

Core of Multicast VPNs: Rationale for Using mLDP in the MVPN Core

Exploring Suitability of Using mLDP Versus P2MP RSVP-TE in the MVPN Core

Multicast Virtual Private Network (MVPN) is a popular technology for transporting IP multicast traffic within a Border Gateway Protocol (BGP)/Multiprotocol Label Switching (MPLS) IP VPN from one VPN site to another. There are a number of different options and procedures for realizing different components of the MVPN solution. Some of these are better suited for MVPNs than others, depending on the type of service being offered, the applications supported, and the amount of traffic transported.

A generic MVPN service delivered by service providers or large enterprises should be able to cater to a large number of different applications such as video distribution, resource discovery and advertisement, software and file distribution, finance, surveillance, and video conferencing, among others. These are some of the most common applications seen in existing MVPN deployments today, and they have very different traffic patterns and characteristics. Most of these applications have dynamic participants and sporadic transmission profiles.

MVPN services should not place any restrictions on customer protocols supported and should have no impact on customer devices. An MVPN service should be simple to maintain, easy to deploy, and be able to scale to future network demands in addition to maintaining existing levels of service and performance when compared to unicast VPNs.

Transport for the MVPN Core

Multicast Label Distribution Protocol (mLDP), Point-to-Multipoint (P2MP) Resource Reservation Protocol—Traffic Engineering (RSVP-TE), and Generic Routing Encapsulation (GRE)/Protocol Independent Multicast (PIM) are three popular variants of transport technologies for transmitting packets in the MVPN core. GRE/PIM has been widely deployed and scales well today. Cisco continues to make enhancements and support GRE/PIM deployments that will meet the demands of future networks. For customers interested in moving to a label-based MVPN solution, there are two different options:

- **mLDP:** Provides extensions to LDP for the setup of P2MP and multipoint-to-multipoint (MP2MP) label switched paths (LSPs) in MPLS networks.
- **P2MP RSVP-TE:** Provides extensions to RSVP-TE for the setup of traffic-engineered P2MP LSPs in MPLS and generalized MPLS (GMPLS) networks.

mLDP and P2MP RSVP-TE Characteristics

mLDP and P2MP RSVP-TE can construct LSPs without relying on or interacting with any multicast tree-construction protocol. Migrating to or deploying a label-based transport for MVPN in the core results in a common data plane for unicast and multicast. Service providers and large enterprises can use their existing MPLS infrastructure for transporting IP multicast data. Existing MPLS benefits such as Fast Reroute (FRR) can now be applied to IP multicast. There is no need for enabling PIM in the core even if the other MVPN components, such as provider edge-to-provider edge (PE-PE) exchange of customer multicast routing information or provider edge-to-customer edge (PE-CE) protocol exchanges, use PIM. This leads to simplification of the core and the control plane.

mLDP and P2MP RSVP-TE are different in the way they construct trees, the types of trees that they can support, and the features they offer. In mLDP, the LSPs are built from the tailend to the headend, while in P2MP RSVP-TE, the LSPs are built from the headend to the tailend. mLDP supports both P2MP and MP2MP LSPs, while P2MP RSVP-TE supports only P2MP LSPs. P2MP RSVP-TE supports FRR inherently, while mLDP supports FRR through unicast

point-to-point (P2P) TE. In mLDP, the control plane is P2MP or MP2MP, whereas in the case of P2MP RSVP-TE, the control plane is P2P. The data plane is P2MP for both mLDP and P2MP RSVP-TE. See Table 1.

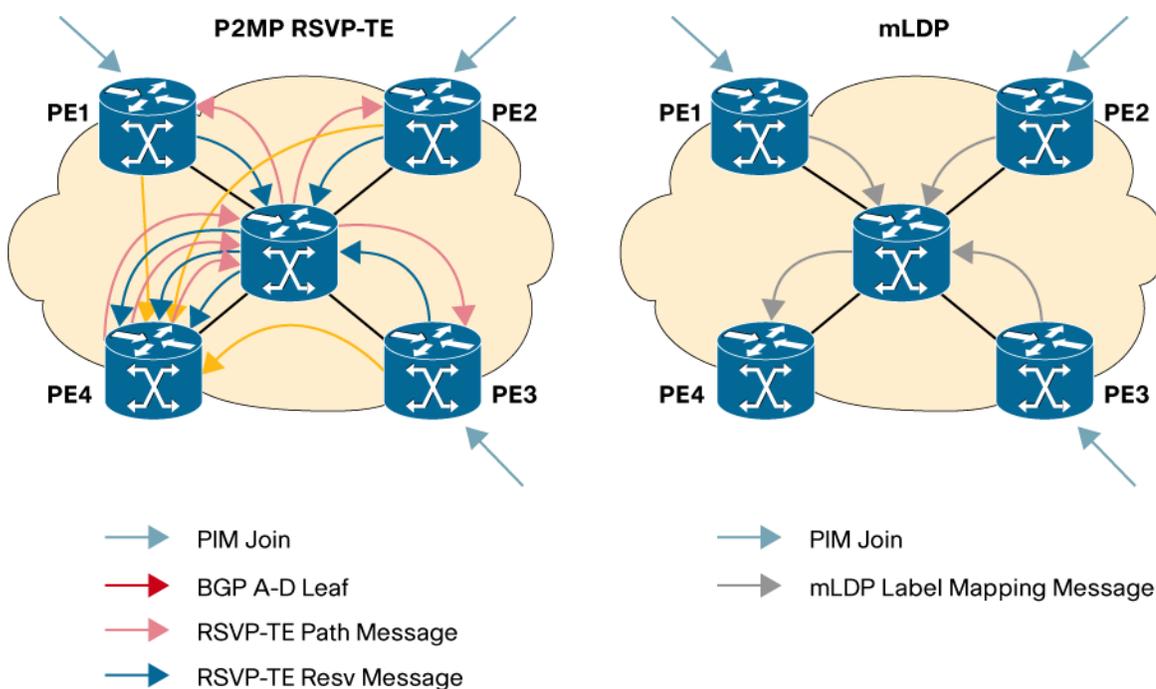
Table 1. Characteristics of P2MP RSVP-TE and mLDP

P2MP RSVP-TE	mLDP
LSPs are build from headend to tailend	LSPs are build from tailend to headend
Supports only P2MP LSPs	Supports P2MP and MP2MP LSPs
Supports traffic engineering	Supports FRR through unicast P2P TE
Signaling is periodic	No periodic signaling
P2P technology at control plane	Control plane is P2MP or MP2MP
P2MP at the data plane	Data plane is P2MP

mLDP and P2MP RSVP-TE Operations in MVPN

These inherent differences between mLDP and P2MP RSVP-TE affect core router states and the number of messages sent and received. Let us consider an example where we have a PIM join coming into the MVPN cloud and the core tree is being established by using mLDP and P2MP RSVP-TE (Figure 1).

Figure 1. Example of mLDP and P2MP RSVP-TE Use in MVPN



In both cases PE1, PE2, and PE3 are the tailend routers, and PE4 is the headend router. Let us consider a single PIM join coming into the network cloud at PE2. In the case of P2MP RSVP-TE, a BGP A–D Leaf message to the headend router (PE4) asks it to create an LSP toward PE2. PE4 then initiates an RSVP-TE PATH message toward PE2. PE2 responds by sending an RSVP-TE RESV message to PE4. The tree is now set up from PE4 to PE2. If PE1 and PE3 need to join the core tree, they need to follow the same procedure and signal PE4 to initiate creation of the tree. Thus for the simple example of three tailends joining a single tree, the core router receives and sends six updates. The ingress PE receives three PATH messages and sends three RESV messages.

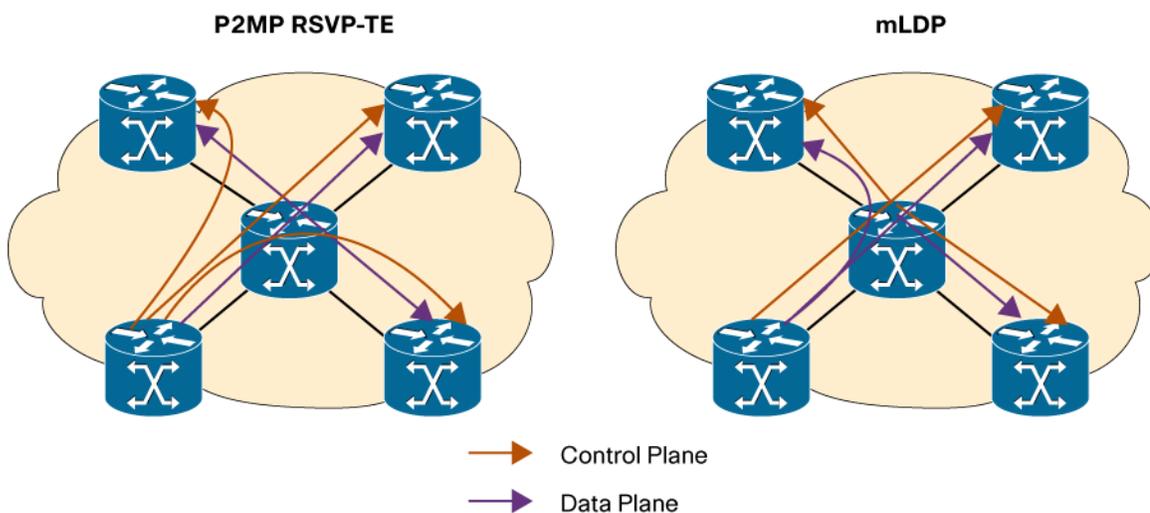
In case of mLDP, when a PIM join comes into the network cloud at PE2, a label mapping message is sent toward the headend (PE4). When other tailends want to join the tree, they only need to send a label mapping message to the

closest branch point of the tree. In this case, for three tailends joining a single tree, the core router receives only three updates. The ingress PE receives only one label mapping message.

From this example it is clear that join and leave are very expensive operations in P2MP RSVP-TE compared to mLDP. For a complex network the number of messages exchanged will be much lower for mLDP than compared to P2MP RSVP-TE. This is due to the inherent nature of the protocols: the headend has to be signaled for setting up the LSP for every join for P2MP RSVP-TE, as the LSP setup is headend initiated, whereas in the case of mLDP, a message has to be sent only to the closest branch point, as the LSP setup is tailend initiated.

Let us consider the same example and compare the amount of state (Figure 2).

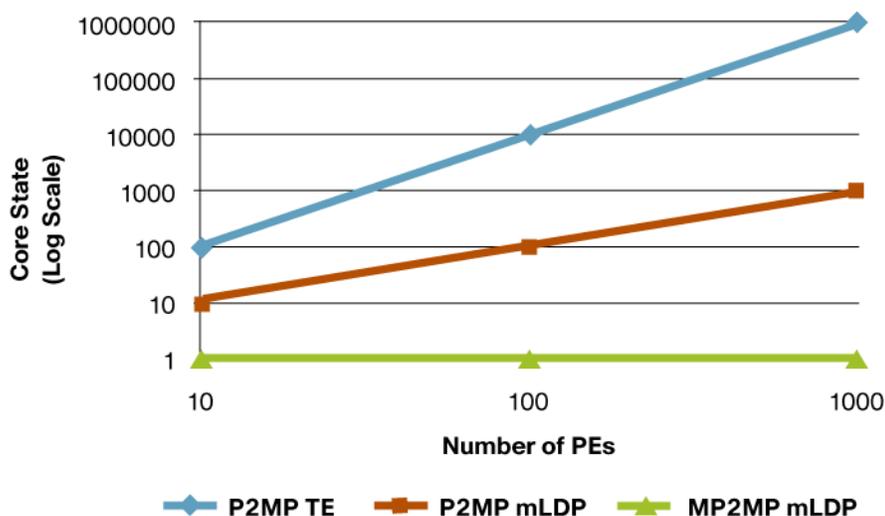
Figure 2. Comparison of State for mLDP and P2MP RSVP-TE Use in MVPN



In case of P2MP RSVP-TE, there are three P2P sub-LSPs built from the ingress PE to the leaves in the control plane. In the data plane, these sub-LSPs are merged to form one P2MP LSP. The core router has to maintain state for each individual leaf. The ingress PE has to maintain three sub-LSPs. In case of mLDP, there is a single P2MP LSP for the control plane and a single P2MP LSP at the data plane. The core and the ingress PE have to maintain only one state independent of the number of MVPN participants.

The number of states and messages in case of P2MP RSVP-TE is related to the number of MVPN participants in the MVPN, whereas in the case of mLDP, the number of states and messages is independent of the number of participants. As the number of participants grow, the number states and messages for P2MP RSVP-TE-based MVPNs increases considerably, making mLDP a better choice for core transport for MVPNs.

P2MP RSVP-TE does not support MP2MP LSPs, so there is a need to construct a large number of trees for supporting MVPNs. A full mesh of P2MP LSPs will be required to connect all the sites. P2MP RSVP-TE also requires periodic signaling to refresh states. mLDP can support MP2MP LSPs, making it the ideal choice for MVPNs. If MP2MP LSPs are used, then only one tree is required to interconnect all the sites. mLDP does not need periodic signaling (Figure 3).

Figure 3. Comparison of State for Label-based MVPN Transport Technologies

There is a mechanism to improve scalability in case of P2MP RSVP-TE by doing aggregation. Doing a lot of aggregation results in data being sent toward sites that are not interested in that data, resulting in wastage of bandwidth. Not doing enough aggregation results in scalability issues. It is very difficult to aggregate LSPs in the case of MVPNs, as the participants are very dynamic, making it extremely difficult to predict a good aggregation scheme.

To support bidirectional PIM customers in an MVPN, a full mesh of P2MP LSPs are needed in the case of P2MP RSVP-TE, while in the case of mLDP, a single MP2MP LSP can be used. A full mesh of P2MP LSPs will require upstream assigned labels to solve the designated forwarder problem, whereas MP2MP LSPs do not require this.

mLDP is better suited for MVPNs. It is simple, robust, and easy to deploy, possesses similar properties to multicast protocols, and can be effectively used for generic MVPN deployments.

Conclusion

There are a number of mechanisms for core transport for MVPNs. GRE/PIM is widely deployed today and is a viable option for core MVPN transport. For label-based MVPN solutions, mLDP presents a better choice for core MVPN transport when compared to P2MP RSVP-TE. mLDP is receiver driven, does not require periodic signaling, supports MP2MP LSPs, and is better suited for MVPN deployments. It is suitable for secondary video and consumer video distribution where participants are dynamic. P2MP RSVP-TE is better suited when there are stringent TE requirements and participants are not dynamic, such as primary video and studio-to-studio distribution. mLDP scales better than P2MP RSVP-TE for MVPNs. MP2MP LSPs provide huge savings for mLDP-based MVPNs. mLDP is the protocol of choice for the MVPN core.



Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

CCDE, CCENT, CCSI, Cisco Eos, Cisco Explorer, Cisco HealthPresence, Cisco IronPort, the Cisco logo, Cisco Nurse Connect, Cisco Pulse, Cisco SensorBase, Cisco StackPower, Cisco StadiumVision, Cisco TelePresence, Cisco TrustSec, Cisco Unified Computing System, Cisco WebEx, DCE, Flip Channels, Flip for Good, Flip Mino, Flipshare (Design), Flip Ultra, Flip Video, Flip Video (Design), Instant Broadband, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn, Cisco Capital, Cisco Capital (Design), Cisco:Financed (Stylized), Cisco Store, Flip Gift Card, and One Million Acts of Green are service marks; and Access Registrar, Aironet, AllTouch, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Lumin, Cisco Nexus, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Collaboration Without Limitation, Continuum, EtherFast, EtherSwitch, Event Center, Explorer, Follow Me Browsing, GainMaker, iLYNX, IOS, iPhone, IronPort, the IronPort logo, Laser Link, LightStream, Linksys, MeetingPlace, MeetingPlace Chime Sound, MGX, Networkers, Networking Academy, PCNow, PIX, PowerKEY, PowerPanels, PowerTV, PowerTV (Design), PowerVu, Prisma, ProConnect, ROSA, SenderBase, SMARTnet, Spectrum Expert, StackWise, WebEx, and the WebEx logo are registered trademarks of Cisco and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1002R)