ovum

The benefits of P2MP label switched paths for MPLS networks

10 November 2008

www.ovum.com

The benefits of P2MP label switched paths for MPLS networks

RSVP-TE-based point-to-multipoint (P2MP) label switched paths (LSPs) provide network operators with the ability to improve quality of experience, increase service availability, and reduce core complexity.

Analysis

Improving the performance and efficiency of network services

Every significant transport technology has supported P2MP (point-to-multipoint) constructs, including ATM (asynchronous transfer mode), IP (Internet protocol), SDH (synchronous digital hierarchy), and WDM (wavelength division multiplexing). Sometimes the capability is referred to as multicast and other times as drop-and-continue. In all cases, it was considered an essential part of the overall capabilities of a networking technology. As MPLS has become more pervasive in the delivery of services, it too has been driven to integrate this essential construct and over the next five years this capability is expected to be supported by all major equipment suppliers and to be adopted first by leading service providers and thereafter throughout the industry.

P2MP LSPs (label switched paths) are the base functionality required to support this capability. P2MP LSPs can be used to deliver video broadcast services, such as IPTV, and also built upon to deliver IP VPNs (virtual private networks) and Ethernet VPNs (for example, VPLS – virtual private LAN service / E-LAN). Using P2MP LSPs, bandwidth usage and costs can be reduced, performance improved, and scalability enhanced.

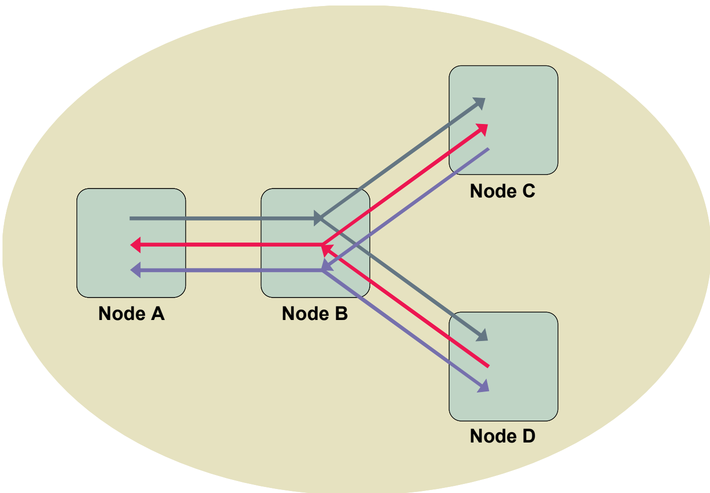
With the addition of P2MP LSPs to an existing MPLS network, MPLS can now be used for all services whether they require point-to-point, point-to-multipoint, multipoint-to-point, or multipoint-to-multipoint connectivity. When all services are based on one technology, like MPLS, a common approach to network management and network architecture can be pursued by network operators in the same way as they are today in the case of SONET/SDH and WDM.

This report describes the benefits of P2MP LSPs, focusing on the technical benefits that lead to bandwidth reduction, enhance scalability, improve quality of experience, and increase service availability while reusing and extending control plane and data plane processing already deployed for point-to-point and multipoint-to-point services.

Networking constructs

Connection-oriented network technologies have traditionally implemented two circuit constructs: point to point and point to multipoint. With these two constructs, any connection-oriented connectivity can be established including point to point (one to one), point to multipoint (one to many), and multipoint to multipoint (many to many). Multipoint to point (many to one) is implemented in connectionless networks but is typically not implemented in connection-oriented networks because of the challenge of providing quality of service when new sources are added and underlying facilities change.

Figure 1 **Any-to-any connectivity with point to multipoint**



Source: Ovum

Using point-to-point and point-to-multipoint constructs, MPLS RSVP-TE can support one-to-one, one-to-many, and many-to-many communications. The two fundamental networking constructs required in connection-oriented networks are point to point and point to multipoint and they are both supported by MPLS RSVP-TE, with the latter being standardized as "P2MP LSP."

The benefits of P2MP LSPs

P2MP LSPs are created using the same signaling mechanisms as point-to-point RSVP-TE LSPs and therefore share many of the same benefits: deterministic routing, bandwidth reservations, fast reroute, and existing OAM mechanisms.

Deterministic routing

As with existing RSVP-TE-based LSPs, P2MP LSPs can be established by specifying some or all of the nodes that must be in a path between a P2MP source and its destinations. This provides an operator with precise control over what paths P2MP LSPs follow, service separation, and traffic balancing. Deterministic routing is not currently implemented in IP multicast.

Network operators have deployed P2MP LSPs in order to achieve disjoint/spatially redundant paths for video distribution. This approach, in combination with dual sourcing of multicast trees, can be deployed with MOS scoring equipment so that video signal selection can be changed not just when there is a hard failure but when there is performance degradation as well.

Bandwidth reservations

P2MP LSPs, again in keeping with existing RSVP-TE-based LSPs, support online and offline capacity management processes through the specification of per-LSP bandwidth reservation. Bandwidth reservation is not currently supported in IP multicast.

Fast reroute

Like existing RSVP-TE-based LSPs, P2MP LSPs can be recovered after a failure by using fast reroute with make before break. While IP multicast depends on IGP convergence and has a recovery range after link failure of hundreds of milliseconds to a few seconds, P2MP LSPs have a recovery range of tens of milliseconds to fifty milliseconds.

Orange UK, a large European operator, has deployed MPLS fast reroute for IPTV distribution without using any underlying optical protection mechanisms.

Core complexity reduction

Some network operators are looking to reduce the complexity, scaling problems, and potential for failure of core nodes by reducing the number of protocols (and hence the state/memory and CPU load) executing on core nodes.

In MPLS networks, BGP4 can be removed from core nodes by using other nodes as BGP4 route reflectors. A further simplification of core nodes is to remove multicast protocols like PIM. With MPLS networks, the same protocols that are already in place, RSVP-TE, for example, can be used to establish P2MP trees and therefore PIM is not required at the core.

Standards progress

The base P2MP LSP specification is already published as an informational IETF (Internet Engineering Task Force) RFC (request for comment) known as RFC 4461 "Signaling Requirements for Point-to-Multipoint Traffic-Engineered MPLS Label Switched Paths (LSPs)."

The IETF Layer-2 VPN working group has accepted as working group documents drafts addressing the use of P2MP LSPs in Ethernet VPNs. The main draft is: draft-ietf-l2vpn-vpls-mcast-04.txt "Multicast in VPLS."

The IETF Layer-3 VPN working group has accepted as working group documents drafts addressing the use of P2MP LSPs in IP VPNs. The main draft is draft-ietf-l3vpn-2547bis-mcast-07.txt "Multicast in MPLS/BGP IP VPNs."

The IETF MPLS working group has accepted as working group documents drafts that address extensions to MPLS OAM mechanisms such as the following: draft-ietf-mpls-p2mp-lsp-ping-06.txt "Detecting Data Plane Failures in Point-to-Multipoint Multiprotocol Label Switching (MPLS) - Extensions to LSP Ping."

In summary, there is a stable informational specification for the base P2MP LSPs capability, and there are drafts being implemented and working their way through standards bodies addressing IP VPNs, Ethernet VPNs, and OAM.

Equipment supplier support

Juniper Networks has been deploying P2MP LSP solutions since 2004 and Ovum anticipates support for P2MP MPLS LSPs by most of the IP/MPLS market share leaders in 2009.

Deployment experience

To date, P2MP LSPs have been deployed in most major geographical regions (Asia, Europe, and North America) including ten deployments in Europe, two in Asia, and one in North America. Many trials are also currently taking place. P2MP LSPs have been deployed by large incumbent network operators, small network operators, and a large cable/MSO operator.

Content distribution

Content distribution can be concerned with any content but two well known types are financial services data and video. In both cases, packet loss has a significant impact on the consumers of the content. In the case of financial services, packet loss can mean the difference between doing a trade that generates millions of dollars in profit or missing the trade. In the case of video distribution, packet loss leads to service interruption, reduced quality of experience, and unhappy customers.

All MPLS mechanisms that contribute to reduced congestion for a service (explicit routing, bandwidth reservation, and policy-based queuing) improve the experience of a service. Fast reroute, however, could make the difference between seeing a small imperfection in video to having a significant interruption. With restoration rates of 10–50 milliseconds, it is theoretically possible that a P2MP LSP could be restored after a link failure in less than a single full image frame (33 milliseconds for NTSC HD and 40 milliseconds for PAL HD). When using IP multicast, the minimum recovery time is likely to be 2–3 full image frames and could be as much as 30 to 60 full image frames. The potential for significant customer dissatisfaction with P2MP LSP is much less than with IP multicast.

For distribution of data content, such as financial services, interruptions of more than 2–3 milliseconds can mean the difference between making a successful trade or not.

For content distribution, P2MP LSPs could provide a significant improvement to performance and quality of experience.

Multicast IP VPNs

Today, the most widely used approach to multicast IP VPNs is draft-rosen-vpn-mcast-08.txt "Multicast in MPLS/BGP VPNs." The industry has recognized this as a first attempt at understanding the requirements of standards in this area, while acknowledging the need to evolve this approach with a new standard such as the current IETF L3VPN working group document draft-ietf-l3vpn-2547bis-mcast-07.txt "Multicast in MPLS/BGP IP VPNs." New multicast IP VPN standards are likely to specify options for both existing approaches to multicast IP VPNs (for example, the Rosen draft) as well as P2MP LSP approaches.

The main concern with existing approaches to multicast IP VPNs (MVPN) is the scaling challenges, especially in the area of control plane processing, as a result of the need for every PE to maintain a PIM adjacency per MVPN with every other PE that is participating in the MVPN.

The scaling challenges with existing approaches to MVPNs are not present with P2MP LSPs because they leverage the adjacencies that already exist in an MPLS network (for example, RSVP-TE and BGP4) and do not create adjacencies on a per-MVPN basis.

Table 1 **Multicast VPN comparison**

Characteristic	PIM-based IP multicast	P2MP MPLS LSP TE
Resource reservation	No	Yes
Same operations framework in core for unicast and multicast	No	Yes
<50 millisecond path restoration	No	Yes
Deterministic path management	No	Yes
PIM adjacency required between each PE in a MVPN	Yes	No

Source: Ovum

The number of IP VPNs a network operator needs to support is currently limited by the fact that operators focus mostly on selling to large enterprises. Even at this scale, existing approaches to multicast IP VPNs may be challenged. Operators are also now targeting or planning to target small and medium-sized businesses, which means future scaling challenges may be much worse.

Ethernet

Ethernet is now being used as both a transport technology (transferring service traffic from aggregation points to service edges) and a service delivery architecture. In both cases, point to point (often referred to as "E-Line"), point to multipoint (often referred to as "E-Tree"), and multipoint to multipoint (often referred to as "E-LAN") are required capabilities. P2MP LSPs can be used to optimize the delivery of multipoint Ethernet capabilities, including E-LAN and E-Tree.

E-LAN

In metro/WAN architectures the most common way to provide multipoint-to-multipoint Ethernet capability is with VPLS (virtual private LAN service) over MPLS, often referred to simply as VPLS.

VPLS is often implemented using a mesh of point-to-point LSPs. The challenge with this approach is that whenever a source PE broadcasts or multicasts traffic, it needs to create a copy of each frame for each VPLS destination and send it separately. This creates traffic-processing scaling challenges as the number of endpoints in a VPLS grows.

An approach to reducing this VPLS scaling problem is to use P2MP LSPs. With this approach, a source PE need only send one frame, dramatically reducing the processing load, and rely on the multicast LSP mechanism within the network to deliver frames to each VPLS endpoint.

E-Tree

An E-Tree is a point-to-multipoint Ethernet construct that can be used for transport or services. In today's Ethernet-over-MPLS service networks, VPLS instances have to be placed strategically in the network, for example, a core ("P") node, so that multicast trees can be optimally created. The approach of using VPLS instances in core nodes results in those nodes having to learn MAC addresses, as well as the need to maintain other VPLS state, both of which are contrary to the goals some operators have of reducing the complexity of the core by reducing the number of control plane protocols, maintenance events, and operations procedures.

An alternative to using VPLS/H-VPLS mechanisms in the core and complicating the core with additional control plane protocols is to establish VPLS instances over multicast MPLS LSPs. There are three ways being considered to do this:

1) BGP-signaled VPLS over RSVP-TE-signaled P2MP LSPs
2) T-LDP (targeted LDP)-signaled VPLS over RSVP-TE-signaled P2MP LSPs
3) T-LDP-signaled VPLS over mLDP (multicast LDP)-signaled P2MP LSPs

In all cases, an underlying "transport" LSP is created using RSVP-TE or mLDP, and then services signaling is established using BGP or T-LDP.

Options 1 and 2 above are currently in draft form within the IETF. Option 3 is not yet in draft form. All options have some potential to relieve the complication of core networks while Option 3 does rely on a variant of LDP, which is not always a protocol used in the core of networks. Option 3 is the simplest option from a path construction perspective, but does not provide the benefits of bandwidth management, disjoint paths, and fast reroute that RSVP-TE-based solutions like Options 1 and 2 do.

Network operator testimony

In researching this subject, Ovum was able to collect information about existing uses of P2MP LSPs in networks, which include over a dozen production deployments globally.

One network operator in particular, referred to here as "Operator X," was able to provide us a significant amount of insight into its deployment. Operator X has built a converged voice, video, and data network. One of the most performance-sensitive services on that network is broadcast streaming video, and it is for this reason that Operator X deployed P2MP LSPs in production over a year ago.

In Operator X's view, one of the most important capabilities enabled by RSVP-TE-based P2MP LSPs is the ability to form disjoint multicast trees sourced from different video headends in different parts of the country but terminating in all the regional distribution points. Disjoint that is trees (multicast topologies) do not have links in common. Link failures can be common in large networks. Packet loss, which degrades performance but does not cause a hard failure, can also occur. In both cases, Operator X can monitor the performance of the two video streams arriving in regional distribution points over disjoint trees/paths and make a real-time decision about which stream has the best quality and distribute that stream. Operator X believed it was unlikely that this could be achieved using traditional IP multicast mechanisms and even if it could, it would be very operationally complex.

While deploying P2MP LSPs did require learning about how multicast trees are constructed and maintained, Operator X found that what had to be learned was incrementally not significant compared to what operations people already knew about MPLS LSPs, similarly for signaling troubleshooting.

Operator X has been pursuing a "service-free" core architecture, already having removed BGP4 from the core. The deployment of P2MP LSPs meant that IP multicast protocols like PIM could also be removed from the core. Operator X told Ovum that a service-free core made the NOC (network operations center) employees happy because it reduced the complexity of troubleshooting, reduced the frequency with which core routers had to be changed, and that in turn also increased the reliability and availability of the core.

In summary, the major benefits Operator X achieved from implementing P2MP LSPs include:

- the ability to form disjoint multicast tree paths and thus reduce the impact of link failures degrading both streams from each video headend at the same time
- the small incremental learning curve above point-to-point MPLS LSPs
- and the ability to continue implementing a service-free core architecture, thereby reducing operational complexity and increasing core availability

Conclusion

There are numerous options available to distribute IP services via a multicast tree. All of these approaches have strengths and weaknesses. In this report, Ovum focused on one approach, RSVP-TE-based P2MP LSPs, which have numerous benefits, including:

- reduction in the likelihood of network failures impacting redundant content streams
- simplification of service maintenance by reduced core complexity
- improved availability of core platforms by reduced maintenance

- small incremental learning curve for network operators that have already operationalized MPLS
- fast service restoration
- improved multicast VPN scaling characteristics
- maturing standards
- maturing technology based on over 12 global deployments

Glossary

BGP4: Border Gateway Protocol Version 4. BGP4 was developed for the exchange of IP routing information between IP networks under separate administrative control but has been extended for the distribution of other control plane information as well as including VPN identifiers and MPLS labels.

E-LAN: Ethernet LAN. An emulated Ethernet local area network.

E-Line: Ethernet Line. An emulated point-to-point Ethernet private-line service.

E-Tree: Ethernet Tree. An emulated Ethernet multicast tree.

LDP: Label distribution protocol. An MPLS control plane protocol for distributing MPLS labels.

MPLS: Multiprotocol label switching. A forwarding plane protocol that enables tunnel stacking, bandwidth management, fast failure restoration, link partitioning, and address space partitioning within an IP network.

Multicast: A unidirectional logical topology that consists of one source and multiple destinations. Often depicted as a tree structure.

P2MP: Point to multipoint. Synonymous with multicast.

RSVP-TE: Resource reservation protocol with traffic engineering extensions. A signaling protocol used to form deterministic paths with or without resource allocations.

Service-free architecture: A logical control plane architecture that results in core platforms only participating in core transport tunnels and not logical service topologies creation and maintenance.

SONET/SDH: Synchronous optical network/synchronous digital hierarchy. A circuit-based transport technology based on time division multiplexing concepts.

T-LDP: Targeted LDP. A variant of LDP that allows exchanges of MPLS labels between IP/MPLS platforms that are not directly connected to each other, but connected over an underlying "transport" MPLS LSP.

VPN: Virtual private network. A logical subset of a topology connecting only endpoints of a specific network service.

WDM: Wavelength division multiplexing. A circuit-based transport technology based on frequency/wavelength division multiplexing concepts.
