# Ethernet Tackles Access, Aggregation And Transport

Kamran Sistanizadeh

## Which does a better job replacing SONET/SDH rings in the metro: Ethernet or MPLS?

Throughout the evolution of local and wide area networks, enterprise customers have been exposed to a variety of service interfaces at different OSI layers—xWDM (wave division multiplexing) at Layer 1, SONET and Ethernet at Layer 2, MPLS at Layer 2.5, and of course, the ubiquitous IP at Layer 3. Each of these interfaces can be appropriate, depending on the customer's application environment, the connectivity requirements among various sites and the need for more or less signaling, control and SLA (service level agreement) management.

Since customer premises networking environments are primarily built on Ethernet LANs, it would be highly desirable for carriers and service providers to find a proper combination of access and transport technologies optimized for transporting Ethernet among geographically dispersed LANs—within either the customers' intranet or extranet service domains. Service providers and carriers require platforms in some parts of their networks that are optimized for Ethernet access, and in some parts of their network they need platforms that are optimized for Ethernet transport.

The evolution of Ethernet from the LAN into the access network (e.g., metro environment) began in earnest in 1999 after standardized Gigabit Ethernet technology interfaces on switches and routers became readily available. However, carriers had been deploying SONET since the mid 1980s, for their inter-office and regional transport backbones, so SONET was the more frequent choice in the metro. For long haul and for backbones in the core, SONET, ATM (asynchronous transfer mode), and frame relay ruled until the late 1990s, when MPLS began to attract more attention from carrier network planners.

Today, Ethernet's proliferation in the LAN environment, and its ongoing speed advancements (e.g., 10-Gbps, and soon 40-Gbps/100-Gbps), have carriers contemplating Ethernet as the technology of choice for replacing SONET and MPLS. Recent initiatives in the IEEE, IETF and ITU are making changes to Ethernet aimed at pushing it from the access network into the transport backbone environment while also providing the operational and network management strengths and features of SONET/SDH. Ethernet is also being adopted as an "aggregation" and "backhaul" technology of choice by carriers for consumer broadband services and cellular data.

At Yipes, we have built our infrastructure to support our Layer 2 managed network service offerings by using Ethernet in the access and aggregation, and MPLS-based VPLS (virtual private LAN service) in the backbone. This article explains our choices and, in this context, compares the value of Ethernet vs. MPLS in the access and aggregation portions of our network. Specifically, we will address attributes of both in terms of implementation and support costs, transmission overhead, degree of difficulty in provisioning, flexibility in service topology and implementation challenges for new services, such as multimedia and multicast.

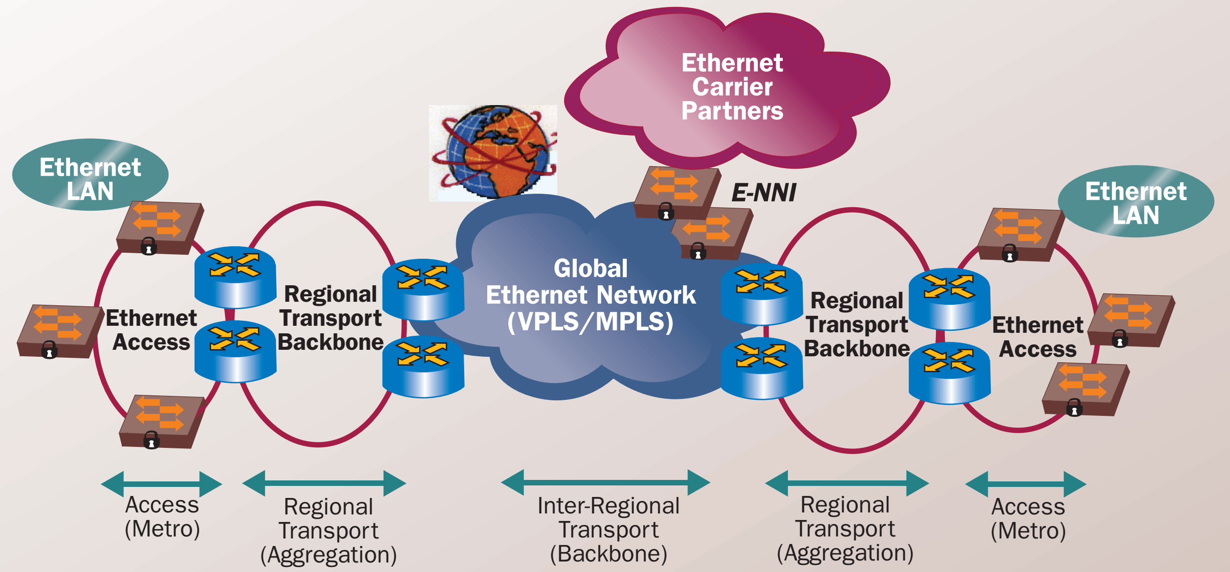
### Metro Architecture Choices

Any service provider's network architecture—be it legacy or greenfield—can be segmented into three portions: the first mile or access, the regional aggregation and the backbone transport. Obviously, the topology and the constituting network elements will vary across service providers.

In Yipes's specific design, as depicted in Figure 1, each metro area that we serve has multiple dark fiber and xWDM fiber rings. Attached to the fiber rings are Ethernet switch/routers, which connect to the enterprise locations via Ethernet, and to the access rings via fiber-based 1-Gbps Ethernet (e.g., native Ethernet over fiber.) The access rings are connected, via higher-density Ethernet switch/routers at metro PoP (point of presence) locations. The metro PoPs also are connected to one another on physical fiber rings using dark

*Dr. Kamran Sistanizadeh is co-founder and CTO with Yipes Enterprise Services, Inc. (www.yipes.com), a provider of managed Ethernet services. He has been a leading contributor to various industry Forums (ANSI-T1, ADSL and MEF) and an ardent advocate of Ethernet services for enterprises. He can be reached at ksistanizadeh@yipes.com.*

**FIGURE 1 Yipes Current Network Architecture**




fiber or xWDM, again via Ethernet, and these are our regional aggregation networks.

Our regional and metro fiber rings are connected to one another over our national and international transport backbone. This backbone is an MPLS-based VPLS that provides a traffic-engineered Layer 2 multipoint-to-multipoint managed service capability.

As mentioned above, traditional metro rings have been built using SONET technology. For a while, about five years ago, it looked like the Resilient Packet Ring (RPR) protocol (802.17) might be an effective alternative, but market adoption has been lagging and the supplier community has been rather small. Although RPR was devised and marketed as a vehicle to combine the best of Ethernet and SONET, more recent innovations—including Rapid Spanning Tree Protocol (RSTP), MPLS fast reroute and Ethernet Automatic Protection Switching (EAPS)—have been proven to deliver sub-second convergence performance on typical access rings (composed of a half-dozen devices), so most of the vendors and carriers lost interest in RPR.

Our choice, like that of many other service providers, has been to implement native Gigabit Ethernet over dark fiber rings, or Ethernet over xWDM in the case of fiber scarcity. To create and interconnect these rings, we have deployed Ethernet switches, routers and transport devices from a variety of suppliers, including Ciena, Extreme, Juniper, Nortel, RAD, Telco Systems and Zhone.

As with any ring topology implementation, the immediate question is the fail-over mechanism and the resiliency in service recovery. The performance of this feature used to be one of the impediments for Ethernet deployment within an access network, due to the lengthy convergence times (30 to 60 seconds) of the spanning tree protocol (STP). However, the IEEE-developed RSTP reduced convergence time to a few seconds, and in 2003, Extreme's EAPS (standardized as IETF RFC 3619) further reduced convergence time to a sub-second. Our internal measurements on operational rings, with up to 12 switches, indicate that the convergence time can be below 100 milliseconds (msec), and usually 50 msec in the majority of cases, thus rendering our convergence performance similar to that of SONET rings.

Our regional aggregation networks also are implemented as rings. Unlike most legacy networks, which use SONET for these inter-office facilities, our implementation uses Ethernet over dark fiber rings, with EAPS for failure recovery.

We limit the number of aggregation PoPs on these regional rings to five or fewer, so the convergence time after failures is usually less than 50 msec. While the number of aggregation PoPs per market is ultimately dictated by the market size, we always deploy at least two, for redundancy. Overall, in both our access and aggregation rings, native Ethernet over fiber has proven to provide resiliency and performance characteristics similar to that of SONET-based rings.

We also have considered limited use of pseudowire technology in our metro and regional networks. A number of pseudowire protocols have been developed, beginning with Luca Martini's eponymous IETF contributions, the Martini drafts (for carrying ATM over an IP network). Today there are numerous IETF RFCs for pseudowires that specify how to carry lower-layer traffic types (e.g., ATM, frame relay, Ethernet, SONET, TDM) over higher-layer networks (e.g., IP, MPLS).

**T-MPLS seems to be aimed more at the backbone, while PBT targets metro networks**

We tested pseudowires to bring functions of the higher layer protocols to some of our metro and regional rings. These functions included variations of the virtual router redundancy protocol (VRRP), such as the Extreme Standby Router Protocol (ESRP) and the MPLS fast reroute mechanism. These capabilities would have been adopted if we had deployed MPLS at the customer premises (using an MPLS UNI handoff). In such a scenario, we would have eliminated the need to have Ether-

## What About PBT And T-MPLS?

Recently, new versions of Ethernet and MPLS have emerged, namely, Provider Backbone Transport (PBT) and Transport-MPLS (T-MPLS.) Both of these technologies are touted to replace and improve upon ATM/SONET in the access and aggregation portions of carrier and service provider networks, based on their merits of operational simplicity and ease of scalability. They are widely perceived to be in contention for the network capacity and capability build-ups that will be needed to satisfy growing demand for all forms of video and multimedia traffic (streaming, uploads, downloads, market data and real-time conferences).

PBT has been introduced over the past two years in the IEEE (P802.1Qay), the IETF (CCAMP working group) and the ITU (Study Group 15, Q12). It promises to solve two concrete problems, namely, VLAN tag exhaustion and forwarding database (FDB) table size. The basic premise is to enable traffic engineering at L2, by allowing carriers to encapsulate customer MAC and VLAN Tag information within the provider MAC/Tag header. In addition, the proposed provisioning tools will make PBT equivalent to MPLS transport provisioning, at a much lower cost. PBT will not require the added complexity of VPLS for multipoint service, since Ethernet is inherently multipoint.

In its primary incarnation, PBT is being designed for point-to-point transport, and could position Ethernet as a replacement for SONET/SDH. It also helps Ethernet's cause as a potential replacement for full IP/MPLS networks within the wide-area transport (backbone) areas, although it cannot achieve this on its own, since (in its current proposed implementation) even PBT will not enable multipoint-to-multipoint traffic.

PBT integrates a set of emerging protocols to enable traffic engineering and operations, administration, management and provisioning (OAM&P). Specifically, it uses the IEEE 802.1ah MAC-in-MAC to reduce FDB (forwarding database) table size, and the IEEE 802.1ad Provider Bridge standard's Q-in-Q for tag space expansion.

The PBT proposed standard also includes the assignment of Ether-types to designate tags for providers, and a standard bridge MAC address range. For OAM&P and performance monitoring functions, PBT has adopted IEEE 802.1ag and ITU-Y.1731 respectively. IEEE is working on the overall harmonization of the constituent standards and will likely rename the standard "Provider Bridge Backbone for Traffic Engineering" (PBB-TE).

Figure 2 and Figure 3 depict possible transitions from a combined Ethernet/MPLS network design to that of Ethernet/PBT (PBB-TE) for access and transport networks. In such a deployment architecture, the customers' Ethernet LAN packets traverse Ethernet-based networks with minimal protocol conversion and packet header transformations. For such an implementation to become a viable alternative to MPLS and gain market adoption, the industry standards community will have to be sure the control/signaling and management planes they are developing with PBT will deliver on their promised lower complexity to the carriers and service providers, in terms of provisioning and network management.

We are working with our current and future suppliers to better understand the capex and opex impact of PBT. Certainly, the delineation of the tagging between customer/access rings (802.1ad) and access/backbone rings (802.1ah) is of great interest for cost-effective scalability of the metro environment—e.g., see Figure 3. In addition, performance measurement capabilities within 802.1ag can potentially decrease our costs by collecting measurement parameters within the compliant device itself rather than installing an out-of-band SLA service management.

### Another Option: T-MPLS

Because PBT could enable carriers to reduce the cost of metro networks, by using it instead of IP/MPLS, the technology is attracting the attention of both supporters and detractors. Some of these detractors are championing another contender to replace SONET/SDH, the Transport MPLS or T-MPLS.

Like PBT, this variation of MPLS also promises greater deployment flexibility and operational efficiency. But T-MPLS seems more directed toward replacing SONET/SDH in the carriers' wide-area backbones than toward the access and aggregation rings.

T-MPLS is basically a control/signaling protocol, a modified version of Generalized net VLANs spanned around the rings, thus limiting Ethernet to point-to-point implementations in these deployments.

However, due to the emergence of very fast loop avoidance protocols, such as EAPS, and the need for further configuration tracking and management complexity that comes with MPLS, we decided to implement pseudowires only for regional transport of VLANs across the metro areas. This is efficient for us, as there are fewer

**FIGURE 2 Possible Future Network Architecture**

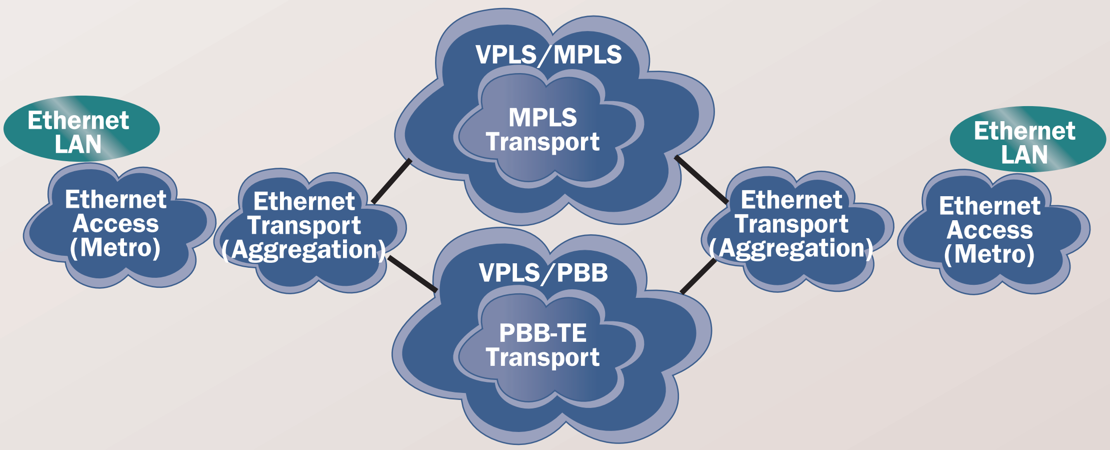
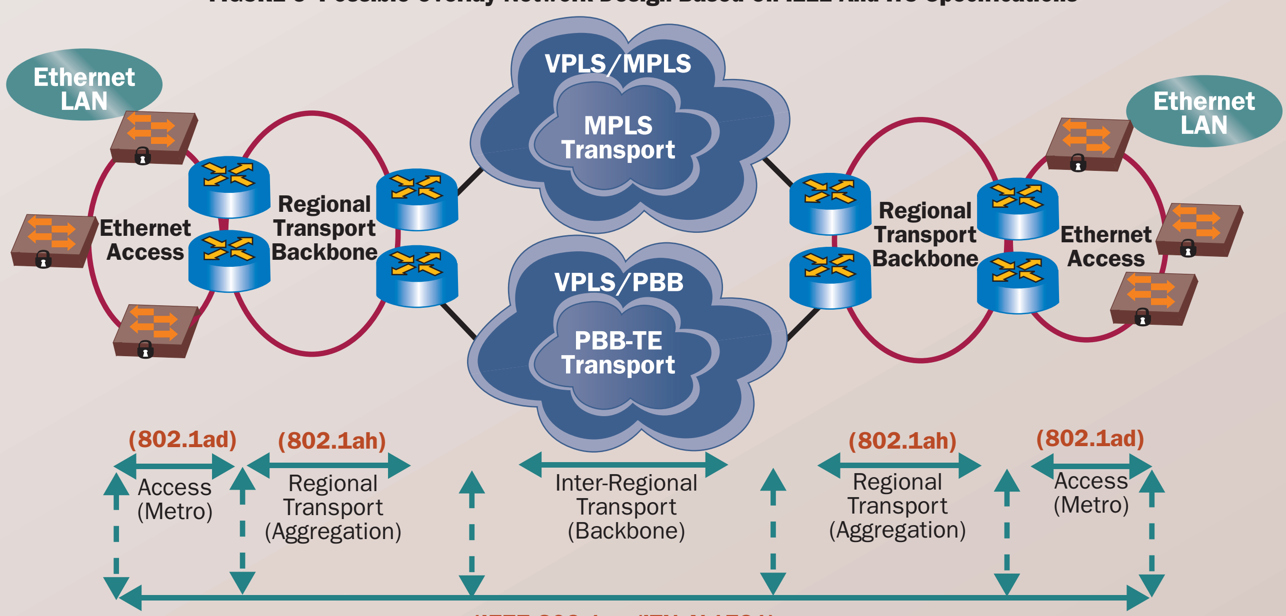



**FIGURE 3 Possible Overlay Network Design Based On IEEE And ITU Specifications**




MPLS (GMPLS) which would provision familiar circuit-type elements, named to correspond with their SONET/SDH forebears—e.g., section, line and path. T-MPLS also will have such attributes as long holding times and bi-directional LSPs (paths), in order to mimic TDM circuits.

At this point, for the legacy carriers to make a comprehensive migration from TDM-based transport to a packet-based transport would require a major paradigm shift in design and deployment philosophy. Such an endeavor will not be an easy feat, considering the total development resources that have been allocated to SONET/SDH over the past 20+ years. Furthermore, in light of the magnitude of the embedded SONET/SDH assets, the economic merits of T-MPLS technology as the next-generation platform would be scrutinized extensively within the carrier community for a long time to come□

**We find Ethernet provides both lower cost and less complexity than MPLS in access networks**

changes in service and configuration in regional transport than in metro networks.

Another fast convergence option that has been available, but which we have not used, are the Optical Transport Networks (OTNs). Just as pseudowires use IP or MPLS convergence protocols instead of Ethernet spanning tree, OTN uses transponders to converge faster than spanning tree. Since we are relying on EAPS or MPLS fast reroute, we have opted not to deploy OTN. We have, however, deployed redundant paths for waves (using xWDM), but the controlling mechanism for failover from one path to another path is handled at Layer 2 by Ethernet EAPS or at Layer 2.5 by MPLS fast reroute.

### Our Wide Area Architecture

Building and deploying a fully-meshed private transport backbone using SONET or wavelengths on a national and international basis is a costly proposition, so we have opted to lease SONET or wave services, while protecting redundancy through Layer-2 or Layer-2.5 protocols (e.g., VRRP, HSRP, ESRP, MPLS Fast-Reroute and variations thereof). Our long-haul carriers and transit providers supply us with various transport links at Layers 1 and 2 (e.g., DS3, OC-3, OC-12, OC-48 and waves) as well as dual-homed, dual-carrier IP transit access using Gig-E interfaces in each of our markets. Thus, we have created a virtual, fully-redundant, meshed international backbone network.

Over this high capacity backbone, we have deployed MPLS for signaling, traffic engineering, service level consistency and predictability across a wide array of interconnected physical layer technologies. MPLS emerged, in the late 1990s, as a migration path to mitigate permanent virtual circuit (PVC)-based ATM complexity. MPLS was touted as a more scalable, traffic-engineered solution in support of IP traffic. It also provided flexibility for provisioning VPNs with different class-of-service metrics.

Unfortunately, MPLS still suffers from operational challenges and high opex relative to the automated provisioning of LSPs (label switched paths), the proliferation of measurement tools for network management metrics, and the troubleshooting of LSP paths. Certainly, operating a network with IP/MPLS and capitalizing on its intended traffic engineering attributes requires a different set of technical skills than the provisioning of SONET paths.

Since we use both SONET transport and routed IP transit, we need to implement MPLS LSPs on both types of links. Also we must be sure our service management functions are configured and tracked using internally developed databases (for customers and for internal LSPs).

We deploy virtual private LAN service (VPLS) based on MPLS to provide our customers—particularly national and global companies in the financial and legal verticals—with a multi-point Layer 2 Ethernet service. VPLS uses the underlying MPLS labels on a full mesh to emulate Ethernet's any-to-any logical topology. This meets our customers' requirements better than the traditional hub and spoke capability provided by frame relay and ATM networks, at higher capacities, and with minimal impact on the individual sites when capacities and configurations require any changes.

Currently, as shown in Figure 1, VPLS/MPLS is deployed primarily across our markets providing any-to-any, Layer 2 reachability from the aggregation points within our metro/regional rings, and we use Ethernet, rather than VPLS, at the edge and within the access and aggregation rings. As a managed service provider, we provide service delivery metrics on a granular basis to our customers through a secure Web portal. There they can see, on a per-service basis, our performance in terms of jitter, latency, packet loss, out-of-sequence packets and other operational-related SLA metrics.

### Current Choices And Future Options

Part of my job as CTO is to separate technologies that are in the research/development phase from those which can be deployed in a revenue-generating environment. We have recently been deliberating the merits of extending MPLS or VPLS into the access network and even directly into the customer premise environment—e.g., deploying MPLS or VPLS at the UNI over a physical Ethernet hand-off. We also have begun evaluating the emerging Provider Based Backbone-Traffic Engineering (PBB-TE, sometimes called PBT) as well as the traffic-engineering enhancements to MPLS (T-MPLS—see the sidebar: "What About PBT and T-MPLS?").

At the present time, however, we believe—and our own experience so far has demonstrated—that native Ethernet not only provides lower cost and operational complexity than MPLS within the access network (e.g., metro and regional aggregation rings), but it also is more flexible for diverse service topologies than MPLS would be. Here are the main points of our reasoning:

■ **Ethernet costs less than MPLS**—Unlike Ethernet, MPLS is not a self-contained solution. It can operate over Ethernet, SONET or other L2 protocol, and in fact, it requires one of these (hence MPLS is referred to as a Layer-2.5 protocol). MPLS also requires an L3 IP protocol like RSVP or LDP, with which to signal over the control plane. These attributes make MPLS a higher-cost solution than Ethernet.

In general, for a given bandwidth and throughput processing capability, the cost of a L2 switch is significantly lower than for a L3 router, as much as 25 to 30 percent, depending on the vendor, volume, contract, etc. Licensing fees also are typically higher for L3 routers than for L2 switches.

By contrast, using native Ethernet saves money for both the service provider and the enterprise customer. Ethernet does not require routers or IP, nor does Ethernet require the control plane signaling or the L3 routing software and hardware, so the equipment costs are lower. Perhaps Ethernet's most important attribute is that the installations do not require the high-level professional support services that IP/MPLS/VPLS installations do.

Support and maintenance costs also are lower with Ethernet, because operating an MPLS-based Layer-3 network requires a more specialized skill set than operating Ethernet. The network management systems (NMS) add-on modules to manage MPLS can be expensive, and the ongoing support cost to maintain proper fault, configuration, accounting, provisioning and security management (FCAPS) is higher for MPLS than for native Ethernet. These NMS modules are important, because they offer LSP path management, signaling statistics, and other MPLS OAM functions. However, the overall complexity of the technology and its operational training and documentation for the network operating center (NOC) often increase R&D costs and time to market as well, incurring opportunity costs for the service provider. We will have to see if the new advances discussed in the sidebar can reduce these costs and improve OAM for Ethernet and/or MPLS.

■ **Ethernet runs with lower overhead and less latency**—With Ethernet, a customer packet will be encapsulated with an 802.1Q VLAN tag (perhaps multiple tags and an encapsulating MAC address as well). With MPLS, additional overhead is required to map the VLAN (or MAC) into an MPLS tag (or label). So there would be an extra 4-byte tag on every packet in addition to that of the standard 802.1Q encapsulation. If we extend VPLS into the metro area as well, then there would be an additional 4-byte MPLS tag for the VPLS pseudowire implementation.

Because of the additional encapsulations, signaling requirements and other L3 processes, MPLS overhead can result in higher packet latency (delay). As packet size decreases, the impact of the additional MPLS tag overhead increases and performance (response time) is adversely impacted. Not only can this raise provisioning costs, but it can also disrupt services with small packet sizes such as voice over IP (VOIP) and message-based financial transactions.

■ **Ethernet is less complex**—Using Ethernet and its modern loop avoidance protocols (e.g., RSTP, EAPS) is less complicated than using MPLS, which requires an underlying L2 (e.g., Ethernet or SONET) and L3 (IP), as well as the control plane signaling protocols (LDP or RSVP). MPLS set-up and device configuration templates are more complex and time consuming to create, understand and execute by the provisioning team.

In our experience, opex costs for MPLS can be from 15 to 25 percent higher than an equivalent Ethernet-based network. The additional protocol dependencies of MPLS also create feature interaction challenges and can lead to difficulty in bringing new technologies and services into the existing network, since all protocols must be retested and re-certified with every strategic change to the network.

■ **Ethernet has a more flexible topology**—Ethernet has an "any-to-any topology" with flooding capability. To emulate Ethernet, MPLS needs VPLS to create a full mesh of LSPs, plus the pseudowires to carry Ethernet. MPLS itself is basically a point-to-point protocol, and it doesn't have specific "hooks" to run directly on fiber rings, hence the need for the mesh. MPLS also offers a point-multipoint, hub-spoke Ethernet emulation called Anything over MPLS (AToM), but we have not adopted this methodology for "any-to-any topology," due to its operational complexities associated with configuration management and troubleshooting.

■ **Ethernet is easier for multicast services**—Ethernet's flooding mechanism makes it an optimal protocol for multicast data distribution within the access ring. The only challenge is controlling the multicasts, and IGMP snooping generally does very well at that.

By contrast, MPLS requires a separate LSP to each destination point, a separate multicast feed to each endpoint (with additive bandwidth), and a separate multicast transmission to each endpoint (replication). Obviously these requirements can be quite processor-intensive for high-volume multicasts—usually resulting in limitations on the number of replications supported and/or high cost. But with Ethernet, simply placing all destination points on the same Ethernet broadcast domain and implementing IGMP snooping, a single feed can be distributed once and reach all endpoints.

### Conclusion

Certainly, MPLS has its value in the backbone environment, especially with VPLS. But our experience has proven that, in the access and aggregation environments, using Ethernet over a physical ring is simple, elegant and inexpensive, particularly compared to running Ethernet emulations over MPLS□

**Companies Mentioned In This Article**

Ciena (www.ciena.com)

Extreme Networks (www.extremenetworks.com)

Juniper Networks (www.juniper.net)

Nortel (www.nortel.com)

RAD (www.rad.com)

Telco Systems (www.telco.com)

Zhone technologies (www.zhone.com)

**Opex costs for MPLS can be from 15 to 25 percent higher than opex costs for an equivalent Ethernet network**