White Paper

Multicast Virtual Private Networks

Executive Summary

The purpose of this paper is to provide an understanding and background to IP Multicast in relation to Virtual Private Networks (VPNs) and to describe in detail the Cisco architecture for a Multicast VPN solution. This is aimed toward service providers providing an MVPN for Enterprise customers only.

Virtual Private Networks

A VPN is network connectivity across a shared infrastructure (such as an ISP). It aims to provide the same policies and “performance” as a private network, at a reduced cost of ownership, thus creating many opportunities for cost savings through operations and infrastructure.

Tunnels through VPNs

Historically, IP in IP / Generic Route Encapsulation (GRE) tunnels was the only way to connect through a service provider network. While such tunnelled networks tend to have scalability issues, they do represent the only means of passing IP multicast traffic through a VPN.

MPLS

Multiprotocol Label Switching (MPLS) was originally derived from Tag Switching, and various other vendor methods of IP-switching support enhancements in the scalability and performance of IP-routed networks by combining the intelligence of routing with the high performance of switching.

MPLS is now used for VPNs, which is an appropriate combination because MPLS decouples information used for forwarding of the IP packet (the label) from the information carried in the IP header.

MPLS VPNs can combine any of the following:

- Globally unique and routable addresses
- Globally unique, non-routable addresses
- Private addresses (RFC1918)
- Addresses that are neither globally unique nor private.

Label Switched Paths are bound to VPN-IP routes and are confined to the VPN Service Provider.

Multicast Overview

Multicast vs. Unicast

IP Multicast is part of the TCP/IP suite of protocols. While IP Unicast uses Class A, B, and C address, IP Multicast uses Class D addresses.

Multicast is an efficient paradigm for transmitting the same data to multiple receivers, because of its concert of a Group address. This allows a group of receivers to listen to the single address.

IP Multicast packets are replicated by routers within the network when there is more than one sub-network requiring a copy of the data. IP Unicast makes the source responsible for creating an individual IP stream for each receiver.

Multicast is a robust and scalable solution for group communication because of this distributed replication of data and because only 1 copy of the packet needs to traverse a link

For example, suppose a company president sends a presentation to all employees.

IP Multicast: bandwidth for one viewer equates bandwidth for all viewers

IP Unicast: it would be impossible, due to the cost of network infrastructure, if each receiver were to have a unique data stream

Applications areas that take advantage of IP Multicast include, but are not limited to, corporate communications, distance learning, and the distribution of software.

Following are some basic IP Multicast concepts (not an exhaustive explanation of IP multicast technology). Please refer to the appendix of this document for details

Trees

Only one packet is transmitted by the source, so the Network is responsible for replicating the packet at each bifurcation point in the network. This results in a multicast distribution tree. There are two main types of trees:

Source Trees/Shortest Path Trees

Uni-directional trees rooted at the data source. They are the only types of tree in a dense mode flood and prune network, but are also used in Sparse Mode networks.

Shared Trees

Uni-directional trees rooted at a common point. Shared trees are used in protocols where receivers are sparsely populated and it would be inefficient to learn about active sources by flooding the data. In Protocol Independent Multicast Sparse Mode (PIM SM), the common root of the shared tree is called the Rendezvous Point. This is the point where receivers join to learn of active sources.

Multicast Routing protocols

Historic

Several routing protocols were designed to work with IP Multicast. These were a “ships in the night” approach, which required a separate Routing table for IP Multicast traffic.

- Distance Vector Multicast Routing Protocol (DVMRP)

  DVMRP was the first Multicast routing Protocol, and is an example of a source tree routing protocol.

- Multicast Open Shortest Path First (MOSPF)

  MOSPF attempted to use OSPF with multicast routing. It is also an example of a Source tree routing protocol.

- Core Base Trees (CBT)

  CBTs were designed to use a shared tree to deliver multicast data, but they were never implemented beyond the experimental networks.

Defacto Standard

Protocol Independent Multicast (PIM)

PIM does not use a “ships in the night” approach; rather, it is designed to forward IP Multicast traffic using the standard Unicast routing table. PIM uses the Unicast routing table to decide if the source of the IP multicast packet has arrived on the optimal path from the source. This process, the RPF check, is Protocol Independent because it is based on the contents of the Unicast Routing table and not any particular routing protocol. There are two types of PIM protocols: Dense Mode (DM) and Sparse Mode (SM).

PIM Dense Mode (DM)

PIM DM is no longer a widely deployed protocol because PIM SM has proven to be the more efficient multicast

PIM SM

PIM sparse mode has been enhanced over the years, evolving from an experimental standard to a draft standard. It is now the most widely deployed multicast protocol. It initially uses a shared tree, but then allows the last hop router to join a Source tree if it so chooses. This is an efficient methodology, as it prevents the flooding of data and associated waste of resources, while forwarding data along the optimal path.

The growing deployment of IP Multicast by applications has necessitated two new distribution paradigms.

Bi-directional PIM (PIM Bi-Dir)

PIM Bi-Dir will create a two-way forwarding tree, which. will allow the efficient transmission of low bandwidth many-to-many communication; for example, a financial trading application.

Source Specific Multicast (SSM)

SSM is a solution where the knowledge of the source is acquired out of band. SSM uses only a source tree, but there is no flooding of data, because learning the source is out of band. SSM is most useful for applications such as Internet broadcasting or corporate communications

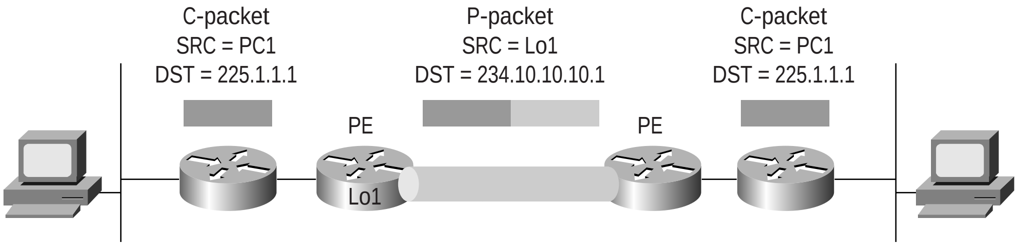
Three Proposed MVPN Solutions

Draft-rosen-vpn-mcast-03.txt extends the IETF VPN specification by describing the protocols and procedures required to support an IP Multicast VPN.

Multicast Domains

This solution requires the provider to enable IP Multicast within its network. On each Provider Edge (PE) router, the provider creates a Multicast Tunnel Interface (MTI) and Multicast VPN routing / forwarding (VRF) for each customer. The MTI encapsulates customers’ Multicast data within its own Multicast packet with a destination group that is unique for a particular customer and to which all PE for that customer belong.

Figure 1
MTI Encapsulation



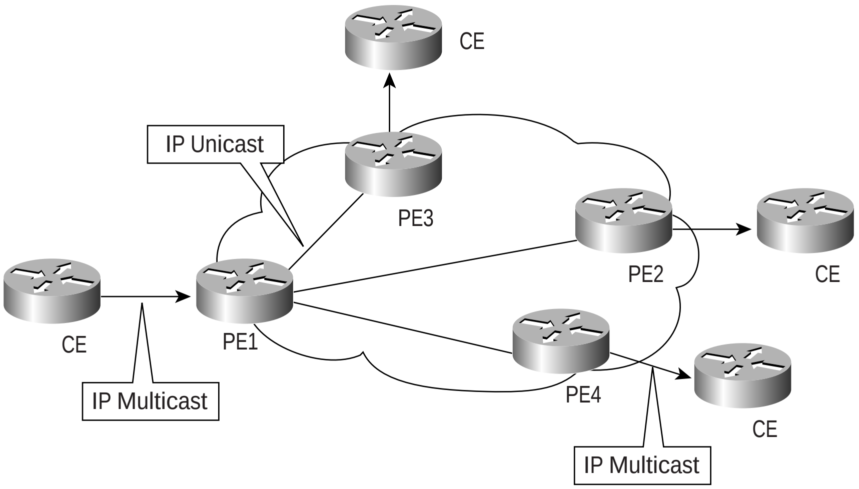
VPN-IP PIM

VPN-IP PM requires that customers' multicast data be delivered natively through the provider's network. Each customer VRF is leaked into a global table. This solution requires that the source of the Multicast tree is noted and the VRF of that tree is used to create a route distinguisher. This becomes "RD:S,G" or "RD: *,G", which makes the mroute entry unique for that customer. While this RD does allow the provider to uniquely identify the multicast data, the packet must be encapsulated with MPLS multicast labels in order to forward the data.

Multicast Domain (MD) Using PIM Non-Broadcast Multi-Access (NBMA) Techniques

This solution uses a tunnel interface on the PE. Unlike GRE tunneling, this is no a point-to-point tunnel. This tunnel interface tracks the remote PE and Unicasts the multicast packets to the remote PEs. This solution is initially attractive, because it keeps multicast state out of the core; however, it requires a large amount of replication by the PE router and creates a great deal of additional to Unicast traffic.

Figure 2
Unicast Forwarding of Multicast using NBMA Technique



Cisco MVPN Details

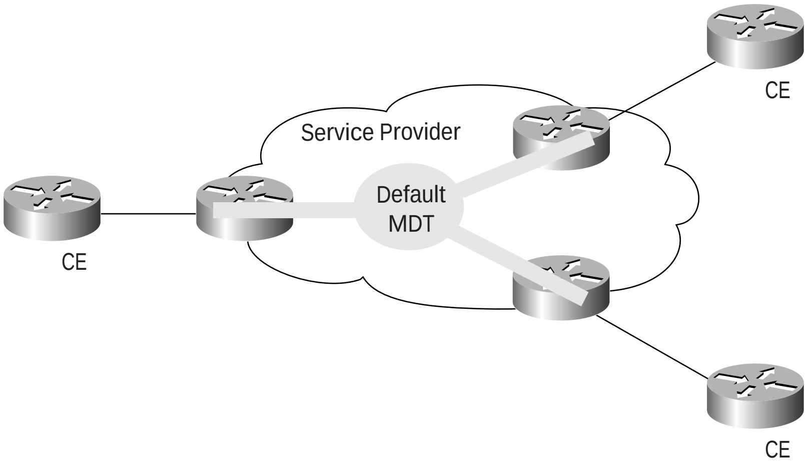
While there are significant deployment obstacles to each of the preceding MVPN solutions, Multicast Domains is the most attractive alternative because:

- The provider must configure a native IP multicast network within their core network; this includes both the P and PE routers.
- IP Multicast is a mature technology that has been deployed since Cisco IOS Software 10.0. This minimizes risk for the provider network, because a new feature will not have to be introduced into its core to support MVPNs.

Multicast Domain Solution

This method originally had less than optimal performance, because it requires that all PE routers connected to a customer receive all of that customer's Multicast data regardless of the presence of an interested receiver in that location. When enhancements resolved this characteristic with a new methodology, it became a truly attractive solution.

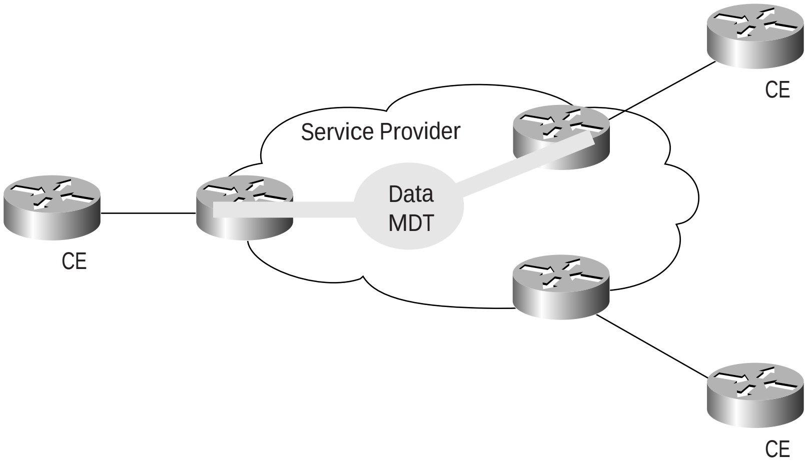
Figure 3
Default MDT Concept



The aforementioned enhancement is the addition of ephemeral trees that are created 'on the fly'. These trees distribute multicast group data that exceeds a certain configured threshold of Bandwidth (BW) to only those PE who have joined this new tree. These are tress are called MDT-data trees. The word data is appended as these groups are designed to be used for groups that will require a higher amount of bandwidth to deliver their data.

Figure 4

Data MDT Concept



This diagram indicates that the Data MDT is only joined by those PE routers that have CE routers who are intended recipients of the data.

PE routers signal use of Data-MDT via a UDP packet on port number 3232, which is sent via the default MDT. This packet contains an all-PIM routers message, indicating the group to be joined if required.

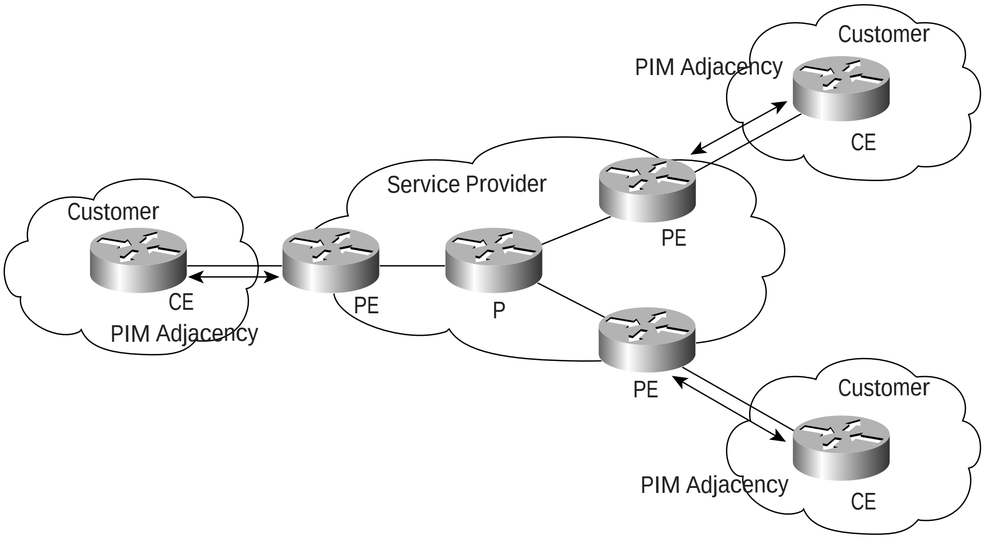
Interaction of Customer and Providers Multicast Network

It is important to remember that the customer's IP Multicast network has no relationship to the provider's multicast network. From the perspective of the provider, the customer's IP Multicast packets are merely data to the provider's distinctive IP Multicast network.

It is important to understand that PIM, and in particular PIM-SM, are the only supported multicast protocols for MVPN. Bi-Dir PIM may be supported in the future, when it is deemed stable enough for the core of a provider network.

Figure 5

Customer PIM Adjacencies



CE routers do not have a PIM adjacency across the provider network with remote CE routers, but rather have an adjacency with their local routers and the PE router.

When the PE router receives an MDT packet, it performs an RPF check. During the transmission of the packet through the Provider network, the normal RPF rules apply. However, at the remote's PE, the router needs to ensure that the originating PE router was the correct one for that CE. It does this by checking the BGP next hop address of the customer's packet's source address. This next hop address should be the source address of the MDT packet. The PE also checks that there is a PIM neighbourship with the remote PE.

Currently, only a single MVRF is supported per customer. This limitation precludes the customer also receiving Internet or any other outside domain's Multicast traffic

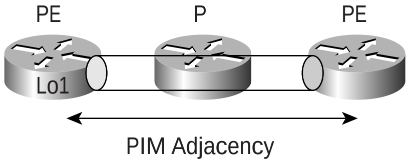
A unique Group address is required to be used as MDT for each particular customer. A unique source address for the Multicast packet in the provider network is also required. This source address is recommended to be the address of the loopback interface, which is used as the source for the IBGP, as this address is used for the RPF check at remote PE.

If the provider uses MDT-data groups, then these will also need to be configured. These MDT-data groups must be unique for each customer.

The PE routers must have a PIM adjacency to each other. No other routing protocols may use these MTIs.

Figure 6

Provider's PIM Adjacencies



BGP Requirements

PE routers are the only routers that need to be MVPN aware and able to signal to remote PE's information regarding the MVPN. It is therefore fundamental that all PE routers have a BGP relationship with each other. Either directly or via a Route Reflector.

The source address of the Default-MDT will be the same address used to source the iBGP sessions with the remote PE routers that belong to the same VPN and MVRF. When PIM-SSM is used for transport inside the provider core, it is via this BGP relationship that the PEs indicate that they are MVPN capable and provide for source discovery. This capability is indicated via the updated BGP message.

When a PE receives a BGP update, which includes the RD and the group information, it joins the root of that tree, thereby joining the MDT.

The RPF check on the PE is satisfied when the following conditions are met:

1. The next hop for the source of the CE data is the BGP neighbor, which is the source of the MDT
2. The Source of the MDT is a PIM neighbor

```
(M)VPN-IPv4 address (12 bytes)
 Route Distinguisher - 8 bytes
type-field: 2 bytes
value-field: 6 bytes
New type for Multicast-VPN: 2
Its value field (AS format must be used):
2 bytes ASN
4 bytes assigned number
IPv4 address - 4 bytes

Extended community attribute - 8 bytes

Type Field: 2 bytes
Value Field: 6 bytes
New type: 0x06 (AS format)
Its Value Field:
2 bytes ASN
4 bytes assigned number  (MDT Group address)
```

OR (currently not supported)

```
New type:  0x0106 (Address format)
Its Value Field:
4 bytes IPv4 address (MDT Group address)
                        2 bytes assigned number
```

Terminology

Virtual Private Network	VPN
Multicast VPN	MVPN
VPN Routing and Forwarding	VRF
Multicast VRF	MVRF
Provider Network Interface	PNI
Customer Network Interface	CNI
Multicast Tunnel Interface	MTI
Provider Edge	PE
Customer Edge	CE
Multicast Distribution Tree	MDT
Tree created statically to connect ALL PE routers	Default MDT
Tree created dynamically to only interested PE routers	Data MDT
Protocol Independent Multicast	PIM
PIM-SM Sparse Mode	PIM SM
PI-DM Dense Mode	PIM DM
Bi-directional PIM	PIM Bi-Dir
Reverse Path Forwarding	RPF
Rendezvous Point	RP

References

http://search.ietf.org/internet-drafts/draft-rosen-vpn-mcast-03.txt

http://www.cisco.com/go/ipmulticast

http://www.ietf.org/html.charters/idmr-charter.html

Developing IP Multicast Networks, Volume I
Beau Williamson, CCIE
ISBN:1578700779

CISCO SYSTEMS

Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the

Cisco Web site at www.cisco.com/go/offices

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia
Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland
Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland
Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe