Network Working Group
Request for Comments: 4605
Category: Standards Track

B. Fenner
AT&T Research
H. He
Nortel
B. Haberman
JHU-APL
H. Sandick
Little River Elementary School

August 2006

Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")

Status of This Memo

Copyright Notice

Abstract

In certain topologies, it is not necessary to run a multicast routing protocol. It is sufficient for a device to learn and proxy group membership information and simply forward multicast packets based upon that information. This document describes a mechanism for forwarding based solely upon Internet Group Management Protocol (IGMP) or Multicast Listener Discovery (MLD) membership information.

1. Introduction

This document applies spanning tree multicast routing [MCAST] to an Internet Group Management Protocol (IGMP) or Multicast Listener Discovery (MLD)-only environment. The topology is limited to a tree, since we specify no protocol to build a spanning tree over a more complex topology. The root of the tree is assumed to be connected to a wider multicast infrastructure.

1.1. Motivation

In a simple tree topology, it is not necessary to run a multicast routing protocol. It is sufficient to learn and proxy group membership information and simply forward multicast packets based upon that information. One typical example of such a tree topology can be found on an edge aggregation box such as a Digital Subscriber Line Access Multiplexer (DSLAM). In most deployment scenarios, an edge box has only one connection to the core network side and has many connections to the customer side.

Using IGMP/MLD-based forwarding to replicate multicast traffic on devices such as the edge boxes can greatly simplify the design and implementation of those devices. By not supporting more complicated multicast routing protocols such as Protocol Independent Multicast (PIM) or Distance Vector Multicast Routing Protocol (DVMRP), it reduces not only the cost of the devices but also the operational overhead. Another advantage is that it makes the proxy devices independent of the multicast routing protocol used by the core network routers. Hence, proxy devices can be easily deployed in any multicast network.

Robustness in an edge box is usually achieved by using a hot spare connection to the core network. When the first connection fails, the edge box fails over to the second connection. IGMP/MLD-based forwarding can benefit from such a mechanism and use the spare connection for its redundant or backup connection to multicast routers. When an edge box fails over to the second connection, the proxy upstream connection can also be updated to the new connection.

1.2. Applicability Statement

The IGMP/MLD-based forwarding only works in a simple tree topology. The tree must be manually configured by designating upstream and downstream interfaces on each proxy device. In addition, the IP addressing scheme applied to the proxying tree topology SHOULD be configured to ensure that a proxy device can win the IGMP/MLD Querier election to be able to forward multicast traffic. There are no other multicast routers except the proxy devices within the tree, and the root of the tree is expected to be connected to a wider multicast infrastructure. This protocol is limited to a single administrative domain.

In more complicated scenarios where the topology is not a tree, a more robust failover mechanism is desired, or more than one administrative domain is involved, a multicast routing protocol should be used.

1.3. Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

This document is a product of the Multicast & Anycast Group Membership (MAGMA) working group within the Internet Engineering Task Force. Comments are solicited and should be addressed to the working group's mailing list at magma@ietf.org and/or the authors.

2. Definitions

2.1. Upstream Interface

A proxy device's interface in the direction of the root of the tree. Also called the "Host interface".

2.2. Downstream Interface

Each of a proxy device's interfaces that is not in the direction of the root of the tree. Also called the "Router interfaces".

2.3. Group Mode

In IPv4 environments, for each multicast group, a group is in IGMP version 1 (IGMPv1) [RFC1112] mode if an IGMPv1 report is heard. A group is in IGMP version 2 (IGMPv2) [RFC2236] mode if an IGMPv2 report is heard but no IGMPv1 report is heard. A group is in IGMP version 3 (IGMPv3) [RFC3376] mode if an IGMPv3 report is heard but no IGMPv1 or IGMPv2 report is heard.

In IPv6 environments, for each multicast group, a group is in MLD version 1 (MLDv1) [RFC2710] mode if an MLDv1 report is heard. MLDv1 is equivalent to IGMPv2. A group is in MLD version 2 (MLDv2) [MLDv2] mode if an MLDv2 report is heard but no MLDv1 report is heard. MLDv2 is equivalent to IGMPv3.

2.4. Subscription

When a group is in IGMPv1 or IGMPv2/MLDv1 mode, the subscription is a group membership on an interface. When a group is in IGMPv3/MLDv2 mode, the subscription is an IGMPv3/MLDv2 state entry, i.e., a (multicast address, group timer, filter-mode, source-element list) tuple, on an interface.

2.5. Membership Database

The database maintained at each proxy device into which the membership information of each of its downstream interfaces is merged. The membership database is a set of membership records of the form:

(multicast-address, filter-mode, source-list)

Please refer to IGMPv3/MLDv2 [RFC3376, MLDv2] specifications for the definition of the fields "filter-mode" and "source-list". The operational behaviors of the membership database is defined in section 4.1.

3. Abstract Protocol Definition

A proxy device performing IGMP/MLD-based forwarding has a single upstream interface and one or more downstream interfaces. These designations are explicitly configured; there is no protocol to determine what type each interface is. It performs the router portion of the IGMP [RFC1112, RFC2236, RFC3376] or MLD [RFC2710, MLDv2] 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.

The proxy device maintains a database consisting of the merger of all subscriptions on any downstream interface. Refer to Section 4 for the details about the construction and maintenance of the membership database.

The proxy device sends IGMP/MLD membership reports on the upstream interface when queried and sends unsolicited reports or leaves when the database changes.

When the proxy device receives a packet destined for a multicast group (channel in Source-Specific Multicast (SSM)), it uses a list consisting of the upstream interface and any downstream interface that has a subscription pertaining to this packet and on which it is the IGMP/MLD Querier. This list may be built dynamically or cached. It removes the interface on which this packet arrived from the list and forwards the packet to the remaining interfaces (this may include the upstream interface).

Note that the rule that a proxy device must be the querier in order to forward packets restricts the IP addressing scheme used; in particular, the IGMP/MLD-based forwarding devices must be given the lowest IP addresses of any potential IGMP/MLD Querier on the link, in order to win the IGMP/MLD Querier election. IGMP/MLD Querier election rule defines that the Querier that has the lowest IP address wins the election. (The IGMP/MLD Querier election rule is defined in IGMP/MLD specifications as part of the IGMP/MLD behavior.) So in an IGMP/MLD-based forwarding-only environment, if non-proxy device wins the IGMP/MLD Querier election, no packets will flow.

For example, the figure below shows an IGMP/MLD-based forwarding-only environment:

```
        LAN 1  --------------------------------------
               Upstream |              | Upstream
                        A(non-proxy)   B(proxy)
             Downstream |(lowest IP)   | Downstream
        LAN 2  --------------------------------------
```

Device A has the lowest IP address on LAN 2, but it is not a proxy device. According to IGMP/MLD Querier election rule, A will win the election on LAN 2 since it has the lowest IP address. Device B will not forward traffic to LAN 2 since it is not the querier on LAN 2.

The election of a single forwarding proxy is necessary to avoid local loops and redundant traffic for links that are considered downstream links by multiple IGMP/MLD-based forwarders. This rule "piggy-backs" forwarder election on IGMP/MLD Querier election. The use of the IGMP/MLD Querier election process to choose the forwarding proxy delivers similar functionality on the local link as the PIM Assert mechanism. On a link with only one IGMP/MLD-based forwarding device, this rule MAY be disabled (i.e., the device MAY be configured to forward packets to an interface on which it is not the querier). However, the default configuration MUST include the querier rule, for example, for redundancy purposes, as shown in the figure below:

```
        LAN 1  --------------------------------------
               Upstream |              | Upstream
                        A              B
             Downstream |              | Downstream
        LAN 2  --------------------------------------
```

LAN 2 can have two proxy devices, A and B. In such a configuration, one proxy device must be elected to forward the packets. This document requires that the forwarder must be the IGMP/MLD querier. So proxy device A will forward packets to LAN 2 only if A is the querier. In the above figure, if A is the only proxy device, A can be configured to forward packets even though B is the querier.

Note that this does not protect against an "upstream loop". For example, see the figure below:

```
        LAN 1  ------------------------------------
                Upstream |              | Downstream
                         A              B
              Downstream |              | Upstream
        LAN 2  ------------------------------------
```

B will unconditionally forward packets from LAN 1 to LAN 2, and A will unconditionally forward packets from LAN 2 to LAN 1. This will cause an upstream loop. A multicast routing protocol that employs a tree building algorithm is required to resolve loops like this.

3.1. Topology Restrictions

This specification describes a protocol that works only in a simple tree topology. The tree must be manually configured by designating upstream and downstream interfaces on each proxy device, and the root of the tree is expected to be connected to a wider multicast infrastructure.

3.2. Supporting Senders

In order for senders to send from inside the proxy tree, all traffic is forwarded towards the root. The multicast router(s) connected to the wider multicast infrastructure should be configured to treat all systems inside the proxy tree as though they were directly connected; e.g., for Protocol Independent Multicast - Sparse Mode (PIM-SM) [PIM-SM], these routers should Register-encapsulate traffic from new sources within the proxy tree just as they would directly-connected sources.

This information is likely to be manually configured; IGMP/MLD-based multicast forwarding provides no way for the routers upstream of the proxy tree to know what networks are connected to the proxy tree. If the proxy topology is congruent with some routing topology, this information MAY be learned from the routing protocol running on the topology; e.g., a router may be configured to treat multicast packets from all prefixes learned from routing protocol X via interface Y as though they were from a directly connected system.

4. Proxy Device Behavior

This section describes an IGMP/MLD-based multicast forwarding device's actions in more detail.

4.1. Membership Database

The proxy device performs the router portion of the IGMP/MLD protocol on each downstream interface. For each interface, the version of IGMP/MLD used is explicitly configured and defaults to the highest version supported by the system.

The output of this protocol is a set of subscriptions; this set is maintained separately on each downstream interface. In addition, the subscriptions on each downstream interface are merged into the membership database.

The membership database is a set of membership records of the form:

(multicast-address, filter-mode, source-list)

Each record is the result of the merge of all subscriptions for that record's multicast-address on downstream interfaces. If some subscriptions are IGMPv1 or IGMPv2/MLDv1 subscriptions, these subscriptions are converted to IGMPv3/MLDv2 subscriptions. The IGMPv3/MLDv2 and the converted subscriptions are first preprocessed to remove the timers in the subscriptions and, if the filter mode is EXCLUDE, to remove every source whose source timer > 0. Then the preprocessed subscriptions are merged using the merging rules for multiple memberships on a single interface (specified in Section 3.2 of the IGMPv3 specification [RFC3376] and in Section 4.2 of the MLDv2 specification [MLDv2]) to create the membership record. For example, there are two downstream interfaces, I1 and I2, that have subscriptions for multicast address G. I1 has an IGMPv2/MLDv1 subscription that is (G). I2 has an IGMPv3/MLDv2 subscription that has membership information (G, INCLUDE, (S1, S2)). The I1's subscription is converted to an IGMPv3/MLDv2 subscription that has membership information (G, EXCLUDE, NULL). Then the subscriptions are preprocessed and merged, and the final membership record is (G, EXCLUDE, NULL).

The proxy device performs the host portion of the IGMP/MLD protocol on the upstream interface. If there is an IGMPv1 or IGMPv2/MLDv1 querier on the upstream network, then the proxy device will perform IGMPv1 or IGMPv2/MLDv1 on the upstream interface accordingly. Otherwise, it will perform IGMPv3/MLDv2.

If the proxy device performs IGMPv3/MLDv2 on the upstream interface, then when the composition of the membership database changes, the change in the database is reported on the upstream interface as though this proxy device were a host performing the action. If the proxy device performs IGMPv1 or IGMPv2/MLDv1 on the upstream interface, then when the membership records are created or deleted, the changes are reported on the upstream interface. All other changes are ignored. When the proxy device reports using IGMPv1 or IGMPv2/MLDv1, only the multicast address field in the membership record is used.

4.2. Forwarding Packets

A proxy device forwards packets received on its upstream interface to each downstream interface based upon the downstream interface's subscriptions and whether or not this proxy device is the IGMP/MLD Querier on each interface. A proxy device forwards packets received on any downstream interface to the upstream interface, and to each downstream interface other than the incoming interface based upon the downstream interfaces' subscriptions and whether or not this proxy device is the IGMP/MLD Querier on each interface. A proxy device MAY use a forwarding cache in order not to make this decision for each packet, but MUST update the cache using these rules any time any of the information used to build it changes.

4.3. SSM Considerations

To support Source-Specific Multicast (SSM), the proxy device should be compliant with the specification about using IGMPv3 for SSM [SSM]. Note that the proxy device should be compliant with both the IGMP Host Requirement and the IGMP Router Requirement for SSM since it performs IGMP Host Portion on the upstream interface and IGMP Router Portion on each downstream interface.

An interface can be configured to perform IGMPv1 or IGMPv2. In this scenario, the SSM semantic will not be maintained for that interface. However, a proxy device that supports this document should ignore those IGMPv1 or IGMPv2 subscriptions sent to SSM addresses. And more importantly, the packets with source-specific addresses SHOULD NOT be forwarded to interfaces with IGMPv2 or IGMPv1 subscriptions for these addresses.

5. Security Considerations

Since only the Querier forwards packets, the IGMP/MLD Querier election process may lead to black holes if a non-forwarder is elected Querier. An attacker on a downstream LAN can cause itself to be elected Querier, and as a result, no packets would be forwarded.

However, there are some operational ways to avoid this problem. It is fairly common for an operator to number the routers starting from the bottom of the subnet. So an operator SHOULD assign the subnet's lowest IP address(es) to a proxy (proxies) in order for the proxy (proxies) to win the querier election.

IGMP/MLD-based forwarding does not provide the "upstream loop" detection mechanism described in Section 3. Hence, to avoid the problems caused by an "upstream loop", it MUST be administratively assured that such loops don't exist when deploying IGMP/MLD Proxying.

The IGMP/MLD information presented by the proxy to its upstream routers is the aggregation of all its downstream group membership information. Any access control applied on the group membership protocol at the upstream router cannot be performed on a single subscriber. That is, the access control will apply equally to all the interested hosts reachable via the proxy device.

6. Acknowledgements

The authors would like to thank Erik Nordmark, Dave Thaler, Pekka Savola, Karen Kimball, and others for reviewing the specification and providing their comments.

7. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3376] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.

[RFC2236] Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, November 1997.

[RFC1112] Deering, S., "Host extensions for IP multicasting", STD 5, RFC 1112, August 1989.

[RFC2710] Deering, S., Fenner, W., and B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, October 1999.

[MLDv2] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.

[SSM] Holbrook, H., Cain, B., and B. Haberman, "Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source-Specific Multicast", RFC 4604, August 2006.

8. Informative References

[MCAST] Deering, S., "Multicast Routing in a Datagram Internetwork", Ph.D. Thesis, Stanford University, December 1991.

[PIM-SM] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006.

Authors' Addresses

Bill Fenner
AT&T Labs - Research
1 River Oaks Place
San Jose, CA 95134

Phone: +1 408 493-8505
EMail: fenner@research.att.com

Haixiang He
Nortel
600 Technology Park Drive
Billerica, MA 01821

EMail: haixiang@nortel.com

Brian Haberman
Johns Hopkins University Applied Physics Lab
11100 Johns Hopkins Road
Laurel, MD 20723-6099

EMail: brian@innovationslab.net

Hal Sandick
Little River Elementary School
2315 Snow Hill Road
Durham, NC 27712

EMail: sandick@nc.rr.com

Acknowledgement

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).