end delay and high link capacity consumption. This traffic overshoot is later mitigated as FDB learns forwarding information from the port and the source address of incoming packets. In the next section, we simulate this phenomenon and show how it affects the performance of network protection by using an OPNET network simulator.

III. PERFORMANCE ANALYSIS OF G.8032

We briefly discussed an overview of the protection mechanism of G.8032 in the previous sections. The current G.8032 protection scheme is too simple to provide optimized protection switching performance against failures. For example, in the case where a link failure occurs in a sub ring, both of sub ring and upper ring flush FDBs and repopulate FDBs by flooding-based MAC source address learning. Consequently, access networks having less link capacity compared with core networks suffer from a great deal of incoming frames, and experience long end-to-end delays.

To show network traffic behavior in the case of a network failure, we implemented a G.8032 module in an OPNET network simulation. Fig. 4 shows the simulation topology and parameters in our simulation model, consisting of three rings having identical link capacity. ERP 1 in the middle of network is the upper ring having ERP 2 and ERP 3 as its sub rings. Each ring node has a subnet of which average utilization is 32% with 200 clients; therefore, the corresponding total network load is 640Mbps. Each ring link has 1Gbps capacity and the access link has 100Mbps link capacity. Destination of each frame is uniformly distributed in the network. Each ring has one RPL owner and one RPL node that control the RPL. The first link failure is activated at the time of 2 seconds in the ERP 2; and at the time of 4 seconds, another link failure is introduced in ERP 3.

Fig. 5 shows the link utilization of a ring link between nodes 5 and 6. Right after link failures at times 2 and 4 seconds, the link utilization of ring links overshoots; however, due to a relatively large capacity, link utilization is still less than the link capacity. Fig. 6 shows link utilization of access links of the upper ring, ERP 1, and that of sub ring, ERP 2. Due to relatively small link capacity of access links, it shows a severe traffic overshooting phenomenon resulting from FDB flush. Upper ring access links show traffic overshooting twice, because the flush request messages originating from each sub ring make ERP 1 to flush FDBs whenever sub rings are failed; and access links in ERP 2 and ERP 3 experience traffic overshooting one
time at time 2 and 4 respectively. High link utilization resulting
from FDB flush increases queue length of intermediate nodes.
Consequently, the average end-to-end delay is increased. In Fig.
7, the simulation results show an 18 times longer end-to-end
delay right after the network failure. In this section, we
evaluated the protection performance of current
recommendation for G.8032.

IV. EFFICIENT ETHERNET MULTI-RING PROTECTION
SCHEMES

A. Minimum Set of Flushing Nodes

In order to find out a solution that can minimize the number
of flushing nodes, we need to clarify the condition making
nodes to flush their FDBs. To make a network topology rather
simpler to deal with, we reorganize a physical network into a
logical network. Fig. 8 shows you a simplified logical topology
of G.8032 Ethernet ring network. The network is composed
of a single logical ring and many trees. A logical ring is composed
of its own ring segments and virtual ring segments consisting
of segments of the other rings. According to the position of
link blocks, the physical topology of logical ring can be several
as shown in Fig. 8. Let’s assume that a link failure happened in
the logical ring A. This failure only affects the topology of the
logical ring. In short, in order to minimize the number of
flushing nodes in the case of a failure in ring A, nodes in the
logical ring A, including nodes in its virtual ring segments,
should be flushed.

B. FDB Port Flush

In this section, we introduce a protection mechanism using
port flush that selectively deletes FDB entries associated with
ports on which nodes receive port flush request messages.

Fig. 10 shows an example of port flush protection. A link
failure happens in ERP 3. Immediate nodes generate R-APS SF
messages. Then messages circulate in the failed ring making
nodes A and interconnection nodes 1 and 2 clear their FDBs as
described in the G.8032 recommendation. Upon receiving
R-APS SF messages from a sub ring, interconnection nodes 1 and 2 start port flush protection process. Instead of sending flush request messages, interconnection nodes send port flush request messages to their upper rings. These messages, then, circulate in the upper rings and flush partial FDB entries associated with ports on which nodes receive port flush messages. Port flush request messages propagate in the whole network via interconnection nodes until they reach the last ring having no upper ring. Note that only shared nodes in the logical ring segments can receive port flush request messages on both ports, and flush FDB entries associated with the east and the west ports. The other nodes would receive port flush request messages through only one port and delete partial forwarding information associated with it. For example, node B is in the logical ring segments of ERP 3; thus, B flushes all FDB entries associated with both the east and the west ports, however, node C not in the logical ring receives port flush request messages on port 1; and node C deletes partially only the FDB entries associated with port 1. However, actually, node C is not a member of a minimum set of flushing nodes. Therefore, it does not need to flush its FDB information. However, comparing with the existing FDB flush mechanism flushing all FDB entries, it is still more efficient.

C. FDB Flush Trigger

In this section, we introduce a novel protection mechanism using flush triggering that is more advanced compared with the port flush. As we showed before, nodes in the logical ring of a specific ring A are minimum nodes affected by failures happened in ring A. Therefore, in this mechanism, FDB flush happens only nodes that receive flush trigger messages on both ports. Because only the shared nodes in the logical ring can receive flush trigger messages on both ports, only these nodes clear their FDBs.

Fig. 11 shows an example case of flush triggering protection. When a link failure happens in ERP 3, R-APS(SF) messages are generated and make nodes in the ERP 3 to clear their FDBs. Interconnection nodes receive R-APS(SF) messages, and generate flush trigger messages; then, flush trigger messages propagate toward the whole network. Upon receiving a flush trigger message, nodes run flush trigger timer and forward the flush trigger message to the next ring node. If node receives a flush trigger message on the other port within timeout, node deletes all FDB entries associated with the east and the west ports. Otherwise, flush trigger timeout is expired and node does not flush FDB. For example, node B in the logical ring receives flush trigger messages on both ports and deletes its FDB information associated with east and west ports. Node C, however, receives a flush trigger message on port 1 and does not flush its FDB. Thus, the flush trigger mechanism efficiently flushes FDBs of nodes in the logical ring of failed ring and reduces the number of flooding frames dramatically.

D. Performance Evaluation

In this section, we evaluate the performance of port flush and flush trigger mechanisms. To evaluate the performance, we will use link utilization and end-to-end delay as the performance criteria. Fig. 12 shows simulation topology and parameters. The simulation network is consisted of three rings having identical capacity. Each node has two 1 Gbps ring ports and one 100Mbps access port. Each access subnet connected to their FDBs. Interconnection nodes receive R-APS(SF) messages, and generate flush trigger messages; then, flush trigger messages propagate toward the whole network. Upon receiving a flush trigger message, nodes run flush trigger timer and forward the flush trigger message to the next ring node. If node receives a flush trigger message on the other port within timeout, node deletes all FDB entries associated with the east and the west ports. Otherwise, flush trigger timeout is expired and node does not flush FDB. For example, node B in the logical ring receives flush trigger messages on both ports and deletes its FDB information associated with east and west ports. Node C, however, receives a flush trigger message on port 1 and does not flush its FDB. Thus, the flush trigger mechanism efficiently flushes FDBs of nodes in the logical ring of failed ring and reduces the number of flooding frames dramatically.
a access port of a ring node generates 32Mbps traffic. With a one-second time interval, starting from ERP 1 each ring fails consecutively. Fig. 13 shows utilization of an access link of ERP 1 of each protection mechanism. Since ERP 1 is an upper ring, whenever link failures happened in the sub rings, flush request messages originating from failed sub rings are propagated to ERP 1. Therefore in the case where the flush mechanism is used, access networks in upper ring experience traffic overshooting three times. In the case of using port flush for protection, we still have three traffic overshoots; however, the duration is shorter than existing flush mechanism. Flush trigger shows the best performance since flush trigger flushes FDBs of nodes in the logical ring only. Thus, the amount of flooding traffic is dramatically reduced and flooding traffic is filtered by the very first node of normal ring as well. For example, nodes 1 and 5 filter out flooded frames from ERP 2 destined to other rings in the case of sub ring failures. Fig. 14 shows average end-to-end delay. The flush scheme shows the longest end-to-end delay among protection schemes. At a time of 3 seconds, it shows the most significant end-to-end delay. This is because a link failure in ERP 3 triggers FDB flushes of all upper rings of ERP 3. Consequently, starting from ERP 3 to ERP 1, all FDBs are flushed and a great deal of frames is flooded. The port flush scheme shows less end-to-end delay compared with the flush mechanism due to partial FDB deletion using port flush. The flush trigger shows the best performance in the end-to-end delay measure due to more efficient FDB management, resulting in the minimum frame flooding.

V. CONCLUSION

In this paper, we study Ethernet ring protection introduced by a recent ITU-T G.8032 Recommendation. An ERP network can provide topology protection switching within 50 ms, competing with SONET/SDH standards. This paper, for the first time, reports the packet transport performances on ERP switching based on simulations. The results show a great potential to replace the SONET/SDH rings. However, the performance of the current ERP scheme can be substantially enhanced by simple additional features.

In the current version of G.8032 Recommendation ver. 1, when network failures happen, nodes delete all FDB
information and re-populate FDB tables by source address learning. During nodes rebuild FDB tables, a great deal of flooding packets in the networks severely congests especially in low capacity access networks. In order to relieve traffic overshooting resulting from FDB flush, we identified a minimum set of nodes to be flushed in the case of failure. Based on this, we introduce two efficient ring protection mechanisms: Port flush and flush trigger schemes. The port flush scheme selectively deletes FDB entries associated with port receiving port flush request messages; the flush trigger scheme clears FDBs only when a node receive flush trigger messages on both ports. Thus, port flush and flush trigger efficiently reduce the number of flushing nodes and localize the traffic congestion. As a result, simulation results show less traffic overshoots and less average end-to-end delay compared with the simple flush mechanism of G.8032.

REFERENCES