

INTERNET DRAFT  
Internet Engineering Task Force  
Document:  
[draft-boddapati-mpls-pim-ldp-p2mp-00.txt](#)  
November 2005  
Expires: May 2006

Suresh Boddapati  
Venu Hemige  
Alcatel

## **P2MP LSP Setup using PIM-SSM and LDP**

Status of this memo

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>.

### **IPR Disclosure Acknowledgement**

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](#).

### **Abstract**

There is emerging interest in the use of MPLS LSPs for the delivery of multicast traffic. Efficient delivery of multicast traffic using MPLS requires P2MP LSPs. This document describes a mechanism to setup P2MP LSPs using PIM-SSM and LDP. PIM-SSM is a widely deployed multicast routing protocol and has well defined procedures for signalling multicast traffic. LDP is a well defined mechanism for distribution of

MPLS labels. Utilizing both PIM-SSM and LDP for setting up P2MP LSPs keeps the clear separation between signalling and label distribution, and minimizes changes to both protocols. PIM-SSM signals when to receive multicast traffic and LDP distributes label information regarding how to receive multicast traffic.

Conventions used in this document

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 [RFC 2119](#).

## Table of Contents

<a href="#">1. Overview.....</a>	<a href="#">2</a>
<a href="#">2. Label allocation and distribution mechanisms for P2MP LSPs.....</a>	<a href="#">3</a>
<a href="#">2.1. Downstream and upstream label distribution.....</a>	<a href="#">3</a>
<a href="#">2.2. Label Distribution Control.....</a>	<a href="#">4</a>
<a href="#">2.3. Liberal Label Retention.....</a>	<a href="#">4</a>
<a href="#">3. Procedures for setting up P2MP LSPs.....</a>	<a href="#">4</a>
<a href="#">3.1. Egress LSR Procedures.....</a>	<a href="#">4</a>
<a href="#">3.2. Branch LSR procedures.....</a>	<a href="#">5</a>
<a href="#">3.3. Ingress LSR procedures.....</a>	<a href="#">6</a>
<a href="#">3.4. Receiving P2MP FEC Advertisements.....</a>	<a href="#">6</a>
<a href="#">4. PIM-SSM-LDP interaction.....</a>	<a href="#">7</a>
<a href="#">5. Detecting and Stopping Duplicate Traffic on a LAN.....</a>	<a href="#">8</a>
<a href="#">6. P2MP FEC Element.....</a>	<a href="#">8</a>
<a href="#">7. An example.....</a>	<a href="#">9</a>
<a href="#">8. Security Considerations.....</a>	<a href="#">10</a>
<a href="#">9. IANA Considerations.....</a>	<a href="#">10</a>
<a href="#">10. Acknowledgements.....</a>	<a href="#">10</a>
<a href="#">11. Normative References.....</a>	<a href="#">10</a>
<a href="#">12. Informative References.....</a>	<a href="#">11</a>
<a href="#">13. Authors' Addresses.....</a>	<a href="#">11</a>
<a href="#">14. Intellectual Property Statement.....</a>	<a href="#">11</a>
<a href="#">15. Full copyright statement.....</a>	<a href="#">12</a>

## **[1. Overview](#)**

Efficient delivery of multicast traffic using MPLS requires P2MP LSPs. A P2MP LSP has one Ingress LSR and one or more Egress LSRs. This document focuses on the setup of leaf-initiated P2MP LSPs. For creating such LSPs, the following options can be considered.

- An existing multicast routing protocol such as PIM-SSM can be extended to also distribute labels required to set up P2MP LSPs.



- An existing label distribution protocol such as LDP can be extended to also build multicast trees. [LDP-P2MP1] and [LDP-P2MP2] are proposals based on this option.
- Without significant modifications to PIM-SSM and LDP, procedures can be defined for both PIM-SSM and LDP to interact with each other such that PIM-SSM builds multicast trees and LDP distributes labels to be used for forwarding labeled traffic on those multicast trees.

This document proposes the use of the third option described above. The advantages of such a mechanism are:

- There is considerable experience built in the development and use of PIM-SSM as a multicast routing protocol. This can be leveraged for building multicast trees for P2MP LSPs as well. This way, the changes to LDP are minimal.
- The changes to PIM-SSM are also minimal. Modifying PIM-SSM to distribute labels would mean that if any other routing protocol replaces PIM-SSM, then that protocol would also have to have procedures defined for label distribution. Besides, PIM-SSM does not lend itself very well for label distribution. Specifically, upstream label allocation is not possible with PIM-SSM as it exists today, since there is no message that goes from an upstream router to a downstream router in response to a Join received from the downstream router.
- PIM-SSM is equipped to deal with duplicate traffic on LANs, for scenarios where multiple upstream routers are forwarding the same traffic to the LAN. Using PIM Assert messages, PIM ensures that only one router forwards traffic to the LAN.

## **2. Label allocation and distribution mechanisms for P2MP LSPs**

### **2.1. Downstream and upstream label distribution**

Downstream unsolicited label distribution suits well for leaf-initiated P2MP LSPs. However, in the case of LAN interfaces, where it is efficient to send only one copy of the multicast traffic, instead of ingress replicating it multiple times, an upstream label allocation scheme is more suitable. [UPSTREAM] discusses one way of assigning upstream labels, where labels come from a context specific label space. Downstream assigned labels can be allocated from the platform wide label space.

This document proposes that downstream unsolicited label distribution **MUST** be used for point-to-point interfaces and upstream unsolicited label distribution **SHOULD** be used for LAN interfaces.



## 2.2. Label Distribution Control

This document also proposes the use of Independent Control for label distribution. If a network has only LAN interfaces and if Ordered Control is used, the P2MP LSP will always end up being constructed from the root. Moreover, traffic could start flowing on the P2MP LSP before the tree is completely set up, and get dropped downstream. On the other hand, if Independent Control is used, the label distribution occurs as soon as the multicast state is available.

## 2.3. Liberal Label Retention

This document also proposes that Liberal Label Retention be used for P2MP LSPs. Liberal label retention enables faster convergence in the following cases:

- When a P2MP LSP is already setup, and new branches are added to it.
- When topology changes occur in the network and the path to the root of the P2MP LSP changes.

## 3. Procedures for setting up P2MP LSPs

In order to setup P2MP LSPs, all LSRs MUST run PIM-SSM as the multicast routing protocol in the domain. All LSRs MUST also run LDP as the label distribution protocol in the domain.

Each P2MP LSP is associated with an (S,G), where S is the IP address of the source of the multicast traffic which will be mapped on to the P2MP LSP. S can be the address of the Ingress LSR, a host directly connected to the Ingress LSR, or a remote host whose traffic is being mapped to a P2MP LSP at the Ingress LSR. G is the multicast group address representing the P2MP LSP. Both S and G can belong to any address family, such as Ipv4, Ipv6 etc. Note that for the purposes of this document, the PIM-SSM requirement does not mean the SSM range of multicast addresses has to be used. It simply means that there is no Rendezvous Point (RP) in the network and no (\*,G) state is built. There is no restriction on the group address that can be used.

The creation of a P2MP LSP is triggered at an Egress LSR when an (S,G) is added to the multicast routing table by PIM-SSM. The router can create P2MP LSPs for all (S,G)s in its multicast routing table or do it selectively based on a local policy.

### 3.1. Egress LSR Procedures

An Egress LSR is a leaf of a P2MP LSP. On an Egress LSR, a labeled multicast packet is received and is forwarded natively (the label is popped) on one or more interfaces.



When a new multicast routing table entry for an (S,G) is added (due to receiving a PIM-SSM Join from a downstream router or due to having local multicast receivers), an LSR does the following:

- It creates an (S,G) multicast forwarding entry which typically consists of an incoming interface (also known as the RPF interface, the interface on which traffic is expected to arrive), and a set of outgoing interfaces. This is done using regular PIM-SSM procedures as defined in [[PIM-SM](#)].
- It allocates a P2MP FEC (see [Section 6](#). ) for the (S,G), and using LDP advertises a label corresponding to the FEC, to each of its directly connected LDP peers. On a point-to-point interface, the label is a downstream assigned label and represents the label which the upstream router should use while sending labeled traffic for the (S,G). On a LAN interface, the label is an upstream assigned label and represents the label which this router would encapsulate the multicast packet in before sending it out on the interface.
- A PIM-SSM Join is sent toward the RPF neighbor corresponding to the (S,G). The RPF neighbor is determined from the unicast routing table. If the interface on which the PIM-SSM Join is sent is a point-to-point interface, an ILM entry is also created for the FEC representing the (S,G), with incoming label set to the label that was advertised on the interface for the (S,G). If the interface on which the PIM-SSM Join is sent is a LAN interface, the ILM entry is created only if the label mapping for the FEC representing the (S,G) is already known (due to the upstream LSR having advertised it).
- The label action in the ILM entry is to pop the label. Further action is determined by what the traffic sent on the P2MP LSP represents. This is outside of the scope of this document.

### 3.2. Branch LSR procedures

A Branch LSR is an LSR in the path between an Ingress LSR and an Egress LSR. On a Branch LSR, a labeled multicast packet is received and is forwarded as a labeled packet (the label undergoes a swap operation) on one or more interfaces. The label encapsulation on one outgoing interface has no relation to the label encapsulation on another.

When a PIM-SSM Join is received by a Branch LSR, if the Join creates a new multicast routing table entry, the procedures are the same as in [Section 3.1](#) except for the ILM entry. The label action for the ILM





entry is to swap the label on each of the outgoing interfaces. The outgoing interface list is as determined by PIM-SSM and the label to encapsulate on each interface is as determined by LDP. If the outgoing interface is a point-to-point, the label encapsulated is the label advertised by the downstream LSR for the FEC associated with the (S,G). If the outgoing interface is a LAN interface, the label encapsulated is the label advertised by this LSR to its downstream LSRs on the LAN for the FEC associated with the (S,G).

If the PIM-SSM Join received by an LSR adds an outgoing interface to an existing (S,G), the following actions are taken.

- The interface is added to the outgoing interface list of the ILM entry.
- The label to be encapsulated for this outgoing interface is also programmed. The label used is as described above, based on whether the interface is a point-to-point interface or a LAN interface.

### 3.3. Ingress LSR procedures

An ingress LSR is the root of the P2MP LSP. The ingress LSR has one or more FTN entries which map incoming multicast traffic to P2MP LSPs. How a traffic flow is mapped to a P2MP LSP is a policy decision and is outside the scope of this document. The label operation at the ingress LSR is to push a label onto the received multicast packet before forwarding it on one or more interfaces. The label pushed for one downstream interface has no relation to the label pushed for another downstream interface.

When an Ingress LSR receives a PIM-SSM Join for an (S,G), if it does not have an FTN entry associated with the (S,G), the following actions are taken.

- An FTN entry is created corresponding the (S,G).
- The interface on which the Join was received is added to the outgoing interface list.
- The label to be encapsulated for this outgoing interface is also programmed. The label used is as described in [Section 3.2](#), based on whether the interface is a point-to-point interface or a LAN interface.

When an Ingress LSR receives a PIM-SSM Join for an (S,G), and it already has an FTN entry associated with the (S,G), the first step is bypassed and the rest of the actions as described above are taken.

### 3.4. Receiving P2MP FEC Advertisements

When an LSR receives a P2MP FEC advertisement in a Label Mapping message from a peer, if the LSR has an (S,G) state corresponding to the FEC, the following actions are taken.



- If the neighbor from which the FEC was received is an RPF neighbor on a LAN interface, an ILM entry for the FEC corresponding to the (S,G) is programmed with incoming label set to the label from the received FEC.
- If the neighbor from which the FEC was received is a downstream LSR that also sent a PIM-SSM Join for the (S,G) corresponding to the FEC, and the interface is a point-to-point interface, the label encapsulation for that interface in the outgoing interface list of the ILM entry is updated with the label from the FEC.

#### **4. PIM-SSM-LDP interaction**

As can be seen from [Section 3](#), the setting up of P2MP LSPs requires interaction between PIM-SSM and LDP. The interactions are as described below.

- When PIM-SSM creates new (S,G), it notifies LDP of the (S,G). This causes LDP to create a FEC for the (S,G) and advertise labels to all its neighbors using Label Mapping messages. PIM-SSM also notifies LDP of the RPF neighbor for the (S,G). LDP programs the ILM entry using the incoming label for the FEC advertised to the RPF neighbor (on point-to-point interfaces) or by the RPF neighbor (on LAN interfaces). Note that on point-to-point interfaces, the label is downstream advertised while on LAN interfaces, the label is upstream advertised.
- When PIM-SSM receives a Join on an interface to an (S,G), it notifies LDP of the outgoing interface. This causes LDP to program the interface in the outgoing interface list for the FEC corresponding to the (S,G) in either the FTN entry or the ILM entry (based on whether the LSR is an Ingress LSR or an Egress or Branch LSR). The label on the outgoing interface for the FEC is also programmed if available.
- When PIM-SSM prunes an interface for an (S,G), it notifies LDP of the prune. This causes LDP to remove the interface from the outgoing interface list for the FEC corresponding to the (S,G) in either the FTN entry or the ILM entry.
- When PIM-SSM removes an (S,G) from its database, it notifies LDP of the removal. This causes LDP to withdraw labels for the FEC associated with the (S,G) from all its neighbors using Label Withdraw messages.
- When PIM-SSM detects an RPF neighbor change either due to change in the unicast routing topology or due to Assert election, PIM-SSM notifies LDP of the change in RPF neighbor.





[illegible]

```

|                               Group Address (Encoded Group Format)                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type: To be assigned by IANA

Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.

Source Address: The source address (S) in (S,G) associated with this P2MP FEC Element. The source address MUST be encoded using the Encoded Source Format defined in [[PIM-SM](#)].

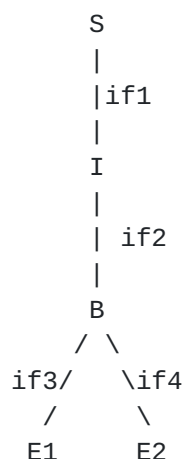
Group Address: The multicast group address (G) in (S,G) associated with this P2MP FEC Element. The group address MUST be encoded using the Encoded Group Format defined in [[PIM-SM](#)].

The P2MP FEC element MUST be advertised in a FEC TLV without any other elements in it. The P2MP FEC element MUST be present only once in a FEC TLV. If present more than once, the FEC TLV MUST NOT be processed and an "Unknown FEC" Notification MUST be sent to the peer that advertised it.

A P2MP FEC element received with an unknown address family or unknown encoding format MUST NOT be processed and an "Unknown FEC" Notification MUST be sent to the peer that advertised it.

## 7. An example

Consider the following network.



I is an Ingress LSR connected to a source S on interface if1.

B is a Branch LSR connected to I on interface if2. if2 is a point-to-point interface.

E1 and E2 are Egress LSRs connected to B on interfaces if3 and if4 respectively. if3 is a point-to-point interface and if4 is a LAN



interface.

Receivers for a group  $G$  are connected to  $E1$  and  $E2$ .

Assuming that a P2MP LSP is desired from I to E1 and E2 for the flow (S,G), the LSP is setup by the following steps.

- Since E1 has receivers for group G and assuming it is known that the source is S, E1 initiates a PIM-SSM Join for (S,G) towards B.
- E1 also advertises a label mapping LE1 for the FEC associated with (S,G) to B, since if3 is a point-to-point interface. LE1 is a downstream allocated label which comes from platform wide label space.
- B forwards the PIM-SSM Join towards I.
- B advertises a downstream assigned platform wide label mapping LB for (S,G) to its upstream neighbor I and its downstream neighbor E1 since both are point to point interfaces. B also advertises a context specific upstream label mapping LBif4 to E2 on interface if4.
- When I receives the (S,G) Join from B, since it is directly connected to S, it creates an FTN entry that maps (S,G) traffic to a set of P2MP NHLFEs. In this case, the set consists of outgoing interface if1 and label LB.
- The P2MP LSP setup is now complete.
- Now if E2 wants to join the P2MP LSP, it simply sends a PIM-SSM Join towards B. E2 will also advertise an upstream assigned label to B. Since it already has the upstream label mapping from B, it programs an ILM entry with incoming label LBif4, that was advertised by B.
- When B receives the PIM-SSM Join from E2, it adds if4 to the outgoing interface list for the ILM entry corresponding to the (S,G), and the outgoing label on if4 is set to LBif4. Thus the branch of the P2MP LSP is added.

## **8. Security Considerations**

**This document does not introduce any new security considerations that are not inherent to PIM and LDP.**

## **9. IANA Considerations**

**This document defines a new FEC Element for which the type has to be assigned by the IANA.**

## **10. Acknowledgements**

**The authors would like to acknowledge in no particular order, Alex Zinin, Joe Regan, Ray Qiu, Vach Kompella, Ron Haberman, Sunil Khandekar, Devendra Raut and Jayant Kotalwar for their input and valuable feedback.**

## **11. Normative References**

[LDP] Andersson, L, et al. "LDP Specification", [RFC 3036](#)



## **12. Informative References**

- [PIM-SM] Fenner, B, et al. "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification", [draft-ietf-pim-sm-v2-new-11.txt](#)
- [P2MP-REQ] Le Roux, J-L et al. "Requirements for point-to-multipoint extensions to the Label Distribution Protocol", [draft-leroux-mpls-mp-ldp-reqs-01.txt](#)
- [UPSTREAM] Aggarwal, R, et al. "MPLS Upstream Label Assignment and Context Specific Label Space", [draft-raggarwa-mpls-upstream-label-00.txt](#)
- [LDP-P2MP1] Minei, I, et al. "Label Distribution Protocol Extensions for Point-to-Multipoint Label Switched Paths", [draft-minei-mpls-ldp-p2mp-01.txt](#)
- [LDP-P2MP2] Wijnands, I, et al. "Multicast Extensions for LDP", [draft-wijnands-mpls-ldp-mcast-ext-00.txt](#)
- [VENU-ASSERT] Hemige, V, et al. "Improved Assert in PIM", [draft-hemige-pim-improved-assert-00.txt](#)

## **13. Authors' Addresses**

Suresh Boddapati  
Alcatel North America  
701 East Middlefield Rd.  
Mountain View, CA 94043  
Suresh.boddapati@alcatel.com

Venu Hemige  
Alcatel North America  
701 East Middlefield Rd.  
Mountain View, CA 94043  
Venu.hemige@alcatel.com

## **14. 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](mailto:ietf-ipr@ietf.org).

#### **15. Full 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.

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.

