Requirements for Multicast in L3 Provider-Provisioned VPNs
draft-ietf-l3vpn-ppvnp-mcast-reqts-01

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on January 14, 2006.

Copyright Notice

Copyright (C) The Internet Society (2005).

Abstract

This document presents a set of functional requirements for network solutions that allow the deployment of IP multicast within L3 Provider Provisioned virtual private networks (PPVPNs). It specifies requirements both from the end user and service provider standpoints. It is intended that potential solutions specifying the support of IP multicast within such VPNs will use these requirements as guidelines.

Working group
This document is a product of the IETF’s Layer 3 Virtual Private Network (l3vpn) working group. Comments should be addressed to WG’s mailing list at <mailto:l3vpn@ietf.org>. The charter for l3vpn may be found at <http://www.ietf.org/html.charters/l3vpn-charter.html>

Contributors

Main contributors to this document are listed below, in alphabetical order:

- Christian Jacquenet
  France Telecom
  3, avenue Francois Chateau
  CS 36901 35069 RENNES Cedex
  France
  Email: christian.jacquenet@francetelecom.com

- Yuji Kamite
  NTT Communications Corporation
  Tokyo Opera City Tower 3-20-2 Nishi Shinjuku, Shinjuku-ku
  Tokyo 163-1421, Japan
  Email: y.kamite@ntt.com [3]

- Jean-Louis Le Roux
  France Telecom R & D
  2, avenue Pierre-Marzin
  22307 Lannion Cedex
  France
Email: jeanlouis.leroux@francetelecom.com [4]

Nicolai Leymann
T-Systems International GmbH
Engineering Networks, Products & Services
Goslarer Ufer 35
10589 Berlin, Germany
Email: nicolai.leymann@t-systems.com [5]

Renaud Moignard
France Telecom R & D
2, avenue Pierre-Marzin
22307 Lannion Cedex
France
Email: renaud.moignard@francetelecom.com [6]

Thomas Morin
France Telecom R & D
2, avenue Pierre-Marzin
22307 Lannion Cedex
France
Email: thomas.morin@francetelecom.com [7]
Table of Contents

1. Introduction ........................................... 6
2. Conventions used in this document .................... 7
   2.1 Terminology ........................................ 7
   2.2 Conventions ......................................... 8
3. Problem Statement ....................................... 9
   3.1 Motivations ......................................... 9
   3.2 General Requirements ................................. 9
   3.3 Scaling vs. Optimizing Resource Utilization ......... 9
4. Use cases .............................................. 11
   4.1 Scenarios ............................................ 11
      4.1.1 Real-time / Unspecified receivers / Large bandwidth 11
      4.1.2 Real-time / Unspecified receivers / Medium bandwidth 11
      4.1.3 Real-time / Unspecified receivers / Small bandwidth 11
      4.1.4 Real-time / Specified receivers .................. 12
      4.1.5 Non-real-time ................................... 12
      4.1.6 Content broadcasting / Unspecified Receivers / Large Bandwidth 12
      4.1.7 Symmetric Low Volume Traffic .................... 12
      4.1.8 Mixed generic multicast VPN ...................... 13
   4.2 Scalability orders of magnitude ..................... 13
5. Requirements for supporting IP multicast within L3 PPVPNs . 15
   5.1 End user/customer standpoint ......................... 15
      5.1.1 Service definition ............................... 15
      5.1.2 CE-PE Multicast routing and management protocols ... 15
      5.1.3 Quality of Service (QoS) ........................ 15
      5.1.4 SLA parameters measurement ....................... 16
      5.1.5 Security Requirements ............................ 17
      5.1.6 Monitoring and Troubleshooting .................... 18
      5.1.7 Extranet ......................................... 18
      5.1.8 Internet Multicast ............................... 19
      5.1.9 Carrier’s carrier ................................. 19
      5.1.10 Multi-homing, load balancing and resiliency .... 20
      5.1.11 RP Engineering .................................. 20
      5.1.12 Addressing ...................................... 20
      5.1.13 Minimum MTU ..................................... 21
   5.2 Service provider standpoint ......................... 21
      5.2.1 Scalability ...................................... 21
      5.2.2 Resource optimization ............................. 23
      5.2.3 Tunneling Requirements ............................ 24
      5.2.4 Control mechanisms ................................ 25
      5.2.5 Quality of Service Differentiation ............... 25
      5.2.6 Infrastructure security ........................... 26
      5.2.7 Robustness ...................................... 26
      5.2.8 Management tools, OAM ............................ 27
      5.2.9 Architectural Considerations ...................... 27
      5.2.10 Compatibility and migration issues ............... 27
1. Introduction

VPN services satisfying requirement defined in [RFC4031] are now being offered by many service providers throughout the world. VPN services are popular because customers need not be aware of VPN technologies deployed in the provider network. They scale well for the following reasons:

- because P-routers need not be aware of VPN service details
- because the addition of a new VPN member requires only limited configuration effort

There is also a growing need for support of IP multicast-based services. Efforts to provide efficient IP multicast routing protocols and multicast group management have been done in standardization bodies which has led, in particular, to the definition of the PIM and IGMP protocols.

However, multicast traffic is not natively supported within existing L3 PPVPN solutions. Deploying multicast over an L3VPN today, with only currently standardized solutions, requires designing customized solutions which will be inherently limited in terms of scalability, operational efficiency and bandwidth usage.

This document complements the generic L3 VPN requirements [RFC4031] document, by specifying additional requirements specific to the deployment of IP multicast-based services within PPVPNs. It clarifies the needs from both VPN client and provider standpoints and formulates the problems that should be addressed by technical solutions with as a key objective to stay solution agnostic. There is no intent to either specify solution-specific details in this document or application-specific requirements. Also this document does NOT aim at expressing multicast-inferred requirements that are not specific to L3 PPVPNs.

It is expected that solutions that specify procedures and protocol extensions for multicast in L3 PPVPNs SHOULD satisfy these requirements.
2. Conventions used in this document

2.1 Terminology

Although the reader is assumed to be familiar with the terminology defined in [RFC4031], [RFC2547] and RFC2547bis [I-D.ietf-l3vpn-rfc2547bis], PIM-SM [RFC2362], PIM-SSM [I-D.ietf-ssm-arch] the following glossary of terms may be worthwhile.

Moreover we also propose here generic terms for concept that naturally appears when multicast in VPNs is discussed.

ASM: Any Source Multicast. One of the two multicast service models that denotes the source/receiver heuristic.

Multicast-enabled VPN, or multicast VPN: a VPN which supports IP multicast capabilities, i.e. for which some PE devices (if not all) are multicast-enabled and whose core architecture support multicast VPN routing and forwarding

PPVPN: Provider-Provisioned Virtual Private Network

PE/CE: Provider/Customer edge Equipment ([RFC4026])

VRF or VR: By this phrase, we refer to the entity defined in a PE dedicated to a specific VPN instance. "VRF" refers to [RFC2547] terminology, and "VR" to the VR [I-D.ietf-l3vpn-vpn-vr] terminology.

MD Tunnel: Multicast Distribution Tunnel, the means by which the customer’s multicast traffic will be conveyed across the SP network. This is meant in a generic way: such tunnels can be either point-to-point or point-to-multipoint. Although this definition may seems to assume that distribution tunnels are unidirectional, but the wording encompasses bi-directional tunnels as well.

G: Denotes a multicast group

Multicast channel: (S,G) in the SSM model

Participating device: Refers to any network device that not only participates to the deployment and the maintenance of the VPN infrastructure, but also to the establishment and the maintenance of the MD Tunnel (see above).
S: Denotes a multicast source

SP: Service provider

SSM: Source Specific Multicast. One of the two multicast service models where each corresponding service relies upon the use of a single source.

RP: Rendez-vous point (PIM-SM [RFC2362])

Please refer to [RFC4026] for details about terminology specifically relevant to VPN aspects, and to [RFC2432] for multicast performance or QoS related terms.

2.2 Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].
3. Problem Statement

3.1 Motivations

More and more L3 VPN customers use IP multicast services within their private infrastructures. Naturally, they want to extend these multicast services to remote sites that are connected via a VPN.

For instance, the customer could be a national TV channel with several geographical locations that wants to broadcast a TV program from a central point to several regional locations within its VPN.

A solution to support multicast traffic would consist in using point-to-point tunnels across the provider network and requiring the PE routers (provider’s routers) to replicate traffic. This is obviously sub-optimal as it places the replication burden on the PE and hence has very poor scaling characteristics. It may also waste bandwidth and control plane resources in the provider’s network.

Thus, to provide multicast services for L3 VPN networks in an efficient manner (that is, with scalable impact on signaling and protocol state as well as bandwidth usage), in a large scale environment, new mechanisms are required to enhance existing L3 VPN solutions for proper support of multicast-based services.

3.2 General Requirements

This document sets out requirements for L3 provider-provisioned VPN solutions designed to carry customers’ multicast traffic. The main requirement is that a solution SHOULD first satisfy requirements documented in [RFC4031]: as far as possible, a multicast service should have the same flavor as the unicast equivalent, including the same simplicity (technology unaware), the same quality of service (if any), the same management (e.g. monitoring of performances), etc.

Moreover, it also has to be clear that a multicast VPN solution MUST interoperate seamlessly with current unicast solutions. It would also make sense that multicast VPN solutions define themselves as extensions to existing L3 provider-provisioned VPN solutions (such as for instance, RFC2547bis [I-D.ietf-l3vpn-rfc2547bis] or VR [I-D.ietf-l3vpn-vpn-vr]) and retain consistency with those, although this is not a core requirement.

3.3 Scaling vs. Optimizing Resource Utilization

When transporting multicast VPN traffic over a service provider network, there intrinsically is tension between scalability and resource optimization, since the latter likely requires maintaining
multicast state in the core network.

Consequently, some trade-off has to be made and this document will express some requirements related to this trade-off.
4. Use cases

The goal of this section is to highlight how different applications and network contexts may have a different impact on how a trade-off is made. We aim here at presenting a few representative examples of multicast VPN deployments, and to express expectations about orders of magnitude of relevant scalability parameters.

4.1 Scenarios

4.1.1 Real-time / Unspecified receivers / Large bandwidth

Broadcasting companies, which want to send their programs in real-time, would need large bandwidth and reachability to many unspecified nodes on VPN. This does expect not only bandwidth guarantee, low delay and low jitter but also rapid following capability of multicast membership changes.

The SP has to take care of the scalability impact about both bandwidth efficiency and number of receivers.

This case is regarded as one-to-many streams.

4.1.2 Real-time / Unspecified receivers / Medium bandwidth

Enterprise customers expect to build video conference environment on their existing VPNs. Because you do not always know which receivers will join each conference, customer’s multicast information might be dynamically added, removed, or changed.

This fact will require that SP should control whether and how MDTunnel topologies are dynamically changed, and their bandwidth usage efficiency in SP core.

Conference applications are often regarded as many-to-many streams.

4.1.3 Real-time / Unspecified receivers / Small bandwidth

Enterprise customers, however, do not always require large bandwidth. For instance, applications like distributing stock market information, will strongly need good real-time QoS, but it does not require huge bandwidth.

This case is not burdened with MDTunnel bandwidth efficiency issues, but it is still necessary to provide low delay, jitter, and high resiliency.
4.1.4 Real-time / Specified receivers

Some customers may want to deploy a multicast VPN where the location of receivers is well known in advance. One example would be a case of real-time programs to fixed locations, such as horse race live coverage to off-course betting shops.

This kind of application is characterized by its static receivers, and by the fact their memberships are not modified so often. In this case MDTunnels might not need rapid changes of their topology, and can be built in a relatively static way.

4.1.5 Non-real-time

Content distributors might want to use a multicast VPN to more efficiently deliver their contents. For example, when a central TV station delivers its recorded contents to lots of local stations, it uses one-to-many streams for transmitting files.

If local stations begin exchanging their contents each other, the situation can be regarded as many-to-many streams. Such file transfer scenario might need large bandwidth but does not require real-time following capability of multicast membership.

Also it does not always need strict delay and guaranteed jitter.

4.1.6 Content broadcasting / Unspecified Receivers / Large Bandwidth

A L3VPN architecture can be leveraged for video (or any content) broadcast distribution to broadband customers.

In such a scenario, the location of receivers of a channel will be unknown, but a high level of aggregation could be expected (high-audience channels are likely to be requested by same sets of PEs). The number of channels will likely be high (hundreds), and the typical bandwidth would be the one of video codecs (somewhere between 1 and 15 Mbps as of today).

In this scenario, the multicast group join delay ([RFC2432], section 3.4) will need to be very low.

In this scenario, contrary to the Section 4.1.1 scenario, the delay and jitter do not need to be very low.

4.1.7 Symmetric Low Volume Traffic

In this scenario, IP-Multicast is used to send heart beats and state information to a number of receivers being member of a single
multicast group. Typical use case are management application monitoring and managing a number of distributed clients. All nodes are senders and receivers at the same time building a many-to-many relationship. If a node is not visible by sending packets to the multicast group it is considered as being down. When IP-Multicast fails in general this is valid for all nodes and the management application fails to work even if unicast connectivity is working (no fallback to unicast available).

Specifics of the scenario:

- One group
- Hundreds of receivers/senders
- Tens of PE devices
- Traffic volume: hundreds of Kb/s, with peaks at a few Mb/s

### 4.1.8 Mixed generic multicast VPN

This is a general deployment scenario where IP-Multicast is used in every VPN: if a customer requests a VPN, then this VPN will support IP-Multicast per default. In this case the number of mVPN equals the number of VPNs in the platform. This implies a quite important scalability requirement (e.g. hundreds of PEs, hundreds of VPNs per PE, with a potential grow by one order of magnitude in the longer term).

The per mVPN traffic behaviour is not predictable because it’s completely up to the customer how the service is used. This results in a traffic mix of the scenarios mentioned in section Scenarios. QoS requirements are similar to typical unicast scenarios, with the need for different classes. Also in a such context, a reasonably large range of protocols should be made available to the customer for use at the PE-CE level.

Also, in such a scenario, customers may want to deploy multicast connectivity between two or more mVPNs as well as access to internet Multicast.

### 4.2 Scalability orders of magnitude

This section proposes orders of magnitude for different scalability metrics relevant for multicast VPN issues. It should be noted that the scalability figures proposed here relate to scalability expectations of future deployments of multicast VPN solutions, as the
The author tried to no restrict the scope to the mere deployments known as of today.

The figures proposed here are the result of an informal survey proposed to ISP in summer 2005.

[ This section will be completed with the result of the "Multicast VPN Survey" posted to the L3VPN WG in July'05 ]
5. Requirements for supporting IP multicast within L3 PPVPNs

Again, the aim of this document is not to specify solutions but to give requirements for supporting IP multicast within L3 PPVPNs.

In order to list these requirements we have taken two different standpoints of two different important entities: the end user (the customer using the VPN) and the service provider.

In the rest of the document, we mean by a "solution", a solution that allows to perform multicast in an L3 provider provisioned VPN, which addresses the requirements listed in this document.

5.1 End user/customer standpoint

5.1.1 Service definition

As for unicast, the multicast service MUST be provider provisioned and SHALL NOT require the customer’s devices (CE) to support any extra feature compared to those required for multicast in a non-VPN context.

5.1.2 CE-PE Multicast routing and management protocols

Consequently to Section 3.1, the CEs and PEs SHOULD employ existing multicast protocols.

Such protocols SHOULD include : PIM-SM [RFC2362] (including PIM-SSM [I-D.ietf-ssm-arch]), bidirectional PIM [I-D.ietf-pim-bidir], PIM-DM [RFC3973], and IGMP (v1 [RFC1112], v2 [RFC2236] and v3 [RFC3376]).

Among those protocols, PIM-SM is considered a MUST.

When IPv6 is supported by a VPN solution, relevant IPv6 corresponding protocols SHOULD also be supported, e.g. Multicast Listener Discovery Protocol (MLD) (v1 [RFC2710], v2 [RFC3810]).

5.1.3 Quality of Service (QoS)

First, general considerations about QoS in L3 VPNs as developed in section 5.5 of [RFC4031] are also relevant to this section.

QoS is measured in terms of delay, jitter, packet loss, and availability. These metrics are already defined for the current unicast PPVPN services, and are included in Service Level Agreements(SLA). In some cases, provided SLA may be different between unicast and multicast, which will need service differentiation mechanisms as such.
The level of availability for the multicast service SHOULD be on par with what exists for unicast traffic. For instance same traffic protection mechanisms SHOULD be available for customer multicast traffic when it is carried over the service provider’s network.

A multicast in VPN solution shall allow to define at least the same level of quality of service than what exists for unicast, and than what exist for multicast in a non-VPN context. From this perspective, the deployment of multicast-based services within an L3 PPVPN environment SHALL benefit from DiffServ [RFC2475] mechanisms that include multicast traffic identification, classification and marking capabilities, as well as multicast traffic policing, scheduling and conditioning capabilities. Such capabilities MUST therefore be supported by any participating device in the establishment and the maintenance of the multicast distribution tunnel within the VPN.

As multicast is often used to deliver high quality services such as TV broadcast, the solution should have additional features to support high QoS such as bandwidth reservation and admission control.

Also, considering that multicast reception is receiver-triggered, group join delay (as defined in [RFC2432]) is also considered one important QoS parameter. It is thus RECOMMENDED that a multicast VPN solution be designed appropriately in this regard.

The group leave delay (as defined in [RFC2432]) may also be important on the CE-PE link for some usage scenarios: in cases where the typical bandwidth of multicast streams is close to the bandwidth a PE-CE link, it will be important to have the ability to stop the emission of a stream on the PE-CE link as soon as it stops being requested by the CE, to allow for fast switching between two different high throughput multicast streams. This implies that it SHOULD be possible to tune the multicast routing or group protocol (e.g. IGMP/MLD or PIM) used on the PE-CE adjacency to reduce the group leave delay to the minimum.

Last, a multicast VPN solution SHOULD as much as possible ensure that client multicast traffic packets are neither lost nor duplicated, even when changes occur in the way a client multicast data stream is carried over the provider network. Packet loss issues have also to be considered when a new source starts to send traffic to a group: any receiver interested in receiving such traffic SHOULD be serviced accordingly.

5.1.4 SLA parameters measurement

As SLA parameters are part of the service that is sold, they are
often monitored. The monitoring is used for technical reasons by the service provider and is often sold to the customer for end-to-end service purposes.

The solution MUST support (SLA) monitoring capabilities, which MAY possibly rely upon similar techniques (than those used by the unicast for the same monitoring purposes).

Multicast specific characteristics that may be monitored are, for instance, multicast statistics per stream, end-to-end delay and group join delay (time to start receiving a multicast group traffic across the VPN, as defined in [RFC2432] section 3).

A generic discussion of SLAs is provided in [RFC3809].

5.1.5 Security Requirements

Security is a key point for a customer who uses subscribes to a VPN service. The RFC2547bis [I-D.ietf-l3vpn-rfc2547bis] model offers some guarantees concerning the security level of data transmission within the VPN.

A multicast VPN solution MUST provide an architecture that can provide the same level of security both for both the unicast and multicast traffics.

Moreover, the activation of multicast features SHOULD be possible:

- with a VRF or VR granularity
- with a CE granularity (when multiple CE of a same VPN are connected to a common VRF)
- with a distinction between multicast reception and emission
- with a multicast group and/or channel granularity

A multicast VPN solution may choose to make the optimality/scalability trade-off stated in Section 3.3 by sometimes distributing multicast traffic of a client group to a larger set of PE routers that may include PEs which are not part of the VPN. From a security standpoint, this may be a problem for some VPN customers, thus a multicast VPN solution using such a scheme MAY offer ways to avoid this for specific customers (and/or specific customer multicast streams).
5.1.6 Monitoring and Troubleshooting

A service provider and its customers MUST be able to manage the capabilities and characteristics of their multicast VPN services. Automated operations and interoperability with standard management platforms SHOULD be supported.

Service management should also include the TMN ‘FCAPS’ functionalities, as follows: Fault, Configuration, Accounting, Provisioning, and Security.

The monitoring of multicast specific parameters and statistics SHOULD include:

- multicast traffic statistics: total traffic conveyed, incoming, outgoing, dropped, etc., by period of time (as a MUST)
- IP Performance Metrics related information (IPPM, [RFC2330]) that is relevant to the multicast traffic usage: such information includes the one-way packet delay, the inter-packet delay variation, etc. (as a MAY)

Apart from statistics on multicast traffic, customers of a multicast VPN will need information concerning the status of their multicast resource usage (state and bandwidth). Indeed, as mentioned in Section 5.2.4, for scalability purposes, a service provider may limit the number (and/or throughput) of multicast streams that are received and produced at a client site, and so a multicast VPN solution SHOULD allow customers to find out their current resource usage (state and throughput), and to receive some kind of feedback if their usage exceed bounds. Whether this issue will be better handled at the protocol level at the PE-CE interface or via the ISP customer support, needs further discussion.

5.1.7 Extranet

In current PP L3VPN models, a customer site may be setup to be part of multiple VPNs and this should still be possible when a VPN is multicast-enabled.

A multicast solution SHOULD offer means so that:

- receivers behind attached CEs can receive multicast traffic sourced in any of the VPNs (if security policy permits)
- sources behind attached CEs can reach multicast traffic receivers located in any of the VPNs
multicast can be independently enabled for the different VPNs (and multicast reception and emission can also be independently enabled)

Proper support for this feature SHOULD not require replicating multicast traffic on a PE-CE link, whether it is a physical or logical link.

For instance, an enterprise using a multicast-enabled VPN should have the ability to receive a multicast stream from, or originate a multicast stream towards, another VPN.

In any case a solution not supporting such a feature MUST be compatible with setups where a VRF or VR is part of multiple VPNs and MUST document how it operates on multicast traffic in such a context.

5.1.8 Internet Multicast

Connectivity with Internet Multicast (as a source or receiver) somehow fits in the context of the previous section.

It should be considered OPTIONAL given additional considerations needed to fulfill requirements for Internet side, such as security treatment.

5.1.9 Carrier’s carrier

Many L3 PPVPN solutions, such as RFC2547bis [I-D.ietf-l3vpn-rfc2547bis] and VR [I-D.ietf-l3vpn-vpn-vr] define the "Carrier’s Carrier" model, where a "carrier’s carrier" service provider supports one or more customer ISP, or "sub-carriers". A multicast VPN solution SHOULD support the carrier’s carrier model in a scalable and efficient manner.

Ideally the range of tunneling protocols available for the sub-carrier ISP should be the same as those available for the carrier’s carrier ISP. This implies that the protocols that may be used at the PE-CE level SHOULD NOT be restricted to protocols required as per Section 5.1.2 and SHOULD include some of the protocols listed in Section 5.2.3.

In the context of MPLS-based L3VPN deployments, such as BGP/MPLS VPNs [I-D.ietf-l3vpn-rfc2547bis], this means that MPLS label distribution SHOULD happen at the PE-CE level, giving the ability to the sub-carrier to use multipoint LSPs as a tunneling mechanism.
5.1.10 Multi-homing, load balancing and resiliency

A multicast VPN solution should be compatible with current solutions that aim at improving the service robustness for customers such as multi-homing, CE-PE link load balancing and failover. A multicast VPN solution SHOULD also be able to offer those same features for multicast traffic. Any solution SHOULD support redundant topology of CE-PE links. It SHOULD minimize multicast traffic disruption and failover.

On the other hand, it is also necessary to care about failover mechanisms that are unique to multicast routing control. For instance, if the customer uses some control mechanism for RP redundancy on PIM-SM (e.g. BSR), it SHOULD work transparently through that VPN.

5.1.11 RP Engineering

When PIM-SM (or bidir-PIM) is used in ASM mode on the VPN customer side, the location of the RP has to be chosen. In some cases this engineering problem is not trivial: for instance, if sources and receivers are located in VPN sites that are different than that of the RP, then traffic may flow twice through the SP network and the CE-PE link of the RP (from source to RP, and then from RP to receivers); this is obviously not ideal. A multicast VPN solution SHOULD propose a way to help on solving this RP engineering issue.

Moreover, some service providers offer to manage customer’s multicast protocol operation on behalf of them. This implies that it is needed to consider cases where the customer’s RPs are outsourced (e.g., on PEs).

5.1.12 Addressing

A multicast provider-provisioned L3VPN SHOULD NOT impose restrictions on multicast group addresses used by VPN customers.

In particular, like unicast traffic, an overlap of multicast group address sets used by different VPN customers MUST be supported.

The use of globally unique means of multicast-based service identification at the scale of the domain where such services are provided SHOULD be recommended. If the ASM model is used, this implies the use of the multicast administratively scoped range, (239/8 as per [RFC2365]) for services which are to be used only inside the VPN, and of globally assigned group addresses for services for which traffic may be transmitted outside the VPN (e.g. GLOP [RFC3180]).
5.1.13 Minimum MTU

For customers, it is often a serious issue whether transmitted packets will be fragmented or not. In particular, some multicast applications might have different requirements than those that make use of unicast, and they may expect services that guarantee available packet length not to be fragmented.

Therefore, a multicast VPN solution SHOULD let customers’ devices be free of any fragmentation or reassembly activity.

A committed minimum path MTU size SHOULD be provided to customers. Moreover, since Ethernet LAN segments are often located at first and last hops, a minimum 1500 bytes IP MTU SHOULD be provided.

It SHOULD also be compatible with Path MTU discovery mechanisms, such as those defined in [RFC1191] or [I-D.mathis-frag-harmful].

5.2 Service provider standpoint

Note: please remember that, to avoid repetition and confusion with terms used in solution specifications, we introduced in Section 2.1 the term MDTunnel (for Multicast Distribution Tunnel), which designates the data plane means used by the service provider to forward customer multicast traffic over the core network.

5.2.1 Scalability

Some currently standardized and deployed L3VPN solutions have the major advantage of being scalable in the core regarding the number of customers and the number of customer routes. For instance, in the RFC2547bis [I-D.ietf-l3vpn-rfc2547bis] and VR [I-D.ietf-l3vpn-vpn-vr] models, a P-router sees a number of MPLS tunnels that is only linked to the number of PEs and not to the number of VPNs, or customers’ sites.

As far as possible, this independence in the core, with respect to the number of customers and to customer activity, is recommended. Yet, it is recognized that in our context scalability and resource usage optimality are competing goals, so this requirement may be reduced to giving the possibility of bounding the quantity of states that the service provider needs to maintain in the core for MDTunnels, with a bound being independent of the multicast activity of VPN customers.

It is expected that multicast VPN solutions will use some kind of point point-to-multipoint technology to efficiently carry multicast
VPN traffic, and that such technologies require maintaining state information, and will use resources in the control plane (memory and processing, and possibly address space).

Scalability is a key requirement for multicast VPN solutions. Solutions MUST be designed to scale well with an increase in the number of any of the following:

- the number of PEs
- the number of customers VPNs (total and per PE)
- the number of PEs and sites in any VPN
- the number of client multicast channels (groups or source-groups)

Scalability of both performance and operation MUST be considered.

Key considerations SHOULD include:

- the processing resources required by the control plane (neighborhood or session maintenance messages, keep-aldives, timers, etc.)
- the memory resources needed for the control plane
- the amount of protocol information transmitted to manage a multicast VPN (e.g. signaling throughput)
- the amount of control plane processing required on PE and P to add/ remove a customer site (or a customer from a multicast session)
- the number of multicast IP addresses used (if IP multicast in ASM mode is proposed as a multicast distribution tunnel)
- other particular elements inherent to each solution that impacts scalability (e.g., if a solution uses some distribution tree inside the core, topology of the tree and number of leaf nodes may be some of them)

It is expected that the applicability of each solution will be evaluated with regards to the aforementioned scalability criteria.

These considerations naturally lead us to believe that proposed solutions SHOULD offer the possibility of sharing such resources between different multicast streams (between different VPNs, between different multicast streams of the same or of different VPNs). This means for instance, if MDTunnels are trees, being able to share an
MDTunnel between several customers.

Those scalability issues are expected to be more significant on P-routers, but a multicast in VPNs solution should address both P and PE routers as far as scalability is concerned.

5.2.2 Resource optimization

5.2.2.1 General goals

One of the aims of the use of multicast instead of unicast is resource optimization in the network.

The two obvious suboptimal behaviors that a multicast VPN solution would want to avoid are needless duplication (when same data travels twice or more on a same link, e.g. when doing ingress PE replication) and needless reception (e.g. a PE receiving traffic that it does not need because there are no downstream receivers).

5.2.2.2 Trade-off and tuning

As previously stated in this document, designing a scalable solution that makes an optimal use of resources is considered difficult. Thus what is expected from a multicast VPN solution is that it addresses the resource optimization issue while taking into account the fact that some trade-off has to be made.

Moreover, it seems that a "one size fits all" trade-off probably does not exist either, and that the most sensible approach is a versatile solution offering the service providers appropriate configuration settings that let them tune the trade-off according to their peculiar constraints (network topology, platforms, customer applications, level of service offered etc.).

As an illustration here are some example bounds of the trade-off space:

Bandwidth optimization: setting up somehow optimal core MDTunnels whose topology (PIM or P2MP LSP trees, etc.) precisely follows customer’s multicast routing changes. This requires managing an important quantity of states in the core, and also quick reactions of the core to customer multicast routing changes. This approach can be advantageous in terms of bandwidth, but it is bad in terms of state management.
State optimization: setting up MDTunnels that aggregate multiple customer multicast streams (all or some of them, across different VPNs or not). This will have better scalability properties, but at the expense of bandwidth since some MDTunnel leaves will very likely receive traffic they don’t need, and because increased constraints will make it harder to find optimal MDTunnels.

5.2.2.3 Traffic engineering

If the VPN service provides traffic engineering features for the connection used between PEs for unicast traffic in the VPN service, the solution SHOULD provide equivalent features for multicast traffic.

A solution should offer means to support key TE objectives as defined in [RFC3272], for the multicast service.

A solution MAY also usefully support means to address multicast-specific traffic engineering issues: it is known that bandwidth resource optimization in the point-to-multipoint case is a NP-hard problem, and that techniques used for unicast TE may not be applicable to multicast traffic.

5.2.3 Tunneling Requirements

5.2.3.1 Tunneling technologies

Following the principle of separation between the control plane and the forwarding plane, a multicast VPN solution SHOULD be designed so that control and forwarding planes are not inter-dependent: the control plane SHALL NOT depend on which forwarding plane is used (and vice versa), and the choice of forwarding plane SHOULD NOT be limited by the design of the solution. The solution SHOULD also NOT be tied to a specific tunneling technology.

In a multicast VPN solution extending a unicast L3 PPVPN solution, consistency in the tunneling technology has to be privileged: such a solution SHOULD allow the use of the same tunneling technology for multicast as for unicast. Migration and operations ease are the main motivations behind this requirement.

For MDTunnels (multicast distribution tunnels, the means used to carry VPNs’ multicast traffic over the provider’s network), a solution SHOULD be able to use a range of tunneling technologies, including point-to-point and point-to-multipoint, such as L2TP (including L2TP for multicast [RFC4045]), IPsec [RFC2401], GRE [RFC2784] (including GRE in multicast IP trees), IP-in-IP [RFC1853],
MPLS [RFC3031] (including MPLS P2MP extensions to RSVP [I-D.ietf-mpls-rsvp-te-p2mp] or LDP [I-D.leroux-mpls-mp-ldp-reqs][I-D.minei-mpls-ldp-p2mp][I-D.wijnands-mpls-ldp-mcast-ext]), etc. Naturally, using the point-to-multipoint variants mentioned here may help improve bandwidth efficiency in our multicast VPN context.

5.2.3.2 MTU and Fragmentation

A solution SHOULD support a method that provides minimum path MTU of the MDTunnel (e.g., to discover MTU, to tell MTU with signaling, etc.) so that:

- fragmentation inside the MDTunnel –even when allowed by the tunneling technology used– does not happen
- proper troubleshooting can be done if packets too big for the MDTunnel happen to be encapsulated in the MDTunnel

5.2.4 Control mechanisms

The solution must provide some mechanisms to control the sources within a VPN. This control includes the number of sources that are entitled to send traffic on the VPN, and/or the total bit rate of all the sources.

At the reception level, the solution must also provide mechanisms to control the number of multicast groups or channels VPN users are entitled to subscribe to and/or the total bit rate represented by the corresponding multicast traffic.

All these mechanisms must be configurable by the service provider in order to control the amount of multicast traffic and state within a VPN.

Moreover it MAY be desirable to be able to impose some bound on the quantity of state used by a VPN in the core network for its multicast traffic, whether on each P or PE router, or globally. The motivation is that it may be needed to avoid out-of-resources situations (e.g. out of memory to maintain PIM state if IP multicast is used in the core for multicast VPN traffic, or out of memory to maintain RSVP state if MPLS P2MP is used, etc.).

5.2.5 Quality of Service Differentiation

A multicast VPN solution SHOULD give a VPN service provider the ability to offer, guarantee and enforce differentiated levels of QoS
to its different customers.

5.2.6 Infrastructure security

The solution shall provide the same level of security for the service provider as what currently exist for unicast VPNs. For instance, that means that the intrinsic protection against DOS and DDOS attacks of the BGP/MPLS VPN solution must be equally supported by the multicast solution.

Moreover, since multicast traffic and routing are intrinsically dynamic (receiver-initiated), some mechanism must be proposed so that the frequency of changes in the way client traffic is carried over the core is bounded and not tightly coupled to dynamic changes of multicast traffic in the customer network. For example, multicast route dampening functions would be one possible mechanism.

Network devices that participate in the deployment and the maintenance of a given L3 VPN MAY represent a superset of the participating devices that are also involved in the establishment and the maintenance of the multicast distribution tunnels. As such the activation of IP multicast capabilities within a VPN SHOULD be device-specific, not only to make sure that only the relevant devices will be multicast-enabled, but also to make sure that multicast (routing) information will be disseminated to the multicast-enabled devices only, hence limiting the risk of multicast-inferred DOS attacks.

Unwanted multicast traffic (e.g. multicast traffic that may be sent by a source located somewhere in the Internet and for which there is no interested receiver connected to a given VPN infrastructure) MUST NOT be propagated within a multicast-enabled VPN.

Last, control mechanisms described in previous section are also to be considered from this infrastructure security point of view.

5.2.7 Robustness

Resiliency is also crucial to infrastructure security, thus a multicast VPN solution shall whether avoid single points of failures or propose some technical solution making possible to implement a failover mechanism.

As an illustration, one can consider the case of a solution that would use PIM-SM as a means to setup MDTunnels. In such a case, the PIM RP might be a single point of failure. Such a solution should thus be compatible with a solution implementing RP resiliency.
5.2.8 Management tools, OAM

The operation of a multicast VPN solution SHALL be as light as possible and providing automatic configuration and discovery SHOULD be prioritized. Particularly the operational cost of setting up multicast on a PE should be as low as possible.

Moreover, monitoring of multicast specific parameters and statistics SHOULD be offered to the service provider.

Most notably the provider SHOULD have access to:

- Multicast traffic statistics (total traffic conveyed, incoming, outgoing, dropped, etc., by period of time) - Information about client multicast resource usage (state and throughput)

- The IPPM (IP Performance Metrics [RFC2330]) -related information that is relevant to the multicast traffic usage: such information includes the one-way packet delay, the inter-packet delay variation, etc.

- Alarms when limits are reached on such resources - Statistics on decisions related to how client traffic is carried on distribution tunnels (e.g. "traffic switched onto a multicast tree dedicated to such groups or channels")

- Statistics on parameters that could help the provider to evaluate its optimality/state trade-off

All or part of this information SHOULD be made available through standardized SNMP ([RFC1157]) MIBs (Management Information Base).

5.2.9 Architectural Considerations

As far as possible, the design of a solution should carefully consider the number of protocols within the core network. If any additional protocols are introduced compared with unicast VPN, the balance between their advantage and operation burden should be examined thoroughly.

5.2.10 Compatibility and migration issues

It is a requirement that unicast and multicast services MUST be able to co-exist within the same VPN.

Likewise, the introduction of IP multicast capabilities in devices that participate to the deployment and the maintenance of a VPN SHOULD be as smooth as possible, i.e. without affecting the overall
quality provided with the services that are already supported by the underlying infrastructure.

A multicast VPN solution SHOULD prevent compatibility and migration issues, for instance by prioritizing mechanisms facilitating forward compatibility. Most notably a solution supporting only a subset of those requirements SHOULD be designed to be compatible with future enhanced revisions.

It SHOULD be an aim of any multicast into VPN solution to offer as much backward compatibility as possible. Ideally a solution would have the ability to offer multicast VPN services across a network containing some legacy routers not supporting any multicast VPN specific features.

In any case a solution SHOULD state a migration policy from possibly existing deployments.

5.2.11 Troubleshooting

A multicast VPN solution that dynamically adapts the way some client multicast traffic is carried over the provider’s network may incur the disadvantage of being hard to troubleshoot. In such a case, to help diagnose multicast network issues, a multicast VPN solution SHOULD provide monitoring information describing how client traffic is carried over the network (e.g. if a solution uses multicast-based MDTunnels, which provider multicast group is used for such and such client multicast stream). A solution MAY also provide configuration options to avoid any dynamic changes, for multicast traffic of a particular VPN or a particular multicast stream.

Moreover, a solution MAY usefully provide some mechanism letting network operators check that all VPN sites that advertised interest in a particular customer multicast stream are properly associated with the corresponding MDTunnel. Providing the operators with means to check the proper setup and operation of MDTunnels MAY also be provided (e.g. when MPLS is used for MDTunnels, integrating mechanisms such as LSPPing[I-D.ietf-mpls-lsp-ping][I-D.yasukawa-mpls-p2mp-lsp-ping] into the L3VPN troubleshooting functionalities will be desirable). Depending on the implementation such verification could be initiated by source-PE or receiver-PE.

5.2.12 Inter-AS, inter-provider

A multicast VPN solution SHOULD support inter-AS and inter-inter-provider VPNs. Considerations about coexistence with unicast inter-AS VPN Options A, B and C (as described in section 10 of
RFC2547bis [I-D.ietf-l3vpn-rfc2547bis]) are strongly encouraged.

A multicast VPN solution SHOULD provide inter-AS mechanisms requiring
the least possible coordination between providers, and keep the need
for detailed knowledge of providers networks to a minimum - all this
being in comparison with corresponding unicast VPN options.

- Within each service provider the service provider SHOULD be able
  on its own to pick the most appropriate tunneling mechanism to
carry (multicast) traffic among PEs (just like what is done today
  for unicast)

- If a solution does require a single tunnel to span P routers in
  multiple ASs, the solution SHOULD provide mechanisms to ensure
  that the inter-provider co-ordination to setup such a tunnel is
  minimized.

Moreover such support should be possible without compromising other
requirements expressed in this requirement document, and should not
incur penalty on scalability and bandwidth usage.
6. Security Considerations

This document does not by itself raise any particular security issue.

A set of security issues have been identified that MUST be addressed when considering the design and deployment of multicast-enabled VPN networks. Such issues have been described in Section 5.1.5 and Section 5.2.6.
7. Acknowledgments

The authors would like to thank, by rough chronological order, Vincent Parfait (Equant), Zubair Ahmad (Equant), Elodie Hemon-Larreur, Sebastien Loye (France Telecom), Rahul Aggarwal (Juniper), Hitoshi Fukuda (NTT Communications), Luyuan Fang (AT&T), Adrian Farrel, Daniel King, Yiqun Cai (Cisco), Ronald Bonica, Len Nieman, Satoru Matsushima (Japan Telecom), Netzahualcoyotl Ornelas (Renater), for their review, valuable input and feedback.
8. References

8.1 Normative references


8.2 Informative references


[I-D.ietf-pim-bidir]
Handle, M., Kouvelas, I., Speakman, T., and L. Vicisano,
"Bi-directional Protocol Independent Multicast (BIDIR-PIM)", draft-ietf-pim-bidir-07 (work in progress),
March 2005.


[RFC3353] Ooms, D., Sales, B., Livens, W., Acharya, A., Griffoul, F., and F. Ansari, 
"Overview of IP Multicast in a Multi-Protocol Label Switching (MPLS) Environment", RFC 3353, 
August 2002.

Xiao, "Overview and Principles of Internet Traffic 

Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, 
March 2000.

[RFC4045] Bourdon, G., "Extensions to Support Efficient Carrying of 
Multicast Traffic in Layer-2 Tunneling Protocol (L2TP)", 
RFC 4045, April 2005.

[RFC3809] Nagarajan, A., "Generic Requirements for Provider 
Provisioned Virtual Private Networks (PPVPN)", RFC 3809, 


[RFC2330] Paxson, V., Almes, G., Mahdavi, J., and M. Mathis, 
"Framework for IP Performance Metrics", RFC 2330, 
May 1998.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., 
and W. Weiss, "An Architecture for Differentiated 


[RFC1157] Case, J., Fedor, M., Schoffstall, M., and J. Davin, 
"Simple Network Management Protocol (SNMP)", STD 15,
RFC 1157, May 1990.

[I-D.ietf-mpls-lsp-ping]
Kompella, K. and G. Swallow, "Detecting MPLS Data Plane
Failures", draft-ietf-mpls-lsp-ping-09 (work in progress),
May 2005.

[RFC1191]  Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191,
November 1990.

[I-D.yasukawa-mpls-p2mp-lsp-ping]
Yasukawa, S., "Detecting Data Plane Failures in Point-to-
Multipoint MPLS Traffic Engineering - Extensions to LSP
Ping", draft-yasukawa-mpls-p2mp-lsp-ping-02 (work in
progress), April 2005.

[I-D.mathis-frag-harmful]
Mathis, M., "Fragmentation Considered Very Harmful",
draft-mathis-frag-harmful-00 (work in progress),
July 2004.

[RFC2629]  Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629,
June 1999.
URIs

[3] <mailto:y.kamite@ntt.com>
[4] <mailto:jeanlouis.leroux@francetelecom.com>
[5] <mailto:nicolai.leymann@t-systems.com>
[6] <mailto:renaud.moignard@francetelecom.com>
[7] <mailto:thomas.morin@francetelecom.com>

Author’s Address

Thomas Morin (editor)
France Telecom R&D
2, avenue Pierre Marzin
Lannion  22307
France

Email: thomas.morin@rd.francetelecom.com
Appendix A. Requirements summary

[This section will contain a summary of all requirements of this document, that were expressed as MUST or SHOULD].
Appendix B. Changelog

This section lists changes made to this document (minor or editorial changes excepted) between major revisions.

It shall be removed before publication as an RFC.

B.1 Changes between -00 and -01

- integrated comments made on L3VPN WG mailing list after -00 submission
- completed Carrier’s carrier section (5.1.9)
- updates in sections 5.1 and 5.2 about minimum MTU
- added a section about "Quality of Service Differentiation" as ISP requirement (section 5.2.5)
- added P2MP LDP extensions as possible MDTunnels techniques (section 5.2.3.1)
- started to build section 4 "Use Case"
- detailed section 5.1.3 "QoS", most notably about group join and leave delays
- additions to section 5.2.12 "Inter-AS, inter-provider"
- added MDTunnel verification requirement to section 5.2.11
- moved "Architectural Considerations" section
- moved contributors to top of document
- made draft content agnostic to unicast L3VPN solutions
- added two appendixes: "Changelog" and "Requirement summary"
- conversion to XML [RFC2629] with the help of some scripting and Bill Fenner’s xml2rfc XMLEnd plugin
- lot’s of editorial changes
Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (C) The Internet Society (2005). This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.