MBONED Working Group
Internet-Draft
Expires: December 31, 2007

H. Liu
W. Cao
Huawei Technologies Co., Ltd.
H. Asaeda
Keio University
June 29, 2007

Lightweight IGMPv3 and MLDv2 Protocols
draft-ietf-mboned-lightweight-igmpv3-mldv2-01

Status of this Memo

Copyright Notice

Abstract

This document describes lightweight IGMPv3 and MLDv2 protocols (LW-IGMPv3 and LW-MLDv2), which simplify the standard (full) versions of IGMPv3 and MLDv2. The interoperability with the full versions and the previous versions of IGMP and MLD is also taken into account.

Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT","SHOULD", "SHOULD NOT", "RECOMMENDED","MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [1].

Table of Contents

1. Introduction

IGMP version 3 [2] and MLD version 2 [3] implement source filtering capabilities that are not suported by their earlier versions, IGMPv1 [4], IGMPv2 [5] and MLDv1 [6]. An IGMPv3 or MLDv2 capable host can tell its upstream router which group it would like to join by specifying which sources it does or does not intend to receive multicast traffic from. IGMPv3 and MLDv2 add the capability for a multicast router to also learn which sources are of interest to neighboring systems, for packets sent to any particular multicast address.

The INCLUDE and EXCLUDE filter-modes are introduced to support the source filtering function. If a host wants to receive from specific sources, it sends an IGMPv3 or MLDv2 report with filter-mode set to INCLUDE. If the host does not want to receive from some sources, it sends a report with filter-mode set to EXCLUDE. A source list for the given sources shall be included in the report message.

INCLUDE and EXCLUDE filter modes are also defined in a multicast router to process the IGMPv3 or MLDv2 reports. When a multicast router receives the report messages from its downstream hosts, it forwards the corresponding multicast traffic by managing requested group and source addresses. Group timers and source timers are used to maintain the forwarding state of desired groups and sources under certain filter modes. When a group report arrives or a certain timer expires, a multicast router may update the desired or undesired source lists, reset related timer values, change filter mode, or trigger group queries. With all of the above factors correlating with each other, the determination rules become relatively complex, as the interface states could be frequently changed.

The multicast filter-mode improves the ability of the multicast receiver to express its desires. It is useful to support Source-Specific Multicast (SSM) [7] by specifying interesting source addresses with INCLUDE mode. However, practical applications do not use EXCLUDE mode to block sources very often, because a user or application usually wants to specify desired source addresses, not undesired source addresses. Even if a user wants to explicitly refuse traffic from some sources in a group, when other users in the same shared network have an interest in these sources, the corresponding multicast traffic is forwarded to the network.

This document proposes simplified versions of IGMPv3 and MLDv2, named Lightweight IGMPv3 and Lightweight MLDv2 (or LW-IGMPv3 and LW-MLDv2), in which EXCLUDE filter-mode is eliminated. Not only are LW-IGMPv3 and LW-MLDv2 compatible with the standard IGMPv3 and MLDv2, but also the protocol operations made by data receiver hosts and routers or

switches (performing IGMPv3/MLDv2 snooping) are simplified in the lightweight protocol, and complicated operations are hence effectively reduced. Since LW-IGMPv3 and LW-MLDv2 are fully compatible with the full version of these protocols (i.e., the standard IGMPv3 and MLDv2), hosts or routers that have implemented the full version do not need to implement or modify anything to cooperate with LW-IGMPv3/LW-MLDv2 hosts or routers.

2. Simplification Method Overview

The principle is to simplify the host and router parts as much as possible to improve efficiency, while guaranteeing interoperability with the full versions, and introducing no side effects on applications.

For convenience, this document mainly discusses IGMPv3, since MLDv2 inherits the same source filtering mechanism, but this document additionally shows MLDv2's unique specifications when needed.

2.1. Behavior of Group Members

In LW-IGMPv3, the same service interface model as that of IGMPv3 is inherited:

```
IPMulticastListen ( socket, interface, multicast-address,
                    filter-mode, source-list )
```

In the lightweight protocol, EXCLUDE mode on the host part is preserved only for EXCLUDE (*,G) join, which denotes a non-source-specific group report (as known as (*,G) join) and is equivalent to the group membership join triggered by IGMPv2/IGMPv1/MLDv1. The detailed host operation of LW-IGMPv3/LW-MLDv2 is described in Section 4.

2.2. Behavior of Multicast Routers

Router filter-mode is defined to optimize the state description of a group [2]. As a rule, once a member report is in EXCLUDE mode, the router filter-mode for the group will be set to EXCLUDE. When all systems cease sending EXCLUDE mode reports, the filter-mode for that group may transit back to INCLUDE mode. Group timer is used to identify such transition.

In LW-IGMPv3, hosts primarily send INCLUDE requests. The only exception is EXLUDE (*,G) join, which can be interpreted by the router as a request to include all sources. Without the more general form of EXCLUDE requests, it is unnecessary for the router to maintain the EXCLUDE filter-mode, and the state model for multicast router can be simplified as:

```
(multicast address, group timer, (source records))
```

Here a group timer is kept to represent (*,G) group join. Its basic behavior is: when a router receives a (*,G) group join, it will set its group timer and keep the source list for sources specified in the source records. When the group timer expires, the router may change

to the reception for the listed sources. The definition of the source record is the same as that of full version.

The elimination of the filter-mode will greatly simplify the router behavior, e.g. the action on reception of reports and the setting of the timers. The detailed operation of router operation is described in Section 4.

3. LW-IGMPv3 Protocol for Group Members

LW-IGMPv3 uses two sets of messages, i.e., Query and Report messages, being the same as the full version protocols. Although most of these message types and corresponding group records are inherited from the full version protocols, an operation that triggers EXCLUDE (S,G) join is omitted and the corresponding record types of the Report are modified on the lightweight protocols.

There are three Group Record Types defined in the full IGMPv3: Current-State Record noted by MODE_IS_INCLUDE (referred to as IS_IN) or MODE_IS_EXCLUDE (IS_EX), Filter-Mode-Change Record noted by CHANGE_TO_INCLUDE_MODE (TO_IN) or CHANGE_TO_EXCLUDE_MODE (TO_EX), and Source-List-Change Record noted by ALLOW_NEW_SOURCES (ALLOW) or BLOCK_OLD_SOURCES (BLOCK).

3.1. Action on Change of Interface State

When the state of an interface of a group member host is changed, a State-Change Report for that interface is immediately transmitted from that interface. The type and contents of the Group Record(s) in that Report are determined by comparing the filter mode and source list for the affected multicast address before and after the change. While the requirements are the same as the full version for the computation, in the lightweight version host, the interface state change rules are simplified due to the reduction of message types. The contents of the new transmitted report are calculated as follows (Group Record Types are described in Section 3.3):

Old State	New State	State-Change Record Sent
INCLUDE (A)	INCLUDE (B)	ALLOW(B-A), BLOCK(A-B)
INCLUDE (A)	EXCLUDE ()	TO_EX()
INCLUDE ()	EXCLUDE ()	TO_EX()
EXCLUDE ()	INCLUDE (B)	TO_IN(B)

To cover the possibility of the State-Change Report being missed by one or more multicast routers, it is retransmitted [Robustness Variable]-1 more times, at intervals chosen at random from the range (0, [Unsolicited Report Interval]). (These values are defined in [2][3].)

In the full version of IGMPv3, as was done with the first report, the interface state for the affected group before and after the latest

change is compared, and the report records expressing the difference are built and merged with the contents of the pending report, to create the new State-Change report. However, for the LW-IGMPv3 host, this merge operation is optional. If the LW-IGMPv3 host does not merge with the contents of the pending report, it transmits each report sequentially. Doing so can greatly simplified the operation for scheduling the reports.

3.2. Action on Reception of a Query

When a lightweight version host receives a Query, it does not respond immediately. Instead, it delays its response by a random amount of time, bounded by the Max Resp Time value derived from the Max Resp Code in the received Query message [2][3]. The system may receive a variety of Queries on different interfaces and of different kinds (e.g., General Queries, Group-Specific Queries, and Group-and-Source-Specific Queries), each of which may require its own delayed response.

Before scheduling a response to a Query, the system must first consider previously scheduled pending responses and in many cases schedule a combined response. Therefore, the lightweight version host must be able to maintain the following state:

- A timer per interface for scheduling responses to General Queries.
- A per-group and interface timer for scheduling responses to Group-Specific and Group-and-Source-Specific Queries.
- A per-group and interface list of sources to be reported in the response to a Group-and-Source-Specific Query.

LW-IGMPv3 inherits most of the rules that are used to determine if a Report needs to be scheduled from the full version. The difference is regarding the simplification of EXCLUDE filter-mode and the type of Report to schedule as detailed in Section 3.3.

While it is optional that a LW-IGMPv3 host merges with the contents of the pending report for unsolicited report (i.e., State-Change report) as mentioned in the previous section, if the received Query is a Group-and-Source-Specific Query and there is a pending response for this group with a non-empty source-list, then the group source list is augmented to contain the list of sources in the new Query and a single response is scheduled using the group timer as with the full version host. The new response is then scheduled to be sent at the earlier of the remaining time for the pending report and the selected delay.

3.3. Applicable Group Record Types

Among Group Record Types defined in the full IGMPv3, several record types are not used in LW-IGMPv3 as some of the processes related to the filter mode change to the EXCLUDE mode are eliminated and some of the report messages are converged with a record having null source address list. All of the record types of report messages used by the full and lightweight version protocols are shown as follows:

IGMPv3	LW-IGMPv3	Comments
IS_EX()	TO_EX()	Query response for (*,G) join
IS_EX(x)	N/A	Query response for EXCLUDE (x,G) join
IS_IN(x)	ALLOW(x)	Query response for INCLUDE (x,G) join
ALLOW(x)	ALLOW(x)	INCLUDE (x,G) join
BLOCK(x)	BLOCK(x)	INCLUDE (x,G) leave
TO_IN(x)	TO_IN(x)	Change to INCLUDE (x,G) join
TO_IN()	TO_IN()	(*,G) leave
TO_EX(x)	N/A	Change to EXCLUDE (x,G) join
TO_EX()	TO_EX()	(*,G) join

where "x" represents a non-null source address list and "()" represents null source address list. For instance, IS_EX() means a report whose record type is IS_EX with null source address list. "N/A" represents not applicable (or no use) because the corresponding operation should not occur in the lightweight version protocols.

LW-IGMPv3 does not use EXCLUDE filter-mode with a non-null source address list. A multicast router creates the same state when it receives a report message containing either IS_EX() or TO_EX() record types. Therefore, LW-IGMPv3 integrates the IS_EX() operation with the TO_EX() operation.

When a LW-IGMPv3 host needs to make a query response for the state of INCLUDE (x,G) join, it makes a response whose message type is expressed with ALLOW(x), instead of using the IS_IN record type. Because the router's processing of the two messages is completely same, the IS_IN(x) type is eliminated for simplification.

A LW-IGMPv3 host does not use EXCLUDE mode, while TO_IN record is used the following situation: the host first launches an application (AP1) that requests INCLUDE (x,G) join, and sends ALLOW(x). Then the host launches another application (AP2) that joins (*,G), and it sends TO_EX(). In this condition, when AP2 terminates but AP1 keeps working on the lightweight version host, the host sends a report with TO_IN(x) record type for [Robustness Variable] times.

4. LW-IGMPv3 Protocol for Multicast Routers

The major difference between the full and lightweight version protocols on the router part is that for the lightweight version filter-mode is discarded and the function of the group timer is redefined. The states maintained by the lightweight router are reduced and the protocol operation is greatly simplified.

4.1. Group Timers and Source Timers in the Lightweight Version

A source timer is kept for each source record and it is updated when the source is present in a received report. It indicates the validity of the sources and needs to be referred when the router takes its forwarding decision.

The group timer being used in the full version of IGMPv3 for transitioning the router's filter-mode from EXCLUDE to INCLUDE, is now redefined to identify the non-source-specific receiving states maintaining for (*,G) join. Once a group record of IS_EX() is received, the group timer is used to represent this (*,G) group join. The expiration of the group timer indicates that there are no listeners on the attached network for this (*,G) group. If there are unexpired sources (whose source timers are greater than zero), the router will change to receiving traffic for those sources. The role of the group timer can be summarized as follows:

Group Timer Value	Actions/Comments
G_Timer > 0	All members in this group.
G_Timer == 0	No more listeners to this (*,G) group. If all source timers have expired then delete group record. If there are still source record timers running, use those source records with running timers as the source record state.

The operation related to the group and source timers has some difference compared with the full IGMPv3. In the full version, if a source timer expires under the EXCLUDE router filter-mode, its corresponding source record is not deleted until the group timer expires for indicating undesired sources. In the lightweight version, since there is no need to keep such records for blocking specific sources, if a source timer expires, its source record should be deleted immediately, not waiting for the time-out of the group timer.

4.2. Source-Specific Forwarding Rules

A full version multicast router needs to consult IGMPv3 state information when it makes decisions on forwarding a datagram from a source or its upstream router to its attached network, based on the router filter-mode and source timer. In LW-IGMPv3, because of the absence of the router filter-mode, the group timer and source timer could be used for such decisions. The forwarding suggestion made by LW-IGMPv3 to the routing protocols is summarized as follows:

Group Timer	Source Timer	Action
G_Timer == 0	S_TIMER > 0	Suggest forwarding traffic from source
G_Timer == 0	S_TIMER == 0	Suggest stopping forwarding traffic from source and remove source record. If there are no more source records for the group, delete group record
G_Timer == 0	No Source Elements	Suggest not to forward traffic from the source
G_Timer > 0	S_TIMER >= 0	Suggest forwarding traffic from source
G_Timer > 0	No Source Elements	Suggest forwarding traffic from source

4.3. Reception of Current-State Records

When receiving Current-State Records, the LW-IGMPv3 router resets its group or source timers and updates its source list within the group. For source-specific group reception state (when G_Timer==0), the source list contains sources whose traffic will be forwarded by the router, while in non-source-specific group reception (when G_Timer>0), the source list remembers the valid sources to receive traffic from after toggling to source-specific reception state.

Although the Lightweight host only sends a subset of the message of that of the full version, the LW-router should be able to process as much messages as possible to be compatible with the full version host. The following table describes the action taken by a multicast router after receiving Current-State Records. The notations have the

same meaning as that in the full IGMPv3 protocol.

Group Timer	Old Source List	Report Rec'd	New Source List	Actions
G_Timer == 0	A	IS_IN(B)	A+B	(B)=GMI
G_Timer == 0	A	IS_EX()	A	G_Timer=GMI
G_Timer > 0	A	IS_IN(B)	A+B	(B)=GMI
G_Timer > 0	A	IS_EX()	A	G_Timer=GMI

The above table could be further simplified for the processes that are completely same for the two values of the G_Timer:

Old Source List	Report Rec'd	New Source List	Actions
A	IS_IN(B)	A+B	(B)=GMI
A	IS_EX()	A	G_Timer=GMI

Without EXCLUDE filter-mode, a router's process on receiving Current-State Record is simple: when a router receives an IS_IN report, it appends the reported source addresses to the previous source list with their source timers set to GMI. Upon receiving an IS_EX() report, the router sets the non-source-specific receiving states by resetting the group timer value and keeps the previous source list without modification.

4.4. Reception of Source-List-Change and Filter-Mode-Change Records

On receiving Source-List-Change and Filter-Mode-Change Records, the LW-IGMPv3 router needs to reset its group and source timers, update its source list within the group, or trigger group queries. The queries are sent by the router for the sources that are requested to be no longer forwarded to a group. The table below describes the state change and the actions that should be taken.

Group Timer	Old Source List	Report Rec'd	New Source List	Actions
G_Timer == 0	A	ALLOW(B)	A+B	(B)=GMI
G_Timer == 0	A	BLOCK(B)	A	Send Q(G,A*B)
G_Timer == 0	A	TO_IN(B)	A+B	(B)=GMI Send Q(G,A-B)
G_Timer > 0	A	ALLOW(B)	A+B	(B)=GMI
G_Timer > 0	A	BLOCK(B)	A	Send Q(G,A*B)
G_Timer > 0	A	TO_IN(B)	A+B	(B)=GMI SendQ(G,A-B) Send Q(G)

The table could be further simplified by merging duplicate lines:

Old Source List	Report Rec'd	New Source List	Actions
A	ALLOW(B)	A+B	(B)=GMI
A	BLOCK(B)	A	Send Q(G,A*B)
A	TO_IN(B)	A+B	(B)=GMI Send Q(G,A-B) If G_Timer>0 Send Q(G)

In this table, TO_EX() report is not included because the processing is exactly the same as that of IS_EX(), as described in the previous section. Section 5.1 gives the lightweight routers's transformation behavior between the two messages.

5. Interoperability

LW-IGMPv3/LW-MLDv2 hosts and routers should interoperate gracefully with the full version protocols [2][3]. Also, LW-IGMPv3/LW-MLDv2 hosts and routers should interoperate gracefully with hosts and routers running IGMPv1/v2 or MLDv1.

5.1. Interoperation with the Full Version of IGMPv3

LW-IGMPv3 does not introduce any change on the message format of the group query and report messages the full version protocols use. With the elimination of the EXLCLUDE filter mode, the LW-IGMPv3 group member sends a subset of IGMPv3 report messages, which can be recognized by a multicast router running the full or the lightweight IGMPv3 protocol on the same LAN.

A LW-IGMPv3 router does not process directly IS_EX(x) and TO_EX(x) records that are used by the full IGMPv3. When a LW-IGMPv3 router receives these report messages from the full version host, it translates them to IS_EX() records and behaves accordingly. All possible record types are compared as follows:

IGMPv3 Report	LW-IGMPv3 Equivalent
IS_IN(x)	IS_IN(x)
IS_EX(x)	IS_EX()
TO_IN(x)	TO_IN(x)
TO_EX(x)	IS_EX()
ALLOW(x)	ALLOW(x)
BLOCK(x)	BLOCK(x)

5.2. Interoperation with IGMPv1/IGMPv2

The LW-IGMPv3 protocol basically adopts the same Host/Group Compatibility Mode and keeps Querier Present timers for IGMPv1 and IGMPv2. Their definition and processing is the same as that of IGMPv3.

5.2.1. Behavior of Group Members

A host's compatibility mode is determined from the Host Compatibility Mode variable which can be in one of three states: IGMPv1, IGMPv2 or

IGMPv3. The Host Compatibility Mode of an interface is set to IGMPv2 and IGMPv2 Querier Present is set to Older Version Querier Present Timeout second (defined in [2]) whenever an IGMPv2 General Query is received on that interface. The Host Compatibility Mode of an interface is set to IGMPv1 and IGMPv1 Querier Present is set to Older Version Querier Present Timeout second whenever an IGMPv1 Membership Query is received on that interface. Based on the Host Compatibility Mode variable, a host acts using the IGMPv3, IGMPv2, or IGMPv1 protocol on that interface.

While above manner is inherited from the definition of [2], LW-IGMPv3 may enable to configure the Host Compatibility Mode variable by other means: when a LW-IGMPv3 host is placed on a link where there are IGMPv1/IGMPv2 hosts, a host may allow its IGMPv3 report message to be suppressed by an IGMPv1 or IGMPv2 report message.

5.2.2. Behavior of Multicast Routers

If a LW-IGMPv3 router is on a network where at least one router running IGMPv1 or IGMPv2 protocols, it is required that the lowest version of querier must be used. This can be administratively assured by supporting IGMPv1, IGMPv2 or IGMPv3 compatibility mode.

An LW-IGMPv3 router may be placed on a network where there are hosts that have not been upgraded to IGMPv3. In order to be compatible with the older version, the lightweight router should keep a Group compatibility mode for each group record, and IGMPv1 and IGMPv2 Host present timers are kept to switch gracefully between different versions of IGMP.

When Group Compatibility mode is IGMPv2 or IGMPv1, a LW-IGMPv3 router translates the following IGMPv2 or IGMPv1 messages for that group to their LW-IGMPv3 equivalents:

IGMP Message	LW-IGMPv3 Equivalent
v1 Report	IS_EX()
v2 Report	IS_EX()
v2 Leave	TO_IN()

6. Implementation Considerations

The lightweight protocols requires no additional procedure on the implementation of the related protocols or systems, e.g. IGMP/MLD snooping, multicast routing protocol, and operation of application sockets, while the processing loads on the switches and routers that running IGMPv3 (snooping) and multicast routing protocols may be greatly decreased.

In the following sections, the implementation related aspects are described for the lightweight version protocols.

6.1. Implementation of Source-Specific Multicast

[8] illustrates the requirements of implementation of Source-Specific Multicast (SSM) on IGMPv3/MLDv2 hosts and routers. The lightweight protocol does not impose any bad influences on an SSM application. The requirements of LW-IGMPv3/LW-MLDv2 for supporting SSM are illustrated below.

A LW-IGMPv3/LW-MLDv2 host should not send a non-source-specific join, i.e., IS_EX(), and IGMPv2 Leave and MLDv1 Done messages for the application whose multicast address is in the SSM address range. The reception of a non-source-specific join with an SSM group address should indicate an error to the application. The SSM-aware router will ignore IS_EX() reports with SSM addresses. Other types of Reports should be processed normally.

6.2. Implementation of Multicast Source Filter (MSF) APIs

Multicast Source Filter (MSF) APIs [9] defines (1) IPv4 Basic MSF API, (2) IPv4 Advanced MSF API, (3) Protocol-Independent Basic MSF API, and (4) Protocol-Independent Advanced MSF API.

According to the MSF APIs definition, a LW-IGMPv3 host should implement at least one of IPv4 Basic MSF API and Protocol-Independent Basic MSF API, and a LW-MLDv2 host should implement Protocol-Independent Basic MSF API. Other APIs, IPv4 Advanced MSF API and Protocol-Independent Advanced MSF API, are optional to implement in LW-IGMPv3/LW-MLDv2 host.

7. Security Considerations

The security considerations are the same as that of the full version of IGMPv3/MLDv2.

8. References

8.1. Normative References

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

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

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

[4] Deering, S., "Host Extensions for IP Multicasting", RFC 1112, August 1989.

[5] Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2373, July 1997.

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

[7] Holbrook, H. and B. Cain, "Source-Specific Multicast for IP", RFC 4607, August 2006.

[8] 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.

8.2. Informative References

[9] Thaler, D., Fenner, B., and B. Quinn, "Socket Interface Extensions for Multicast Source Filters", RFC 3678, January 2004.

Authors' Addresses

Hui Liu
Huawei Technologies Co., Ltd.
Huawei Bld., No.3 Xinxi Rd.
Shang-Di Information Industry Base
Hai-Dian Distinct, Beijing 100085
China

Email: Liuhui47967@huawei.com

Wei Cao
Huawei Technologies Co., Ltd.
Huawei Bld., No.3 Xinxi Rd.
Shang-Di Information Industry Base
Hai-Dian Distinct, Beijing 100085
China

Email: caowayne@huawei.com

Hitoshi Asaeda
Keio University
Graduate School of Media and Governance
5322 Endo
Fujisawa, Kanagawa 252-8520
Japan

Email: asaeda@wide.ad.jp

Full Copyright Statement

Intellectual Property

Acknowledgment

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