Ethernet as a Carrier Grade Technology: Developments and Innovations

Rafael Sánchez (*rsfuente@it.uc3m.es*)

Universidad Carlos III de Madrid, Spain

Lampros Raptis (*Lampros.Raptis@hol.net*), Kostas Vaxevanakis (*Kostas.Vaxevanakis@hol.net*)

Hellas-On-Line, Corporate Services Division

Abstract — Recent innovations in the Ethernet networking technolgy are enhancing both the scalability and capability of Ethernet as a carrier-grade and transport technology. This article explains four main innovations recently added to Ethernet, namely improvements related to scalability, OAM functionality and enhanced forwarding capability in order to permit Ethernet to assume a much larger role in carrier networks with substantial economic and operational benefits.

***Index Terms* — Optical Ethernet, Networks, Protocols, Communication systems, Optical communication**

I. Introduction

Ethernet-based networking technology has become ubiquitous in both the enterprise and home broadband arenas. The combination of simplicity and rigorous specification has permitted a degree of integration and commoditization that other networking technologies have been unable to achieve.

However, some service providers' infrastructure is based on a legacy circuit-based infrastructure, using technologies like SDH, frame relay and ATM to provide private lines services and interconnection. This has placed service providers in a difficult position, as they face both the costs of supporting multiple technologies and a service arbitrage situation — they sell the same service on multiple technology platforms.

Ethernet is the technology of choice in the customer domain and is therefore a desirable choice in the service-provider domain to eliminate potential interworking problems and leverage the customer-driven investment. However, every technology transformation in the service provider space is time-consuming and also represents major commitment; consequently, comprehensive functionality is required as a prerequisite to mass deployment. From a carrier's perspective, Ethernet still has deficiencies with respect to OAM, reliability, traffic management, and scalability.

It turns out that many of the fundamental issues with Ethernet are well understood, and are currently being addressed with the same rigor and drive for simplicity that has been the objective of Ethernet to date. This article dives into the challenges faced, and how existing Ethernet behaviors can be combined with standards in progress in order to provide a comprehensive network infrastructure that will address the carrier's concerns.

After a summary of the challenges to Ethernet in section II, the remainder of this article is structured as follows: section III describes new Ethernet technologies and how this technologies resolved some of the key challenges; section IV discusses traffic engineering applied to Ethernet; and finally section V covers OAM capabilities. The article concludes with the main findings, which justify the maturity of Ethernet as a carrier grade transport networking technologies.

II. Challenges to ethernet

While end customers are convinced of Ethernet's cost benefits, they are demanding the same levels of performance they had from leased lines, Frame Relay and ATM services. For Ethernet to reach the kind of penetration predicted by analysts, it is required that Ethernet should evolve to display the same properties of current WAN technologies.

The Metro Ethernet Forum (MEF) has defined this evolution as "Carrier Ethernet", which should have the following attributes:

1. **Scalability** — Providers require that the network scale to support the 100,000s of customers to adequately address metropolitan and regional served areas.
2. **Protection** — This really implies reliability and resiliency, as service providers typically boast "five 9's" or 99.999 percent network availability. One of the benchmark tools for achieving this has been

Manuscript finished March 28th, 2008.

SONET/SDH’s ability to provide 50ms link recovery, as well as protection mechanisms for nodal and end-to-end path failures. For Carrier Ethernet to be adopted — especially in support of converged, real-time applications — it must match these performance levels seen by traditional WAN technologies.

3. **Hard Quality of Service (QoS**) — Service providers must be able to offer customers differentiated levels of service to match application requirements. QoS mechanisms provide the functionality to prioritize different traffic streams, but Hard QoS ensures that service level parameters agreed for each level of service are guaranteed and enforced across the network. This provides customers with the guaranteed, deterministic performance they receive from their existing leased line services.
4. **Service management** — Service providers require mature network and service management systems that firstly allow quick services provisioning in order to delivery existing and new services and secondly monitoring different parameters of the provided services. Such monitoring is used against an SLA and the service provider must have the performance measurements to back up any service level claims. And if a fault does occur, the service provider needs to have the troubleshooting functionality to locate the fault, identify which services have been impacted and react appropriately.
5. **TDM support —** While service providers see substantial growth potential in Ethernet services, existing leased lines are still a significant revenue source for them which they must be able to retain and seamlessly interwork with existing leased lines services as they migrate to a Carrier Ethernet network

Equipment vendors are challenged with how to add this carrier-grade functionality to Ethernet equipment without losing the cost-effectiveness and simplicity that make it attractive in the first place. In the next chapters, we will examine the different technologies that are designed to achieve this.

III. ETHERNET TECHNOLOGIES

The Metro Ethernet Forum has defined Ethernet services using the concept of Ethernet Virtual Connections (EVC) established across an Ethernet Network. Customer Equipment (CE) attaches to the network at the User-Network Interface (UNI) using standard 10Mbps, 100Mbps, 1Gbps or 10Gbps Ethernet interfaces. There are three types of EVCs defined:

1. Point-to-Point, called E-Line
2. Point to Multipoint, called E-Tree
3. Multipoint-to-Multipoint, called E-LAN

In order to provide such services, different Ethernet technologies have been proposed and are used for the delivery of the previous services.

A. IEEE 802.1Q Virtual LAN (VLAN)

The basic technology standard used for delivering an E-LAN service is the IEEE 802.1Q standard [2] for Virtual LANs (VLANs). This standard creates VLANs across a common LAN infrastructure to enable enterprises to support and separate traffic from different departments within a company (for example finance, legal and general administration). Each VLAN is identified by a Q-tag (also known as a VLAN tag or VLAN ID) that identifies a logical partitioning of the network to serve the different communities of interest.

IEEE 802.1Q works fine within the boundaries of a single organization, but is found inadequate when service providers attempt to deliver Ethernet services to multiple end users over a shared network infrastructure. Issues arise because enterprises need to retain control over their own VLAN administration (such as assigning Q-tags to VLANs), and over a shared infrastructure the service provider must control this to ensure that one customer’s Q-tags do not overlap with another’s. Also, because the Q-tag consists of a 12-bit tag, up to 4,094 possible service instances can be created. (Note: 4,096 service IDs are available, but two of these are reserved for administration.) Although this is sufficient for an enterprise’s LANs, it does not offer the scalability required to support Ethernet services in a large metropolitan area. What is needed is a method for defining secure Ethernet services to individual customers within which the customer can create further LANs for departments or groups of users. There are two developing standards that support this approach: IEEE 802.1ad Provider Bridges [3] (also known as Q-in-Q or VLAN stacking) and IEEE 802.1ah Provider Backbone Bridges [4] (also known as MAC-in- MAC).

The standardization of these technologies is being driven by the IEEE 802.1 working group. The Provider Bridges standard was officially approved in December 2005, while Provider Backbone Bridges was formally introduced as draft standard in March 2005 and it is expected to be officially approved in the second quarter of 2008.

B. IEEE 802.1ad Provider Bridges (Q-in-Q)

Provider Bridges work by simply adding an additional service provider VLAN ID (S-VID) to the customer’s Ethernet frame. This new S-VID tag is used to identify the service in the provider network while the customer’s VLAN ID (C-VID) remains intact and is not altered by the service provider anywhere within the provider’s network as shown in Figure 1. This solves the transparency problem experienced by IEEE 802.1Q.

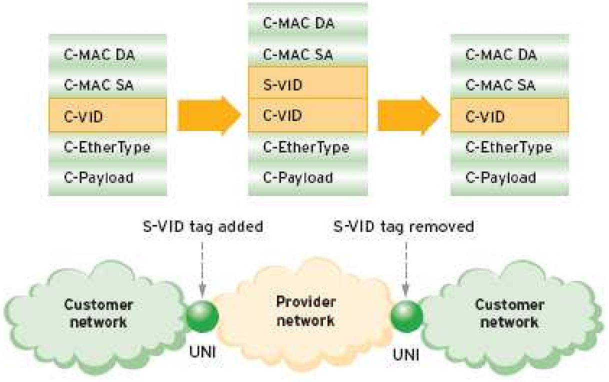

Fig. 1. S-VID added to the customer frame

Provider Bridges use the S-VID to identify the service to which a customer's Ethernet frames belongs to and therefore each service instance requires a separate S-VID. Because the S-VID consists of a 12-bit tag, Provider Bridges has the same scalability limitation of IEEE 802.1Q and only 4,094 services instances can be created.

In addition, Provider Bridges uses the same MAC address for the provider's and customers' networks. This makes both networks appear as one large network to the provider's switches, as shown in Figure 2.

In the scenario depicted in Figure 2, the provider's and customers' MAC addresses are visible to all network elements and this creates a significant burden for core switches, as they must maintain a forwarding table for every MAC address in the service provider and customer networks. Also, any changes to the customer network will have an impact on the provider core. For example, when a new host is added in the customer's network, the new MAC address must be learned by the provider's switches, or when a failure occurs in the customer network, the resulting action taken by Spanning Tree Protocol (STP) can impact the provider network. Although such changes are outside the service providers network, yet they impact their network and create instability. From the customers' perspective, a potential security concern emerges from the fact that their addressing information is now visible outside of their secure network domain.

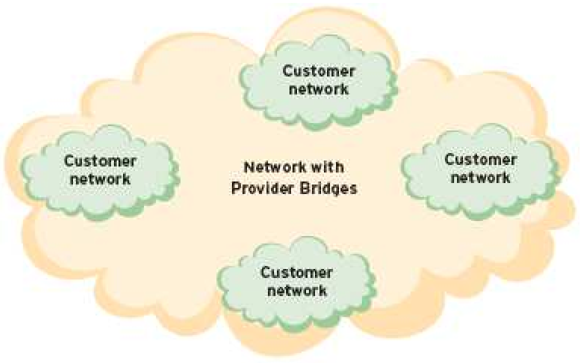

Fig. 2. Provider's and customer's MAC addresses visible to all networks

Provider Bridges does not provide separation between the provider and customer networks and this creates problems where control protocols are concerned. Most Ethernet control protocols, such as Bridged Protocol Data Units (B-PDUs) used by customer networks, must not interact with the provider's networking equipment. For example, STP used in the customer network must not interact with STP used in the provider network. B-PDUs are identified by their destination MAC address and do not have a VLAN tag associated with them. For example, the Spanning Tree Protocol is identified by destination MAC address 01-80-C2-00-00-00. Provider Bridges cannot provide differentiation between customer and provider B-PDUs because each entity's B-PDUs have the same MAC address, and duplicate MAC addresses cannot be supported. This will cause unpredictable network behavior because the provider's networking equipment cannot distinguish between customer and provider B-PDUs. IEEE standard solves this limitation by introducing a different set of destination MAC addresses for B-PDUs in the provider's network. However, to support these new provider B-PDU MAC addresses, the service providers must replace the existing Ethernet switches, because B-PDU MAC addresses are not configurable. For this reason, Provider Bridges technology has significant limitations for E-LAN services that must support multiple customer control protocols.

C. *IEEE 802.1ah Provider Backbone Bridges (PBB)*

Provider Backbone Bridges (IEEE 802.1ah) evolves the Ethernet frame by adding a MAC header dedicated to the service provider and, in doing so, adds a Backbone source and destination MAC address, a Backbone VLAN ID (B-VID) and a Backbone Service ID (I-SID) to the customer's Ethernet frame. Figure 3 illustrates the Provider Backbone Bridges frame and shows how this compares to the standard Ethernet frame (IEEE 802.1), Virtual LANs (IEEE 802.1Q) and Provider Bridges (IEEE 802.1ad).

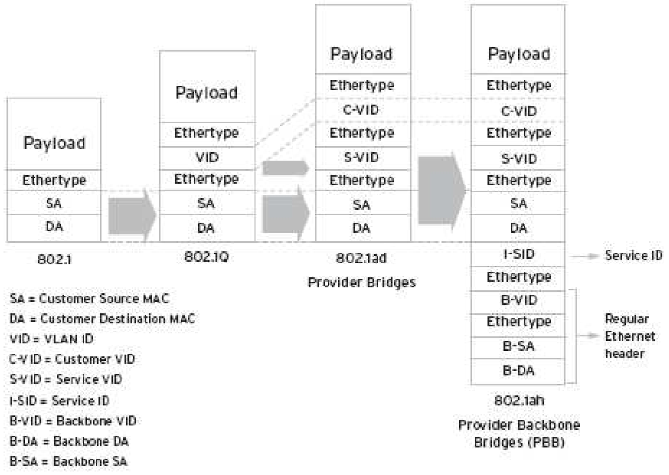

Fig. 3. PBB Ethernet frame

The main benefit of Provider Backbone Bridges is **that the 24-bit I-SID identifies the service in the provider's network.** This means Provider Backbone Bridges provides **up to 16 million services**, completely removing the scalability problems of Provider Bridges.

In addition, Provider Backbone Bridges provides clear separation between the service provider and customer networks, because each has a dedicated set of MAC addresses as shown in Figure 4. When an Ethernet frame reaches the Ethernet UNI , the service provider MAC address is added to the customer's Ethernet frame, and the service provider network switches check this MAC address against their forwarding tables. This is an added advantage in that only switches at the edge of the provider network need to be Provider Backbone Bridges enabled. Switches in the core of the network switch on a standard MAC header (in this case, the service provider header) and so any IEEE 802.1 Ethernet switch will suffice.

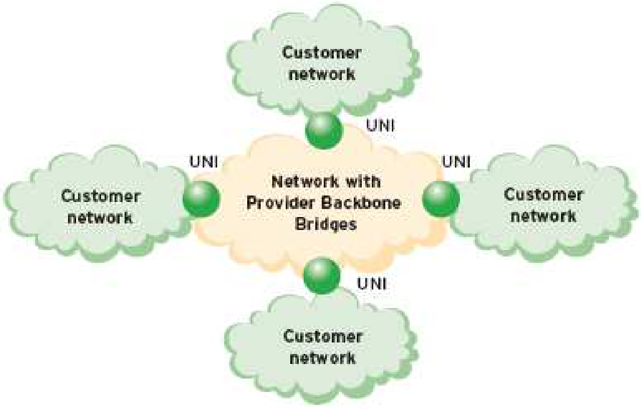

Fig. 4. Provider/Customer MAC addresses separated at the UNI

This solution allows customers' MAC addresses to overlap with the provider's MAC addresses, because the customers' Service Frames are tunneled by Provider Backbone Bridges and are not used when switching frames inside the provider's network. As a result, customers are free to assign identifier and class of service values to their VLANs without any concern that those VLANs will be altered by the service provider. Meanwhile, the service provider does not need to worry about coordinating VLAN administration with its customers.

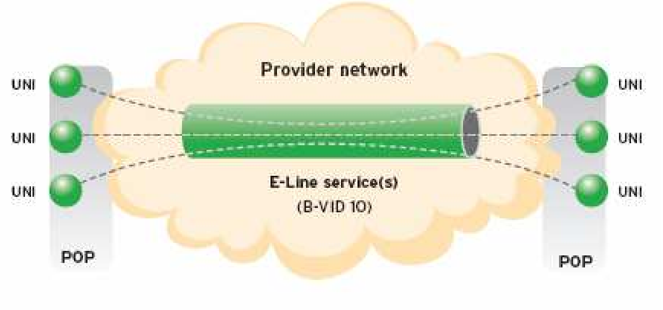

Fig. 5. Single B-VLAN for multiple services

Also, because the service provider's core switches only use the provider MAC header, there is no need for them to maintain visibility of customers' MAC addresses, reducing the burden on the forwarding tables in the provider's network. This also ensures that changes to the customers' networks do not impact the service provider network, improving the stability of the service provider's network. Finally, customer security is improved, because the service provider switches are no longer inspecting the customer MAC header.

Another benefit of Provider Backbone Bridges is that because the I-SID is used for service identification, **the Backbone VLAN ID (B-VID) can be used to segregate the service provider's network into regions** or "zones" to simplify traffic engineering. Backbone VLANs enable the support of multiple customer services instances; for example, a B-VID can be engineered to support 1,000 10 Mbps E-Line services between POPs, as in Figure 5.

This means the service provider engineers the network once when the B-VID is set up. Individual services can be then activated at the source and destination nodes and supported over the B-VID according to its engineered limitations. With Provider Bridges, each individual service needs to be configured across the network node-by-node, creating a substantial operational burden.

Since Provider Backbone Bridges tunnels customers' Service Frames, all customer Ethernet Control Protocols (B-PDUs) are tunneled transparently across the service provider's network. This allows Ethernet Control Protocols to be used independently by the customers' networks and the service provider's network. As discussed, Spanning Tree Protocol (STP) in the customers' networks must not interact with STP used in the service provider's network. STP is identified by its destination MAC address 01-80-C2-00-00-00 and with Provider Backbone Bridges, the customers' STP B-PDUs are tunneled through the provider's network. Therefore, both the provider and customers can simultaneously use the standard STP destination MAC address with no additional provisioning required on the provider's switches. This allows the provider to use the standard B-PDU MAC addresses on the existing switches in the network.

IV. Adding traffic engineering to Ethernet

It is now possible to support connection-oriented forwarding using native Ethernet with a new technology called Provider Backbone Bridges – Traffic Engineering (PBB-TE). PBB-TE is an innovative Ethernet technology, invented by Nortel with the former name PBT [15], currently being standardized as part of IEEE 802.1Qay and that proposes only minor addition to the existing Ethernet standards. In its simplest form, PBB-TE provides Ethernet tunnels that enable deterministic service delivery with the traffic engineering, QoS, resiliency and

OAM requirements that service providers demand.

PBB-TE takes advantage of the fact that by simply turning off some Ethernet functionality, the existing Ethernet hardware is capable of a new forwarding behavior. This means that a **connection-oriented forwarding mode can be introduced to current Ethernet networks** without complex and expensive network technologies.

Currently, Ethernet switches forward on the basis of a full 60-bit lookup of both the VLAN tag (12 bits) and the destination MAC address (48 bits) in each Ethernet frame. In conventional operation, both the VLAN ID (VID) and MAC address are globally unique, but this doesn’t have to be the case. Where a VID typically identifies a loop free multicast domain in which MAC addresses can be flooded, if we choose to configure loop free MAC paths instead, the VID is freed up to can be used to denote something else. In the case of PBB-TE, it will use a range of VIDs to identify specific paths through the network to a given destination MAC address. Each VID is then locally significant to the destination MAC address only, and since the MAC address is still globally significant, the combination of VID + MAC (60 bits) becomes globally unique.

PBB-TE allocates a range of VID/MAC addresses whose forwarding tables are populated via the management or control plane instead of through the traditional flooding and learning techniques. In this case. Spanning Tree and all its associated constraints and problems disappear. The switches still behave fundamentally as with traditional Ethernet: forwarding data to its intended destination. What is different is the fact that the forwarding information is no longer based on the MAC learning mechanisms of the switches, but is provided directly by the management plane, resulting in a prescribed, pre-determined path through the network and totally predictable network behavior under all circumstances.

In the example shown in Figure 6, two uni-directional paths have been configured between Provider Edge (PE) 1 and 2 (a pair of links in opposite directions is required for bi-directional connectivity). Each PE is IEEE 802.1ah enabled, allowing the service provider to clearly separate the service provider and customer MAC domains, thus allowing the service provider to apply PBB-TE within the core of the network. Within the service provider domain, a number of VIDs have been reserved for PBB-TE — these include VID 44 and 45 in our example. As explained, within the group of VIDs reserved for PBB-TE behavior, the VID is no longer globally unique, but locally significant to each MAC. Instead, VID 44 and 45 are used to separately identify the two paths between PE 1 and 2. Both of these VIDs can be reused to create paths between a different pair of PEs because it is the combination of MAC and VID that uniquely identifies each of these paths.

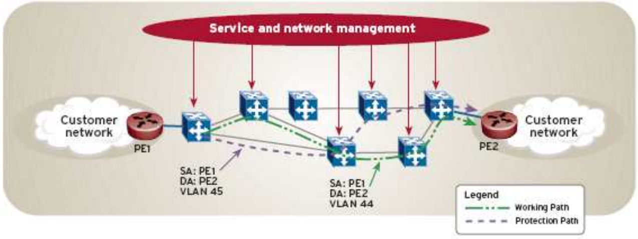

Fig. 6. PBB-TE configuration

PBB-TE preserves the destination-based forwarding attributes of Ethernet, which means multiple sources can use a VID+MAC destination. If 16 VIDs were reserved for PBB-TE in this network, the network could be fully meshed 16 times. This would provide massive scalability for the PBB-TE links and still leave 4,078 VIDs for normal connectionless Ethernet behavior, operating on the same network. It should be noted that each frame still carries a source MAC address that uniquely identifies its origin; so PBB-TE offers the scaling of destination-based forwarding in the core (order “N”) while preserving the operational attributes of point-to-point at the edges.

In the example given in Figure 6, a pair of bi-directional Ethernet links has been configured across the network to create working and protection paths (they would typically be diverse routed, however in our example, they were chose to cross in a core switch to shown how different VIDs may be used to identify different routes). PBB-TE derives connection monitoring from IEEE 802.1ag (Connectivity Fault Management) messages. A Connectivity Check (CC) session is established on both paths. Both ends of the link send CC frames at regular (configurable) 10ms intervals and listen to the messages that arrive. If three CC messages do not arrive, the link is deemed to be down and a protection switch is initiated. Alternatively, Alarm Indication Signal (AIS) messages defined by the ITU-T Y.1731 standard could be used to trigger a protection switch.

Protection switching [12] is implemented by applying the new VLAN tag (that of the protection path) to each frame at the encapsulation point. The control plane is used to configure and monitor the paths, but isn’t involved in the actual switching, so sub-50ms protection switching (similar to SONET/SDH) can be achieved.

V. Adding OAM to Ethernet

OAM functionality in traditional TDM networks is well-defined and is an important building block in ensuring that operators can deliver “carrier grade” performance services.Traditional Ethernet in the LAN environment does not have the OAM functionality required by network operators in Metropolitan and Wide Area Networks environment.

If Carrier Ethernet is to fulfill its promise as the next-generation packet-based infrastructure for metropolitan and wide area networks, OAM capabilities must be added to Ethernet.

New standards that provide Ethernet with OAM capabilities is described in the next chapters.

A. *Fault management*

There are two main areas of OAM: fault management and performance monitoring. Fault management ensures that when a defect occurs in the network, it is reported to the operator, who can then take the appropriate action. This is divided into the following functions:

1. **Fault Detection** — IEEE 802.1ag [10] and ITU-T Y.1731 [9] support fault detection through Continuity Check Messages (CCM). These allow endpoints to detect an interruption in service. CCMs are sent from the source to destination node at periodic intervals; if either end does not receive a CCM within a specified duration, then a fault is detected against the service.
2. **Fault verification** — IEEE 802.1ag and ITU-T Y.1731 support fault verification through Loopback Messages (LBM) and Loopback Reply (LBR). These can be used during initial set-up or after a fault has been detected to verify that the fault has occurred between two end points.
3. **Fault isolation** — IEEE 802.1ag and ITU-T Y.1731 support fault isolation through Linktrace Messages (LTM) and Linktrace Reply (LTR). In the example (see Figure 7), node A initiates an LTM, each intermediate node along the path (B and E) sends an LTR back and forwards the LTM towards node F. Under normal conditions, it allows the operator to determine the path used by the service through the network, whereas under fault conditions, it allows the operator to isolate the fault location without making a site visit.

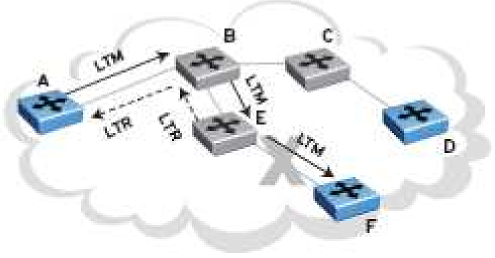

Fig. 7. Fault isolation

4. **Fault notification** — ITU-T Y.1731 supports fault notification through Alarm Indication Signal (AIS). In the example (see Figure 8), a failure between nodes B and E triggers AIS packets in both directions towards the service end points. This functionality alerts the operator for a fault in the network, before it is reported by customers. At nodes A and F, the service end points, the alarm can be replicated across all services supported at that UNI (User Network Interface) that are impacted by the fault. The AIS packets are issued periodically by nodes B and E, to ensure that while the fault still exists, a failure state is maintained. Additionally, the AIS packets can be used to trigger the survivability mechanisms.

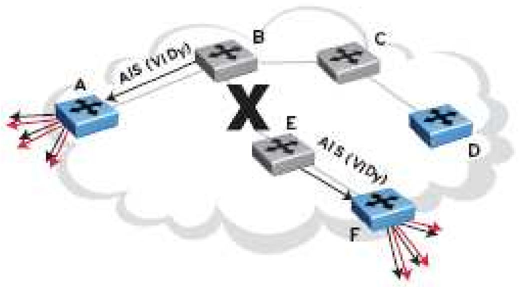

Fig. 8. Fault notification

B. *Performance monitoring*

In many respects the fault management concepts above have been adopted from existing practices in traditional TDM networks. However, while connection-orientated TDM services offer customers a predictable and guaranteed service, packet or frame-based services are connectionless and can have varying performance levels. This is because each individual frame in a service can suffer varying delays due to possible queuing, while network congestion can result in actual loss of frames. Specially, video and voice services, which are part of a residential triple play bouquet, are particularly susceptible to the effects of latency and jitter. As a result, Carrier Ethernet networks require advanced performance monitoring to enforce customer SLAs and this functionality is introduced by ITU-T Y.1731. The following functionality is included:

1. **Frame Loss Ratio** — ITU-T Y.1731 calculates frame loss by sending transmit and receive counters within the CCM for dual-ended measurements. The far end counters can then be compared with those produced locally to derive frame loss as a percentage.
2. **Frame Delay** — Similarly, ITU-T Y.1731 calculates frame delay (or latency). The receiving end can derive the time delay experienced across the network. This requires each service end point to have synchronized clocks.
3. **Frame Delay Variation** — Finally, ITU-T Y.1731 calculates frame delay variation (or jitter) by tracking frame delay measurements.

The emergence of carrier-grade Ethernet has driven the need for improved Ethernet OAM functionality. Ethernet OAM allows the exchange of management information from the network elements to the management layer. Without this capability, it is impossible to provide the comprehensive network management functionality that operators have today in their TDM networks.

VI. CONCLUSIONS

Traditionally, Ethernet lacks of some capabilities to become a technology deployed in the Metropolitan and Wide Area Network environment. However, recent innovations like PBB, PBB-TE and OAM, allow operators to consider Ethernet as a carrier grade networking technology alternative to the traditional technologies like SONET/SDH, ATM or MPLS.

Provider Backbone Bridges (IEEE 802.1ah) provides carrier-grade scalability, resiliency and security between the service provider and customer. Provider Backbone Bridging – Traffic Engineering is then employed in the service provider domain, creating the ability to configure resilient, SLA-driven point-to-point Ethernet trunks. Finally, the combination of IEEE 802.1ag and ITU-T Y.1731 provides powerful fault management and performance monitoring capabilities to Ethernet.

These developments allow service providers to offer scalable, differentiated Ethernet services while retaining Ethernet's cost points and operational simplicity.

However, even with all the improvements that have described in the previous chapters, Ethernet still has one significant weakness. Ethernet still relies on Spanning Tree protocol for any-to-any connectivity, which brings with it several undesirable behaviors. For example, in order to maintain loop free topologies, links are blocked and not used, therefore compromising network capacity. In addition, changes in the network topology such as those resulting from link failures can have a significant impact on the state of the entire network during re-convergence.

New innovations are addressing this key issue. A new technology called Provider Link State Bridging (PLSB) [14], which is based on Link State, can resolve the problems described above. PLSB removes the need of Spanning Tree, and through the use of a Link State Protocol, allows much faster network convergence (hundreds of milliseconds versus several seconds), as well as broadcast containment for PBB E-LAN services.

REFERENCES

[1] IEEE Std. 802.1D, "Local and Metropolitan Area Networks, Media Access Control (MAC) Bridges," 2004

[2] IEEE Std. 802.1Q, "Local and Metropolitan Area Networks, Virtual Bridged Local Area Networks," 2003

[3] IEEE 802.1ad, "IEEE Draft Standard for Local and Metropolitan Area Networks, Virtual Bridged Local Area Networks, Amendment 4: Provider Bridges."

[4] IEEE 802.1ah, "IEEE Draft Standard for Local and Metropolitan Area Networks, Virtual Bridged Local Area Networks, Amendment 6: Provider Backbone Bridges."

[5] IETF Internet draft, "GMPLS control of Ethernet PBB-TE," Nov. 2007 (work in progress), available at draft-fedyk-gmpls-ethernet-pbb-te-02

[6] IETF RFC 4379, "Detecting MPLS Data Plane Failures," Feb 2006

[7] ITU-T Rec. Y.1711 (2004), "Operation and Maintenance Mechanism for MPLS Networks."

[8] ITU-T Rec. Y.1730 (2004), "Requirements for OAM functions in Ethernet-based networks and Ethernet services."

[9] ITU-T Rec. Y.1731 (2006), "OAM functions and mechanisms for Ethernet based networks."

[10] IEEE 802.1ag, "Local and Metropolitan Area Networks, Virtual Bridged Local Area Networks, Amendment 5: Connectivity Fault Management."

[11] IEEE 802.1AB, "Station and Media Access Control Connectivity Discovery."

[12] ITU-T Rec. G.8031/Y.1342 (2006), "Ethernet Protection Switching."

[13] IETF Internet draft, "Carrying PWE3 Pseudo Wires over Provider Backbone Transport," July 2007 (work in progress) available at draft-allan-pw-o-pbt-03.txt.

[14] Don Fedyk and Paul Bottorf, "Provider Link State Bridging (PLSB)", January 2007. http://www.ieee802.org/1/files/public/docs2007/aq-fedyk-provider-link-state-bridging-0107-01.pdf

[15] http://www.nortel.com/pbt