

Ethernet Ring Protection Using Filtering Database Flip Scheme For Minimum Capacity Requirement

June-Koo Kevin Rhee, Jinsung Im, and Jeong-dong Ryoo

ABSTRACT—Ethernet ring protection is a new technology introduced in ITU-T Recommendation G.8032, which utilizes the generic Ethernet MAC functions. We introduce an alternative enhanced protection switching scheme to suppress penalty in the switching transient, in which the Ethernet MAC filtering database (FDB) is actively and directly modified by information disseminated from the nodes adjacent to failure. The modified FDB at all nodes are guaranteed to be consistent to form a complete new ring network topology immediately. This scheme can reduce the capacity requirement of the G.8032 by several times. This proposed scheme can be also applied in IP protection rings.

Keywords—Ethernet ring protection, source address learning, APS, traffic overshoot phenomenon.

I. Introduction

Ethernet ring protection has gained much interest as it can avoid the requirement of physical layer protection, which is rather costly. IEEE has developed the IEEE 802.17 resilient packet ring protection switching scheme adopted from the SONET/SDH operation principle, and ITU-T has introduced Recommendation G.8032 on the ring automatic protection switch (R-APS) technique [1], [2] which is compatible with OAM functions recommended by ITU-T Y.1731 [3], [4].

As recommended by G.8032, an Ethernet ring has a ring protection link (RPL), where the link is blocked in order to avoid forming a loop in the working state. On detection of a

failure on a link or port, a signal failure (SF) message is multicast to inform other ring nodes of the failure condition. On protection switching, the RPL is unblocked, and the blocks are moved to the nodes adjacent to the failure, forming a new traffic pattern on the ring. To rebuild the filtering databases (FDBs), an SF message activates all nodes to “flush” the FDBs so that all FDB entries are cleared. In turn, the FDBs undergo a source address (SA) learning process to rebuild the FDB entries. Meanwhile, all data frames with unknown destination addresses (DAs) are broadcast, resulting in a large traffic overshoot. This flush method may provide protection within 50 ms only when there is a large link-capacity overprovision beyond the nominal traffic volume [5]. Frame loss occurs if the link capacity is not sufficient, and the overshoot period is elongated in the high-traffic links.

To overcome these problems of the flush method, we propose a new scheme for fast FDB updates by an *FDB flip* which utilizes forwarding failure information encountered at the nodes adjacent to the failure. Such information can be delivered to other nodes by use of an expanded R-APS frame payload [5]. An R-APS (SF, flip) message contains a list of MAC addresses which indicate troubled DAs. This simple process contains enough information to fix FDB entries immediately and provide protection switching.

II. Ethernet Ring Protection Using APS Frame

The FDB flip scheme in Ethernet ring protection is shown in Fig. 1 with an example of a four-node protected Ethernet ring network. In the working state, an RPL is selected as the link between nodes A and D, thereby forming a linear logical network of a chain of nodes A-B-C-D. Host AX is attached to node A as shown in Fig. 1(a). Consider that all FDBs have learned forwarding information, and the health of each link

Manuscript received Aug. 17, 2008; revised Oct. 20, 2008; accepted Oct. 28, 2008.

This work at ICU is supported in part by the IT R&D program of MKE/ITTA [2008-F017-01] and the ERC program of MOST-KOSEF [R11-2000-074-02006-0].

June-Koo Kevin Rhee (phone: +82 42 866 6124, email: rhee.jk@jeec.org) and Jinsung Im (email: jinsung_im@tmax.co.kr) are with the School of Engineering, Information and Communications University, Daejeon, Rep. of Korea.

Jeong-dong Ryoo (email: ryoo@etri.re.kr) is with the Broadcasting & Telecommunications Convergence Research Laboratory, ETRI, Daejeon, Rep. of Korea.

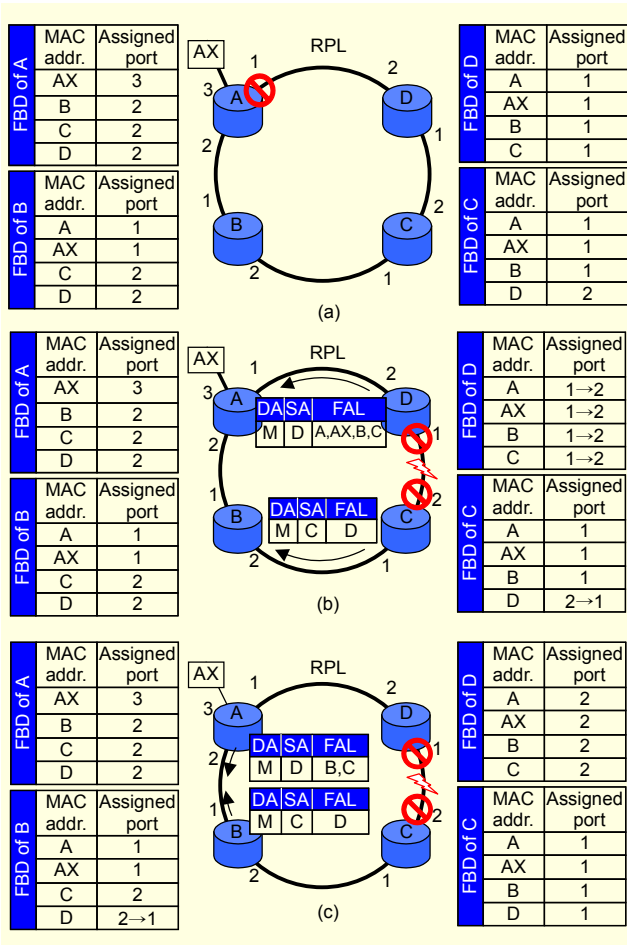


Fig. 1. Schematic description of FDB flip procedures: (a) working state, (b) selective FDB flip by FAL, and (c) FAL deletion. Entry M at DA fields of the messages indicates a multicast address.

is monitored [4]. When a link or port failure occurs, SF messages are multicast from both nodes that have detected the failure. We refer to such a node as a node adjacent to failure (NAF). The SF messages remove the existing port blocks, and the ports of the failed link automatically become new blocking ports, as shown in Fig. 1(b). Subsequently, FDBs in all nodes need to be updated for the newly configured network.

The FDB flush method is the simplest solution to update the FDB. Upon receiving an SF message with FDB flush notification, all FDB entries are deleted and the new topology is learned from the SA of incoming Ethernet frames. On the other hand, the FDB flip scheme actively fixes FDB entries based on the information provided by the NAFs. The NAFs initiate the FDB flip for protection by the following process. First, *protection ports* (port 1 of node C and port 2 of node D in Fig.1(b)) are determined in the opposite direction to the failed link. Second, the NAFs search for MAC addresses associated with *failure ports* (port 2 of node C and port 1 of node D in

Fig. 1(b)) in the FDB. These MAC addresses of each NAF form a *flip-address list* (FAL). Third, each NAF changes the port association of the FAL to the protection port. Last, each NAF multicasts R-APS (SF, flip) messages, which include the FAL in the payload of the message through the protection port.

On reception of an R-APS (SF, flip) message by the other nodes, each node flips the FDB address entries indicated in the FAL. For example, at node B in Fig. 1(b), the R-APS (SF, flip) message from node C arrives through port 2, meaning that node D is unreachable through this port. The R-APS process at node B searches the FDB entries for port 2 to find an FAL address matching D. The R-APS process of node B changes the FDB entry of address D to port 1 as indicated in Fig. 1(c).

Meanwhile, when a received FAL contains the addresses of the received node or its client hosts, these addresses need to be eliminated from the FAL. When the R-APS message from node D arrives at node A as shown in Fig. 1(b), FDB entries for A and AX are found in the FAL of the received message, and FAL entries of A and AX are deleted for transmission to the next node as in Fig. 1(c). This process is important to prevent repeated FDB flipping as a node receives two R-APS (SF, flip) messages from both sides. This faulty flipping can destroy the consistency of FDBs on the ring. However, the deletion requirement makes implementation costly because the R-ASP (SF, flip) message has to be terminated and regenerated hop-by-hop. By processing port flipping and deletion at all nodes by circling R-APS (SF, flip) messages, the FDBs can immediately have valid information for the new ring block position, which prevents broadcast traffic overshoot. Considering the generic similarity of routing and switching tables of IP and MPLS networks, respectively, this scheme can be extended to IP and MPLS protection rings.

III. Performance Evaluation

The performance of protection switching by FDB flipping and its benefits are evaluated in comparison with those of the flush method. For computer simulation, we assume 6 Ethernet ring nodes with 10,000 client Ethernet hosts attached to each ring node (see Fig. 2). The links of the ring are assumed to have a 10 Gbps capacity, which is large enough not to limit the traffic volume generated by each client subnet. Each host is supposed to generate and receive Ethernet frames at an average rate of 500 kiloframes per second (kfps) measured at the ingress port of each ring node. We assume that both frame length and inter-frame distance are exponentially distributed with the mean of 1,000 bits for the ingress traffic at each ring node of the corresponding local subnet. The destinations of the ingress traffic at one ring node are equally distributed among all the subnets. The link propagation delay is considered to be 0.125 ms, which corresponds to approximately 25 km fiber links between two adjacent nodes.

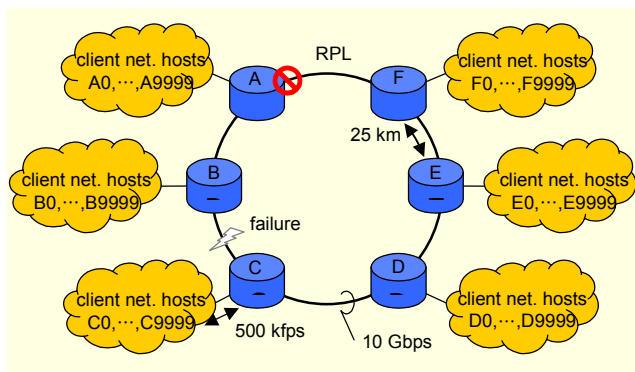


Fig. 2. Simulation scenario in normal Ethernet ring topology.

A link failure is introduced at 10 ms between nodes B and C. Then, nodes B and C are assumed to detect the failure immediately and transmit SF or R-APS (SF, flip) messages. In the flush method, an SF message will flush all FDBs as it propagates over the ring and removes the RPL block. In our FDB flip scheme, R-APS (SF, flip) messages are generated, and receiving nodes change the FDB entries by the previously described algorithm, in addition to the removal of the RPL block.

Figure 3 compares the link traffic volume of the flush method and that of our scheme. In the FDB flush scheme, flooding due to the broadcast of all first arriving frames occurs as no DA is known to FDBs. At this initial point, the traffic volume peaks to form an overshoot shown as the “D→C flush” case in Fig. 3. This overshoot lasts for 46 ms in this example until all frames are generated from all ring nodes and local subnet nodes. The peak value of the traffic volume overshoot is 2,400 kfps, while the equilibrium state traffic volume is 750 kfps at the maximum traffic link E-F. The transient traffic does not relax to an equilibrium state until 80 ms. Note that in this simulation, the link capacity of 10 Gbps is large enough compared with the peak bit rate of 2.4 Gbps, which causes only negligible node-delay. However, if the node and link capacities are less than the peak traffic, the frames in the overshoot period are queued and transmitted later, which makes the broadcast flooding period longer. Under this situation, many frames can be lost even after 50 ms. Consequently, the 50 ms protection transition time cannot be guaranteed without capacity that can handle the large traffic overshoot. In addition, such over-provision is required for every link because a fault can occur anywhere in the ring.

The proposed FDB flip method demonstrates an immediate transition to the steady state on protection. As shown in Fig. 3, FDB flipping at all Ethernet ring nodes takes only 12 ms to reach a steady state, with no traffic overshoot. This scheme can guarantee a 50 ms protection switching time. However, large FALs are included in an R-APS (SF, flip) message or in many fragmented R-APS (SF, flip) messages. FDB flipping may

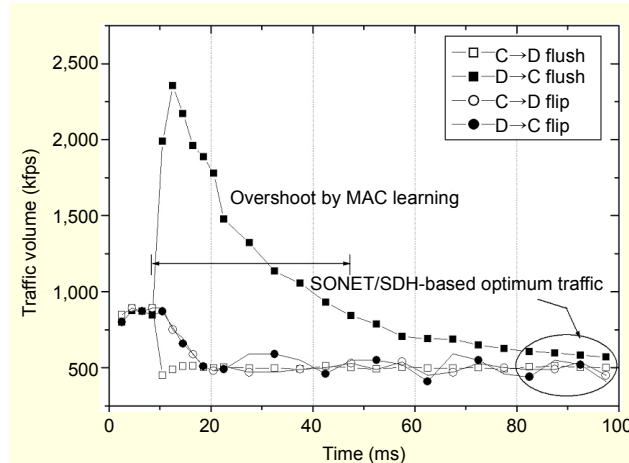


Fig. 3. Transient link traffic volume at link C-D with flush and flip schemes. During protection switching, link C-D turns to the minimum load link from the maximum load link.

require a lot of computing time when many client subnets are served. However, we found the processing time for 10,000 entries in the FAL takes less than a few milliseconds in a 2 GHz CPU system, which is fast enough to mitigate the switching transient.

IV. Conclusion

The proposed FDB flip scheme for Ethernet ring protection can guarantee a 50 ms protection transition time, with minimal link capacity requirements. No over-provisioning above the steady state traffic is required, in contrast to the currently adopted FDB flush scheme at ITU-T, which requires up to 500% over-provisioning with respect to the steady state traffic in the worst case of a six-node ring network servicing a large number of local client subnet hosts. This improvement of Ethernet ring protection technology makes the carrier-class Ethernet the strongest candidate to replace costly SDH/SONET infrastructure for access and metro networks.

References

- [1] ITU-T Rec. G.8032, *Ethernet Rings Protection Switching*, ITU-T, Geneva, 2008.
- [2] J. Ryoo et al., “Ethernet Ring Protection for Carrier Ethernet Networks,” *IEEE Comm. Mag.*, vol. 46, no. 9, 2008, pp. 136-143.
- [3] ITU-T Rec.Y.1731, *OAM Functions and Mechanisms for Ethernet Based Networks*, ITU-T, Geneva, 2006.
- [4] J. Ryoo et al., “OAM and Its Performance Monitoring Mechanisms for Carrier Ethernet Transport Networks,” *IEEE Comm. Mag.*, vol. 46, no. 3, 2008, pp. 97-103.
- [5] J. Im et al., “Ethernet Ring Protection with Managed FDB Using APS Payload,” *Proc. of APOC*, Paper 6784-26, 2007.