› ADDING SCALE, QoS AND OPERATIONAL SIMPLICITY TO ETHERNET

NORTEL

White Paper

Provider Backbone Transport

Market overview

For many years, Ethernet has been the dominant networking protocol in the LAN. Its simplicity not only made it easier to operate, but allowed significant commoditization to make it extremely cost-effective. However, in the MAN and in the WAN, it was a different story, with service providers offering leased lines, Frame Relay and ATM services — all significantly more complicated and offering less bandwidth.

What enterprises really want today is to connect their sites together without the complexity of traditional MAN and WAN technologies. Not only does this retain network simplicity, but it allows higher, more cost-effective bandwidth connectivity that supports multimedia applications and delivers storage services and server consolidation. As a result, annual worldwide revenues in Ethernet services topped $5.9 billion in 2005 and are expected to reach $22.5 billion in 2009[1].

Faced with this demand, service providers must transform their metropolitan and regional networks to cost-effectively and profitably deliver the next generation of services. Service providers now realize that by making this investment, they not only benefit from the cost savings and operational improvements of Ethernet, but they gain a converged infrastructure capable of addressing today's hottest applications:

- **Broadband Triple Play:** Ethernet used as a network infrastructure for broadband service aggregation (i.e. backhaul of traffic from DSLAM sites) concurrently with large-scale delivery of multicast/broadcast video content.
- **Wireless Backhaul:** Ethernet used as a network infrastructure for aggregation of data traffic emanating from 3G wireless networks.
- **Ethernet connectivity services:** A set of carrier-based services that deliver wholesale and retail Ethernet-based interconnect.

The cumulative effect of these infrastructure builds in support of service provider backhaul and Ethernet services is a strong demand for Metro Ethernet equipment that reached $3.1 billion annually in 2004 and is projected to double to $7.6 billion by 2008[2] (see Figure 1).

---

[1] Infonetics Research: Ethernet Services Annual Worldwide Market Size and Forecast, April 2006

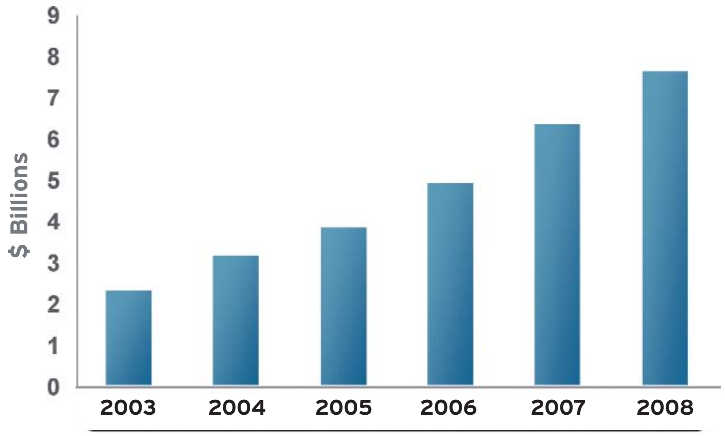
[2] Infonetics Research: Worldwide Metro Ethernet Equipment Forecast, September 2005

Challenges

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 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:

- **Scalability** — Providers require that the network scale to support the 100,000s of customers to adequately address metropolitan and regional served areas.
- **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 SONET/SDH's ability to provide 50ms link recovery, as well as protection mechanisms for nodal and end-to-end path failures. For Metro Ethernet to be adopted — especially in support of converged, real-time applications — it must match these performance levels seen by traditional WAN technologies.
- **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 adhered to across the network. This provides customers with the guaranteed, deterministic performance they receive from their existing leased line services.

Figure 1. Worldwide Metro Ethernet equipment forecast



Source:
Infonectics
Research,
September
2005

- **Service management** — Service providers require mature network and service management systems that allow quick configuration of the network to support new services. Also, just as it is important to keep the customer's service running, service providers must be able to prove it is doing so. Typically this is measured 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.
- **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 as they migrate to a Metro 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. Throughout this paper, we will examine the different technologies that are designed to achieve this.

Scalability and a hierarchical view of Ethernet

One of the original concepts of Ethernet is to provide every user with fair and equitable access to the network with a minimalist implementation in hubs, bridges and switches using a flat network addressing structure. This is fine in the LAN, but to use Ethernet to offer differentiated services, providers need some way of dividing the network into private networks for each user.

This is nothing new. In order to support and separate different departments' traffic (e.g. finance, legal and general administration), enterprises create Virtual LANs (VLANs) across a common LAN infrastructure. Each VLAN is identified by a Q-tag (a 12-bit field, or tag, added to the frame header and referred to as the "Q-tag" because it is defined in the **IEEE 802.1Q** standard) which identifies a logical partitioning of the network to serve the different communities of interest. This technique is the first example of a hierarchy being introduced to an otherwise flat network structure in order to ease management and improve performance.

Delivering business connectivity services to more end users over a shared Ethernet infrastructure that covers greater distances requires adding to this hierarchy and allowing the service provider to provide secure VLANs to individual customers. Within these separate service instances the customer can create further VLANs for departments or groups of users. This three-tier hierarchy allows separate domains for the service provider, customer and individual departments.

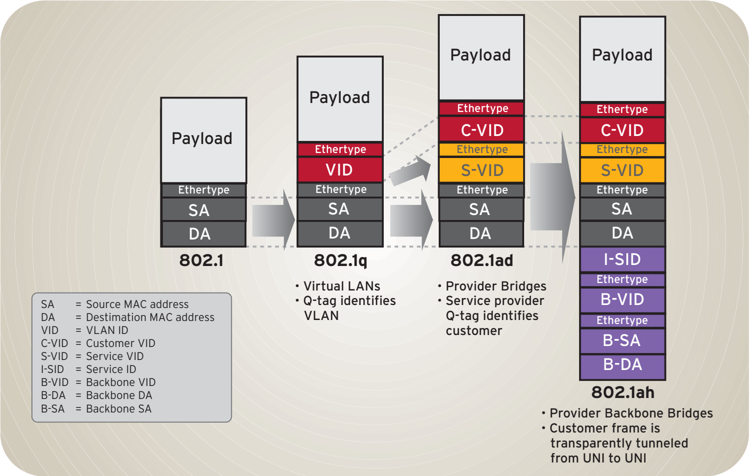
There are two standards to support this hierarchical approach (Figure 2). The first, **IEEE 802.1ad** (also known as Q-in-Q, stacked VLANs or Provider Bridges), extends the original concept of VLANs. IEEE 802.1ad simply adds a new Q-tag that allows the service provider to administer their own tags to identify individual customer networks, while the first (original) Q-tag is used to identify VLANs within the customer's network (i.e. departments in our example). Although Q-in-Q supports a three-tiered hierarchy, the service provider can still only create 4,094 customer VLANs, which is insufficient for large metropolitan and regional networks.

This shortcoming is addressed by the second standard, **IEEE 802.1ah** (also known as MAC-in-MAC or Provider Backbone Bridges), which encapsulates the customer MAC header with a service provider MAC header. Instead of using the additional Q-tags to separate end customers, a 24-bit service tag in the service provider MAC header is used, enabling a theoretical maximum of 16 million service instances to be supported, completely removing the scalability issue.

With IEEE 802.1ah, the overall network is treated as separate service provider and end customer domains. In the service provider domain, the network switches on the service provider MAC header and the customer MAC is not visible. This introduces strict demarcation between the customer and service provider, enabling a truly hierarchical approach to the network. In addition to tackling the scalability issue by removing the 4,094 customers per tier limit, IEEE 802.1ah also has the following benefits:

Figure 2. Evolution of Ethernet hierarchy

Separating infrastructure end points from service end points has profound implications as to what Ethernet is capable of.



› **Security** — Because there is a clear demarcation point between the customer and service provider networks, there is no requirement for either party to have any knowledge of each other's addressing scheme. The provider's network only switches on provider administered addressing information, significantly increasing the security of their network, services and applications.

› **Simpler operations** — The service provider can plan their network without the need to worry about overlapping VLAN or MAC addresses with their customers creating conflict in the service provider's network.

› **Robustness** — The service provider's network is now more robust as it is isolated from broadcast storms and potential forwarding loops created in the end customers' networks.

› **Lower capital expenditure** — The switches in the service provider portion of the network only need to learn the service provider MAC addresses (and not the customer addresses), thereby reducing the memory and processing power required and ultimately the cost of the Ethernet switches in the service provider's network.

The emergence of IEEE 802.1ah and the strict hierarchy it enforces has gone a long way to making Ethernet scalable and carrier-grade. Although alone it does not meet all of the carrier-grade Ethernet criteria, by introducing clear separation of the customer and service provider networks, the service provider now has the ability to tackle the remaining issues within their network domain. All forwarding identifiers used within the provider network are under the provider's control.

Connectionless Ethernet

Many of the problems associated with Ethernet in the WAN are as a result of its connectionless behavior. To understand why this is a problem, we need to take a closer look at how Ethernet works. The primary function of an Ethernet switch is to forward data to its intended destination in the network — a simple enough task when the switch knows where a given address resides in the network. But when a switch receives data destined for an unknown destination, its only option is to copy the data to all of its outgoing ports. This process is known as "flooding". Eventually, the intended destination will be reached via one of the ports, and a reply will be returned. This reply is used by each switch to note which specific port corresponds to the destination. This is the "learning" process. In essence, Ethernet is a broadcast medium that "learns" to conserve bandwidth by observing passing traffic and adjusting the appropriate forwarding tables accordingly.

This approach works well in small networks, but as networks grow larger and more complex — in the MAN, for instance — the flooding and learning processes generate considerable network congestion and can create security concerns. In addition, the learning process is only reliable where there is one, and only one, path towards a given destination. To eliminate the possibility of multiple paths that could lead to forwarding loops, Spanning Tree Protocol (STP) is used to selectively disable switch ports and thus block one or more of the redundant physical paths. This leaves a lot of expensive network assets sitting idle.

Spanning Tree Protocol can also find a new route between two nodes in the network when one fails, creating a simplistic protection mechanism. But while spanning tree is converging on the best alternate path, service is interrupted across the network. The problem with STP is that it is simply too slow at protection switching — and although Rapid Spanning Tree Protocol (RSTP) and Multiple Spanning Tree Protocol (MSTP) help to address this, they only provide incremental improvements in restoration times. Spanning Tree was designed for the tree topology that naturally exists in the LAN, not the complex mesh topology that exists in the MAN. In the LAN, routes can take tens of seconds to be re-established. If you then extend into the MAN, convergence can take even longer. These long restoration times will not support today's voice, video and other real-time services that are quickly converging on the burgeoning Metro Ethernet network. STP also compounds the congestion problem by turning off the very links which could be used to move traffic off the most heavily laden links. Finally, because the network controls the path that data takes across the network, it is difficult to predict network performance and to provide guaranteed QoS and SLAs.

So — what is required? Service providers need a way to:

- Deliver guaranteed, deterministic services over Ethernet infrastructure on a wider scale
- Ensure reliability, management and scalability in order to deliver the multimedia services that enterprises demand
- Take advantage of the operational and cost efficiencies that Metro Ethernet can generate while moving towards a converged infrastructure

Network evolution

Many of the problems described above stem from the connectionless behavior seen in Ethernet. These problems are similar to the ones encountered several years ago in service providers' wide area IP networks — which they addressed by deploying MPLS as an IP-helper to provide connection-oriented "tunnels" through the network. These "MPLS tunnels" provided scalability, traffic engineering, QoS and resiliency over a single IP/MPLS network. On top of this infrastructure of tunnels, service providers began offering new converged Layer 3 (IP-VPN) and Layer 2 (VPWS, VPLS) "MPLS services".

In today's MAN, conventional wisdom has been that connectionless Ethernet needs a "helper", just as IP did. So it is only natural that many have looked to solve the problem by extending MPLS into the MAN. This strategy worked well for the relatively few number of nodes (i.e. hundreds) in the WAN — but it quickly becomes unmanageable and expensive for the relatively large number of nodes (i.e. thousands) in the MAN. The problem is that deploying MPLS requires implementing many new protocols and standards (LDP, RSVP-TE, OSPF, BFD, FRR, etc.) which adds not only operational complexity and costs, but also increased network equipment capital costs due to the likely control plane and data plane upgrades required to support them.

Service providers are seeking to migrate their networks onto a purely packet infrastructure and their goal is to combine the flexibility of packet processing with the determinism, OAM and operational attributes which they are used to from circuit infrastructure — all at Ethernet's cost points. Service providers have had to deploy "MPLS everywhere" as there has been no other acceptable or practical alternative for their metro networks — until now.

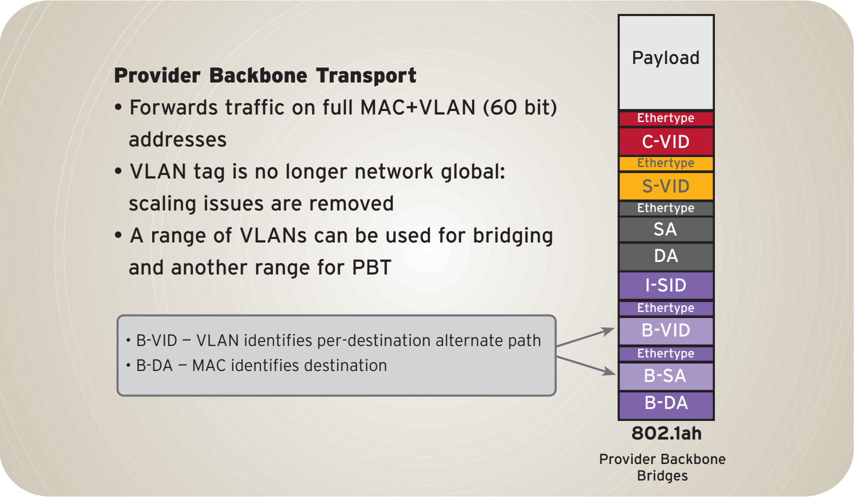
Provide Backbone Transport

It is now possible to support connection-oriented forwarding using native Ethernet with a Nortel-pioneered technology called Provider Backbone Transport (PBT). PBT is an innovative Ethernet technology currently being introduced into the relevant industry standards bodies (IEEE, ITU, IETF, etc.), which proposes only minor additions to the existing Ethernet standards. In its simplest form, PBT provides Ethernet tunnels that enable deterministic service delivery with the traffic engineering, QoS, resiliency and OAM requirements service providers demand.

Provider Backbone Transport 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.

Figure 3. Provider Backbone Transport

PBT functions by turning off MAC learning, broadcast unknown and STP and using a management plane (or optionally a GMPLS control plane) to populate the switch bridging tables for a specific range of VID/MAC addresses.



Provider Backbone Transport

- Forwards traffic on full MAC+VLAN (60 bit) addresses
- VLAN tag is no longer network global: scaling issues are removed
- A range of VLANs can be used for bridging and another range for PBT

- B-VID – VLAN identifies per-destination alternate path
- B-DA – MAC identifies destination

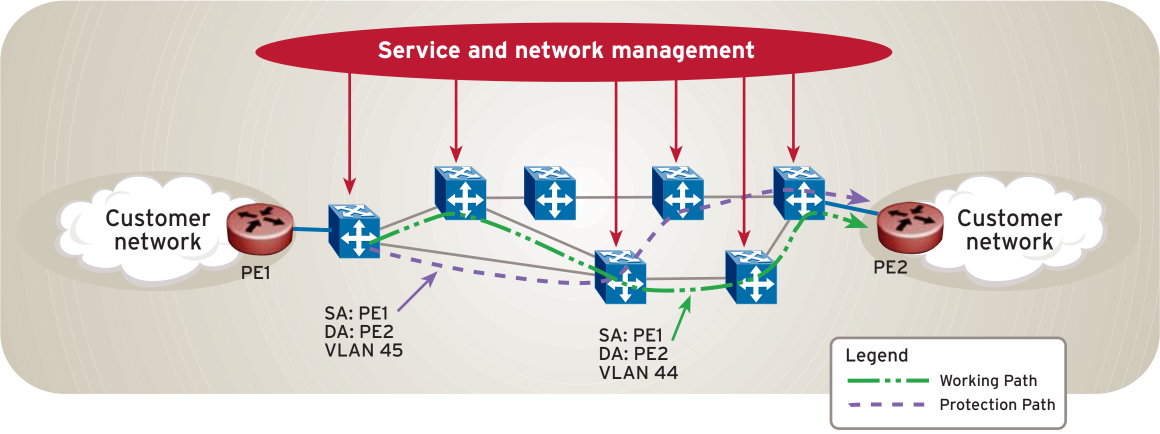
Currently, Ethernet switches forward on the basis of a full 60-bit lookup of both the VLAN tag (12 bits) and the destination MAC address (48 bits) in each Ethernet frame. In conventional operation both the VLAN ID (VID) and MAC address are globally unique, but this doesn't have to be the case. Where a VID typically identifies a loop free domain in which MAC addresses can be flooded, if we choose to configure loop free MAC paths instead of using flooding and learning, the VID is freed up to denote something else. PBT employs this concept by allocating 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 (Figure 3).

PBT 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. Suddenly Spanning Tree and all its associated constraints and problems disappear. The switch still behaves largely as with traditional Ethernet: forwarding data to its intended destination. All that has changed is the forwarding information is no longer learned by 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 4, 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 PBT within the core of the network. Within the service provider domain, a number of VIDs have been reserved for PBT — these include VID 44 and 45 in our example. As explained, within the group of VIDs reserved for PBT 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.

PBT preserves the destination-based forwarding attributes of Ethernet, which means multiple sources can use a VID+MAC destination. If 16 VIDs were reserved for PBT in this network, the network could be fully meshed

Figure 4. Configuring Ethernet trunks with PBT



16 times. This would provide massive scalability for the PBT 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 PBT 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 4, a pair of bi-directional Ethernet links has been configured across the network to create working and protection paths. PBT 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.

The protection switch 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.

Ethernet offers the lowest cost per bit but has several challenges:	**Nortel's Metro Ethernet Solution with PBB and PBT**
Services scalability and customer segregation/hierarchy/security	• 16 million service instances with 802.1ah PBB • 60-byte point-to-point tunnel scalability with PBT • Complete separation with true multi-tier hierarchy: - 802.1Q VLAN - 802.1ad Provider Bridges - 802.1ah Provider Backbone Bridges - Provider Backbone Transport
Protection and Spanning Tree challenges: • Poor convergence/protection/reliability • Stranded bandwidth	• Network resiliency • 50ms restoration with traffic engineering and backup paths • Efficient network utilization supporting both: - Standard connectionless Ethernet services - Traffic engineered Ethernet services – in native Ethernet
Hard QoS and traffic engineering	• Traffic Engineering/Hard QoS • Resource reservation • Differentiated Services
Service management and OAM	• 802.3ah – Ethernet in the First Mile OAM • 802.1ag – Ethernet Connectivity Fault Management • ITU Y.1731 – Ethernet OAM and performance management • Dynamic provisioning
Services supported	• Service flexibility - Ethernet Services – E-LINE, E-LAN, E-TREE - MPLS Services – VPWS, VPLS, IP-VPN

Operations, Administration and Management (OAM)

As discussed, for Ethernet to be accepted as a vehicle for metropolitan and regional connectivity it needs to display the same OAM capabilities as SONET/SDH networks. The following standards are being developed to provide Ethernet with the necessary functionality.

- IEEE 802.1ag (Connectivity Fault Management) — Specifies support for proactive alarming of service faults and assists with the detection, verification and isolation of connectivity failures.
- IEEE 802.3ah (Ethernet in the First Mile) – Defining Ethernet PHYs for first mile access, be it copper, PON (Passive Optical Network) or point-to-point fiber, to allow data communication to Ethernet customer premises equipment.
- IEEE 802.1AB (Station and Media Access Control Connectivity Discovery) – Is required to enable standardized topology discovery by management/OSS.
- ITU G.8031 – SG15 Ethernet Protection.
- ITU Y.1731 – SG13 Ethernet OAM, which augments 802.1ag with additional performance monitoring capability.
- MEF Ethernet Performance Monitoring

As a result of these developments, Ethernet will acquire the service management functionality that service providers expect from today's SONET/SDH networks. By moving the OAM functionality into the data link layer, the network no longer has to rely on the underlying physical layer or overlaid network layer, allowing greater simplification at the network edge and providing further cost savings.

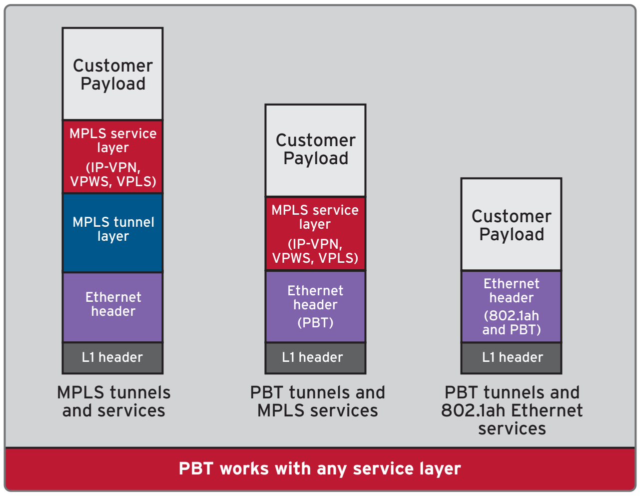
PBT can reuse many of these developments. The fact that PBT populates forwarding tables using the management/control plane instead of traditional flooding and learning techniques has no effect on the service management and OAM capabilities of the network. Only *how* the tables get populated has been changed, not the actual transfer function.

Benefits of PBT

PBT enables the creation of connection-oriented Ethernet tunnels that allow service providers to offer dedicated Ethernet links with guaranteed, deterministic performance levels. PBT is designed to meet or exceed the functionality of MPLS RSVP-TE tunnels, but at Ethernet cost points and with virtually no learning curve for the existing Ethernet operations team. With these capabilities, PBT offers service providers several new alternatives to deploying a next-generation metro network in terms of both the "tunneling" technology and the "services" that it supports (Figure 5).

Figure 5. MAN tunneling and services technology options

PBT simplifies the label stack while providing an alternative to the MPLS "tunnel layer" that can work with any "service layer" — whether MPLS services or native Ethernet services.



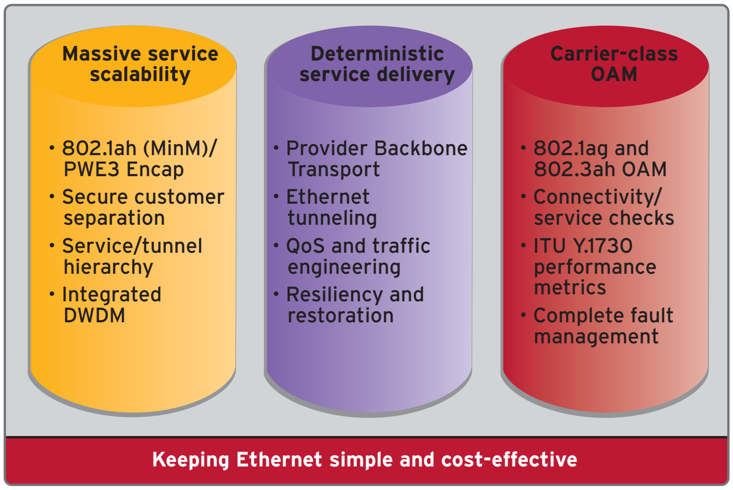
As a traffic-engineered tunneling technology, PBT provides an alternative to deploying MPLS tunnels (such as RSVP-TE) in the metro and supports multiplexing of any Ethernet or MPLS service inside a PBT tunnel. Therefore, service providers can deliver native Ethernet, 802.1Q, 802.1ad or 802.1ah in addition to MPLS-based services — such as VPWS or VPLS — over PBT tunnels. This flexibility allows service providers to deploy native Ethernet services initially, and MPLS services (such as pseudowires over PBT) if and when they need to. As both a tunneling and services infrastructure technology, PBT delivers the following benefits to service providers:

- **Scalability** — By turning off the MAC learning features we remove the undesirable broadcast functionality that creates MAC flooding and limits the size of the network. Additionally, with full 60-bit addressing and destination-based forwarding, PBT enables a virtually limitless (2^{60}) number of tunnels in the service provider network.
- **Protection** — PBT not only allows the service provider to provision a point-to-point Ethernet connection across the network, but to provision an additional backup route to provide resiliency and reliability. In combination with IEEE 802.1ag, these working and protection paths enable PBT to provide <50ms recovery — similar to TDM, SONET/SDH or MPLS Fast Reroute technologies (although without the requirement for additional protocols).
- **Hard QoS** — By specifying the path a packet takes across the network, service providers can now traffic engineer their Ethernet networks. PBT delivers Hard QoS, allowing bandwidth reservation and customer SLAs to be met without over-provisioning network capacity. This in turn allows the service provider to maximize network utilization — and hence minimize the cost per bit carried. In addition, security is increased, as any misconfiguration or packet leakage becomes immediately obvious when using point-to-point Ethernet across the network. This means traffic is protected from finger trouble, malicious intent or unintentional leakage of packets to end-points for which they were not intended due to standard Ethernet flooding.
- **Service management** — The fact the OSS is aware of the route taken by each service enables alarm correlation, service-fault correlation and service-performance correlation. It also enables protection switching for maintenance purposes to be performed in a controlled manner that guarantees performance against the SLA.
- **TDM support** — As a Layer 2 tunneling technology, PBT can interwork with existing WAN technologies, supporting Ethernet E-LINE services as well as MPLS-based services such as VPWS (Figure 5). However, the very low latency of Ethernet switches, combined with the deterministic traffic flow of PBT, provides an ideal platform on which to build to emulate traditional TDM/circuit emulation services.

PBT delivers the scalability, traffic engineering, QoS, reliability and manageability that have been missing from Ethernet to allow service providers to fully leverage it as an infrastructure for converged, next-generation metro networks that support business and residential voice, video and data services. The fact that PBT is enabled by making a small alteration to the normal Ethernet behavior means that this technology can be easily implemented on existing Ethernet hardware. As a result, there is no requirement to introduce complex and expensive network overlay technologies (e.g. MPLS) in the MAN. PBT combines the best of Ethernet with the best of MPLS: delivering a richer converged infrastructure capability with a simpler de-layered network built with the lowest cost commodity components. The result is reduced initial Capex costs on simpler devices that translates into recurring savings as the operational burden is correspondingly reduced.

Figure 6. Nortel's Metro Ethernet solutions

Key differentiators



Summary

The market for carrier-grade Ethernet equipment is growing as enterprises look to Ethernet services for more cost-effective bandwidth and service providers look to reduce their own network infrastructure costs to increase service profitability. Essential components for Ethernet to support these applications are: scalability, reliability, hard QoS/traffic management, service management and support for TDM services. To date, MPLS-based solutions have been the only technology option that solved all of these requirements. Nortel's Metro Ethernet solutions change this landscape by providing new alternatives for Metro networks that are based on Ethernet forwarding, simplicity and cost curves, without any loss of functionality.

Nortel achieves this by driving standards-based innovations across three key pillars of differentiation in our Metro Ethernet solutions: massive service scalability, deterministic service delivery and carrier-class OAM (Figure 6). Provider Backbone Bridges (IEEE 802.1ah) provides carrier-grade scalability and security between the service provider and customer. Provider Backbone Transport is then employed in the service provider domain, creating the ability to configure resilient, SLA-driven point-to-point Ethernet trunks. Finally, this architecture is supported by strong and comprehensive OAM functionality, currently being defined by industry standards bodies.

These developments allow service providers to offer scalable, differentiated Ethernet services and support existing Layer 2 (VPWS, VPLS) and Layer 3 (IP-VPN) MPLS-based services — all while retaining Ethernet's cost points and operational simplicity. The result is a cost-efficient Metro Ethernet network supporting converged residential and business services that will provide the high levels of quality of service, security and resiliency demanded by enterprise customers.

For more information, contact your Nortel representative, or call 1-800-4 NORTEL or 1-800-466-7835 from anywhere in North America.

In the United States:
Nortel
35 Davis Drive
Research Triangle Park, NC 27709 USA

In Canada:
Nortel
8200 Dixie Road, Suite 100
Brampton, Ontario L6T 5P6 Canada

NN115500-041707

>BUSINESS MADE **SIMPLE**