An Ethernet Ring Protection Method to Minimize Transient Traffic by Selective FDB Advertisement

Kwang-Koog Lee, Jeong-dong Ryoo, and Seungwook Min

ABSTRACT—We introduce an improved Ethernet ring protection method, selective filtering database (FDB) advertisement, to minimize traffic overshoot in the event of a failure or recovery. On the protection switching event, the proposed method makes all ring nodes perform an FDB flush except the FDB entries associated with their client subnets. Then, ring nodes rapidly exchange their client MAC address lists so that their FDBs are immediately updated by indirect MAC address learning. The proposed scheme guarantees fast and reliable protection switching over the standard scheme.

Keywords—Ethernet ring protection switching, FDB flush, FDB flip.

I. Introduction

Ethernet ring protection (ERP) was designed by ITU-T to achieve carrier-grade resiliency in Ethernet transport networks [1]. ERP offers full compatibility with Ethernet OAM [2] as well as standard Ethernet MAC and relay. It also supports fast automatic protection switching (APS) for Ethernet ring networks [3].

The basic idea of ERP is to use one specified link, the ring protection link (RPL), one end of which is connected to the RPL owner. When a ring network is in normal condition, the RPL owner blocks the port attached to the RPL to create a logical loop-free structure. To coordinate the activities of protection switching, ERP uses control messages called ring-APS (R-APS). When a signal fail (SF) occurs, nodes adjacent to the failed link (NAFs) block the port attached to the failed link and issue R-APS(SF) messages over both ring ports. Upon reception of an R-APS(SF) message, the RPL owner unblocks the port connected to the RPL.

Whenever the position of the block in a ring changes due to a failure or the recovery of a failure, all ring nodes should remove all learned MAC addresses from their filtering databases (FDBs). This action, called an *FDB flush*, guarantees FDB consistency for a new topology. From this point on, all ring nodes broadcast data frames until source MAC learning is completed. Duplicated frames cause a ring network to suffer from a large amount of traffic several times greater than the steady state traffic. When the traffic volume is far greater than the link capacity, the majority of frames are lost or delayed due to queuing in the buffer.

To solve this flooding problem, Rhee and others proposed a protection switching scheme called *FDB flip* [4]. The method features an immediate transition to the steady state using the flipped port information. However, as the forwarding of flipped messages should be carried out in a hop-by-hop manner, a large number of misrouted frames can be produced by the forwarding and processing delay of the flipped messages. Moreover, in order to achieve such pursuable protection switching, all ring nodes must have identical FDB entries.

We propose an enhanced protection method called *selective FDB advertisement*. In this protection scheme, all ring nodes perform an FDB flush except the entries associated with their client subnets. After the FDB flush, each node exchanges its client MAC address list with other ring nodes. Then, their FDBs are immediately updated by indirect MAC address learning which lets a ring node update its FDB as if it receives multiple individual data frames.

II. Selective FDB Advertisement Mechanism

The selective FDB advertisement mechanism under failure

Manuscript received June 8, 2009; revised July 16, 2009; accepted July 30, 2009.

Kwang-Koog Lee (phone: +82 42 860 6723, email: kwangkoog@etri.re.kr) and Jeong-dong Ryoo (phone: +82 42 860 5384, email: ryoo@etri.re.kr) are with the Broadcasting & Telecommunications Convergence Research Laboratory, ETRI, Daejeon, Rep. of Korea, and also with the Department of Engineering, University of Science and Technology, Daejeon, Rep. of Korea.

Seungwook Min (email: swmin@smu.ac.kr) is with the Division of Computer Science, Sangmyung University, Seoul, Rep. of Korea.

doi:10.4218/etrij.09.0209.0244

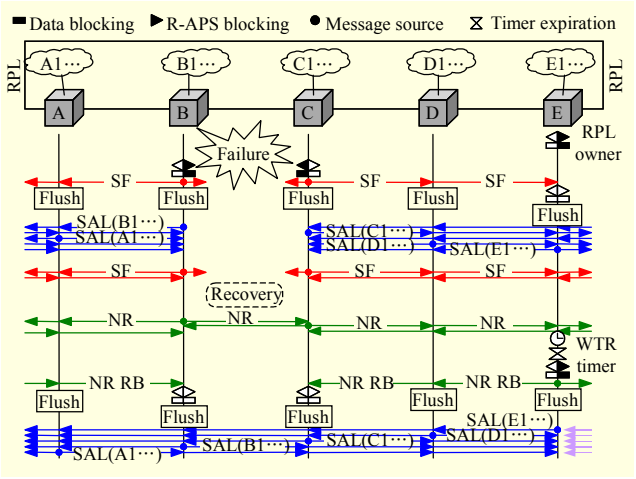

Fig. 1. Failure/recovery with selective FDB advertisement.

and recovery conditions is illustrated in Fig. 1. For loop avoidance in normal state, the RPL owner (node E) blocks its RPL port. When a failure occurs, NAFs, nodes B and C, block the port attached to the failed link and perform an FDB flush except the FDB entries associated with their subnets. Then, they multicast R-APS(SF) messages to inform all ring nodes of the failure event. Upon reception of R-APS(SF) messages, other ring nodes also recognize the failure condition and flush their FDB entries except their own client subnet MAC addresses. In addition, the RPL owner (node E) opens its blocked RPL port.

After each node receives the first R-APS(SF) message, it makes a subnet MAC address list (SAL) from its FDB table. The generated list is included in the payload of R-APS(SAL), which is newly defined to support the proposed scheme. If the SAL exceeds the maximum transmission unit size, it is fragmented resulting in multiple R-APS(SAL) messages. The R-APS(SAL) is then multicast on both ring ports or one ring port depending on whether it is an NAF. An NAF sends such a message only to a ring port opposite to the failed port to prevent FDB inconsistency by a unidirectional failure. When a node receives an R-APS(SAL), it first copies the message and sends the original to the next node as defined in the node model of the G.8032 recommendation. The node then performs MAC address learning from the SAL of the copied frame.

After recovery from the link failure, nodes B and C transmit R-APS(NR) messages to both ring ports. When the RPL owner receives the messages, it recognizes a recovery event and starts the wait-to-restore (WTR) timer to prevent frequent operation of protection switching due to an intermittent defect. When the WTR timer expires, the RPL owner blocks its RPL port and performs an FDB flush except FDB entries associated with its subnets. It then multicasts R-APS(NR, RB) messages to inform all ring nodes that the RPL is blocked. Upon reception of the first R-APS(NR, RB) message, the ring nodes flush their FDB entries except their subnet MAC addresses. Additionally, nodes B and C remove their blocks on the blocked ports. As in the case of failure, all ring nodes rapidly exchange their SAL information after they process the first R-APS(NR, RB) message and update their FDBs by performing indirect MAC address learning. The RPL owner does not send or relay an R-APS(SAL) message to the RPL port.

III. Performance Evaluation

The performance of protection switching by FDB flush, FDB flip, and the proposed selective FDB advertisement schemes was evaluated using an OPNET simulator [5]. To observe how transient traffic under a failure or recovery event affects the performance of a ring network, link utilization and the number of misrouted frames are measured every 4 ms. In Fig. 2, a simulation scenario is modeled as a 1,200 km Ethernet ring network with 16 Ethernet ring nodes (A to P). Each ring node is connected to two adjacent nodes with an 80 km 10 Gbps link. The RPL is assigned to link A-P, and the RPL port of node A (RPL owner) is blocked in normal condition. Each ring node has one subnet in which 3,000 clients reside. Each client exponentially generates 40 kbps traffic with 2,000 bits/frame toward destinations equally distributed among all subnets. To evaluate the protection switching behaviors, the link H-I is caused to fail at 2.0 s.

One R-APS(SAL) message is assumed to contain 200 MAC addresses. Fifteen R-APS(SAL) messages are generated by each node in a protection event. It is also assumed that each switch processes data frames and R-APS frames separately. The service rates of data and R-APS frames were set to 6.5 mpps (million packets per second) and 0.1 mpps, respectively. The data service rate is set according to the commercial metro Ethernet switch product. As an R-APS(SAL) frame contains

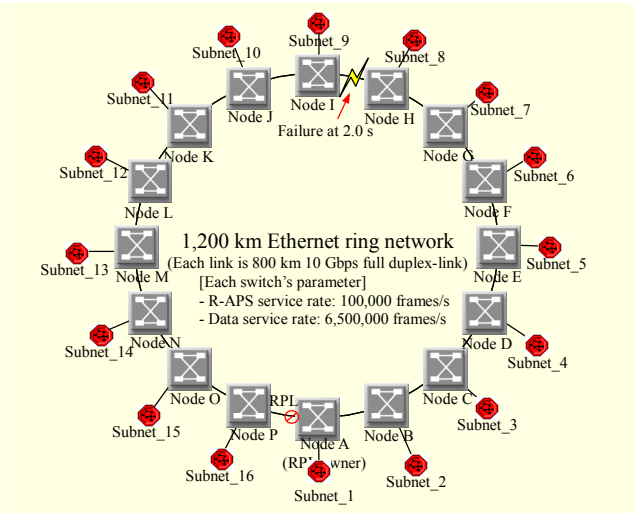

Fig. 2. Simulation scenario.

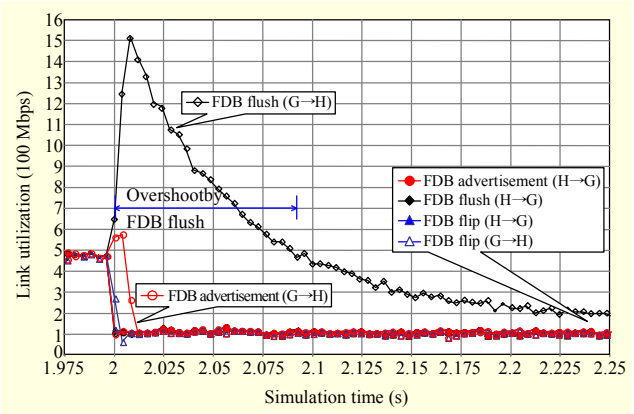

Fig. 3. Link utilization at link H-G under failure condition.

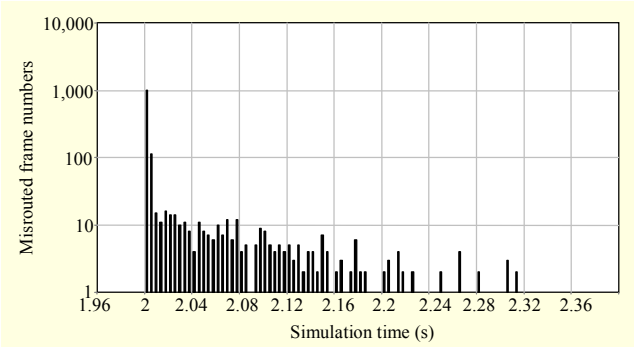

Fig. 4. Number of misrouted (lost) frames in FDB Flip.

lengthy SAL information compared to other normal R-APS frames, the processing time of one R-APS(SAL) message is set to be the sum of the service time of one R-APS frame and the processing time of all the address entries in the SAL. The processing time of one entry in the SAL is set to half of the *service time per data frame* as a data frame normally experiences two FDB accesses, DA lookup and SA learning. The processing time for FDB flip is set in the same manner.

As shown in Fig. 3, the FDB flush scheme introduces a large traffic overshoot due to flooding. At the peak, the link utilization of link G→H at protection switching surges up to three times more than that in normal state. Before failure, link G→H carries the heaviest traffic among all links. The overshoot lasts for 100 ms, and transient traffic cannot reach steady state even after 200 ms. FDB flip shows the fast protection switching time (only a few milliseconds) over other schemes. However, as seen in Fig. 4, FDB flip introduces frame misrouting due to the hop-by-hop processing of R-APS (SF, Flip) messages. Because a node forwards the R-APS(SF, Flip) message to the next node after completing modification of the flip list, frames are misrouted until all the nodes finish updating their FDBs. A misrouted frame is filtered and lost at the node with a blocked port. At the time of 2.0 s, more than a thousand frames are misrouted, and the misrouted frames occur even in 300 ms after failure. If all ring nodes do not maintain identical MAC addresses in their FDB table, the misrouted frames can persist until the mismatched FDB entries expire.

The proposed scheme, FDB advertisement, shows that transient traffic is stabilized in about 10 ms at link G→H, and the amount of overshot traffic is far less than with FDB flush. There are no misrouted frames with FDB advertisement because a selective flush operation is performed as soon as the failure is propagated over the ring.

When a ring node does not have the FDB entry for the destination of a data frame due to any late R-APS(SAL) message, it can always deliver the frame by flooding. Moreover, even though all ring nodes do not maintain identical MAC addresses in their FDB tables, the proposed scheme prevents FDB inconsistency at all times because it performs the FDB flush by the first received R-APS(SF) message. When there is any lost R-APS(SAL) message, flooding of data frames occurs for those addresses in the lost message. However, the loss does not introduce any FDB inconsistency.

In general, the entry created by an upper-layer protocol is not aged out of FDB table. It should be noted for implementation that the FDB entries obtained from an R-APS(SAL) message need to have a finite expiration time as in the MAC learning from a data frame.

IV. Conclusion

In this letter, a transient traffic overshoot issue in Ethernet ring protection switching was discussed. To minimize traffic overshoot caused by FDB flush, we proposed an approach called selective FDB advertisement. Simulation results demonstrated that the proposed scheme guarantees fast protection switching. Since the proposed method fully complies with the node architecture and R-APS protocol message relay model defined in the ITU-T G.8032 recommendation, it can be an effective protection solution for Ethernet ring networks.

References

[1] ITU-T Rec. G.8032, *Ethernet Ring Protection Switching*, ITU-T, Geneva, 2008

[2] J. Ryoo et al., "OAM and Its Performance Monitoring Mechanisms for Carrier Ethernet Transport Networks," *IEEE Commun. Mag.*, vol. 46, no. 3, Mar. 2008, pp. 97-103.

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

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

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