Etherlay: An Overlay Enhancement for Metro Ethernet Networks

Minh Huynh and Prasant Mohapatra
Computer Science Department
University of California at Davis
Davis, CA. USA.
{mahuynh, pmohapatra}@ucdavis.edu

***Abstract*— The ubiquitous Ethernet technology has propelled itself into a wide-scale adoption for Metro Ethernet Networks (MEN). Despite recent advancements in Ethernet and commercialization of the first generation of MEN, the fundamental technology does not meet the expectations that carriers have traditionally held in terms of network resiliency and load management. This paper addresses these two issues. We propose a new concept of overlay network in the Ethernet layer, called Etherlay, that increases the resiliency of the MEN while provisioning the support for load balancing. As a result, the capacity in terms of network throughput is greatly enhanced while almost avoiding performance hits for any re-convergence in case of failures. Compared to the standard protocols, Etherlay's total throughput gain ranges from 5.93% to 20.7% in the face of failure; while load balancing capability increases an additional 16% to 60% of the total throughput.**

Keywords- Etherlay, Metro Ethernet Network, MSTP, Overlay, Resilience, RSTP.

I. INTRODUCTION

For over thirty years, Ethernet has been the predominant technology for Local Area Network (LAN). Ethernet is a simple and cost-effective protocol that provides a variety of services. The recent advancements for Ethernet protocol have propelled it for consideration in the scope of Metropolitan Area Networks (MAN). Moreover, due to increasing customer demands, several companies are focusing their developments, products, and services for Metro Ethernet Networks (MEN) [13].

MENs [12] have several access networks connecting to a core metro network, and the customers' networks are connected to the access networks. In effect, metro core helps in interconnecting the customers' remote sites as if they are in the same LAN. Packets hop through multiple switches in both access and metro core networks. Redundant links are used both in the core as well as the access networks.

The main challenges in the context of MEN include resiliency, load balancing, and support for QoS [5][6][12]. Current Ethernet solutions deploy the Spanning Tree Protocol and its variants to manage the topology autonomously. However, they are inadequate in all three areas. In this work, we address the resiliency and load balancing aspects of MEN using the concept of overlays.

Overlaying is a useful technique in supporting new services without modifying the underlying infrastructure. Since there are numerous devices currently running Ethernet, it would not be cost effective to redesign it completely. Complications also arise with compatibility between different technologies. Overlaying fits best in this case when we want to incorporate new and better services that are transparent to the current devices.

In an overlay scheme, a subset of nodes from all the nodes is chosen. Transparent to the underlying physical connections, the connections between overlay nodes are formed by overlay links that operate in the application layer. Each overlay link can be composed of one or more physical connections. Examples of overlay networks at the IP layer are Resilient Overlay Network (RON) [16], Service Overlay Network (SON) [17], QoS-aware routing for Overlay Networks (QRON) [18], and OverQoS [19].

We have introduced an overlay approach, called Etherlay, in the Ethernet layer. This feature enhances the resiliency as well as facilitates load balancing. In addition to fast recovery, it also increases the capacity of the network in terms of the achievable throughput.

The encouraging experimental results from Etherlay presented in this paper were obtained using the OPNET [11] simulation product to quantify the resiliency and the gain in terms of the network throughput. The behaviors of the Ethernet switches within OPNET Modeler were modified to imbue the Etherlay approach. In the resilience test scenarios, Etherlay yields an increase of 5.93% to 20.7% over Multiple Spanning Tree Protocol (MSTP) and Rapid Spanning Tree Protocol (RSTP). Likewise, Etherlay gains an additional of 16% to 60% of the total throughput comparing to MSTP and RSTP.

The organization of the paper is as follows: Section II describes the spanning tree protocol family that is currently deployed in Ethernet and their limitations. Section III explains the concept of Etherlay. Section IV illustrates the simulation setup. Section V and VI show Etherlay in action against failures and load balancing compared to RSTP and MSTP. Related works are presented in section VII, and the paper is concluded in Section VIII.

II. BACKGROUND

Traditionally, Ethernet-based networks use the IEEE 802.1D Spanning Tree Protocol (STP) [1] for forwarding packets in the network. The standard STP is a layer2 protocol that can be implemented in switches and bridges. STP essentially uses a shortest-path approach in forming a tree that

is overlaid on top of the mesh-oriented Ethernet networks. Primarily, spanning tree is used to avoid the formation of cycles or loops in the network. Unlike IP packets, Ethernet data frames do not have a time-to-live field. To prevent loops in the network, STP blocks redundant links. Therefore, the load is concentrated on a single link which leaves it at risk of failures and with no load balancing mechanism. The root of the tree is chosen based on the bridge priority; and the path cost to the root is propagated throughout so that each switch can determine the state of its ports. Only the ports that are in the forwarding state can forward data frames ensuring a shortest single path to the root. Whenever there is a change in the topology, switches rerun the protocol that can take 30 to 60 seconds. At any time, only one spanning tree dictates the network.

Although STP has been used for most Ethernet networks, it has several shortcomings in the context of its use for MEN. These shortcomings are enumerated as follows:

1. Spanning trees restrict the number of ports being used. In high-capacity Ethernets, this restriction translates to a very low utilization of the network.
2. STP has poor resiliency: a very high convergence time (30s to 60s) after a link failure.
3. STP does not have any mechanisms to balance load across the network.
4. STP does not support QoS.

An improvement of STP is the Rapid Spanning Tree Protocol RSTP [2] standardized in IEEE 802.1w. RSTP reduces the number of port states from five to three: discarding, learning, and forwarding. Through faster aging time and rapid transition to forwarding state, RSTP is able to reduce the convergence time to between 1 and 3 seconds. It is understood that depending on the network topology, this value varies. In addition, the topology change notification is propagated throughout the network simultaneously, unlike STP, in which a switch first notifies the root, then the root broadcast the changes. Similar to STP, only one spanning tree dictates the whole network. RSTP still blocks redundant links to ensure loop free paths leaving the network underutilized, vulnerable to failures, and with no load balancing.

MSTP or Multiple Spanning Tree Protocol [4] is defined in IEEE 802.1s. Using a common spanning tree, MSTP connects all of the regions in the topology. The regions in MSTP are multiple instances of the spanning tree where each instance is an instance of the RSTP. An instance of RSTP governs a region with its own regional root. The regional roots are in turn connected to the common root that belongs to the common spanning tree. Since MSTP runs pure RSTP as the underlying protocol, it inherits some drawbacks of RSTP as well. However, a failure in MSTP can be isolated into a separate region leaving the traffic flows in other regions untouched. In addition, the administrators can perform static load balancing manually by assigning certain flows to a specific spanning tree.

III. Conceptual Approach To Etherlay

In this section, we describe the basic philosophy behind the Etherlay protocol and its potential for provisioning enhanced performance and services.

A. *Etherlay Overview*

Our approach will be to create an overlay over Ethernet through the virtual peering concept for enhancing resiliency and enriching flexibility in metro Ethernet services. The proposed framework, called Etherlay, requires changes only at the edge switches leaving the core switch architecture unchanged. Therefore, it is transparent to the underlying Ethernet forwarding approach. Using the concept of virtual peering, we can enhance the network utilization by providing multiple configurable paths instead of a dedicated one. Hence, virtual peering allows the creation of dynamic back-up soft pipes to provide resilience for the network. In addition, the network load is distributed, lessening the load on the critical links. Etherlay provides smart edge switches that will be used as traffic policing. Furthermore, it can be incrementally deploy and is scalable.

Figure 1 demonstrates an example of this approach. The solid path indicates the primary path between A and E that is provided by the spanning tree protocol. The dashed paths are the overlay paths: A-D-E and A-G-E. Assume that the solid path is faulty; an appropriate overlay path is chosen to enhance the resiliency. D and G act as virtual peers for sending packets to E on behalf of A. Overlay paths use the transparent underlying forwarding scheme. The overlay path AD, DE, AG, and GE are spanning tree paths that could belong to different spanning trees so that the positive properties of the spanning tree protocol (loop-free) are still preserved. Encapsulation technique can be used for tunneling the frames through the overlay.

B. *Implementation Details*

When implementing Etherlay, there are other issues that must be taken into consideration. First, Etherlay must be backward compatible with the current protocol. To do this, we leverage the MSTP protocol to implement the functionality and operations needed by Etherlay. Since MSTP is backward compatible with RSTP and STP, Etherlay can interoperate with them and MSTP as well. Thus, Etherlay retains the advantages of MSTP while providing enhanced features in terms of resiliency, and load balancing.

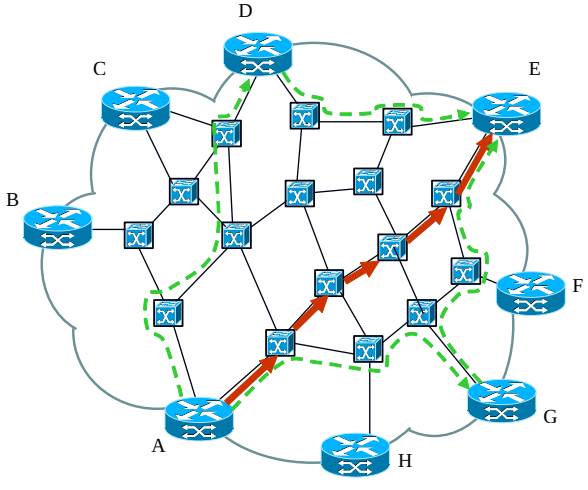

Figure 1 Etherlay approach with the solid line being the primary path provided by spanning tree protocol and the dotted lines are the overlay paths.

The decision of when to use overlaying is made at the ingress nodes. In Etherlay, each overlay path can be configured as a separate VLAN. Therefore, it is possible to optimize the path to an overlay node. Another configuration is to collect a set of overlay paths into a VLAN. The ingress nodes keep track of which VLAN is operational. Initially, the ingress nodes send all traffic on the primary VLAN. They update the availability of the VLANs through the Bridge Protocol Data Unit (BPDU) packets. If the primary VLAN is unsatisfied for the requested traffic, the ingress nodes encapsulate each data frame with a new header that includes the VLAN ID of the overlay VLAN, the MAC of the overlay peer node as the destination, and a new value for the type field in the Ethernet header distinguishing it from the regular Ethernet frame. When the overlay peer receives the encapsulated frame, it strips the encapsulated header, modifies the VLAN ID of the original frames, and sends the frame to the destination based on this new overlay VLAN.

Unlike overlay approaches at the IP layer, Etherlay uses a local estimator instead of packet probing. Local estimator is used to determine the congestion condition of the network in order to perform load balancing if necessary. Each ingress node measures the load percentage on the intended outgoing link for a given frame. If the load percentage reaches the link load threshold, the ingress node assumes that the primary spanning tree is congested. Then, it will send the frame using the overlay path. In contrast to probing, the local estimator does not know the exact condition of the network end to end. Using the local links that are attached to it, it will try to estimate the load congestion overall. The assumption is that since Etherlay starts in one primary spanning tree, the traffic eventually will go through some bottleneck link(s) where many sources aggregated into it. This is the nature of the spanning tree algorithm as exhibited by STP and RSTP. Therefore, by measuring the load at the source, a switch estimates the load of the entire tree.

The original intension of having VLAN [3] tags is for isolation of traffic. When Etherlay uses VLAN tags as id for overlay path, VLAN's initial objective is still preserved. Instead of mapping a VLAN ID to a traffic group, Etherlay now uses more than one VLAN to map to a traffic group: one VLAN for the primary spanning tree path and the subsequence VLANs is for the overlay paths. The VLAN partition is implementation dependent. The shortage of VLAN ids can be an issue but there are proposals to perform VLAN stacking or Q-in-Q [14][15]. This technique increases the number of VLAN tag from 2^{12} to 2^{24}. From here on, the term VLAN and spanning tree are used interchangeably since each VLAN will run a separate spanning tree.

IV. SIMULATION DESIGN

The OPNET [11] simulator tool was chosen because of its comprehensive implementation of Ethernet. OPNET includes implementation for RSTP, MSTP, and VLAN which are crucial to the evaluation of Etherlay.

Etherlay will be evaluated on a 6x6 grid topology as seen in Figure 2. RSTP has one spanning tree operating in the 6x6 grid topology. The root of the tree is located at **node_14** which is the center of the topology. Conversely, MSTP and Etherlay are configured with six spanning trees. The common root is at **node_0**. The regional roots for spanning tree 1 through spanning tree 6 are set diagonally in the topology from top left to bottom right, respectively, indicated by a diamond shape background shown in Figure 2. As specified in the Spanning Tree Protocol standard, there is a maximum of 7 hops. In order to form a stable spanning tree for a topology of this size, the "hop count" parameter is increased. The source switches are indicated by the attachment of the end hosts in Figure 2. The overlay peers are indicated in Figure 2 by a circle shape background.

Within this topology, resilience and load balance will be evaluated. Although they are being evaluated separately to prevent any interference, they can work well together. The description and specific parameters used are included. The resilience test in this topology includes both node failures and link failures. Whenever there is a node failure, all the links attached to the node also fail.

A. Failures Scenario

There are four flows to each of the server in Figure 2. Each flow is a bi-directional UDP stream of approximately 4.3Mbps. All links have capacity of 100Mbps. This capacity is enough to carry the traffic without causing any congestion. The simulation runs for 200s and all traffics start at 100s. The link failures and node failures/recovery are scheduled as follows (<–> is used to denote a link connecting two nodes):

110s: node_13 fails
: node_8 <–> node_9 fails
: node_7 <–> node_31 fails
: node_1 <–> node_7 fails
140s: node_8 fails
: node_0 <–> node_1 fails
: node_20 <–> node_26 fails
: node_17 <–> node_16 fails
170s: node_13 recovers
: node_15 <–> node_16 fails
: node_21 <–> node_22 fails

In MSTP scenario, the flows are grouped by the destination to put into the corresponding spanning tree. For example, since S1, S2, S3, and S4 are going to server2, they run on the same spanning tree. Similar arrangement is made for B{1,2,3,4} and G{1,2,3,4}. However, all sources started in the same tree in Etherlay.

B. Unbalanced Load Scenario

The setting in this load balancing experiment is similar to that which was used for the resilience experiment (seen in Figure 2). However, the links are now 10Mbps. The 10Mbps links efficiently congest the network without exhausting the available computer resources needed to run the simulation. Each flow is a bi-directional UDP stream of approximately 4.3Mbps. The simulation ran for 200s and all traffic starts at 100s. Failures are not scheduled in this scenario.

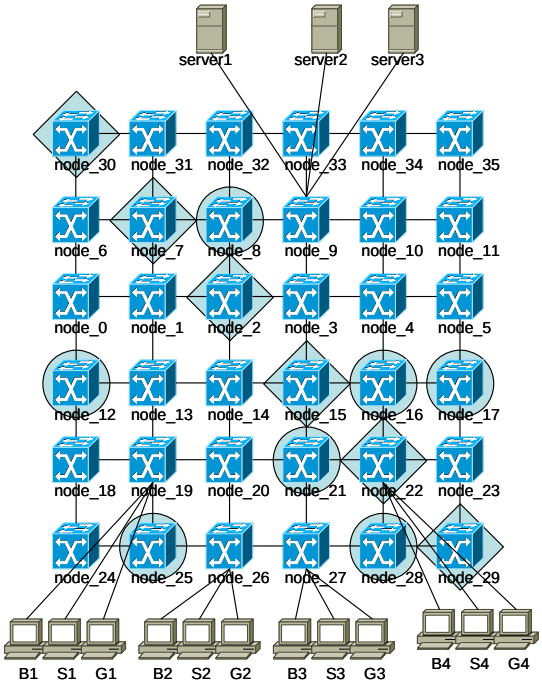

Figure 2 A 6x6 grid topology. Cicrle indicates the overlay peers. Diamond indicates the regional roots.

V. ENHANCED ETHERNET RESILIENCE

As alluded to earlier, resilience is of particular importance for carriers; and this is one area for which Ethernet is well recognized as being very weak. Etherlay was specifically formulated to address the inherent weakness of Ethernet resilience. An experiment is presented in this section in which RSTP, MSTP and Etherlay are evaluated for their resilience in the face of link failures and recoveries.

A. Performance in Metro Area Network Topology

This subsection shows the performance of each protocol in the face of failures separately. Then a superimposed graph shows how they are stacked up against each other. Each graph shows the combined throughput receiving at the three servers.

1) RSTP

In RSTP, each dip in the graph in Figure 3 corresponds to the failure or recovery events at time 110s, 140s, and 170s, respectively. Since there is only one spanning tree operating in the network, each performance hit reduces the throughput to zero. On average, each turn-around takes 8 seconds. The turn-around time includes failure detection and reconvergence.

2) MSTP

Figure 4 shows the performance of MSTP. Since there are six active spanning trees, the failures can be isolated to affect certain spanning trees. By distributing the traffic as described in section IV.A, the performance hit is limited to affected spanning trees allowing other traffic to go through. As shown in Figure 4, the performance hits do not go to zero as in the case of RSTP. The turn around time is still eight seconds.

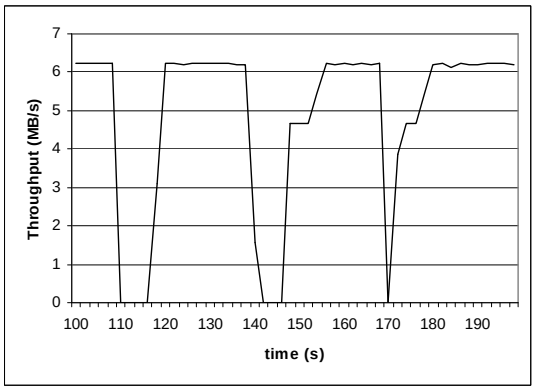

Figure 3 The throughput as observed by the receiving servers during the failures for RSTP

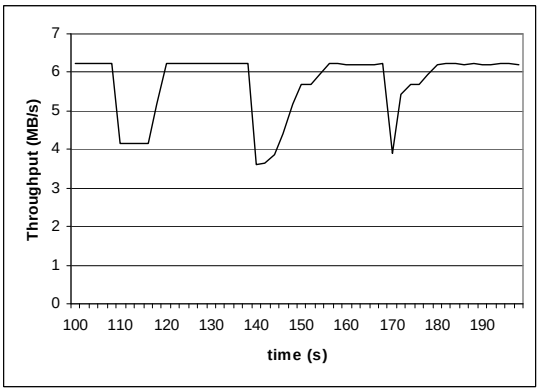

Figure 4 The throughput as observed by the receiving servers during the failures for MSTP

3) Etherlay

Etherlay displays only one performance hit in all scheduled failure events. The only dip on the graph is the result of the failures at 110s. Since the ingress nodes handle the decision of whether to use overlaying, it needs information on the stability of the current VLAN. Therefore, there is a propagation delay associated with the turn around time. For example, at 110s, the failed components are further away from ingress nodes: **node_26**, **node_27**, and **node_22**. Thus, there is a delay before the ingress nodes receive the failures notification and send to alternative overlay peers. Since **node_19** is closer to the failed components, it does not contribute significant delay to the turn around time; hence, the dip at 110s was not impacted by the loss of traffic from **node_19** as the majority of traffic from **node_19** was not lost. That is why the dip did not go to zero. For the failed events at 140s and 170s, the failed components are closer to the ingress nodes. Therefore, they are able to handle the failures reverting to the overlay path before any loss of traffic interrupts the flows as shown in Figure 5.

The turn around time for Etherlay is also shorter because the overlay routing overlaps and hides the reconvergence time. The performance hit is caused only by the latency of the failure detection. This is shown in Figure 6 where the gap of Etherlay is not as wide as the gap of MSTP and RSTP.

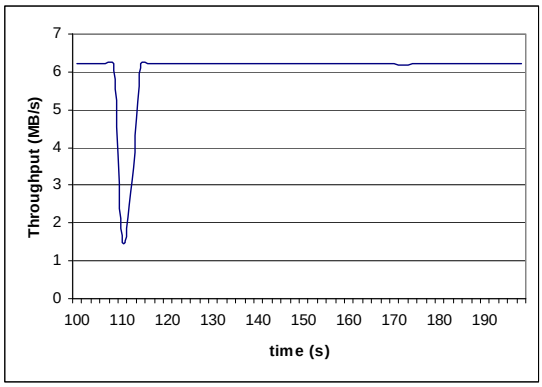

Figure 5 The throughput as observed by the receiving servers during the failures for Etherlay

4) Comparison of RSTP, MSTP and Etherlay

Figure 6 shows the superimposed cumulative throughput graph of the three protocols. Because we do not want congestion to interfere, the maximum throughput is achieved by all three protocols under normal condition. However, during the period of failures, Etherlay performs better than RSTP and MSTP as shown by the smaller total area displayed by the dips. To observe the impact of the dip in the throughput graph, the loss percentage, which is the area of the dipped region, of RSTP and MSTP is calculated against Etherlay. The results show that RSTP loses 20.7% of the total received traffic compared to Etherlay; and MSTP loses 5.93% of the total received traffic compared to Etherlay.

VI. ETHERNET SWITCH LOAD BALANCING

By being able to distribute the traffic, or to load balance, across various links in a network, it is possible to increase the capacity and utility of the network. However, load balancing is not possible under RSTP. MSTP allows load balancing by static assignment of traffic to different VLANs. The problem with the static assignment is that it is not efficient given the dynamic nature of packet switching paradigm. On the other hand, Etherlay's overlay peers will facilitate load balancing across links in the MEN based on the condition of the network.

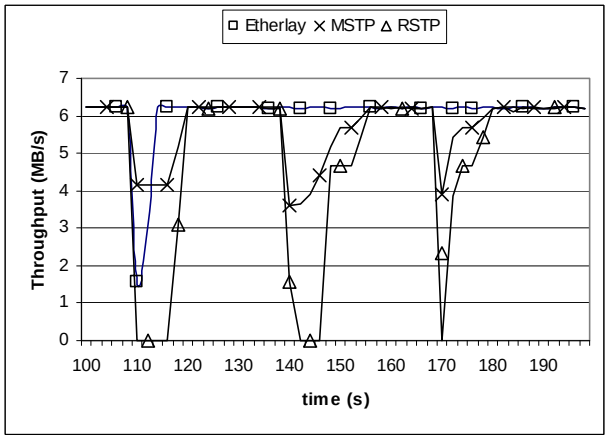

Figure 6 The superimposed throughput graph of RSTP, MSTP, and Etherlay during failures.

A. Comparison of RSTP, MSTP, and Etherlay

For this scenario, the comparative performance of RSTP, MSTP, and Etherlay can be visualized by superimposing the cumulative throughput graphs as shown in Figure 7. The configuration for Etherlay in this case uses 80% for the local estimator. This means, upon sending frames to an outgoing link, the ingress nodes check to see if this link is 80% loaded. If the threshold is reached, the ingress nodes will assume that the network is congested and will use the overlay peers instead. From Figure 7, it is clear that Etherlay's performance exceeds MSTP and RSTP. Compared to Etherlay's total received traffic, MSTP loses about 16% and RSTP loses about 60%. The next section will show that the gain for Etherlay depends on the threshold. However, for all of the thresholds, Etherlay still has a positive net gain over MSTP and RSTP.

B. Local Estimator Trend

The local estimator tries to estimate the congestion of the whole primary spanning tree with an adjustable threshold for the link load percentage of the locally attached links. When this threshold is reached, it will turn the traffic over to the overlay peers. The ingress nodes resume sending traffic on the primary spanning tree again when the outgoing links is below the threshold. Figure 8 shows the trend in the different thresholds. The optimum threshold is at 30%. At 30% threshold, Etherlay's total throughput gains 30.1% compared to MSTP and 66.6% compared to RSTP. When the threshold is between 20% and 50%, the throughput capacity is greatly enhanced. This means that each source sends only 20% to 50% of its traffic on the primary spanning tree. The remaining traffic runs on the overlay paths. By sending a small amount of traffic is enough to congest the network supports the assumption that multiple sources aggregate into some bottleneck link(s). Therefore, for a large topology with numerous sources, at most 50% traffic uses the primary spanning tree.

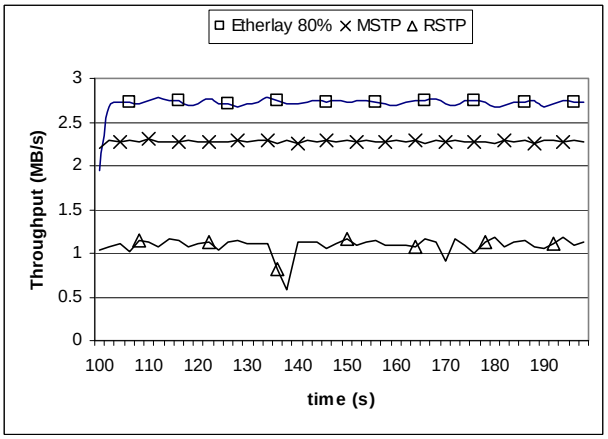

Figure 7 The throughput for RSTP, MSTP, and Etherlay in a congested network

VII. RELATED WORKS

Although overlaying has been moderately worked on, most researches tend to focus on the application layer such as [20], [21], [16], [17], and [18].

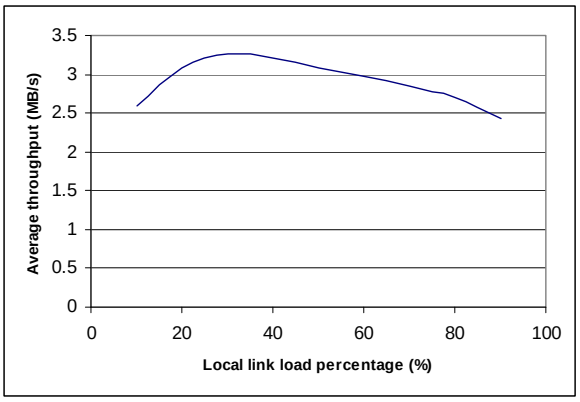

Figure 8 The throughput of Etherlay at various link load threshold

STAR [9] is an overlay approach to improve upon the spanning tree protocol similar to Etherlay. It finds an alternate route that is shorter than the corresponding path on the Spanning Tree. STAR-aware switches are the overlay nodes of the topology. STAR calculates the shortest path from a super node to the next using the distance vector. It does not take advantage of the MSTP or VLAN capability.

Viking is a Multiple Spanning Tree architecture proposed by Sharma et al [6]. Viking precomputes multiple spanning trees so that it can change to a backup spanning tree in the event of a failure. The paths are computed based on the weight that is assigned to each link. The weight is derived from the criticality of the corresponding link. Other approaches to build an efficient spanning tree is Tree-Based Turn-Prohibition (TBTP) [8] and Ethereal [7]. TBTP constructs a less restrictive spanning tree by blocking a small number of pairs of links around nodes, called turn, so that all cycles in a network can be broken. Ethereal, a real time connection oriented architecture supporting best effort and assured service traffic at the link layer, proposes to use the propagation order spanning tree for faster re-convergence of the spanning tree once a failure has been detected.

Lim et al. [10] address the underutilization of the standard Spanning Tree. They also recognize that the simple priority queuing of 802.1 potentially starves low priority traffic when the high priority traffic dominates a significant fraction of the traffic. Each multimedia traffic flow uses the Spanning Tree that is built for the tuple <traffic type, VLAN>. While non-multimedia traffic flows use the Spanning Tree that is built for a traffic type. In contrast to Etherlay, each flow stays in the designated spanning tree and no crossing over is allowed.

VIII. Conclusion

In this paper, we proposed a new concept of overlay in Ethernet, Etherlay, for routing packets in the MEN. We have presented results from a preliminary study that demonstrates the potential benefits that can be offered to the carriers by Etherlay. The implementation of Etherlay has a low complexity overhead and can leverage MSTP support already commonly available in Ethernet chipsets. In this work we have focused primarily on the resiliency and load balancing aspects of Etherlay.

References

[1] IEEE Information technology - telecommunications and information exchange between systems - local and metropolitan area networks - common specifications. Part 3: Media Access Control (MAC) bridges, ISO/IEC 15802-3, ANSI/IEEE Std 802.1D, 1998.

[2] IEEE Standard for Local and Metropolitan Area Networks — Common specifications Part 3: Media Access Control (MAC) Bridges — Amendment 2: Rapid Reconfiguration Amendment to IEEE Std 802.1D, 1998 Edition. IEEE Std 802.1w-2001

[3] IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks. IEEE Std 802.1Q-1998

[4] IEEE Standards for Local and metropolitan area networks Virtual Bridged Local Area Networks — Amendment 3: Multiple Spanning Trees Amendment to IEEE Std 802.1Q™, 1998 Edition. IEEE Std 802.1s-2002

[5] G. Chiruvolu, A. Ge, D. Elle-Dit-Cosaque, M. All, and J. Rouyer "Issues and Approach on Extending Ethernet Beyond LANs" IEEE Communications Magazine March 2004

[6] S. Sharma, K. Gopalan, S. Nanda, T. Chiueh "Viking: A Multi-Spanning-Tree Ethernet Architecture for Metropolitan Area and Cluster Networks" Proceedings of IEEE INFOCOM 2004.

[7] S. Varadarajan, T. Chiueh "Automatic Fault Detection and Recovery in Real Time Switched Ethernet Networks" Proceedings of IEEE INFOCOM 1999.

[8] F. De Pellegrini, D. Starobinski, M. G. Karpovsky, and L. B. Levitin. "Scalable Cycle-Breaking Algorithms for Gigabit Ethernet Backbones" Proceedings IEEE INFOCOM 2004

[9] K. Lui, W. C. Lee, K. Nahrstedt. "STAR: A Transparent Spanning Tree Bridge Protocol with Alternate Routing" ACM SIGCOMM Computer Communications Review Volume 32, Number 3: July 2002.

[10] Y. Lim, H. Yu, S. Das, S. S. Lee, M. Gerla "QoS-aware Multiple Spanning Tree Mechanism over a Bridged LAN Environment" Proceedings IEEE GLOBECOM 2003

[11] OPNET simulator, http://www.opnet.com

[12] MEF, "Metro Ethernet Networks – A Technical Overview" http://www.metroethernetforum.org

[13] Light Reading, "Reports: Carriers Accelrating Ethernet" April 20, 2005 http://www.lightreading.com/document.asp?doc_id=72441

[14] Nortel Networks "Service Delivery Technologies for Metro Ethernet Networks" Nortel Networks Whitepaper Sept. 19 2003 http://www.nortel.com/solutions/optical/collateral/nn-105600-0919-03.pdf

[15] Riverstone Networks "Scalability of Ethernet Services Networks" http://www.riverstonenet.com/solutions/ethernet_scalability.shtml

[16] D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, R. Morris, "Resilient Overlay Network," In Proc. ACM SOSP'01, pp.131-185, Oct. 2001

[17] Z. Duan, Z. Zhang, and Y. T. Hou, "Bandwidth Provisioning for Service Overlay Networks," In Proc. SPIE ITCOM Scalbility and Trafiic Control in IP Network (II) '02, July 2002.

[18] Z. Li and P. Mohapatra, "QRON: QoS-aware Routing in Overlay Network," Special issue on Service Overlay Network in IEEE Journal on Selected Area in Communications, Vol. 22, No.1, January 2004 pp.29-40.

[19] L. Subramanian, I. Stoica, H. Balakrishnan and R.H. Katz, "OverQoS: Offering Internet QoS using Overlays," In Proc. HotNet-I Workshop, October 2002.

[20] S. Ratnasamy,P. Francis, M. Handley, R. Karp and Scott Shenker. "A Scalable Content Addressable Network," In Proc of ACM SIGCOMM 2001, Aug. 2001

[21] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan. "Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications," In Proc. ACM SIGCOMM 2001, Aug. 2001, pp. 149-160