Hammerhead SYSTEMS

White Paper

MPLS in Access Networks

A Practical Way to Transport Ethernet Traffic

Introduction

Multiprotocol Label Switching (MPLS) has been steamrolling its way across carrier core IP networks for some time, and has already reached into metro networks. Is it now the turn of the access network to embrace MPLS?

There are quite a lot of reasons why it might be. In very simplistic terms, a big attraction of MPLS is that it allows carriers to make useful sense of connectionless Layer 3 packet networks by thinking and operating in familiar connection-oriented terms. Ultimately, carriers sell – because customers buy – the ability to connect A to B at a quality of service (QOS) specified by a service level agreement (SLA) of some sort. MPLS virtual circuits have emerged pretty much as the standard way of doing this in IP-based core networks, because they are conceptually simple and, crucially, can be handled by the IP control plane and thus fully integrated into the IP scheme of things.

So MPLS simplifies the provisioning of carrier IP virtual private networks (VPNs), allows carriers to perform traffic engineering, simplifies QOS, and allows Layer 2 Pseudowires (emulations of native services that behave as far as possible like a simple wire connection) to be transported across IP networks.

It's the last point about Pseudowires that has recently got a lot of people interested in thinking seriously about what it would take to push MPLS out into the access network. Pseudowires can run end-to-end from one customer premises equipment (CPE) location to another, and can transport all the important user protocols – ATM, Frame Relay, Ethernet, TDM, and even SONET/SDH and IP. This suggests a massive potential simplification in the access network, where today's mish-mash of protocols and access technologies is replaced by MPLS-enabled Pseudowires, increasingly integrated with Ethernet as the latter steadily becomes the Layer 2 technology of choice.

Such a simplification clearly has appeal. Over a third of the respondents to an online poll during the Light Reading Webinar on which this report is based agreed with the propositions that end-to-end provisioning, access to the core, and the adaptation of legacy protocols to a single access connection would drive the adoption of MPLS in the access network. But a sizeable minority (nearly 13%) thought that the cost and complexity of MPLS would outweigh any of its positive values.

So can it be done? Is MPLS really the answer for bringing converged services to the end user, or could it be a rerun of ATM all over again?

This white paper addresses these questions.

Table of Contents

Access Network Evolution and Challenges

Two countervailing trends have dominated the development of access networks over the last few years. At the physical layer, heterogeneity has become the rule, and different technologies are competing and complementing each other to bring the highest-speed services possible for the end user. As a result, almost every operator connects customers over a variety of access media, such as TDM, DSL, Ethernet, and FTTx, and further options, like WiMax, are already on the horizon.

But, as Figure 1 illustrates, this physical diversity is being matched by increasing unification at the services and data-link layers around, respectively, IP and Ethernet as access networks evolve. IP has won out as the mechanism for converging all services – voice, data, and video – on a single network, and Ethernet is being increasingly recognized as a robust, cost effective, and flexible method of delivering these services at high speeds to end customers. Ethernet also has a wide range of physical layers and matches the huge installed Ethernet base at end-customer locations. So it should be a very seamless way to bring the value of converged IP services all the way to the end customer.

Access networks are thus undergoing a lot of change. Traditionally, leased lines have been dominated by DS-1/3s or OC-3/12s running Frame Relay and ATM, and voice, via PBX trunks. Today those same links would likely be changed to operate with 10/100-Mbit/s or Gigabit Ethernet. Residential DSL broadband access is currently being driven by ATM DSLAMs, but in the future will move to IP-based DSLAMs and Ethernet-based (or

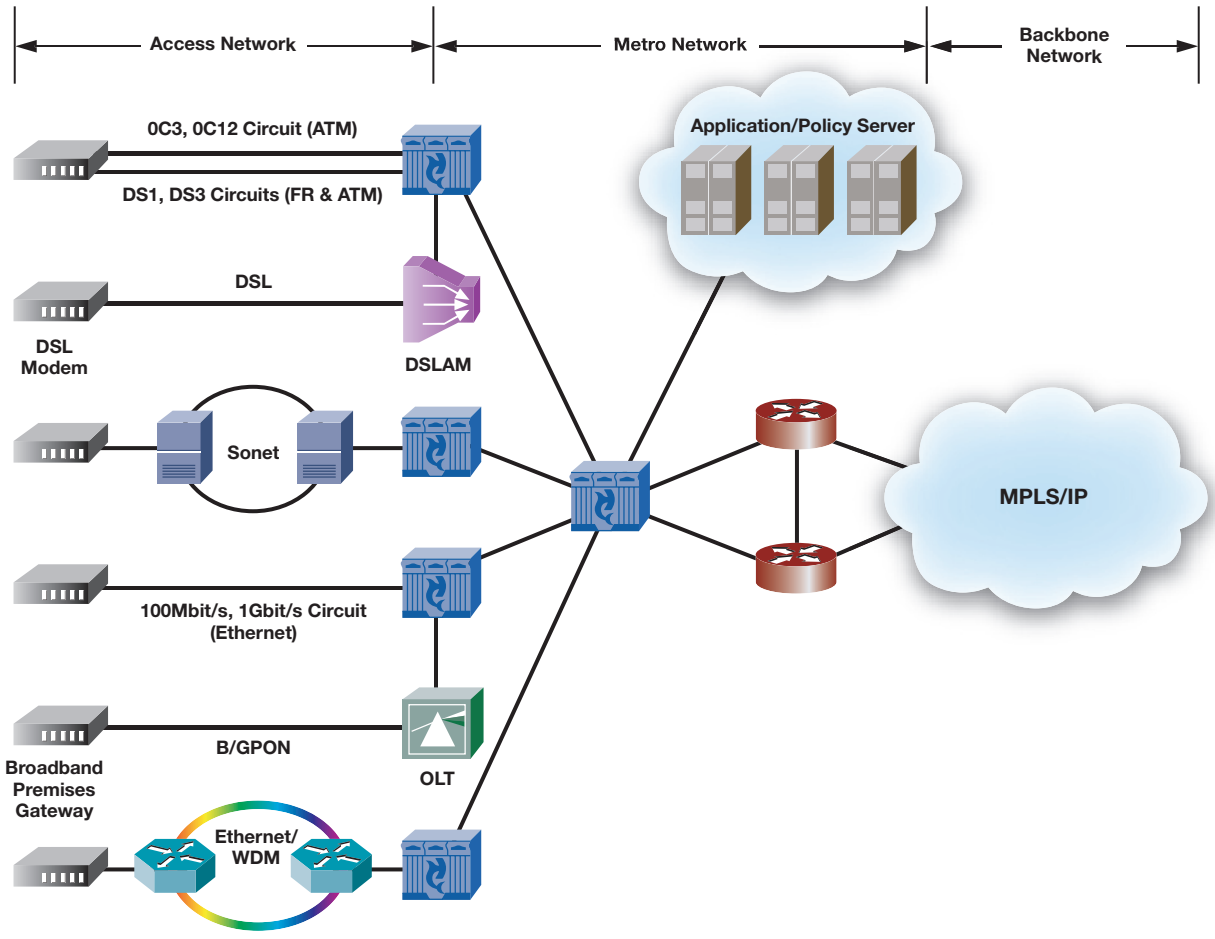

Figure 1. Access Network Evolution

enabled) PONs like EPON and GPON. Even within the metro area, the standard SONET/SDH ring is beginning to be challenged by Ethernet meshes and WDM/Ethernet transport systems.

And the access network cannot escape the changes that are taking place in other parts of carrier networks. The growth of multimedia services and the looming gold rush into IPTV by RBOCs and other telcos means that a lot more application servers will be put in closer to the access network, with all the implications for QoS, traffic engineering, provisioning, and management. Carriers are universally moving towards MPLS/IP core networks, which have to support the end-to-end services delivered by the access networks.

For carriers, the basic appeal of the Ethernet/IP combination is that it seems to satisfy best the classic service-provider wish list of:

- Operational- and capital-expenditure (opex and capex) reductions
- Optimized total cost of ownership (TCO)
- Rapid service creation
- Network architecture independence
- Differentiated/flexible service delivery
- Proactive service maintenance

But it isn't that simple, and there are still many challenges to Ethernet/IP as a blanket solution in real carrier networks, many stemming from the fact that Ethernet, as a LAN protocol, was not originally conceived for such applications. So, with Ethernet, carrier-grade QoS and SLA guarantees are hard to come by, OAM capabilities can be insufficient, and the manageability of an Ethernet/IP combination is problematic and potentially complex and costly. All these issues are compounded in access networks because of the huge number of users, CPE, and locations involved. Scaleability is crucial to carriers in this environment.

This is where MPLS might step in. MPLS – a shim-layer technology lying between Layers 2 and 3 (often referred to as Layer 2.5) but subject to the Layer 3 control plane – is usually thought of as appropriate mainly to the IP core network, where it is used for QoS, traffic engineering, and IP-VPN tunneling. But MPLS virtual connections could, in principle, be extended into the access network to address Ethernet's QoS and management weaknesses.

Technically, the challenge is concentrated in QoS, robustness, manageability, and cost. The available technologies are Ethernet and MPLS. It is not a case of one technology versus the other – Ethernet versus MPLS, rather, it is more MPLS plus Ethernet that is the best solution for access networks.

MPLS to the Rescue?

MPLS today is increasingly seen as a unifying network technology. It provides a framework for managed service convergence within all types of packet networks, and has moved from its original role in the core network to adoption in metro networks. As a result, regardless of an end user's access connection or networks, circuit and packet traffic is increasingly being transported over MPLS-enabled packet networks. In principle, converging traffic onto an extensible MPLS framework into the access network should provide substantial capex and opex savings by allowing the use of existing infrastructures and management paradigms. Existing management toolsets, coupled with current and emerging vendor-independent OAM frameworks, should give better end-to-end management and higher network availabilities.

To make such an MPLS-based, end-to-end architecture work, premium services will require MPLS extension to customer premises. Hierarchical Virtual Private LAN Service (HVPLS) – which divides a VPN into a hierarchical structure of meshed hubs fed by spokes – will be needed for Pseudowire-spoke scaleability; and real-time services over non-over-engineered lines will need RSVP-TE.

Ethernet vs. MPLS, or Ethernet + MPLS

If Ethernet and MPLS are to be combined to provide an access-network solution, it's useful to compare some of their respective pros and cons to see where problems might lie and what can be done about them. Table 1 provides a basic list.

Both MPLS and Ethernet are still evolving rapidly, and quite a lot of standards bodies are involved, as Table 1 indicates. Broader market development, too, is having an impact. The enormous recent interest surrounding triple play and IPTV has suddenly brought Ethernet's multicasting capabilities into prominence, transforming them into a key technology strength.

Table 1: Ethernet and MPLS Compared

Characteristics	Ethernet	MPLS
Drivers	IEEE 802.1, MEF, ITU	IETF, MFA, ITU
Key Strengths	Low cost; Multicasting	Manageability (all IP); Traffic engineering; OAM; Fast reroute; Layer 2/Layer 3 VPN
Key Weaknesses	Control plane: Spanning tree cuts off redundant path —> suboptimal routes —> no load balancing; Data path: No hop count in packet header —> looping	Not a link layer (shim layer) —> the cost is driven by other parts of the system
New Developments	VLAN aggregation (e.g. PBT); QoS; OAM	Point-to-multipoint

Source: Light Reading, 2006

An important new Ethernet development currently underway is VLAN aggregation, where the issues are how to improve the allocation of VLANs to provide better scaleability and QoS, and also to accommodate this with ITU proposals for OAM. This is a new development for the Ethernet control plane, and is the subject of the Provider Backbone Transport (PBT) proposal to the International Telecommunication Union (ITU) .

Nevertheless, real Ethernet weaknesses remain. In comparison to the IP/MPLS control plane, for example, which runs routing protocols such as OSPF and IS-IS, the Ethernet control plane's routing is pretty feeble, being limited to the various Spanning Tree Protocols (STPs). When multiple VLANs are involved (the norm for carriers), Multiple STP has to be used to create a separate spanning tree for each VLAN. However, STP also removes redundant links within each VLAN, and this usually causes a suboptimal route – and there is also no load balancing.

IP routing, in contrast, will provide multiple routes per destination, thereby enabling load balancing, a critical part of IP forwarding in today's networks. Similarly, with MPLS, knowing that there are multiple routes to the same destination, the edge nodes can direct user traffic into different Label Switched Paths (LSPs) – again, something that STP will not support.

Further, there is no "Time to Live" field in the Ethernet packet, and the consequent absence of hop counting can lead to packets circulating indefinitely if a loop occurs – which can happen, as STP is not very fast at responding to loops in some cases. This is a known problem in Ethernet bridging.

So Ethernet still has some way to go in terms of QoS, OAM, and scaleability. MPLS, on the other hand, looks more obviously like a carrier-class technology, with its features of IP control plane, traffic engineering, OAM, fast reroute, and VPN technology. But it still needs development. One key area for the access network, for example, is point-to-multipoint.

Finally, perhaps the biggest practical drawback to MPLS is that it is not a real link layer, but only a shim layer grafted onto an existing packet structure. This means that MPLS is always implemented as part of something else, so the cost of running MPLS is driven by other parts of the system.

What problems need to be solved? Overall, it looks from a comparison of Ethernet and MPLS that Ethernet is basically cheap, while MPLS is rich in carrier features. The issue is whether these two can be successfully brought together to solve the following problems in the access network:

- **QoS:** The network needs to establish edge-to-edge QoS tunnels and enforce per-flow rates and delays. Without good per-flow policing, user traffic and applications can become unusable. This is also not just a traffic/performance issue, but a matter of the business model of leased lines. And such QoS tunnels are essential for triple-play services over infrastructure shared by many end users.
- **OAM:** The network needs to monitor and detect failures on every user connection. This is a fundamental requirement for carriers, and has always been a challenge throughout the network.

Table 2: MPLS/Pseudowire Over Ethernet

Requirement	MPLS Approach
QoS tunnels and per-flow rate/delay enforcement	RSVP-TE, PW3 QoS
Monitor and detect failure on every user connection	LSP-ping, BFD
All nodes visible to each other	OSPF-TE, ISIS-TE
Low-cost nodes (similar to Ethernet-switch costs)	Purpose-built MPLS/PW3 switch at aggregation points

Source: Light Reading, 2006

- **Manageability:** All nodes in the network need to be visible to each other. Opex is a key issue here, because carriers cannot really afford to put a Layer 2 and a Layer 3 operating system together in the same network. It has to be one or the other.
- **Cost:** Network nodes need to be in the price range of Ethernet switches, simply because there are so many access nodes in real carrier access networks. This compares starkly with backbone networks, which may use only a few score of large routers.

It is not an issue of an Ethernet switch versus an MPLS router. The issue is: At what cost? And how to manage it? So why don't we use MPLS and IP as a control plane, and Ethernet as a physical plane?

The basic point from Table 2 is that MPLS-based Pseudowires already have the capabilities to satisfy the main requirements of the access network. Even where capabilities are less complete – principally OAM – progress has already been made. For example, the existing MPLS OAM mechanisms, such as LSP and Bidirectional Forwarding Detection (BFD) for link fault detection, will, in combination with Ethernet 802.3ah OAM, help. But there still remains much to be done on OAM interworking, an area where there is currently much standards activity.

With a purpose-built MPLS and Pseudowire switch, you can run IP and MPLS as the control plane at the cost of an Ethernet switch.

Ethernet Traffic over Access Networks

Pseudowire is a key term that comes up frequently in discussions about MPLS in the access network, because carriers are not necessarily talking about bringing full routers out to customers' locations, and having a full MPLS/IP-routed infrastructure that goes all the way to these locations. Instead, the point is to use Pseudowires to encapsulate customer traffic by using MPLS tools, and bringing that encapsulated traffic back into the service edge. This is simpler and cheaper than placing routers at every customer location.

A Pseudowire is effectively a simplified version of MPLS. It takes a Layer 2 flow and adds a header in front of it, thereby managing the flow as a virtual circuit. From the provisioning point of view, this is simpler than encapsulating the flow in an RFC 2547 VPN. From the access point of view, all that happens is that a Layer 2 flow (say, Frame Relay or ATM) is mapped into a Pseudowire.

The attractions of a Pseudowire are that it is a circuit with a point-to-point connection and uses the MPLS/IP control plane to manage the flow. Most importantly, it gives edge-to-edge data transport and is transparent to the underlying network.

There are a number of implications of using Pseudowires in the access network:

- **Access devices can be cheap and simple:** There is a simple control plane, and the access device can aggregate user flows with minor packet-forwarding add-ons. The simplicity of Pseudowires enables access devices, such as PONs, CPE, and MSPPs, to be less expensive.
- **Aggregators can interface with any MPLS router at the control plane:** This follows from the support of IP/MPLS control-plane features, and means that user flows can be aggregated toward IP routers for VPN services and the like. Such purpose-built aggregators can be a lot less expensive than MPLS Label Edge Routers (LERs).

- **Pseudowires can be used as demarcation points:** This is highly important, providing security advantages compared to the use of Layer 3 routers.

Figure 2 shows how an MPLS-enabled access network interface using Pseudowires would operate. On the left-hand side, an existing MSPP or a new Ethernet CPE or GPON aggregates incoming Ethernet flows (shown in different colors). On the right-hand side is the access aggregation switch, linked by a single trunk to the access device. Over the trunk are aggregated as many Pseudowires or Layer 2 flows as are needed. The access aggregation switch then forwards the flows appropriately towards the MPLS core.

There are two ways of managing this setup. Currently, some carriers will use out-of-band Pseudowire setup via a proxy. The proxy essentially can talk to both access devices (near and far end) as well as to the access aggregation switch to set up the Pseudowire. However, this may not scale well, and a longer-term solution would be to use a lightweight signaling protocol to negotiate QoS and OAM directly between the access device and aggregator. The Internet Engineering Task Force (IETF) is still working on such a signaling system (known, perhaps inevitably, as the Dry Martini specification).

To clarify what is going on in the data path in an MPLS-enabled access network interface, Figure 3 (shown on page 9) shows how Pseudowire and MPLS tunnel labels are added to a data packet. The data coming in could be Ethernet flows or VLAN flows, for example. At the access device, an MPLS or Pseudowire label would be attached – say, Pseudowire 1000 for one flow and 2000 for another. Any form of transport (such as SONET/SDH ring or Gigabit Ethernet) could be used between the access and edge devices. The edge device then maps the Pseudowires into an MPLS tunnel – say, tunnel 100. IP routing is then used to set up a traditional IP tunnel to the far edge. At the egress side, the process is reversed.

A key point is that the access device is essentially an Ethernet aggregator whose only additional task is to add a Pseudowire label – it does not get involved in any IP routing. So it would share Ethernet's low costs.

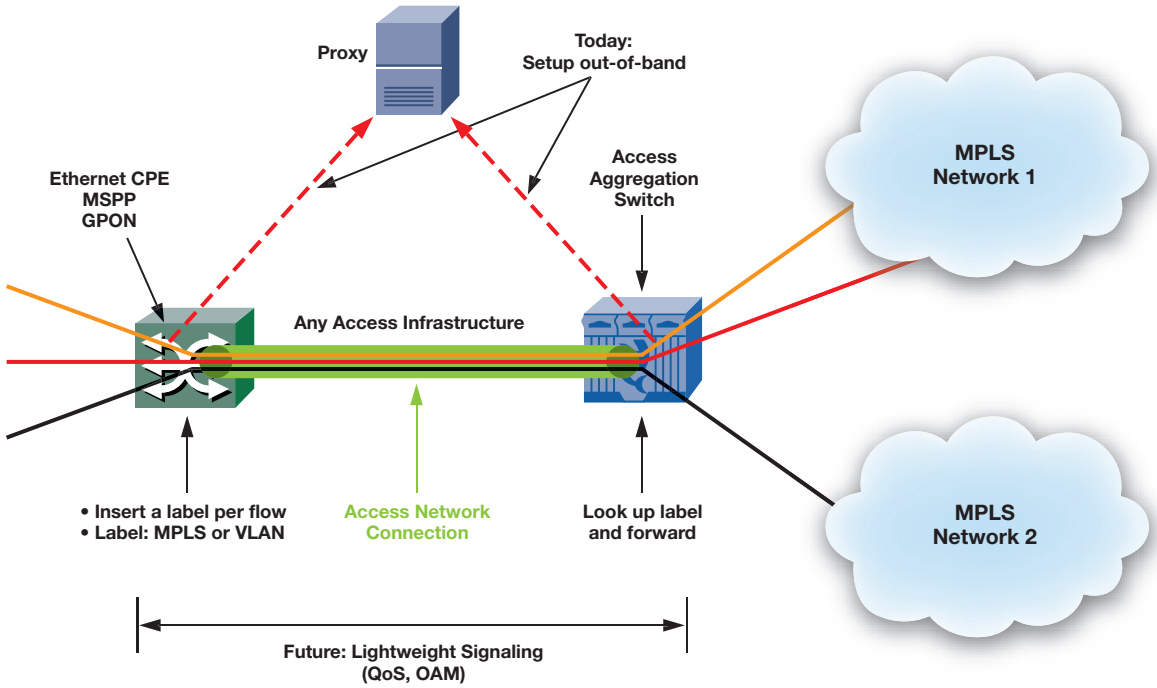

Figure 2. MPLS-Enabled Access Network Interface

Pseudowire Termination Point in Deployment

Figure 4 (shown on page 10) shows the deployment of a possible MPLS/Pseudowire solution in an access network, and indicates how each component works. On the left-hand side are access networks, such as Ethernet or ATM, but also some Pseudowire access networks. Carriers could consider Pseudowire access like this when, for example, they have a lot of remote locations. Instead of upgrading the remote devices individually to Ethernet, it is simpler to backhaul them as native traffic in Pseudowires.

The aggregation point processes the incoming Pseudowires by performing Pseudowire Switching (called segment Pseudowire switching in the IETF Draft Working Group document, Segmented Pseudowire). Essentially, this just switches an incoming Pseudowire from the customer premises into another Pseudowire. This has the attraction of avoiding demarcation issues and the leakage of trunk information between carriers (which carriers generally try to avoid).

The switched Pseudowires then cross the metro access network. Currently this is likely to be a SONET/SDH ring, but, increasingly, some carriers treat this network as being Gigabit Ethernet or 10-Gigabit Ethernet rings. As all metro nodes are taking in customer traffic, every node must be able to operate with the full-specification MPLS control plane (in this case it is running OSPF/IS-IS). It is not necessary for all LSPs inside the metro access network to use MPLS Layer 3 reroute.

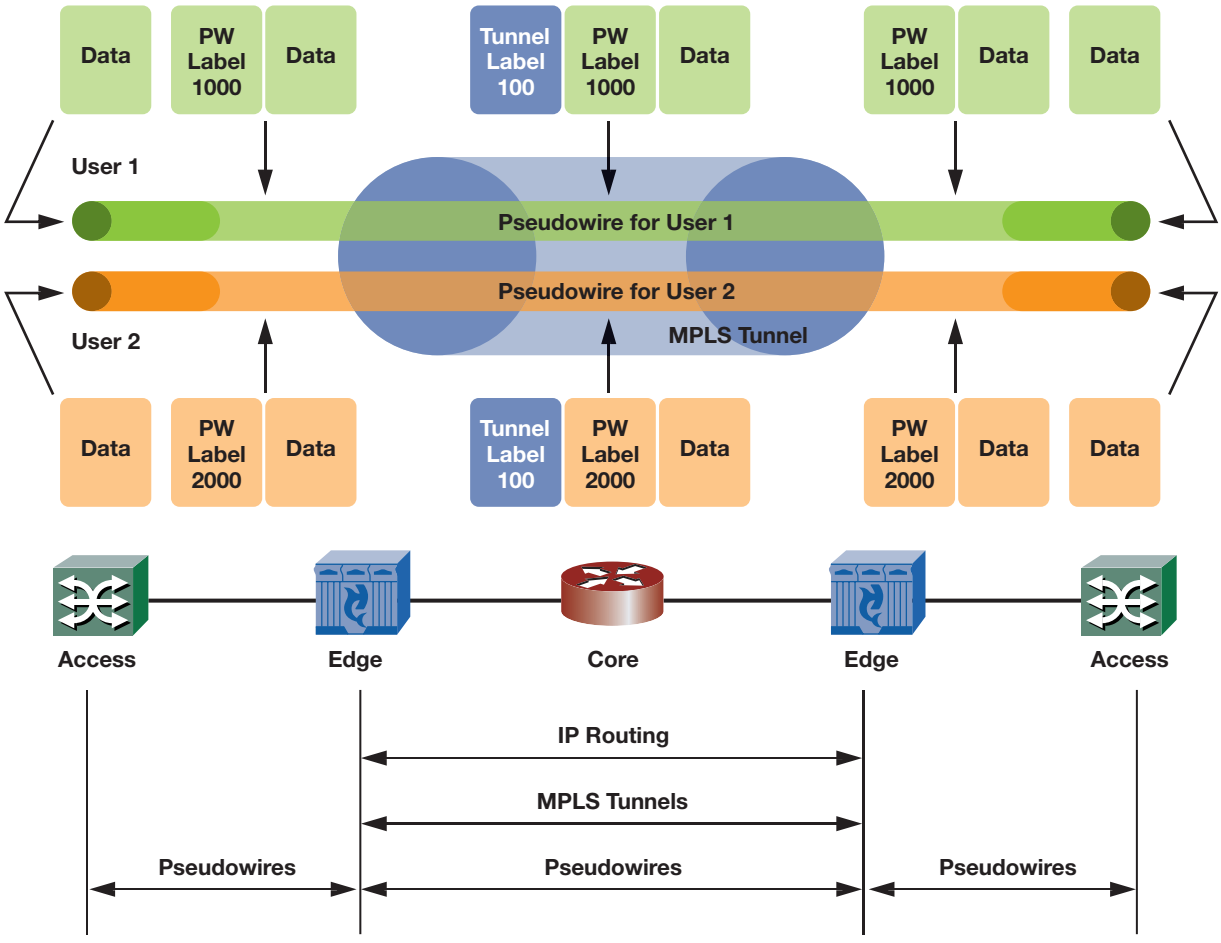

Figure 3. 'Life of a Packet' in an MPLS-Enabled Access Network

Notice the use of a protection Pseudowire. This is another topic currently under consideration by the IETF – 1:1 or 1:n protection is possible, and even a standard MPLS LSP can act as a backup for another. Pseudowire protection has been proposed by Hammerhead Systems, based on its interaction with carriers. A newer version of this draft – Pseudowire Protection – was submitted in February 2006 to the IETF.

A practical issue is Pseudowire QoS. This has been proposed to the IETF as part of the Pseudowire Multihop Draft (see below), and some vendors, including Hammerhead Systems, are actively implementing it. It defines the bandwidth available to each Pseudowire, and the access node will run admission control and aggregate those flows into one of the LSPs.

Another very practical issue with real Pseudowire deployments is their vulnerability to the *N*-squared effect. That is, with N CPE's communicating across a network, the number of connections the network has to support tends to increase as *N*-squared (as a result of full meshing). There are two problems with this:

1. Control plane scaleability – carriers cannot manage the required number of connections that correspond directly to the vast number of CPE devices in real networks – the control plane would crash.
2. Carriers need to manage those flows from the provider edge, as the flow cannot be initiated independently from the CPE.

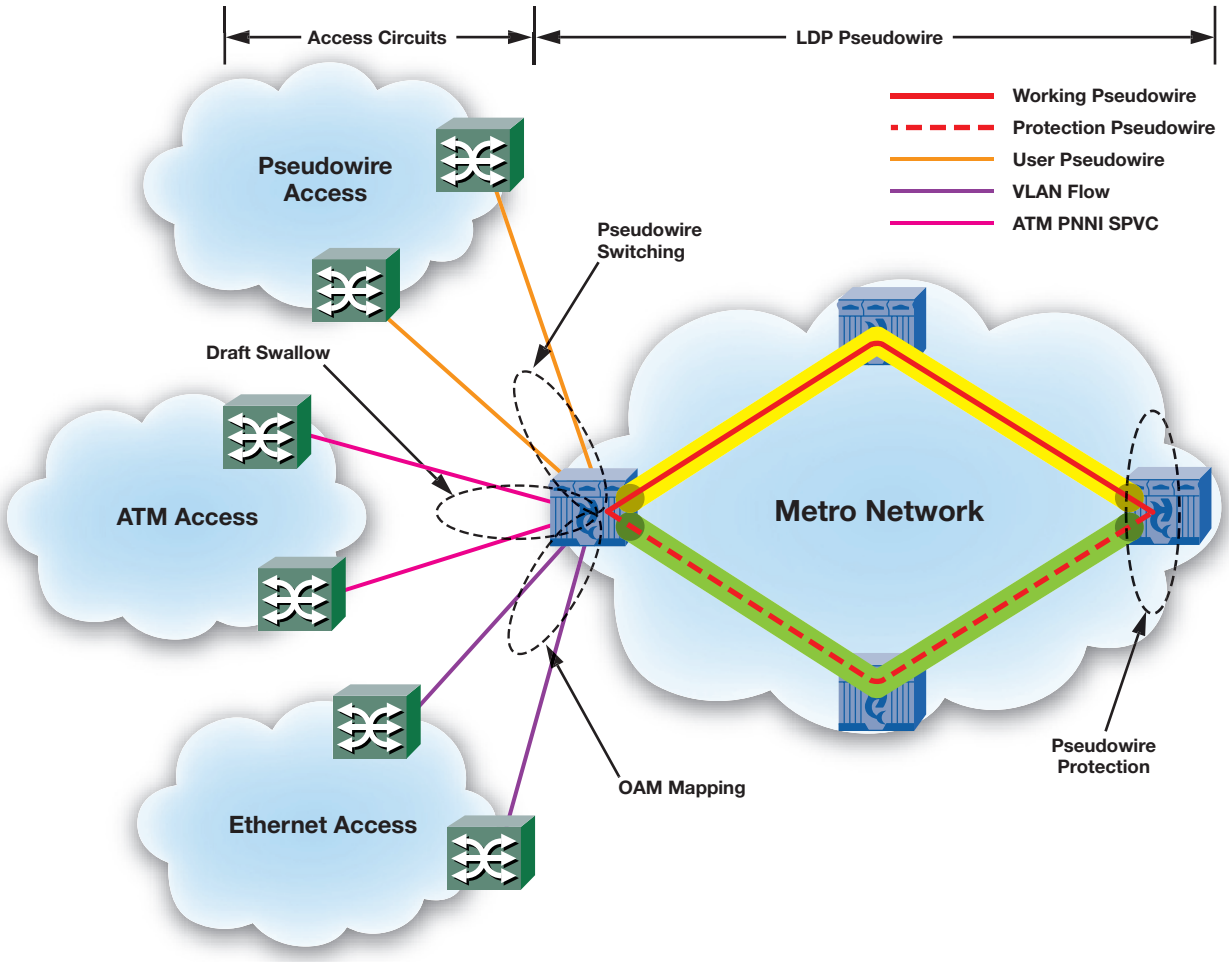

Figure 4. Deployment of an MPLS/Pseudowire Solution

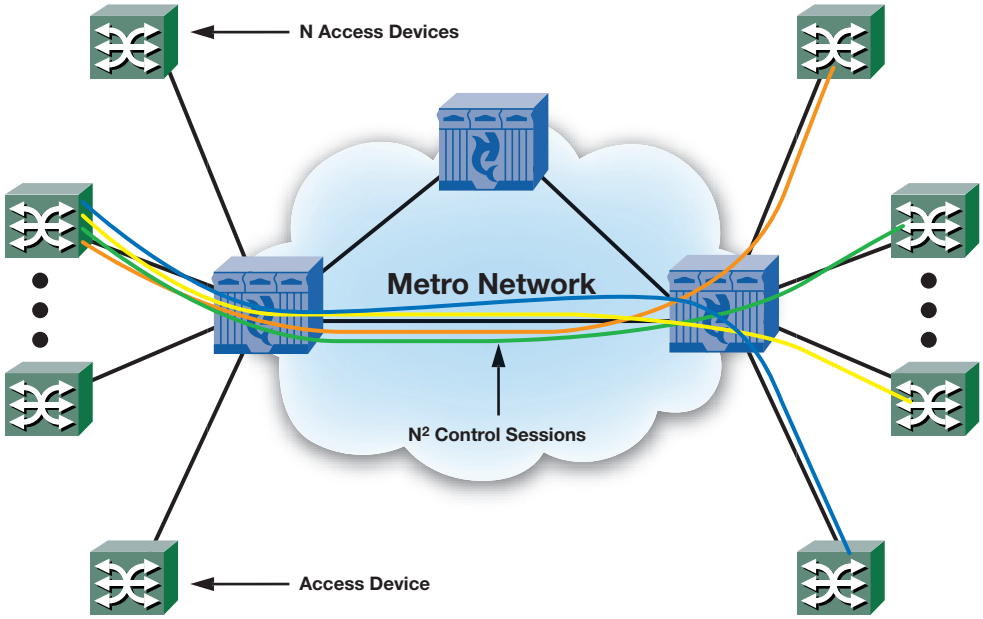

Figure 5. Traditional Point-to-Point Approach

Fortunately, there is a feasible technical solution to this problem, which has created a lot of recent interest within the IETF. This is called Pseudowire Multihop, which is very simple conceptually and basically allows a user on one CPE to initiate a Pseudowire to another CPE across the network. This specification has been explicitly requested by a number of Tier 1 carriers in North America, with multiple vendors currently implementing the specification.

The technique works by switching Pseudowires in a series of hops from switch to switch across the network. Logically, it remains the same connection, but the

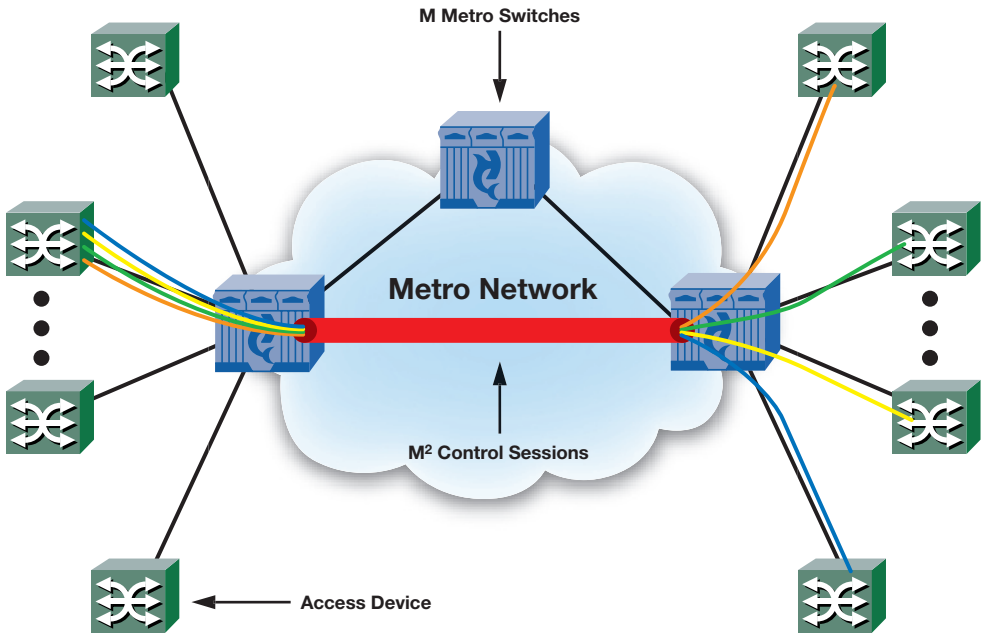

Figure 6. Multi-Hop Pseudowire Approach

data flow will go through multiple hops. Effectively, this introduces a hierarchical switching system where multiple P2P sessions entering aggregation points are mapped into a single LDP session between the edge nodes. This means that the approach scales. Instead of the *N*-squared problem, the number of flows the carrier needs to manage is really only the number of LDP sessions between the provider edge nodes – and this is of order *M*-squared, where M is the number of metro edge nodes, which is significantly less than the number of CPE.

In Figure 5 (shown on page 11) the traditional point-to-point approach will create N^2 control sessions to manage all of the Pseudowires in the network. Each CPS node may need to manage up to (N-1) separate sessions. For a network with 1,000 CPS's, that's 1,000 LDP sessions for each CPE to manage, which is way beyond the means for a simple and inexpensive CPE to handle.

In Figure 6 (shown on page 11) the Multi-Hop Pseudowire approach is to manage the control sessions from the aggregators which will introduce at most M^2 control sessions for all the Pseudowires in the network. Each CPE node will have only one LDP session to each aggregator which is used to manage all of the Pseudowires.

Conclusion

The Ethernet-MPLS combination for access networks continues to attract the attention of carriers and vendors alike. Ethernet provides robust, cost-effective, flexible, high-speed delivery of emerging triple-play IP services. An MPLS framework can extend carrier-class features – QoS, OAM, and scalability – to the access network.

More importantly, MPLS enables the use of Layer 2 Pseudowires to overcome many of the shortcomings of the Ethernet-MPLS combination and to enable cost-effective solutions. Purpose-built Pseudowire-based switches, already proven to be price-equivalent to Ethernet switches, deliver many benefits to access networks:

- **Affordable access devices:** The simple control plane and packet-forwarding techniques of Pseudowire switches lend themselves to low-cost PONs, CPE, and MSPPs.
- **Efficient aggregation for and access to the core:** For much lower cost than MPLS Label Edge Routers (LERs), purpose-built Pseudowire switches manage user flows between the access layer and the core IP/MPLS routers.
- **Increased security and reliability:** Pseudowires can act as demarcation points, using Pseudowire Switching techniques. Another feature, Pseudowire Protection, lets a standard MPLS LSP act as a backup for another.

The most recent advances in Pseudowire implementations and related industry standards – Pseudowire QoS and Pseudowire Multihop – are further strengthening the argument in favor of Ethernet and MPLS in the access network. These and other innovations have already been implemented by Hammerhead Systems. Carriers can easily integrate Pseudowire switches into existing networks today (as described in this paper), without disruption of existing services, and be positioned to fully exploit the capex and opex advantages of these evolving trends in access networks.

Glossary

ATM	Asynchronous Transfer Mode
DSLAM	Digital Subscriber Line Access Multiplexer
IETF	Internet Engineering Task Force
IS-IS	Intermediate System-Intermediate System (IP)
LSP	Label-Switched Path (MPLS)
LDP	Label Distribution Protocol (MPLS)
LER	Label Edge Router (MPLS)
MFA Forum	MPLS/Frame Relay Alliance Forum
MPLS	Multi-Protocol Labeled Switching
OAM	Operation, Administration, and Management
OSPF	Open Shortest Path First (IP)
PON	Passive Optical Network
PW	MPLS Pseudowire (same as PWE3)
PWE3 Emulation	MPLS Pseudowire Edge-to-Edge
QoS	Quality of Service
RSVP-TE	Resource Reservation Protocol for Traffic Engineering
TDM	Time-Division Multiplexing
VLAN	Virtual LAN (Ethernet)

Hammerhead Systems, Inc.
640 Clyde Court
Mountain View, CA 94043
650-210-3300 phone
650-210-3303 fax
www.hammerheadsystems.com

610-0010-003 050406