

Local Restoration in Metro Ethernet Networks for Multiple Link Failures

Shibu. V

Department of Computer Applications,
Cochin University College of Engineering, Pulincunoo.
Alappuzha, Kerala, India

Preetha Mathew K

Department of Computer Applications,
Cochin University College of Engineering, Pulincunoo.
Alappuzha, Kerala, India

Jabir.K.V.T

Department of Information Technology,
Cochin University College of Engineering, Pulincunoo.
Alappuzha, Kerala, India

Abstract: Ethernet a popular choice for metropolitan-area networks (MAN) due to simplicity, cost effectiveness and scalability. The Spanning-Tree based switching mechanism, which is considered to be very efficient at avoiding switching loops in LAN environment, is a performance bottleneck in Metro network context. Handling of link failure is an important issue in metro Ethernet networks. A link failure may result in serious service disruptions. A local restoration method for metro Ethernet with multiple spanning trees, which aims at fast handling of single link failures in a distributed manner, have been proposed in the literature. In this paper, we propose a local restoration mechanism that uses MULTILINK algorithm for solving multiple link failures

Keywords: Metropolitan Area Networks (MAN), Ethernet, Spanning Tree Protocol, RSTP.

1. INTRODUCTION

Ethernet is a family of computer networking technologies for local area networks (LANs). Systems communicating over Ethernet divide a stream of data into individual packets called frames. Each frame contains source and destination addresses and error-checking data so that damaged data can be detected and re-transmitted. Ethernet has evolved over the past decade from a simple shared medium access protocol to a full-duplex switched network. Ethernet dominates current local area network (LAN) realizations. It has been estimated that more than 90 percent of IP traffic originates from Ethernet LANs. Efforts are underway to make Ethernet an end-to-end technology spanning across LANs, metropolitan area networks (MANs), and possibly wide area networks (WANs) [2].

A Metro Ethernet [1] is a computer network that covers a metropolitan area and that is based on the Ethernet standard. It is commonly used as a metropolitan access network to connect subscribers and businesses to a larger service network or the Internet. Metro Ethernet network is a set of interconnected LANs and access networks that work together using Ethernet technologies to provide access and services within a metro region. Metro Ethernet networks are built from Ethernet switches/ bridges interconnected by fiber links. A spanning tree protocol is used to establish one or more trees spanning every access point that connects LANs.

Failure handling is a key issue in metro Ethernet networks. A component failure may result in serious service disruptions. To support carrier-grade services in MANs using Ethernet, it is a critical requirement to have a fast, reliable, and efficient failure-handling mechanism [3]. Current Ethernet switched networks use the spanning tree protocol family without any fast failure recovery mechanism. The IEEE 802.1d Spanning Tree Protocol (STP) [4] establishes a single spanning tree to guarantee a unique path between any two switches. It suffers

from low convergence speed and inefficient bandwidth usage in case of a failure.

The spanning tree approach fails to exploit all the physical network resources, because in any network of N nodes there are at most N-1 links actively forwarding traffic. This produces an imbalance of load in the network. This scenario is impractical in large scale networks like metro networks. Further, switch and link failures require rebuilding of the spanning tree, which is a lengthy process. IEEE 802.1w [5], the rapid spanning tree configuration protocol (RSTP), mitigates this problem by providing mechanisms to detect failures and quickly reconfigure the spanning tree. However, the recovery period can still range from an optimistic 10 milliseconds to more realistic multiple seconds *after* failure detection, which is still not adequate for many applications.

In this paper, we propose a local restoration mechanism for metro Ethernet, which aims at fast handling of multiple link failures in an efficient manner. Multiple link failures come from the fact that when a connection is switched to a backup spanning tree, it has no record of the original working spanning tree. Therefore, when the connection encounters another failure on the backup spanning tree, there is a possibility that it would be switched back to the original working spanning tree and form a loop in the network. However, when multiple link failure happens in the network and both the primary and backup spanning tree fail simultaneously, some packets would be dropped when they encounter the failure on the backup spanning tree. We propose a possible approach to handle these multiple link failures. This approach is to allow multiple VLAN switching and add more information in the header of the frames, e.g., VLAN ID of the original working spanning tree, when the frames are switched to a backup tree. Thus, they are able to select a backup tree without forming a loop when they are affected by the second failure in the network.

2. PROBLEM DEFINITION

The simplicity and the low cost provided by Ethernet makes it an attractive network technology choice in networking application deployments. The Spanning-Tree Protocol (STP), which is proposed in initial version of IEEE 802.1D [4], is responsible for building a loop-free logical forwarding topology over the physical one providing connectivity among all nodes. The links that are not part of this tree are blocked. In case of a failure, the blocked links are activated providing a self-healing restoration mechanism. All information propagated between the switches is embedded in Bridge Protocol Data Units (BPDUs). These packets are exchanged only between adjacent bridges, and protocol events (e.g., port state changes) are invoked by timers, thus rebuilding the topology takes considerable time. This timer based operation, which is an STP property, results in reconfiguration times up to 50 seconds and, thus, affects network performance. The existing system defines a local restoration mechanism for metro Ethernet using multiple spanning trees, which is distributed and fast and does not need failure notification. Upon failure of a single link, the upstream switch locally restores traffic to preconfigured backup spanning trees. There are two restoration approaches, connection-based and destination-based, to select backup trees. When multiple link failure happens in the network and both the primary and backup spanning tree fail simultaneously, some packets would be dropped when they encounter the failure on the backup spanning tree. The proposed system defines a possible approach to handle these multiple link failures. This approach is to allow multiple VLAN switching and add more information in the header of the frames, e.g., VLAN ID of the original working spanning tree, when the frames are switched to a backup tree. Thus, they are able to select a backup tree without forming a loop when they are affected by the second failure in the network.

3. EXISTING METHODS

3.1 Metro Ethernet Local Restoration Framework

The existing system define a local restoration mechanism[1] in metro Ethernet that selects appropriate backup spanning trees for rerouting traffic on a working spanning tree. Then restores the traffic to the backup spanning trees in case of failure locally. The path on a backup tree to reroute traffic is from the immediate upstream node of the failed link to the destination node and should exclude the failed link. Using the current Ethernet protocols the local restoration mechanism in metro Ethernet can be implemented. For this, an additional module should be maintained in the Ethernet switch for checking the appropriate backup spanning tree and restoring frames to the backup tree after failure.

The existing system has implemented the local restoration mechanism in the following three steps.

3.1.1 Per VLAN Spanning Tree:

Local restoration from one spanning tree to another can be implemented by assuming in each spanning tree is assigned a dedicated virtual LAN (VLAN) ID [6]. The pre-calculated spanning tree topologies are implemented in the network by means of VLANs, which do not change during network operation and ensure that there are no loops in the Ethernet network. Therefore, STP is disabled, as it is not needed to provide loop free topology. A unique VLAN ID is assigned to each spanning tree, which is used by the edge routers to forward traffic over the appropriate trees [3]. By changing the

VLAN ID on the Ethernet header, Ethernet frames can be switched among spanning trees. Frames that frequently switch among spanning trees may form unexpected loops. To allow VLAN switching once, set a bit in the frame's class-of-service (CoS) field as the restoration bit. The Ethernet switch will check a frames restoration bit before restoration and drop those frames that have been earlier restored once.

3.1.2 Local Restoration Mechanism:

Local restoration doesn't require the convergence of spanning trees after failure. To inform that the switch is alive, each switch periodically sends a message to its neighbors. Within a predefined interval, if a switch does not receive any message from a port, it changes the port's status as "failed". The restoration module is activated when an incoming frame is forwarded to the failed port; a preconfigured backup VLAN ID replaces the frame's original VLAN ID. At the same time, its restoration bit is set to 1. Then, the modified frame is forwarded to the alternative output port according to its new VLAN ID.

3.1.3 Local Restoration Mechanism:

The network manager will perform the pre configuration operation which includes three parts: multiple spanning trees generation, working spanning tree assignment, and backup spanning tree configuration.

Multiple Spanning Trees Generation: The network manager is responsible for generating multiple spanning trees [8] *a priori*. The trees should satisfy the condition to handle a link failure: For each link, there is at least one spanning tree that does not include that particular link [7]. Commonly, more spanning trees should be generated to utilize network resources efficiently.

Working Spanning Tree Assignment: The network manager should assign a VLAN ID to each source and destination (s-d) pair based on the long-term traffic demand matrix. The frames entering the network are attached with VLAN IDs at the ingress switches according to their source and destination addresses and are forwarded to the proper working spanning trees.

Backup Spanning Tree Configuration: Frames traversing the failed link should be restored to a preconfigured backup spanning tree locally when a failure happens. Backup trees at each switch should be carefully configured according to the traffic demand such that there is enough spare capacity on the backup tree for restoration.

Backup Tree Selection Strategy

The Ethernet switch selects the backup tree for each frame according to the backup tree configuration by the network manager. The existing system uses two backup spanning tree selection strategies: connection-based and destination-based. The traffic between a source-destination pair is termed as a connection. In connection-based strategy an Ethernet switch determines the incoming frame's backup VLAN ID according to its source address, destination address, and original VLAN ID. Therefore, traffic between different source-destination pairs traversing the same failed link may be restored to different backup trees

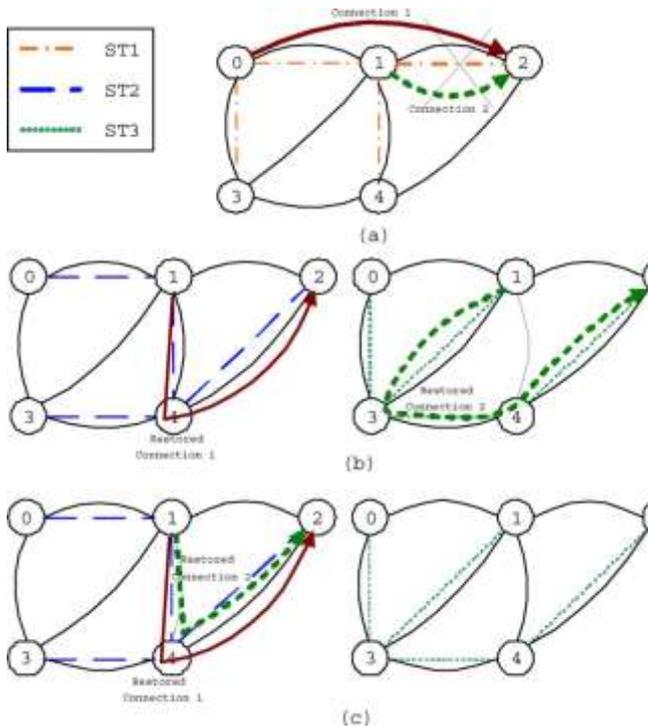


Fig.1 Backup tree selection strategy. (a) Two connections on ST1 before failure of link 1–2. (b) Two connections are restored to different STs after failure in connection-based strategy. (c) Two connections are restored to the same ST in destination-based strategy.

Connection 1 from node 0 to 2 and connection 2 from node 1 to 2 uses ST1 as the working spanning tree before failure [Fig. 1(a)]. When the link between 1–2 fails, node 1 restores connection 1 to ST2 and connection 2 to ST3 according to the pre configuration. According to their source and destination MAC addresses frames are restored and different connections are assigned independent backup spanning trees. Connection-based backup tree selection requires a complex computation during pre configuration and per-(source–destination) pair information should be maintained by each switch.

The existing system uses a destination-based backup tree selection strategy, in which the frame’s backup VLAN ID is determined by its destination address and original VLAN ID, regardless of its source address. Frames with the same VLAN ID and destination address would use the same backup tree in a local Ethernet switch. Fig. 1(c) shows an example of the destination-based backup tree selection strategy. Connections 1 and 2 have to use the same backup spanning tree in node 1 upon failure of link 1–2 since they have the same destination which is different from the connection-based strategy. Node 1 can only restore the two connections with the same destination to the same spanning tree.

4. PROPOSED METHODOLOGY

When a connection is switched to a backup spanning tree and when the connection encounters another failure on the backup spanning tree, multiple link failure happens in the network. Both the primary and backup spanning tree fail simultaneously, some packets would be dropped when they encounter the failure on the backup spanning tree.

The existing system only handles single link failures in the metro Ethernet. The proposed system is an enhancement of the existing system. This system defines a possible approach to handle these multiple link failures. The approach is to allow

multiple VLAN switching and add more information in the header of the frames, e.g., VLAN ID of the original working spanning tree, when the frames are switched to a backup tree. Thus, they are able to select a backup tree without forming a loop when they are affected by the second failure in the network. It also uses the concept of local restoration mechanism provided in the existing system in its restoration module. The restoration module uses an algorithm namely MULTILINK. The working of this proposed algorithm is demonstrated in the next section.

The MULTILINK algorithm

This algorithm considers three connections. We named these connections as TLink, Blink and RLink respectively. In the first case consider all the three links are up. Two bits in the Ethernet IP packet header are used in this algorithm. The first is the restoration bit (RB) that is used in the existing system. The second bit is termed as new bit (NB) in the proposed restoration module, which will be assigned by a value 1 or 0 according to various conditions in the algorithm. After considering that all the three links are up, assign value zero to RB and NB in the second step. If the TLink is down then set the value of RB to one and set the value of NB to zero. Then switch to the Blink by activating the restoration module as in the existing system. Now the traffic is going through the Blink. Check whether the TLink is up frequently within a fixed time interval of 5Ms. If the TLink is up then set the value of the bit NB to one. Otherwise set the value of the bit NB to zero. If the Blink is up, repeat the said operations by checking TLink is down or not. If the Blink is down then check whether the value of NB. If the value of NB is equal to one then route the packet through TLink otherwise route the packet through RLink by using Rapid Spanning Tree Protocol (RSTP). The working of MULTILINK algorithm is demonstrated in the following Figures.

5. EXPECTED RESULTS

The most important design objectives of a failure handling mechanism are fast failover, simplicity, robustness and low protocol processing and transport overheads. Ethernet has built-in functionalities for failure handling developed in standardization bodies. When a connection is switched to a backup spanning tree and when the connection encounters another failure on the backup spanning tree, multiple link failure happens in the network. Both the primary and backup spanning tree fail simultaneously, some packets would be dropped when they encounter the failure on the backup spanning tree. When an Ethernet switch finds a packet that has been rerouted once and its output port on the backup tree also fails, the switch should notify the network manager or broadcast the failure message asking for spanning tree re convergence by RSTP. The algorithm invokes RSTP only when the last condition NB is equal to one doesn’t satisfy. The MULTILINK algorithm described in the proposed system, on implementation, can solve the multiple link failure efficiently

Algorithm is defined as follows,

Algorithm MULTILINK

1. Initially all the three links TLink, Blink and RLink are considered as in up state.
2. If TLink is down then
 - A) set RB=0 and NB=0
 - B) Switch to Blink by activating the restoration module.
 - C) Check whether TLink is up frequently by setting a time limit of 5ms.
 - (i) if TLink is up then set NB=1
 - else
 - (a) set NB=0
 - (b) if Blink is down then go to step 3
 - else go to step 2.
3. If NB=1 then
 - Route the packet through TLink
- Else
 - Route the packet through RLink.

Fig.2 Algorithm

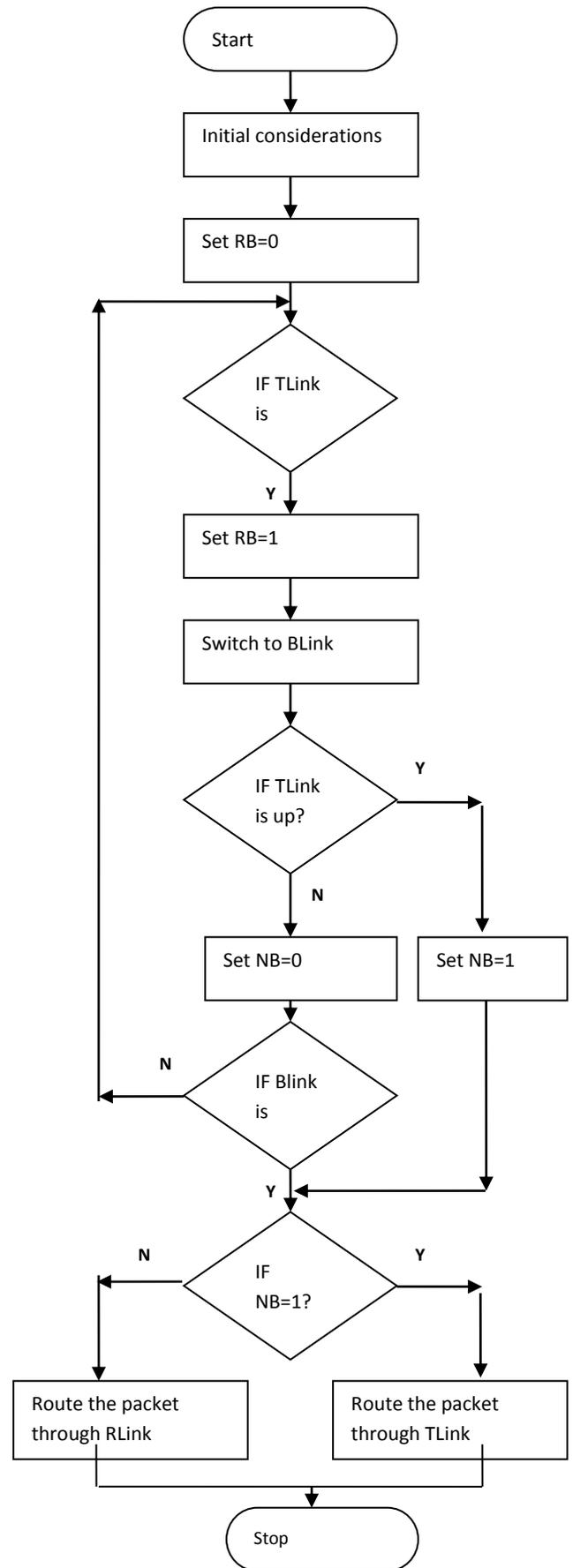


Fig.3 Flowchart working of multi link algorithm

6. CONCLUSION

The existing system only handles single link failures in the metro Ethernet. The proposed system is an enhancement of the existing system. This system defines a possible approach to handle these multiple link failures. It also uses the concept of local restoration mechanism provided in the existing system in its restoration module. The restoration module uses an algorithm namely MULTILINK. Two bits in the Ethernet IP packet header are used in this algorithm. The occurrence of multiple link failure is a rare event. When implemented properly, the proposed system solves the problem of multiple link failures in metro Ethernet network.

7. REFERENCES

- [1] Jian Qiu, Mohan Gurusamy, “Local Restoration With Multiple Spanning Trees in Metro Ethernet Networks”, *IEEE/ACM Transactions On Networking*, Vol. 19, No. 2, April 2011
- [2] A. Meddeb, “Why Ethernet WAN transport,” *IEEE Commun. Mag.*, vol. 43, no. 11, pp. 136–141, Nov. 2005.
- [3] C. Antal, L. Westberg, A. Paradisi, T. R. Tronco, and V. G. Oliveira, “Fast failure handling in Ethernet network,” in *Proc IEEE ICC*, 2006, vol. 2, pp. 841–846.
- [4] Standard for Local and Metropolitan Area Networks—Media Access Control (MAC) Bridges, IEEE 802.1d, 1998.
- [5] Standard for Local and Metropolitan Area Networks—Rapid Reconfiguration of Spanning Tree, IEEE 802.1w, 2001.
- [6] Standard for Local and Metropolitan Area Networks—Virtual Bridged Local Area Networks, IEEE 802.1q, 1999.
- [7] J. Farkas, C. Antal, G. Toth, and L. Westberg, “Distributed resilient architecture for Ethernet networks,” in *Proc. DRCN*, 2005, pp. 515–522.
- [8] K. Goplan, S. Nanda, and T. Chiueh, “Viking: A multiple-spanning tree Ethernet architecture for metropolitan area and cluster networks,” in *Proc. IEEE INFOCOM*, 2004, pp. 2283–2294.