

PIM Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 18, 2013

H. Asaeda
NICT
S. Jeon
Institute de Telecomunicacoes
October 15, 2012

Multiple Upstream Interfaces Support for IGMP/MLD Proxy
draft-asaeda-pim-mlproxy-multif-00

Abstract

This document describes the way of supporting multiple upstream interfaces for an IGMP/MLD proxy device. The proposed extension enables that an IGMP/MLD proxy device receives multicast packets through multiple upstream interfaces, so that it is useful for multihoming support.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on April 18, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as

described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Terminology	3
3. Upstream Interface Selection	4
3.1. Basic Operation	4
3.2. Supported Address Prefix	5
3.3. Interface Priority	6
4. IANA Considerations	6
5. Security Considerations	6
6. Normative References	6
Authors' Addresses	7

1. Introduction

The Internet Group Management Protocol (IGMP) [1][2][3][4] for IPv4 and the Multicast Listener Discovery Protocol (MLD) [5][6][4] for IPv6 are the standard protocols for hosts to initiate joining or leaving of multicast sessions. The IGMP/MLD proxy [7] maintains multicast membership information by IGMP/MLD protocols on the downstream interfaces and forwards IGMP/MLD report via the upstream interface to the upstream multicast routers.

According to the specification of [7], a proxy device performing IGMP/MLD-based forwarding (as known as IGMP/MLD proxy) has *a single* upstream interface and one or more downstream interfaces. It performs the router portion of the IGMP or MLD protocol on its downstream interfaces, and the host portion of IGMP/MLD on its upstream interface. The proxy device must not perform the router portion of IGMP/MLD on its upstream interface.

On the other hand, there are requirements that an IGMP/MLD proxy device allows to use multiple upstream interfaces. For example, a proxy device having more than two interfaces may want to access to different networks, such as Internet and Intranet. Or, a proxy device having wired link (e.g., ethernet) and high-speed wireless link (e.g., WiMAX or LTE) may want to have the capability to connect to the Internet through both links. These proxy devices shall receive multicast packets from the different upstream interfaces and forward to the downstream interface(s).

This document describes the way of supporting the scenario in which an IGMP/MLD proxy device enables to configure "multiple upstream interfaces" and receives multicast packets through these interfaces. An IGMP/MLD proxy device selects a single upstream interface from configured upstream interfaces per IGMP/MLD records; same IGMP/MLD records MUST NOT be transmitted from different upstream interfaces simultaneously. This document does not make any changes to the IGMPv3 and MLDv2 protocols, and only adds the functionality to configure multiple upstream interfaces on an IGMP/MLD proxy device by operation. Therefore, this document does not provide any mechanism to "dynamically configure" multiple upstream interfaces, and provides a mechanism to "manually configure" an upstream interface by operation.

2. Terminology

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 [8].

In addition, the following terms are used in this document.

Upstream interface (or selected upstream interface):

A proxy device's interface in the direction of the root of the multicast forwarding tree.

Downstream interface:

Each of a proxy device's interfaces that is not in the direction of the root of the multicast forwarding tree.

Configured upstream interface:

An interface that potentially becomes an upstream interface of the proxy device.

3. Upstream Interface Selection

3.1. Basic Operation

An IGMP/MLD proxy device maintains a database consisting of the merger of all subscriptions on any downstream interface. It sends IGMP/MLD membership report messages on the upstream interface when the database changes (e.g., by receiving solicited/unsolicited report messages). The proxy device then forwards appropriate multicast packets received on its upstream interface to each downstream interface based on the downstream interface's subscriptions.

The multicast forwarding tree must be manually configured by designating upstream and downstream interfaces on an IGMP/MLD proxy device, and the root of the tree is expected to be connected to a wider multicast infrastructure. This document provides the way of supporting the scenario in which an IGMP/MLD proxy device enables to configure multiple upstream interfaces and receives multicast packets through these interfaces.

Configured upstream interfaces MUST be manually set up by operation. An IGMP/MLD proxy device MUST NOT select multiple upstream interfaces for the same IGMP/MLD records, and hence the same IGMP/MLD records MUST NOT be transmitted through different upstream interfaces.

Regarding the case that a proxy device receives multicast packets on its downstream interface, it forwards the packets to each downstream interface based on the downstream interface's subscriptions. A proxy device forwards packets received on any downstream interface to the configured upstream interfaces, and to each downstream interface other than the incoming interface based upon the downstream interfaces' subscriptions.

3.2. Supported Address Prefix

The "supported address prefixes" MAY be configured for each configured upstream interface by operation. The supported address prefix is expressed by the following information:

(multicast address prefix, source address prefix)

An IGMP/MLD proxy device selects an upstream interface from its configured upstream interfaces based on the configuration of the supported address prefixes. When the proxy device transmits an IGMP/MLD report message, it examines the source and multicast addresses in the IGMP/MLD records of the report message and transmits the appropriate IGMP/MLD report message(s) from the selected upstream interface(s) that are configured with the range of the supported source and multicast address prefixes.

The default values of both source and multicast address prefixes are a wildcard. If no address prefix value is configured on a configured upstream interface, a wildcard value (i.e., default value) is implicitly set up for the configured upstream interface. The wildcard multicast address prefix is represented by the entire multicast address range (i.e., '224.0.0.0/4' for IPv4 or 'ff00::/8' for IPv6). The wildcard source address prefix is represented by any host. If the default value is set up on a configured upstream interface, the decision whether the configured upstream interface is selected as the upstream interface or not is made by the "interface priority" value defined in Section 3.3.

There may be the case that one configured upstream interface is configured with specific multicast address prefixes (i.e., non wildcard value) and the other configured upstream interface is configured with specific source address prefixes. In this case, the proxy device may need to transmit an IGMP/MLD record whose source address, say S, is in the range of the supported source address prefix of the configured upstream interface A, and whose multicast address, say G, is in the range of the supported multicast address prefix of the configured upstream interface B. For such case, the proxy device selects the configured upstream interface A, which supports the source address prefix, as the upstream interface, and then the (S,G) record is transmitted via the interface A.

The same address prefix MUST NOT be configured on different configured upstream interfaces. If the same address prefix is configured on different configured upstream interfaces, that address prefix configuration is ignored and warned the mis-configuration.

3.3. Interface Priority

Each configured upstream interface SHOULD have the "interface priority" value. The priority value is configured by operation. The configured upstream interface with the highest priority is chosen as the upstream interface. If there is more than one configured upstream interfaces and all of the priorities are identical, the configured upstream interface having lower IP address is selected as the upstream interface.

The default value of the interface priority is 0.

4. IANA Considerations

This document has no actions for IANA.

5. Security Considerations

This document neither provides new functions nor modifies the standard functions defined in [1][2][3][4][5][6]. Therefore there is no additional security consideration provided.

6. Normative References

- [1] Deering, S., "Host Extensions for IP Multicasting", RFC 1112, August 1989.
- [2] Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, July 1997.
- [3] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.
- [4] Liu, H., Cao, W., and H. Asaeda, "Lightweight Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Version 2 (MLDv2) Protocols", RFC 5790, February 2010.
- [5] Deering, S., Fenner, W., and B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, October 1999.
- [6] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.
- [7] Fenner, B., He, H., Haberman, B., and H. Sandick, "Internet

Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying"), RFC 4605, August 2006.

- [8] Bradner, S., "Key words for use in RFCs to indicate requirement levels", RFC 2119, March 1997.

Authors' Addresses

Hitoshi Asaeda
National Institute of Information and Communications Technology
4-2-1 Nukui-Kitamachi
Koganei, Tokyo 184-8795
Japan

Email: asaeda@nict.go.jp

Seil Jeon
Institute de Telecomunicacoes
Campus Universitario de Santiago
Aveiro 3810-193
Portugal

Email: seiljeon@av.it.pt

Versions: 00

PIM WG
Internet-Draft
Intended status: Informational
Expires: April 15, 2013

J. Asghar
IJ. Wijnands
S. Krishnaswamy
Cisco Systems, Inc.
V. Arya
Directv, Inc.
October 15, 2012

Explicit RPF Vector
draft-asghar-pim-explicit-rpf-vector-00

Abstract

This document describes a use of the Reverse Path Forwarding (RPF) Vector TLV as defined in [RPC 5496] to build multicast trees via an explicitly configured path sent in the PIM join.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and 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/lid-abstracts.html>

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

This Internet-Draft will expire April 15, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents

(<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Specification of Requirements	3
3. Use of the Explicit RPF Vector	3
4. Explicit RPF Vector Attribute TLV Format	4
5. Interoperability	4
6. IANA Considerations	4
7. Security Considerations	5
8. Acknowledgments	5
9. Normative References	5
Authors' Addresses	6

1. Introduction

In some applications it might be useful to have a way to specify the explicit path along which the PIM join is propagated.

This document defines a new TLV in the PIM Join Attribute message [RFC5384] for specifying the explicit path.

The procedures in [RFC5496] define how a RPF vector can be used to influence the path selection in the absence of a route to the Source. However, the same procedures can be used to override a route to the Source when it exists. It is possible to include multiple RPF vectors in the stack where each router along the path will perform a unicast route lookup on the first vector in the attribute list. Once the router owning the address of the RPF vector is reached, following the procedures in [RFC5496], the RPF vector will be removed from the attribute list. This will result in a 'loosely' routed path based on the unicast reachability of the RPF vector(s). We call this loosely because we still depend on unicast routing reachability to the RPF Vector.

In some scenario's we don't want to rely on the unicast reachability to the RPF vector address and we want to build a path strictly based on the RPF vectors. In that case the RPF vector(s) represent a list of directly connected PIM neighbors along the path. For these vectors we MUST NOT do a unicast route lookup. We call these 'explicit' RPF vector addresses. If a router receiving an explicit RPF vector does not have a PIM neighbor matching the explicit RPF vector address it MUST NOT fall back to loosely routing the JOIN. Since the behavior of the explicit RPF vector differs from the loose RPF vector as defined [RFC5496], we're defining a new attribute called the explicit RPF Vector.

2. Specification of Requirements

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. Use of the PIM Explicit RPF Vector

Normally PIM builds a receiver driven multicast forwarding tree by sending PIM Joins from the leaf router towards root based on unicast route lookup to source.

Figure 1 provides an example multicast join path R4->R3->R2->R1, where the forwarding states are installed hop-by-hop dynamically.

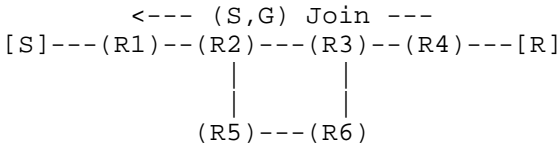


Figure 1

Figure 2 provides an example multicast join path R4->R3->R6->R5->R2->R1, where the multicast JOIN is explicitly routed to the source hop-by-hop using the explicit RPF vector list.

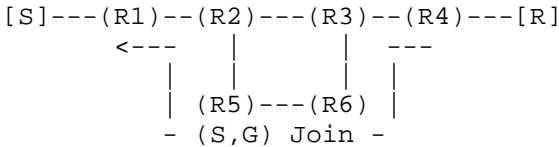
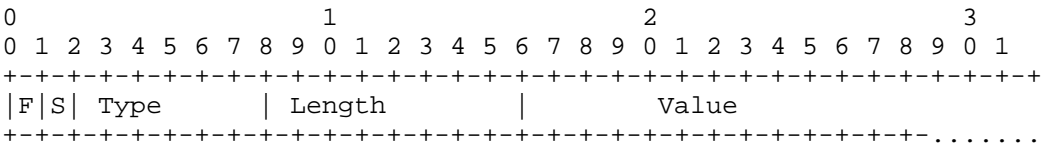


Figure 2

4. Explicit RPF Vector Attribute

This draft uses vector attribute 1 for specifying an explicit rpf vector.

5. Explicit RPF Vector Attribute TLV Format



F bit

Forward Unknown TLV. If this bit is set the TLV is forwarded regardless of whether the router understands the Type. If the TLV is known the F bit is ignored.

S bit

Bottom of Stack. If this bit is set then this is the last TLV in the stack.

Type

The Vector Attribute type is 1.

Length

Length depending on Address Family of Encoded-Unicast address.

Value

Encoded-Unicast address.

6. Interoperability

The behaviour is dictated by the F flag as specified in RFC 5496.

7. IANA Considerations

An new attribute type from the "PIM Join Attribute Types" registry needs to be assigned by IANA for the RPF Vector. The proposed value is 1.

8. Security Considerations

Security of the RPF Vector Attribute is only guaranteed by the security of the PIM packet, so the security considerations for PIM join packets as described in PIM-SM [RFC4601] apply here.

9. Acknowledgments

The authors would like to thank Vatsa Kumar for the comments on the draft.

10. Normative References

- [RFC5496] Wijnands, IJ., Boers, A., Rosen, E., "The Reverse Path Forwarding (RPF) Vector TLV", RFC 5496, March 2009.
- [RFC5384] Boers, A., Wijnands, IJ., Rosen, E., "The Protocol Independent Multicast (PIM) Join Attribute Format", RFC 5384, Nov 2008.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006.
- [RFC5384] Boers, A., Wijnands, I., and E. Rosen, "The Protocol Independent Multicast (PIM) Join Attribute Format", RFC 5384, November 2008.

Authors' Addresses

Javed Asghar
Cisco Systems, Inc.
725, Alder Drive
Milpitas, CA 95035

Email: jasghar@cisco.com

IJsbrand Wijnands
Cisco Systems, Inc.
De kleetlaan 6a
Diegem 1831
Belgium

Email: ice@cisco.com

Sowmya Krishnaswamy
Cisco Systems, Inc.
3750 Cisco Way
San Jose, CA 95134

Email: sowkrish@cisco.com

Vishal Arya
DIRECTV Inc.
2230 E Imperial Hwy
El Segundo, CA 90245

Email: varya@directv.com

Html markup produced by rfcmarkup 1.98, available from <http://tools.ietf.org/tools/rfcmarkup/>

MULTIMOB Working Group
INTERNET-DRAFT
Intended Status: Proposed Standard
Expires: April 18, 2013

Luis M. Contreras
Telefonica I+D
Carlos J. Bernardos
Universidad Carlos III de Madrid
October 15, 2012

Extension of the MLD proxy functionality to support multiple
upstream interfaces
draft-contreras-multimob-multiple-upstreams-00

Abstract

This document presents different scenarios of applicability for an MLD proxy running more than one upstream interface. Since those scenarios impose different requirements on the MLD proxy with multiple upstream interfaces, it is important to ensure that the proxy functionality address all of them for compatibility.

The purpose of this document is to define the requirements in an MLD proxy with multiple interfaces covering a variety of applicability scenarios, and to specify the proxy functionality to satisfy all of them.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and 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/lid-abstracts.html>

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

Copyright and License Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1	Introduction	4
2.	Terminology	4
3.	Problem statement	4
4.	Scenarios of applicability	4
4.1	Applicability to multicast listener mobility	4
4.1.1	Single MLD proxy instance on MAG	4
4.1.1.1	Requirements	5
4.1.2	Remote and local multicast subscription	5
4.1.2.1	Requirements	6
4.1.3	Dual subscription to multicast groups during handover	6
4.1.3.1	Requirements	7
4.2	Applicability to multicast source mobility	7
4.2.1	Support of remote and direct subscription in basic source mobility	7
4.2.1.1	Requirements	8
4.2.2	Direct communication between source and listener associated with distinct LMAs but on the same MAG	8
4.2.3.1	Requirements	9
4.2.3	Route optimization support in source mobility for remote subscribers	9
4.2.3.1	Requirements	9
4.3	Summary of the requirements needed	10
5	Functional specification of an MLD proxy with multiple interfaces	12
6	Security Considerations	12
7	IANA Considerations	12
8	Conclusions	12
9	Acknowledgements	12
10	References	12

10.1	Normative References	12
10.2	Informative References	13
	Authors' Addresses	13

1 Introduction

The aim of this document is to specify the functionality that an MLD proxy with multiple upstream interfaces should have in order to support a set of different scenarios of applicability that have been identified on the MULTIMOB working group. Such functional specification is required to ensure the compatibility of the MLD proxy instances deployed in PMIPv6 domains.

To do that, a set of requirements are firstly identified to satisfy the different scenarios where an MLD proxy instance with multiple upstream interfaces can be potentially applied.

2. Terminology

<To be completed>.

3. Problem statement

The concept of MLD proxy with several upstream interfaces has emerged within the MULTIMOB working group as a way of optimizing (and in some cases enabling) service delivery scenarios in both multicast listener and source mobility cases.

Since those scenarios can motivate distinct needs in terms of MLD proxy functionality, it is necessary to consider a comprehensive approach, looking at the possible scenarios, and establishing a minimum set of requirements which can allow the operation of a versatile MLD proxy with multiple upstream interfaces as a common entity to all of them (i.e., no different kinds of proxies depending on the scenario, but a common proxy applicable to all the potential scenarios).

4. Scenarios of applicability

The use of an MLD proxy supporting multiple upstream interfaces can improve the performance and the scalability of multicast-capable PMIPv6 domains.

4.1 Applicability to multicast listener mobility

Three sub-cases can be identified for the multicast listener mobility.

4.1.1 Single MLD proxy instance on MAG

The base solution for multicast service in PMIPv6 [2] assumes that any MN subscribed to multicast services receive the multicast traffic through the associated LMA, as in the unicast case. As standard MLD proxy functionality only supports one upstream interface, the MAG should implement several separated MLD proxy instances, one per LMA, in order to serve the multicast traffic to the MNs, according to any particular LMA-MN association.

A way of avoiding the multiplicity of MLD proxy instance in a MAG is to deploy a unique MLD proxy instance with multiple upstream interfaces, one per LMA, without any change in the multicast traffic distribution.

4.1.1.1 Requirements

These are the requirements identified so far:

- The MLD proxy should be able of delivering the multicast control messages sent by the MNs to the associated LMA.
- The MLD proxy should be able of delivering the multicast control messages sent by each of the connected LMAs to the corresponding MN.
- The MLD proxy should be able of routing the multicast data coming from different LMAs to the corresponding MNs according to the MN to LMA association.
- The MLD proxy should be able of maintaining a 1:1 association between an MN and LMA (or downstream to upstream).

4.1.2 Remote and local multicast subscription

Standard MLD proxy definition, with a unique upstream interface per proxy, does not allow the reception of multicast traffic from distinct upstream multicast routers. In other words, all the multicast traffic being sent to the MLD proxy in downstream traverses a concrete, unique router before reaching the MAG. There are, however, situations where different multicast content could reach the MLD proxy through distinct next-hop routers.

For instance, the solution adopted to avoid the tunnel convergence problem in basic multicast PMIPv6 deployments [3] considers the possibility of subscription to a multicast source local to the PMIPv6 domain. In that situation, some multicast content will be accesses remotely, through the home network via the multicast tree mobility anchor, while some other multicast content will reach the proxy

directly, via a local router in the domain.

4.1.2.1 Requirements

These are the requirements identified so far:

- The MLD proxy should be able of delivering the multicast control messages sent by the MNs to the associated upstream interface based on the location of the source, remote or local, for a certain multicast group.
- The MLD proxy should be able of delivering the multicast control messages sent either local or remotely to the corresponding MNs.
- The MLD proxy should be able of routing the multicast data coming from different upstream interfaces to a certain MN according to the MN subscription, either local or remote. Note that it is assumed that a multicast group can be subscribed either locally or remotely, but not simultaneously. However more than one subscription could happen, being local or remote independently.
- The MLD proxy should be able of maintaining a 1:N association between an MN and the remote and local multicast router (or downstream to upstream).
- The MLD proxy should be able of switching between local or remote subscription for per multicast group according to specific configuration parameters (out of the scope of this document).

4.1.3 Dual subscription to multicast groups during handover

In the event of an MN handover, once an MN moves from a previous MAG (pMAG) to a new MAG (nMAG), the nMAG needs to set up the multicast status for the incoming MN, and subscribe the multicast channels it was receiving before the handover event. The MN will then experience a certain delay until it receives again the subscribed content.

A generic solution is being defined in [4] to speed up the knowledge of the ongoing subscription by the nMAG. However, for the particular case that the underlying radio access technology supports layer-2 triggers (thus requiring extra capabilities on the mobile node), there could be inter-MAG cooperation for handover support if pMAG and nMAG are known in advance.

This could be the case, for instance for those contents not already arriving to the nMAG, where the nMAG temporally subscribes the multicast groups of the ongoing MN's subscription via the pMAG, while the multicast delivery tree among the nMAG and the mobility anchor is

being established.

A similar approach is followed in [5] despite the solution proposed there differs from this approach (i.e., there is no consideration of an MLD proxy with multiple interfaces).

4.1.3.1 Requirements

These are the requirements identified so far:

- The MLD proxy should be able of delivering the multicast control messages sent by the MNs to the associated upstream interface based on the handover specific moment, for a certain multicast group.
- The MLD proxy should be able of delivering the multicast control messages sent either from pMAG or the multicast anchor to the corresponding MNs, based on the handover specific moment.
- The MLD proxy should be able of handle the incoming packet flows from the two simultaneous upstream interfaces, in order to not duplicate traffic delivered on the point-to-point link to the MN.
- The MLD proxy should be able of maintaining a 1:N association between an MN and both the remote multicast router and the pMAG (or downstream to upstream).
- The MLD proxy should be able of switching between local or remote subscription for all the multicast groups (from pMAG to multicast anchor) according to specific configuration parameters (out of the scope of this document).

4.2 Applicability to multicast source mobility

A couple of sub-cases can be identified for the multicast source mobility.

4.2.1 Support of remote and direct subscription in basic source mobility

In the basic case of source mobility, the multicast source is connected to one of the downstream interfaces of an MLD proxy. According to the standard specification [1] every packet sent by the multicast source will be forwarded towards the root of the multicast tree.

However, linked to the mobility listener problem, there could be the case of simultaneous remote subscribers, subscribing to the multicast

content through the home network, and local subscribers, requesting the contents directly via a multicast router residing on the same PMIPv6 domain where the source is attached to.

Then, in order to provide the co-existence of both types of subscribers, an MLD proxy with two upstream interfaces could simultaneously serve all kind of multicast subscribers.

Basic source mobility is being defined in [6] but the solution proposed there does not allow simultaneous co-existence of remote and local subscribers (i.e., the content sent by the source is either distributed locally to a multicast router in the PMIPv6 domain, or remotely by using the bi-directional tunnel towards the mobility anchor, but not both simultaneously).

4.2.1.1 Requirements

These are the requirements identified so far:

- The MLD proxy should be able of forwarding (replicating) the multicast content to both upstream interfaces, in case of simultaneous remote and local distribution.
- The MLD proxy should be able of handling control information incoming through any of the two upstream interfaces, providing the expected behavior for each of the multicast trees.
- The MLD proxy should be able of routing the multicast data towards different upstream interfaces for both remote and local subscriptions that could happen simultaneously.
- The MLD proxy should be able of maintaining a 1:N association between an MN and both the remote and local multicast router (or downstream to upstream).

4.2.2 Direct communication between source and listener associated with distinct LMAs but on the same MAG

In a certain PMIPv6 domain can be MNs associated to distinct LMAs using the same MAG to get access to their corresponding home networks. For multicast communication, according to the base solution [2], each MN <-> LMA association implies a distinct MLD proxy instance to be invoked in the MAG.

In these conditions, when a mobile source is serving multicast content to a mobile listener, both attached to the same MAG but each of them associated to different LMAs, the multicast flow must

traverse the PMIPv6 domain from the MAG to the LMA where the source maintains an association, then from that LMA to the LMA where the listener is associated to, and finally come back to the same MAG from where the flow departed. This routing is extremely inefficient.

An MLD proxy with multiple upstream interfaces avoids this behavior since it allows to invoke a unique MLD proxy instance in the MAG. In this case, the multicast source can directly communicate with the multicast listener, without need for delivering the multicast traffic to the LMAs.

4.2.3.1 Requirements

These are the requirements identified so far:

- The MLD proxy should be able of forwarding (replicating) the multicast content to different upstream or downstream interfaces where subscribers are present.
- The MLD proxy should be able of handling control information incoming through any of the upstream or downstream interfaces requesting a multicast flow being injected in another downstream interface.
- The MLD proxy should be able of maintaining a 1:N association between an MN and any of the upstream or downstream interfaces demanding the multicast content.

4.2.3 Route optimization support in source mobility for remote subscribers

Even in a scenario of remote subscription, there could be the case where both the source and the listener are attached to the same PMIPv6-Domain (for instance, no possibility of direct routing within the PMIPv6, or source and listener pertaining to distinct home networks). In this situation there is a possibility of route optimization if inter-MAG communication is enabled, in such a way that the listeners in the PMIPv6 domain are served through the tunnels between MAGs, while the rest of remote listeners are served through the mobility anchor.

A multi-upstream MLD proxy would allow the simultaneous delivery of traffic to such kind of remote listeners.

A similar route optimization approach is proposed in [7].

4.2.3.1 Requirements

These are the requirements identified so far:

- The MLD proxy should be able of forwarding (replicating) the multicast content to both kinds of upstream interfaces, inter-MAG tunnel interfaces and MAG to mobility anchor tunnel interface.
- The MLD proxy should be able of handling control information incoming through any of the two types of upstream interfaces, providing the expected behavior for each of the multicast trees (e.g., no forwarding traffic on one inter-MAG link once there are not more listeners requesting the content).
- The MLD proxy should be able of routing the multicast data towards different upstream interfaces for both remote and route optimized subscriptions that could happen simultaneously.
- The MLD proxy should be able of maintaining a 1:N association between an MN and both the remote and local MAGs (or downstream to upstream).

4.3 Summary of the requirements needed

After the previous analysis, a number of different requirements can be identified by the MLD proxy to support multiple upstream interfaces. The following table summarizes these requirements.

	Scenarios					
	Multicast Listener			Multicast Source		
	Single MLD Proxy (4.1.1)	Remote & local subscr. (4.1.2)	Dual subscr. in HO (4.1.3)	Direct & remote subscr. (4.2.1)	Listener & source on MAG (4.2.2)	Route optimi. (4.2.3)
Upstream Control Delivery	X	X	X	X	X	X
Downstr. Control Delivery	X	X	X		X	
Upstream Data Delivery				X		X
Downstr. Data Delivery	X	X	X		X	
1:1 MN to upstream assoc.	X					
1:N MN to upstream assoc.		X	X	X	X	X
Upstr i/f selection per group		X				
Upstr i/f selection all group			X			
Upstream traffic replicat.				X		X

Table I. Functionality needed on MLD proxy with multiple upstream interfaces per application scenario

5 Functional specification of an MLD proxy with multiple interfaces

<To be completed>.

6 Security Considerations

<To be completed>.

7 IANA Considerations

<IANA considerations text>.

8 Conclusions

Through this document several scenarios of applicability of an MLD proxy with multiple upstream interfaces have been presented.

<To be completed>.

9 Acknowledgements

The authors thank Stig Venaas for his valuable comments and suggestions.

The research of Carlos J. Bernardos leading to these results has received funding from the European Community's Seventh Framework Programme (FP7-ICT-2009-5) under grant agreement n. 258053 (MEDIEVAL project), being also partially supported by the Ministry of Science and Innovation (MICINN) of Spain under the QUARTET project (TIN2009-13992-C02-01).

10 References

10.1 Normative References

- [1] B. Fenner, H. He, B. Haberman, and H. Sandick, "Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")", RFC 4605, August 2006.

10.2 Informative References

- [2] T.C. Schmidt, M. Waehlich, and S. Krishnan, "A Minimal Deployment Option for Multicast Listeners in PMIPv6 Domains", RFC6224, April 2011.

- [3] J.C. Zuniga, L.M. Contreras, C.J. Bernardos, S. Jeon, Y. Kim, "Multicast Mobility Routing Optimizations for Proxy Mobile IPv6", work in progress, draft-ietf-multimob-pmipv6-ropt-01, September 2012.
- [4] L.M. Contreras, C.J. Bernardos, I. Soto, "PMIPv6 multicast handover optimization by the Subscription Information Acquisition through the LMA (SIAL)", work in progress, draft-ietf-multimob-fast-handover-01, July 2012.
- [5] T.C. Schmidt, M. Waehlich, R. Koodli, G. Fairhurst, "Multicast Listener Extensions for MIPv6 and PMIPv6 Fast Handovers", work in progress, draft-schmidt-multimob-fmipv6-pfmipv6-multicast-06, May 2012
- [6] T.C. Schmidt, S. Gao, H. Zhang, M. Waehlich, "Mobile Multicast Sender Support in Proxy Mobile IPv6 (PMIPv6) Domains", work in progress, draft-ietf-multimob-pmipv6-source-01, July 2012.
- [7] J. Liu, W. Luo, "Routes Optimization for Multicast Sender in Proxy Mobile IPv6 Domain", work in progress, draft-liu-multimob-pmipv6-multicast-ro-02, July 2012.

Authors' Addresses

Luis M. Contreras
Telefonica I+D
EMail: lmcm@tid.es

Carlos J. Bernardos
Universidad Carlos III de Madrid
EMail: cjbc@it.uc3m.es

PIM Working Group
Internet-Draft
Intended status: Informational
Expires: October 12, 2012

H. Asaeda
Keio University
N. Leymann
Deutsche Telekom AG
April 10, 2012

IGMP/MLD-Based Explicit Membership Tracking Function for Multicast
Routers
draft-ietf-pim-explicit-tracking-01

Abstract

This document describes the IGMP/MLD-based explicit membership tracking function for multicast routers. The explicit tracking function is useful for accounting and contributes to saving network resource and fast leaves (i.e. shortened leave latency).

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on October 12, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as

described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Explicit Tracking Function	4
2.1. Reducing the Number of Specific Queries	4
2.2. Shortening Leave Latencies	5
2.3. Considerations	5
3. Membership State Information	6
4. Multicast Router Behavior	7
5. Interoperability and Compatibility	8
6. Security Considerations	8
7. Acknowledgements	8
8. References	9
8.1. Normative References	9
8.2. Informative References	9
Authors' Addresses	9

1. Introduction

The Internet Group Management Protocol (IGMP) [1] for IPv4 and the Multicast Listener Discovery Protocol (MLD) [2] for IPv6 are the standard protocols used by listener hosts and multicast routers. When a host starts listening particular multicast channels, it sends IGMP/MLD State-Change Report messages specifying the corresponding channel information as the join/leave request to its upstream router (i.e., an adjacent multicast router or IGMP/MLD proxy [4]). This "unsolicited" Report is sent only once upon reception.

IGMP/MLD are non-reliable protocols; the unsolicited Report messages may be lost or not be reached to upstream routers. To recover the problem, the routers need to update membership information by sending IGMP/MLD General Query messages periodically. Member hosts then reply with "solicited" Report messages whenever they receive the Query messages.

Multicast routers are able to periodically maintain the multicast listener (or membership) state of downstream hosts attached on the same link by getting unsolicited Report messages and synchronize the actual membership state within the General Query timer interval (i.e., [Query Interval] value defined in [1][2].) However, this approach does not guarantee that the membership state is always perfectly synchronized. To minimize the possibility of having the outdated membership information, routers may shorten the periodic General Query timer interval. Unfortunately, this would increase the number of transmitted solicited Report messages and induce network congestion. And the more the network congestion is occurred, the more IGMP/MLD Report messages may be lost and the membership state information may be outdated in the router.

The IGMPv3 [1] and MLDv2 [2] protocols can provide the capability of keeping track of downstream (adjacent) multicast listener state to multicast routers. This document describes the "IGMP/MLD-based explicit member tracking function" for multicast routers and details the way for routers to implement the function. By enabling the explicit tracking function, routers can keep track of the downstream multicast membership state. This function implements the following requirements:

- o Per-host accounting
- o Reducing the number of transmitted Query and Report messages
- o Shortening leave latencies

- o Maintaining multicast channel characteristics (or statistics)

where this document mainly focuses on the above second and third bullets in the following sections.

The explicit tracking function does not change message formats used by the standard IGMPv3 [1] and MLDv2 [2], and their lightweight version protocols [3]. It does not change a multicast data sender's and receiver's behavior as well.

2. Explicit Tracking Function

2.1. Reducing the Number of Specific Queries

The explicit tracking function reduces the number of Group-Specific or Group-and-Source Specific Query messages transmitted from a router, and then the number of Current-State Report messages transmitted from member hosts. As the result, network resources used for IGMP/MLD query-and-reply communications between a router and member hosts can be saved.

According to [1] and [2], whenever a router receives the State-Change Report, it sends the corresponding Group-Specific or Group-and-Source Specific Query messages to confirm whether the Report sender is the last member host or not. After getting these Query messages, all member hosts joining the corresponding channel reply with own Current-State Report messages. This condition requires transmitting a number of Current-State Report messages and consumes network resources especially when many hosts have been joining the same channel.

On the other hand, if a router enables the explicit tracking function, it does not need to always ask Current-State Report message transmission to the member hosts whenever it receives the State-Change Report. This is because the explicit tracking function works with the expectation that the State-Change Report sender is the last remaining member of the channel. Even if this expectation is wrong (i.e., the State-Change Report sender was not the sole member), other members remaining in the same channel will reply with identical Report messages, so the end result is the same and no problem occurs. (Section 3 details the point.)

In addition, the processing of IGMP membership or MLD listener reports consumes CPU resources on the IGMP/MLD querier devices itself. Therefore, the explicit tracking function reduces not only the network load but also the CPU load on the querier devices as well.

snooping switch [5]. If the timer to refresh membership record on snooping switch is shorter than the General Query timer interval (i.e. [Query Interval]),

2.2. Shortening Leave Latencies

The explicit tracking function works with the expectation that the State-Change Report sender is the last remaining member of the channel. Thanks to this functionality, a router can tune timers and values related to decide that the State-Change Report sender was the sole member.

The [Last Member Query Interval] (LMQI) and [Last Listener Query Interval] (LLQI) values specify the maximum time allowed before sending a responding Report. The [Last Member Query Count] (LMQC) and [Last Listener Query Count] (LLQC) are the number of Group-Specific Queries or Group-and-Source Specific Queries sent before the router assumes there are no local members. The [Last Member Query Time] (LMQT) and [Last Listener Query Time] (LLQT) values are the total time the router should wait for a report, after the Querier has sent the first query.

The default values for LMQI/LLQI defined in the standard specifications [1][2] are 1 second. For the router enabling the explicit tracking function, LMQI/LLQI would be set to 1 second or shorter. The LMQC/LLQC may be set to "1" for the router, whereas their default values are the [Robustness Variable] value whose default value is "2". Smaller LMQC/LLQC give smaller LMQT/LLQT; this condition shortens the leave latencies.

2.3. Considerations

As with the basic concepts of IGMP and MLD, the explicit tracking function does not guarantee the membership state is always perfectly synchronized; routers enabling the explicit tracking function still need to send IGMPv3/MLDv2 Query messages and inquire solicited IGMPv3/MLDv2 Report messages from downstream members to maintain downstream membership state.

- o IGMP/MLD messages are non-reliable and may be lost in the transmission, therefore routers need to confirm the membership by sending Query messages.
- o To preserve compatibility with older versions of IGMP/MLD, routers need to support downstream hosts that are not upgraded to the latest versions of IGMP/MLD and run the report suppression mechanism.

- o It is impossible to identify hosts when hosts send IGMP reports with a source address of 0.0.0.0.

Regarding the last bullet, the IGMPv3 specification [1] mentions that an IGMPv3 Report is usually sent with a valid IP source address, although it permits that a host uses the 0.0.0.0 source address (as it happens that the host has not yet acquired an IP address), and routers MUST accept a report with a source address of 0.0.0.0. The MLDv2 specification [2] mentions that an MLDv2 Report MUST be sent with a valid IPv6 link-local source address, although an MLDv2 Report can be sent with the unspecified address (::), if the sending interface has not acquired a valid link-local address yet. [2] also mentions that routers silently discard a message that is not sent with a valid link-local address or sent with the unspecified address, without taking any action, because of the security consideration.

Another concern is that the explicit tracking function requires additional processing capability and a possibly large memory for routers to keep all membership states. Especially when a router needs to maintain a large number of member hosts, this resource requirement may be potentially-impacted. Operators may decide to disable this function when their routers do not have enough memory resources.

3. Membership State Information

The explicit tracking function is implemented with the following membership state information:

(S, G, number of receivers, (receiver records))

where each receiver record is of the form:

(IGMP/MLD Membership/Listener Report sender's address)

This state information must work properly when a receiver (i.e., Report sender) sends the same Report messages multiple times.

In the state information, each "S" and "G" indicates a single IPv4/IPv6 address. "S" is set to "Null" for an Any-Source Multicast (ASM) communication (i.e., (*,G) join reception). In order to simplify the implementation, the explicit tracking function does not keep the state of (S,G) join with EXCLUDE filter mode. If a router receives (S,G) join/leave request with EXCLUDE filter mode from the downstream hosts, it translates the join/leave request to (*,G) join state/leave request and records the state and the receivers' addresses into the maintained membership state information. Note that this membership

state translation does not change the routing protocol behavior; the routing protocol must deal with the original join/leave request and translate the request only for the membership state information.

4. Multicast Router Behavior

The explicit tracking function makes routers expect whether the State-Change Report sender is the last remaining member of the channel. Therefore the router transmits a corresponding Current-State Report message only when the router thinks that the State-Change Report sender is the last remaining member of the channel. This contributes to saving the network resources and also shortening leave latency.

To synchronize the membership state information, when a multicast router receives a Current-State or State-Change Report message, it adds the receiver IP address to or delete from the receiver records or creates the corresponding membership state information. If there are no more receiver records left, the membership state information is deleted from the router.

However, the membership state information may be still outdated in the router. It may be happened especially in a mobile multicast environment that some member hosts have joined to or left from the network without sending State-Change Report messages. Or, some State-Change Report messages are lost due to network congestion. Therefore, the router enabling the explicit tracking function ought to send the periodic General Query regularly.

Regarding the leave latency, as specified in Section 2.2, the explicit tracking function contributes to the fast leave by setting LMQUI/LLQI to "1" second or shorter and LMQC/LLQC to "1". However, if LMQC/LLQC is configured "2" or bigger value, then the router's behavior may be changed from the standard specification. According to [1] and [2], a router sends a Group- (and-Source) Specific Query [LMQC - 1] or [LLQC - 1] times when it receives State Change Report message (e.g. leave request) from a member host, in order to confirm whether or not the host is the only remaining member. However, this document RECOMMENDS that if the router enabling the explicit tracking function receives the corresponding Current State Report before the Specific Query retransmission, it cancels sending the same Specific Query for other [LMQC - 1] or [LLQC - 1] times.

Note that there is some risk that a router misses or loses Report messages sent from remaining members if the router adopts small LMQC/LLQC; however the wrong expectation would be lower happened for the router enabling the explicit tracking function. And to avoid the

problem, a router can start sending a Group- (and-Source) Specific Query message when it expects the number of the remaining members is small, such as 5, but not 0.

5. Interoperability and Compatibility

The explicit tracking function does not work with the older versions of IGMP or MLD, IGMPv1 [6], IGMPv2 [7] or MLDv1 [8], because a member host using these protocols adopts a report suppression mechanism by which a host would cancel sending a pending membership Reports if a similar Report was observed from another member on the network.

If a multicast router enabling the explicit tracking function changes its compatibility mode to the older versions of IGMP or MLD, the router should turn off the explicit tracking function but should not flush the maintained membership state information (i.e., keep the current membership state information as is). When the router changes back to IGMPv3 or MLDv2 mode, it would resume the function with the kept membership state information, even if the state information is outdated. This manner would give "smooth state transition" that does not initiate the membership state from scratch and synchronizes the actual membership state smoothly.

There are several points TBD in the further discussions regarding the interoperability and compatibility issues. At first, it is necessary whether a multicast router enabling the explicit tracking function needs to detect adjacent routers that do not support the explicit tracking function on the link or not. After the clarification, this document will describe the method how to detect them. It would be done by a new signaling message, but the new message leads compatibility problems for older routers or other routing protocols such as PIM-DM. All of these discussions are TBD.

6. Security Considerations

There is no additional security considerations.

7. Acknowledgements

Toerless Eckert, Stig Venaas, and others provided many constructive and insightful comments.

8. References

8.1. Normative References

- [1] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.
- [2] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.
- [3] Liu, H., Cao, W., and H. Asaeda, "Lightweight Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Version 2 (MLDv2) Protocols", RFC 5790, February 2010.

8.2. Informative References

- [4] Fenner, B., He, H., Haberman, B., and H. Sandick, "Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")", RFC 4605, August 2006.
- [5] Christensen, M., Kimball, K., and F. Solensky, "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches", RFC 4541, May 2006.
- [6] Deering, S., "Host Extensions for IP Multicasting", RFC 1112, August 1989.
- [7] Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2373, July 1997.
- [8] Deering, S., Fenner, W., and B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, October 1999.

Authors' Addresses

Hitoshi Asaeda
Keio University
Graduate School of Media and Governance
5322 Endo
Fujisawa, Kanagawa 252-0882
Japan

Email: asaeda@wide.ad.jp
URI: <http://web.sfc.wide.ad.jp/~asaeda/>

Nicolai Leymann
Deutsche Telekom AG
Winterfeldtstrasse 21-27
Berlin 10781
Germany

Email: n.leymann@telekom.de

Multimob Working Group
Internet-Draft
Intended status: Informational
Expires: January 13, 2013

H. Liu
M. McBride
Huawei Technologies
July 12, 2012

IGMP/MLD Optimizations in Wireless and Mobile Networks
draft-liu-multimob-igmp-mld-wireless-mobile-02

Abstract

This document proposes a variety of optimization approaches for IGMP and MLD in wireless and mobile networks. It aims to provide useful guideline to allow efficient multicast communication in these networks using IGMP or MLD protocols.

Requirements Language

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].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on January 13, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Requirements	3
2.1. Characteristics of Wireless and Mobile Multicast	3
2.2. Wireless Link Model	4
2.3. Requirements on IGMP and MLD	5
3. IGMP/MLD Optimization for Wireless and Mobile Networks	6
3.1. Switching Between Unicast and Multicast Queries	6
3.2. General Query Supplemented with Unicast Query	6
3.3. Retransmission of Queries	7
3.4. General Query Suppression	7
3.5. Tuning Response Delay According to Link Type and Status	8
3.6. Triggering Reports and Queries Quickly During Handover	9
4. Applicability and Interoperability Considerations	9
5. IANA Considerations	10
6. Security Considerations	10
7. Acknowledgements	10
8. Normative References	10
Authors' Addresses	11

1. Introduction

With the wide deployment of various wireless access techniques and the tendency to support video applications on wireless networks, wireless and mobile multicast come to attract more and more interests from content and service providers, but still face great challenges when considering dynamic group membership management under constant update of delivery path introduced by node movement, and high probability of loss and congestion due to limited reliability and capacity of wireless links.

Multicast network is generally constructed by IGMP and MLD group management protocol (respectively for IPv4 and IPv6 networks) to track valid receivers and by multicast routing protocol to build multicast delivery paths. This document focuses only on IGMP and MLD, which are used by a host to subscribe a multicast group and are most possibly to be exposed to wireless link to support terminal mobility. As IGMP and MLD were designed for fixed users using wired link, they do not necessarily work well for different wireless link types and mobile scenarios, thus should be considered to be enhanced to be more applicable in these environments.

This memo proposes a variety of optimizations for IGMP and MLD in wireless and mobile networks to improve network performance, with minimum changes on the protocol behavior and without introducing interoperability issues. These solutions can also be applied in wired network when efficiency or reliability is required.

For generality, this memo does not put limitations on the type of wireless techniques running below IGMP or MLD. They could be cellular, WiMAX, WiFi and etc, and are modeled as different abstract link models as described in section 2.2. Even though some of them (such as WiFi) have multicast limitations, it is probable that IGMP/MLD is enabled on the wireless terminal and multicast is supported across the network. The mobile IP protocol adopted on the core side, upstream from the access router, could be PMIP, MIPv4, or MIPv6.

2. Requirements

2.1. Characteristics of Wireless and Mobile Multicast

Several limitations should be considered when supporting IP multicast in wireless and mobile networks, including:

O Limited link bandwidth: wireless link usually has limited bandwidth, and the situation will be made even worse if high volume video multicast data has to be carried. Also the bandwidth available

in the upstream and downstream directions may be asymmetrical.

O High loss rate: wireless link usually has packet loss ranging from 1% to 30% according to different links types and conditions. Also when packets have to travel between home and access networks (e.g. through tunnel), they are prone to loss if the two networks are distant from each other.

O Frequent membership change: in fixed multicast, membership change only happens when a user leaves or joins a group, while in mobile scenario membership may also change when a user changes its location.

O Prone to performance degradation: the possible increased interaction of protocols across layers for mobility management, and the limitation of link capacity, may lead to network performance degradation and even to complete connection loss.

O Increased Leave Latency: the leave latency in mobile multicast might be increased due to user movement, especially if the traffic has to be transmitted between access and home networks, or if there is a handshake between networks.

2.2. Wireless Link Model

Wireless links can be categorized by their different transmission modes into three typical models: point-to-point (PTP), point-to-multipoint (PTMP), and broadcast link models.

In PTP model, one link is dedicated for two communication facilities. For multicast transmission, each PTP link normally has only one receiver and the bandwidth is dedicated for that receiver. Such link model may be implemented by running PPP on the link or having separate VLAN assignment for each receiver. In mobile network, tunnel between entities of home and foreign networks should be recognized as a PTP link.

PTMP is the model for multipoint transmission wherein there is one centralized transmitter and multiple distributed receivers. PTMP provides common downlink channels for all receivers and dedicated uplink channel for each receiver. Bandwidth downstream is shared by all receivers on the same link.

Broadcast link can connect two or more nodes and supports broadcast transmission. It is quite similar to fixed Ethernet link model and its link resource is shared in both uplink and downlink directions.

2.3. Requirements on IGMP and MLD

IGMP and MLD are usually run between mobile or wireless terminals and their first-hop access routers (i.e. home or foreign routers) to subscribe an IP multicast channel. Currently the version in-use includes IGMPv2 [RFC2236] and its IPv6 counterpart MLDv1 [RFC2710], IGMPv3 [RFC3376] and its IPv6 counterpart MLDv2 [RFC3810], and LW-IGMPv3/MLDv2 [RFC5790]. All these versions have basic group management capability required by a multicast subscription. The differences lie in that IGMPv2 and MLDv1 can only join and leave a non-source-specific group, while IGMPv3 and MLDv2 can select including and excluding specific sources for their join and leave operation, and LW-IGMPv3/MLDv2 simplifies IGMPv3/MLDv2 procedures by discarding excluding-source function. Among these versions, (LW-) IGMPv3/MLDv2 has the capability of explicitly tracking each host member.

From the illustration given in section 2.1 and 2.2, it is desirable for IGMP and MLD to have the following characteristics when used in wireless and mobile networks:

- o Adaptive to link conditions: wireless network has various link types, each with different bandwidth and performance features. IGMP or MLD should be able to be adaptive to different link model and link conditions to optimize its protocol operation.
- o Minimal group join/leave latency: because mobility and handover may cause a user to join and leave a multicast group frequently, fast join and leave by the user helps to accelerate service activation and to release unnecessary resources quickly to optimize resource utilization.
- o Robust to packet loss: the unreliable packet transmission due to instable wireless link conditions and limited bandwidth, or long distance transmission in mobile network put more strict robustness requirement on delivery of IGMP and MLD protocol messages.
- o Reducing packet exchange: wireless link resources are usually more limited, precious, and congested compared to their wired counterpart. This requires packet exchange be minimized without degrading protocol performance.
- o Packet burst avoidance: large number of packets generated in a short time interval may have the tendency to deteriorate wireless network conditions. IGMP and MLD should be optimized to reduce the probability of packet burst.

3. IGMP/MLD Optimization for Wireless and Mobile Networks

This section introduces several optimization methods for IGMP and MLD in wireless or mobile environment. The aim is to meet the requirements described in section 2.3. It should be noted that because an enhancement in one direction might result in weakening effect in another, balances should be taken cautiously to realize overall performance elevation.

3.1. Switching Between Unicast and Multicast Queries

IGMP/MLD protocol uses multicast Queries whose destinations are multicast addresses and also allows use of unicast Query with unicast destination to be sent only to one host. Unicast Query has the advantage of not affecting other hosts on the same link, and is desirable for wireless communication because a mobile terminal often has limited battery power [RFC6636]. But if the number of valid receivers is large, using unicast Query for each receiver is inefficient because large number of Unicast Queries have to be generated, in which situation normal multicast Query will be a good choice because only one General Query is needed. If the number of receivers to be queried is small, unicast Query is advantageous over the multicast one.

More flexibly, the router can choose to switch between unicast and multicast Queries according to the practical network conditions. For example, if the receiver number is small, the router could send unicast Queries respectively to each receiver, without arousing other non-member terminal which is in dormant state. When the receiver number reaches a predefined level, the router could change to use multicast Queries. To have the knowledge of the number of the valid receivers, a router is required to enable explicit tracking, and because Group-Specific Query and Group-and-Source-Specific Query are usually not used under explicit tracking [RFC6636], the switching operation mostly applies to General Queries.

3.2. General Query Supplemented with Unicast Query

Unicast Query also can be used in assistance to General Query to improve the robustness of solicited reports when General Query fails to collect all of its valid members. It requires the explicit tracking to be enabled and can be used when a router after sending a periodical General Query collects successfully most of the valid members' responses while losing some of which are still valid in its database. This may be because these reports are not generated or generated but lost for some unknown reasons. The router could choose to unicast a Query respectively to each non-respondent valid receiver to check whether they are still alive for the multicast reception,

without affecting the majority of receivers that have already responded. Unicast Queries under this condition could be sent at the end of the [Maximum Response Delay] after posting a General Query, and be retransmitted for [Last Member Query Count] times, at an interval of [Last Member Query Interval].

3.3. Retransmission of Queries

In IGMP and MLD, apart from the continuously periodical transmission, General Query is also transmitted during a router's startup. It is transmitted for [Startup Query Count] times by [Startup Query Interval]. There are some other cases where retransmission of General Query is beneficial which are not covered by current IGMP and MLD protocols as shown as following.

For example, a router which keeps track of all its active receivers, if after sending a General Query, fails to get any response from the receivers which are still valid in its membership database. This may be because all the responses of the receivers happen to be lost, or the sent Query does not arrive at the other side of the link to the receivers. The router could compensate this situation by retransmitting the General Query to solicit its active members. The retransmission can also be applied to Group-Specific or Group-and-Source-Specific Query on a router without explicit tracking capability, when these Specific Queries cannot collect valid response, to prevent missing valid members caused by lost Queries and Reports.

The above compensating Queries could be sent [Last Member Query Count] times, at the interval of [Last Member Query Interval], if the router cannot get any feedback from the receivers.

3.4. General Query Suppression

In IGMP and MLD, General Query is sent periodically and continuously without any limitation. It helps soliciting the state of current valid member but has to be processed by all hosts on the link, whether they are valid multicast receivers or not. When there is no receiver, the transmission of the General Query is a waste of resources for both the host and the router.

An IGMP/MLD router could suppress its transmission of General Query if it knows there is no valid multicast receiver on an interface, e.g. in the following cases:

O When the last member reports its leave for a group. This could be judged by an explicit tracking router checking its membership database, or by a non-explicit-tracking router getting no response

after sending Group-Specific or Group-and-Source-Specific Query.

O When the only member on a PTP link reports its leaving

O When a router after retransmitting General Queries on startup fails to get any response

O When a router previously has valid members but fails to get any response after several rounds of General Queries.

In these cases the router could make the decision that no member is on the interface and totally stop its transmission of periodical General Queries. If afterwards there is any valid member joins a group, the router could resume the original cycle of general Querying. Because General Query has influences on all hosts on a link, suppressing it when it is not needed is beneficial for both the link efficiency and terminal power saving.

3.5. Tuning Response Delay According to Link Type and Status

IGMP and MLD use delayed response to spread unsolicited Reports from different hosts to reduce possibility of packet burst. This is implemented by a host responding to a Query in a specific time randomly chosen between 0 and [Maximum Response Delay], the latter of which is determined by the router and is carried in Query messages to inform the hosts for calculation of the response delay. A larger value will lessen the burst better but will increase leave latency (the time taken to cease the traffic flowing after the last member requests the escaping of a channel).

In order to avoid message burst and reduce leave latency, the Response Delay may be dynamically calculated based on the expected number of responders, and link type and status, as shown in the following:

O If the expected number of reporters is large and link condition is bad, longer Maximum Response Delay is recommended; if the expected number of reporters is small and the link condition is good, smaller Maximum response Delay should be set.

o If the link type is PTP, the Maximum Response Delay can be chosen smaller, whereas if the link is PTMP or broadcast medium, the Maximum Response Delay can be configured larger.

The Maximum Response Delay could be configured by the administrator as mentioned above, or be calculated automatically by a software tool implemented according to experiential model for different link modes. The measures to determine the instant value of Maximum Response Delay

are out of this document's scope.

3.6. Triggering Reports and Queries Quickly During Handover

When a mobile terminal is moving from one network to another, if it is receiving multicast content, its new access network should try to deliver the content to the receiver without disruption or performance deterioration. In order to implement smooth handover between networks, the terminal's membership should be acquired as quickly as possible by the new access network.

An access router could trigger a Query to a terminal as soon as it detects the terminal's attaching on its link. This could be a General Query if the number of the entering terminals is not small (e.g when they are simultaneously in a moving train). Or this Query could also be a unicast Query for this incoming terminal to prevent unnecessary action of other terminals in the switching area.

For the terminal, it could send a report immediately if it is currently in the multicast reception state, when it begins to connect the new network. This helps establishing more quickly the membership state and enable faster multicast stream injection, because with the active report the router does not need to wait for the query period to acquire the terminal's newest state.

4. Applicability and Interoperability Considerations

Among the optimizations listed above, 'Switching between unicast and multicast Queries'(3.1) and 'General Query Supplemented with Unicast Query'(3.2) require a router to know beforehand the valid members connected through an interface, thus require explicit tracking capability. An IGMP/MLD implementation could choose any combination of the methods listed from 3.1 to 3.6 to optimize multicast communication on a specific wireless or mobile network.

For example, an explicit-tracking IGMPv3 router, can switch to unicast General Queries if the number of members on a link is small (3.1), can trigger unicast Query to a previously valid receiver if failing to get expected responses from it (3.2), can retransmit a General Query if after the previous one cannot collect reports from all valid members (3.3), and can stop sending a General Query when the last member leaves the group (3.4), and etc.

For interoperability, it is required if multiple multicast routers are connected to the same network for redundancy, each router are configured with the same optimization policy to synchronize the membership states among the routers.

5. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

6. Security Considerations

Since the methods only involve the tuning of protocol behavior by e.g. retransmission, changing delay parameter, or other compensating operations, they do not introduce additional security weaknesses. The security considerations described in [RFC2236], [RFC3376], [RFC2710] and [RFC3810] can be reused. And to achieve some security level in insecure wireless network, it is possible to take stronger security procedures during IGMP/MLD message exchange, which are out of the scope of this memo.

7. Acknowledgements

The authors would like to thank Qin Wu, Stig Venaas, Gorrry Fairhurst, Thomas C. Schmidt, Marshall Eubanks, Suresh Krishnan, J. William Atwood, WeeSan Lee, Imed Romdhani, Hitoshi Asaeda, Liu Yisong and Wei Yong for their valuable comments and suggestions on this document.

8. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2236] Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, November 1997.
- [RFC2710] Deering, S., Fenner, W., and B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, October 1999.
- [RFC3376] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.
- [RFC3810] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.
- [RFC5790] Liu, H., Cao, W., and H. Asaeda, "Lightweight Internet

Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Version 2 (MLDv2) Protocols", RFC 5790, February 2010.

[RFC6636] Asaeda, H., Liu, H., and Q. Wu, "Tuning the Behavior of the Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) for Routers in Mobile and Wireless Networks", RFC 6636, May 2012.

Authors' Addresses

Hui Liu
Huawei Technologies
Building Q14, No.156, Beiqing Rd.
Beijing 100095
China

Email: helen.liu@huawei.com

Mike McBride
Huawei Technologies
2330 Central Expressway
Santa Clara CA 95050
USA

Email: michael.mcbride@huawei.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 18, 2013

W. Zhou
Cisco Systems
October 15, 2012

VRRP PIM Interoperability
draft-zhou-pim-vrrp-00.txt

Abstract

This document introduces VRRP Aware PIM, a redundancy mechanism for the Protocol Independent Multicast (PIM) to interoperate with Virtual Router Redundancy Protocol (VRRP). It allows PIM to track VRRP state and to preserve multicast traffic upon failover in a redundant network with virtual routing groups enabled.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on April 18, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Tracking and Failover	4
3. PIM Assert Metric Auto-Adjustment	5
4. DF Election for BiDir Group	6
5. Tracking Multiple VRRP Groups on an Interface	7
6. Support of HSRP	8
7. Security Considerations	9
8. Acknowledgments	10
9. Informative References	11
Author's Address	12

1. Introduction

Virtual Router Redundancy Protocol (VRRP) [RFC5798] is a redundancy protocol for establishing a fault-tolerant default gateway. The protocol establishes a framework between network devices in order to achieve default gateway failover if the primary gateway becomes inaccessible .

PIM has no inherent redundancy capabilities and its operation is completely independent of VRRP group states. As a result, IP multicast traffic is forwarded not necessarily by the same device as is elected by VRRP. The VRRP Aware PIM feature provides consistent IP multicast forwarding in a redundant network with virtual routing groups enabled.

In a multi-access segment (such as LAN), PIM designated router (DR) election is unaware of the redundancy configuration, and the elected DR and VRRP master router (MR) may not be the same router. In order to ensure that the PIM DR is always able to forward PIM Join/Prune message towards RP or FHR, the VRRP MR becomes the PIM DR (if there is only one VRRP group). PIM is responsible for adjusting DR priority based on the group state. When a failover occurs, multicast states are created on the new MR elected by the VRRP group and the MR assumes responsibility for the routing and forwarding of all the traffic addressed to the VRRP virtual IP address. This ensures the PIM DR runs on the same gateway as the VRRP MR and maintains mroute states. It enables multicast traffic to be forwarded through the VRRP MR, allowing PIM to leverage VRRP redundancy, avoid potential duplicate traffic, and enable failover, depending on the VRRP states in the device.

2. Tracking and Failover

With VRRP Aware PIM enabled, PIM listens to the state change notifications from VRRP and automatically adjusts the priority of the PIM DR based on the VRRP state, and ensures VRRP MR (if there is only one VRRP group) becomes the DR of the LAN. If there are multiple VRRP groups, the DR is determined by user-configured priority.

PIM triggers communication between upstream and downstream devices upon failover in order to create mroute states on the new MR. PIM sends additional PIM Hello message using the VRRP virtual IP addresses as the source address for each active VRRP group when a device becomes VRRP Active. The PIM Hello will carry a new GenID in order to trigger other routers to respond to the failover. When a downstream device receives this PIM Hello, it will add the virtual address to its PIM neighbor list. The new GenID carried in the PIM Hello will trigger downstream routers to resend PIM Join messages towards the virtual address. Upstream routers will process PIM Join/Prunes (J/P) based on VRRP group state.

If the J/P destination matches the VRRP group virtual address and if the destination device is in VRRP active state, the new MR processes the PIM Join because it is now the acting PIM DR. This allows all PIM Join/Prunes to reach the VRRP group virtual address and minimizes changes and configurations at the downstream routers side.

3. PIM Assert Metric Auto-Adjustment

It is possible that, after VRRP active switched from A to B; A is still forwarding multicast traffic which will result in duplicate traffic and PIM Assert mechanism will kick in. PIM Assert with redundancy is enabled.

- o If only one VRRP group, passive routers will send a large penalty metric preference (PIM_ASSERT_INFINITY - 1) and make MR the Assert winner.
- o If there are multiples VRRP groups configured on an interface, Assert metric preference will be (PIM_ASSERT_INFINITY - 1) if and only if all VRRP groups are in passive.
- o If there is at least one VRRP group is in Active, then original Assert metric preference will be used. That is, winner will be selected between routers using their real Assert metric preference with at least one active VRRP Group, just like no VRRP is involved.

4. DF Election for BiDir Group

Change to DF offer/winner metric is handled similarly to PIM Assert handling with VRRP.

- o If only one VRRP group, passive routers will send a large penalty metric preference in Offer (`PIM_BIDIR_INFINITY_PREF- 1`) and make MR the DF winner.
- o If there are multiples VRRP groups configured on an interface, Offer metric preference will be (`PIM_BIDIR_INFINITY_PREF- 1`) if and only if all VRRP groups are in passive.
- o If there is at least one VRRP group is in Active, then original Offer metric preference to RP will be used. That is, winner will be selected between routers using their real Offer metric with at least one active VRRP Group, just like no VRRP is involved.

5. Tracking Multiple VRRP Groups on an Interface

User can configure PIM to track more than one VRRP groups on an interface. This allows other applications to exploit the PIM/VRRP interoperability to achieve various goals (e.g., load balancing). Since each VRRP groups configured on an interface could be in different states at any moment, the DR priority is adjusted. PIM Assert metric and PIM Bidir DF metric if and only if all VRRP groups configured on an interface are in passive (non-Active) states to ensure that interfaces with all-passive VRRP groups will not win in DR, Assert and DF election. In other words, DR, Assert, DF winner will be elected among the interfaces with at least one Active VRRP group.

6. Support of HSRP

Although there are differences between VRRP and Hot Standby Router Protocol (HSRP) [RFC2281] including number of backup (standby) routers, virtual IP address and timer intervals, the proposed scheme can also enable HSRP aware PIM with similar switchover and tracking mechanism described in this draft.

7. Security Considerations

The proposed tracking mechanism has no negative impact on security.

8. Acknowledgments

I would like to give a special thank you and appreciation to Stig Venaas for his ideas and comments in this draft.

9. Informative References

- [RFC2281] Li, T., Cole, B., Morton, P., and D. Li, "Cisco Hot Standby Router Protocol (HSRP)", RFC 2281, March 1998.
- [RFC5798] Nadas, S., "Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6", RFC 5798, March 2010.

Author's Address

Wei Zhou
cisco Systems
Tasman Drive
San Jose, CA 95134
USA

Email: weizho2@cisco.com

