

Network Working Group
Request for Comments: 2710
Category: Standards Track

S. Deering
Cisco Systems
W. Fenner
AT&T Research
B. Haberman
IBM
October 1999

Multicast Listener Discovery (MLD) for IPv6

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (1999). All Rights Reserved.

Abstract

This document specifies the protocol used by an IPv6 router to discover the presence of multicast listeners (that is, nodes wishing to receive multicast packets) on its directly attached links, and to discover specifically which multicast addresses are of interest to those neighboring nodes. This protocol is referred to as Multicast Listener Discovery or MLD. MLD is derived from version 2 of IPv4's Internet Group Management Protocol, IGMPv2. One important difference to note is that MLD uses ICMPv6 (IP Protocol 58) message types, rather than IGMP (IP Protocol 2) message types.

1. Definitions

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 [KEYWORDS].

2. Introduction

The purpose of Multicast Listener Discovery (MLD) is to enable each IPv6 router to discover the presence of multicast listeners (that is, nodes wishing to receive multicast packets) on its directly attached links, and to discover specifically which multicast addresses are of interest to those neighboring nodes. This information is then

3.1. Type

There are three types of MLD messages:

Multicast Listener Query (Type = decimal 130)

There are two subtypes of Multicast Listener Query messages:

- General Query, used to learn which multicast addresses have listeners on an attached link.
- Multicast-Address-Specific Query, used to learn if a particular multicast address has any listeners on an attached link.

These two subtypes are differentiated by the contents of the Multicast Address field, as described in section 3.6.

Multicast Listener Report (Type = decimal 131)

Multicast Listener Done (Type = decimal 132)

In the rest of this document, the above messages types are referred to simply as "Query", "Report", and "Done".

3.2. Code

Initialized to zero by the sender; ignored by receivers.

3.3. Checksum

The standard ICMPv6 checksum, covering the entire MLD message plus a "pseudo-header" of IPv6 header fields [ICMPv6,IPv6].

3.4. Maximum Response Delay

The Maximum Response Delay field is meaningful only in Query messages, and specifies the maximum allowed delay before sending a responding Report, in units of milliseconds. In all other messages, it is set to zero by the sender and ignored by receivers.

Varying this value allows the routers to tune the "leave latency" (the time between the moment the last node on a link ceases listening to a particular multicast address and moment the routing protocol is notified that there are no longer any listeners for that address), as discussed in section 7.8. It also allows tuning of the burstiness of MLD traffic on a link, as discussed in section 7.3.

3.5. Reserved

Initialized to zero by the sender; ignored by receivers.

3.6. Multicast Address

In a Query message, the Multicast Address field is set to zero when sending a General Query, and set to a specific IPv6 multicast address when sending a Multicast-Address-Specific Query.

In a Report or Done message, the Multicast Address field holds a specific IPv6 multicast address to which the message sender is listening or is ceasing to listen, respectively.

3.7. Other fields

The length of a received MLD message is computed by taking the IPv6 Payload Length value and subtracting the length of any IPv6 extension headers present between the IPv6 header and the MLD message. If that length is greater than 24 octets, that indicates that there are other fields present beyond the fields described above, perhaps belonging to a future backwards-compatible version of MLD. An implementation of the version of MLD specified in this document MUST NOT send an MLD message longer than 24 octets and MUST ignore anything past the first 24 octets of a received MLD message. In all cases, the MLD checksum MUST be computed over the entire MLD message, not just the first 24 octets.

4. Protocol Description

Note that defaults for timer values are described later in this document. Timer and counter names appear in square brackets.

Routers use MLD to learn which multicast addresses have listeners on each of their attached links. Each router keeps a list, for each attached link, of which multicast addresses have listeners on that link, and a timer associated with each of those addresses. Note that the router needs to learn only that listeners for a given multicast address are present on a link; it does NOT need to learn the identity (e.g., unicast address) of those listeners or even how many listeners are present.

For each attached link, a router selects one of its link-local unicast addresses on that link to be used as the IPv6 Source Address in all MLD packets it transmits on that link.

For each interface over which the router is operating the MLD protocol, the router must configure that interface to listen to all link-layer multicast address that can be generated by IPv6 multicasts. For example, an Ethernet-attached router must set its Ethernet address reception filter to accept all Ethernet multicast addresses that start with the hexadecimal value 3333 [IPv6-ETHER]; in the case of an Ethernet interface that does not support the filtering of such a range of multicast address, it must be configured to accept ALL Ethernet multicast addresses, in order to meet the requirements of MLD.

With respect to each of its attached links, a router may assume one of two roles: Querier or Non-Querier. There is normally only one Querier per link. All routers start up as a Querier on each of their attached links. If a router hears a Query message whose IPv6 Source Address is numerically less than its own selected address for that link, it MUST become a Non-Querier on that link. If [Other Querier Present Interval] passes without receiving, from a particular attached link, any Queries from a router with an address less than its own, a router resumes the role of Querier on that link.

A Querier for a link periodically [Query Interval] sends a General Query on that link, to solicit reports of all multicast addresses of interest on that link. On startup, a router SHOULD send [Startup Query Count] General Queries spaced closely together [Startup Query Interval] on all attached links in order to quickly and reliably discover the presence of multicast listeners on those links.

General Queries are sent to the link-scope all-nodes multicast address (FF02::1), with a Multicast Address field of 0, and a Maximum Response Delay of [Query Response Interval].

When a node receives a General Query, it sets a delay timer for each multicast address to which it is listening on the interface from which it received the Query, EXCLUDING the link-scope all-nodes address and any multicast addresses of scope 0 (reserved) or 1 (node-local). Each timer is set to a different random value, using the highest clock granularity available on the node, selected from the range [0, Maximum Response Delay] with Maximum Response Delay as specified in the Query packet. If a timer for any address is already running, it is reset to the new random value only if the requested Maximum Response Delay is less than the remaining value of the running timer. If the Query packet specifies a Maximum Response Delay of zero, each timer is effectively set to zero, and the action specified below for timer expiration is performed immediately.

When a node receives a Multicast-Address-Specific Query, if it is listening to the queried Multicast Address on the interface from which the Query was received, it sets a delay timer for that address to a random value selected from the range [0, Maximum Response Delay], as above. If a timer for the address is already running, it is reset to the new random value only if the requested Maximum Response Delay is less than the remaining value of the running timer. If the Query packet specifies a Maximum Response Delay of zero, the timer is effectively set to zero, and the action specified below for timer expiration is performed immediately.

If a node's timer for a particular multicast address on a particular interface expires, the node transmits a Report to that address via that interface; the address being reported is carried in both the IPv6 Destination Address field and the MLD Multicast Address field of the Report packet. The IPv6 Hop Limit of 1 (as well as the presence of a link-local IPv6 Source Address) prevent the packet from traveling beyond the link to which the reporting interface is attached.

If a node receives another node's Report from an interface for a multicast address while it has a timer running for that same address on that interface, it stops its timer and does not send a Report for that address, thus suppressing duplicate reports on the link.

When a router receives a Report from a link, if the reported address is not already present in the router's list of multicast address having listeners on that link, the reported address is added to the list, its timer is set to [Multicast Listener Interval], and its appearance is made known to the router's multicast routing component. If a Report is received for a multicast address that is already present in the router's list, the timer for that address is reset to [Multicast Listener Interval]. If an address's timer expires, it is assumed that there are no longer any listeners for that address present on the link, so it is deleted from the list and its disappearance is made known to the multicast routing component.

When a node starts listening to a multicast address on an interface, it should immediately transmit an unsolicited Report for that address on that interface, in case it is the first listener on the link. To cover the possibility of the initial Report being lost or damaged, it is recommended that it be repeated once or twice after short delays [Unsolicited Report Interval]. (A simple way to accomplish this is to send the initial Report and then act as if a Multicast-Address-Specific Query was received for that address, and set a timer appropriately).

When a node ceases to listen to a multicast address on an interface, it SHOULD send a single Done message to the link-scope all-routers multicast address (FF02::2), carrying in its Multicast Address field the address to which it is ceasing to listen. If the node's most recent Report message was suppressed by hearing another Report message, it MAY send nothing, as it is highly likely that there is another listener for that address still present on the same link. If this optimization is implemented, it MUST be able to be turned off but SHOULD default to on.

When a router in Querier state receives a Done message from a link, if the Multicast Address identified in the message is present in the Querier's list of addresses having listeners on that link, the Querier sends [Last Listener Query Count] Multicast-Address-Specific Queries, one every [Last Listener Query Interval] to that multicast address. These Multicast-Address-Specific Queries have their Maximum Response Delay set to [Last Listener Query Interval]. If no Reports for the address are received from the link after the response delay of the last query has passed, the routers on the link assume that the address no longer has any listeners there; the address is therefore deleted from the list and its disappearance is made known to the multicast routing component. This process is continued to its resolution (i.e. until a Report is received or the last Multicast-Address-Specific Query is sent with no response) despite any transition from Querier to Non-Querier on this link.

Routers in Non-Querier state MUST ignore Done messages.

When a router in Non-Querier state receives a Multicast-Address-Specific Query, if its timer value for the identified multicast address is greater than [Last Listener Query Count] times the Maximum Response Delay specified in the message, it sets the address's timer to that latter value.

5. Node State Transition Diagram

Node behavior is more formally specified by the state transition diagram below. A node may be in one of three possible states with respect to any single IPv6 multicast address on any single interface:

- "Non-Listener" state, when the node is not listening to the address on the interface (i.e., no upper-layer protocol or application has requested reception of packets to that multicast address). This is the initial state for all multicast addresses on all interfaces; it requires no storage in the node.

- "Delaying Listener" state, when the node is listening to the address on the interface and has a report delay timer running for that address.
- "Idle Listener" state, when the node is listening to the address on the interface and does not have a report delay timer running for that address.

There are five significant events that can cause MLD state transitions:

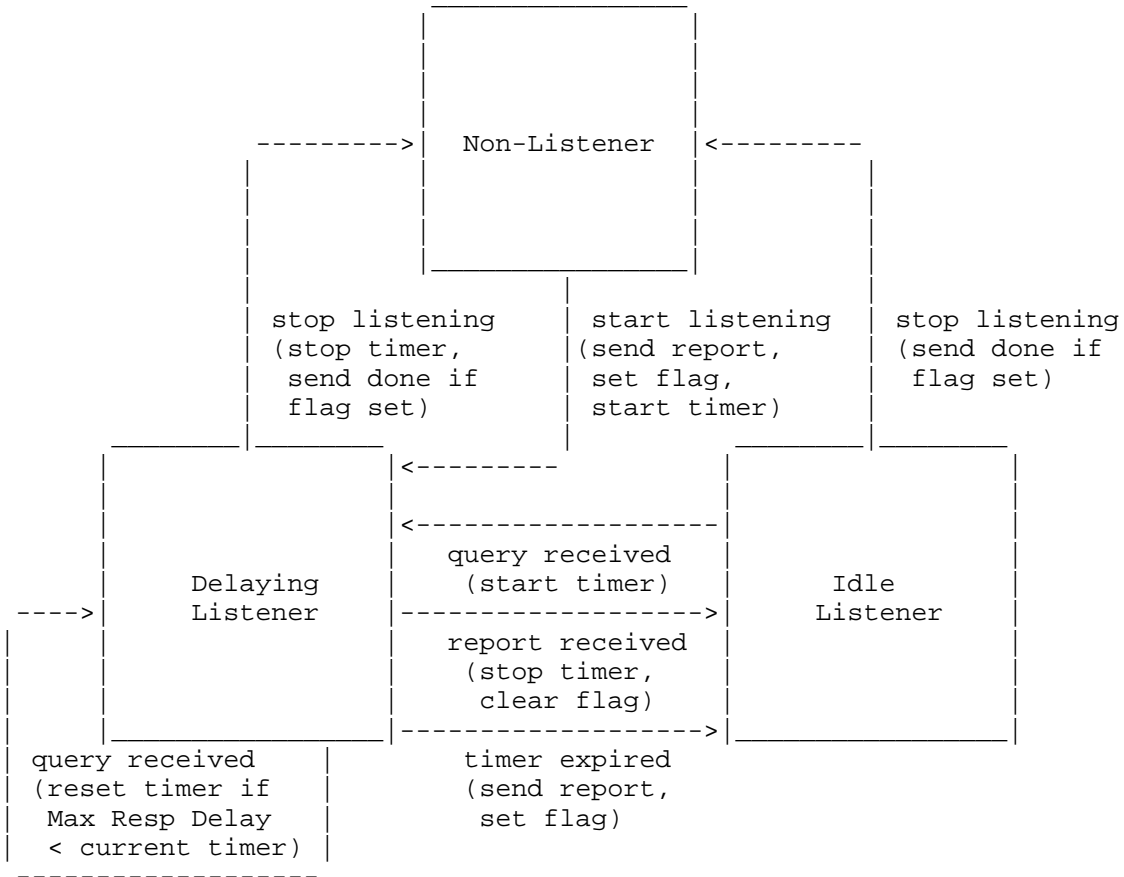
- "start listening" occurs when the node starts listening to the address on the interface. It may occur only in the Non-Listener state.
- "stop listening" occurs when the node stops listening to the address on the interface. It may occur only in the Delaying Listener and Idle Listener states.
- "query received" occurs when the node receives either a valid General Query message, or a valid Multicast-Address-Specific Query message. To be valid, the Query message MUST come from a link-local IPv6 Source Address, be at least 24 octets long, and have a correct MLD checksum. The Multicast Address field in the MLD message must contain either zero (a General Query) or a valid multicast address (a Multicast-Address-Specific Query). A General Query applies to all multicast addresses on the interface from which the Query is received. A Multicast-Address-Specific Query applies to a single multicast address on the interface from which the Query is received. Queries are ignored for addresses in the Non-Listener state.
- "report received" occurs when the node receives a valid MLD Report message. To be valid, the Report message MUST come from a link-local IPv6 Source Address, be at least 24 octets long, and have a correct MLD checksum. A Report applies only to the address identified in the Multicast Address field of the Report, on the interface from which the Report is received. It is ignored in the Non-Listener or Idle Listener state.
- "timer expired" occurs when the report delay timer for the address on the interface expires. It may occur only in the Delaying Listener state.

All other events, such as receiving invalid MLD messages or MLD message types other than Query or Report, are ignored in all states.

There are seven possible actions that may be taken in response to the above events:

- "send report" for the address on the interface. The Report message is sent to the address being reported.
- "send done" for the address on the interface. If the flag saying we were the last node to report is cleared, this action MAY be skipped. The Done message is sent to the link-scope all-routers address (FF02::2).
- "set flag" that we were the last node to send a report for this address.
- "clear flag" since we were not the last node to send a report for this address.
- "start timer" for the address on the interface, using a delay value chosen uniformly from the interval [0, Maximum Response Delay], where Maximum Response Delay is specified in the Query. If this is an unsolicited Report, the timer is set to a delay value chosen uniformly from the interval [0, [Unsolicited Report Interval]].
- "reset timer" for the address on the interface to a new value, using a delay value chosen uniformly from the interval [0, Maximum Response Delay], as described in "start timer".
- "stop timer" for the address on the interface.

In all of the following state transition diagrams, each state transition arc is labeled with the event that causes the transition, and, in parentheses, any actions taken during the transition. Note that the transition is always triggered by the event; even if the action is conditional, the transition still occurs.



The link-scope all-nodes address (FF02::1) is handled as a special case. The node starts in Idle Listener state for that address on every interface, never transitions to another state, and never sends a Report or Done for that address.

MLD messages are never sent for multicast addresses whose scope is 0 (reserved) or 1 (node-local).

MLD messages ARE sent for multicast addresses whose scope is 2 (link-local), including Solicited-Node multicast addresses [ADDR-ARCH], except for the link-scope, all-nodes address (FF02::1).

6. Router State Transition Diagram

Router behavior is more formally specified by the state transition diagrams below.

A router may be in one of two possible states with respect to any single attached link:

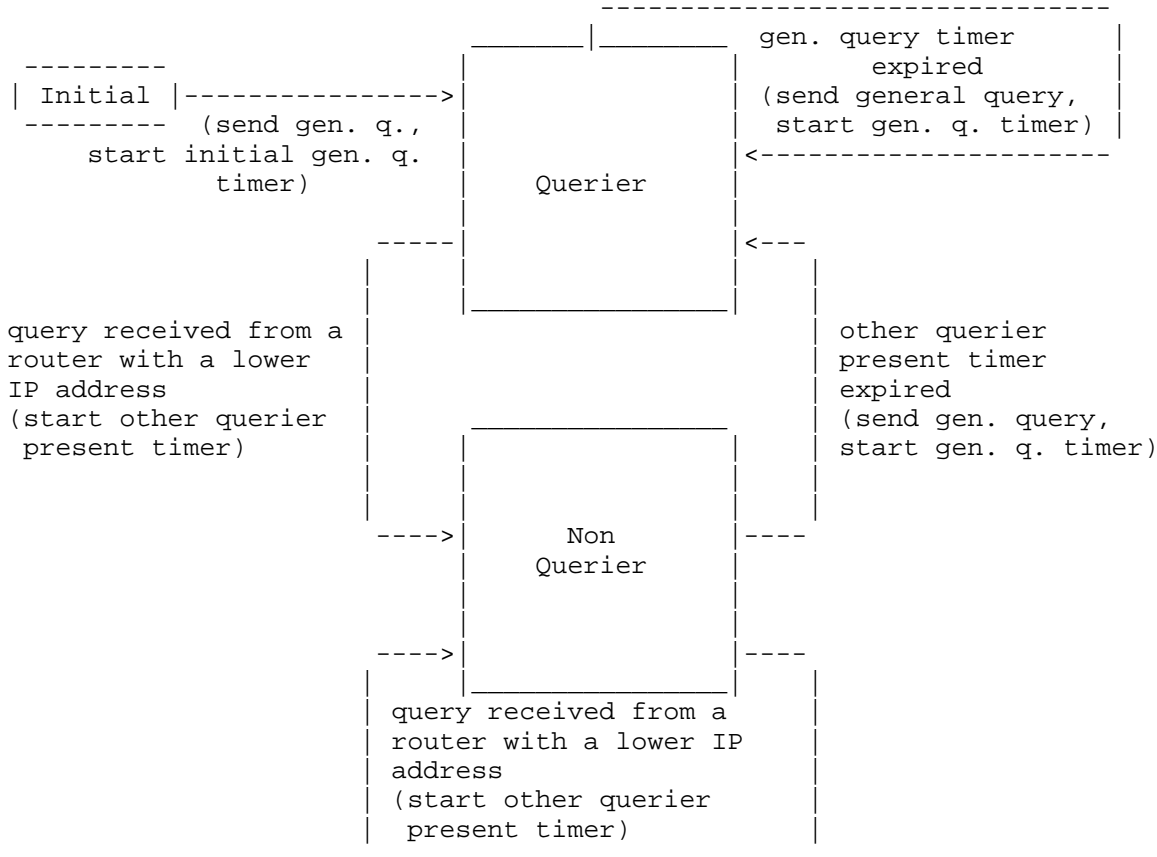
- "Querier", when this router is designated to transmit MLD Queries on this link.
- "Non-Querier", when there is another router designated to transmit MLD Queries on this link.

The following three events can cause the router to change states:

- "query timer expired" occurs when the timer set for query transmission expires. This event is significant only when in the Querier state.
- "query received from a router with a lower IP address" occurs when a valid MLD Query is received from a router on the same link with a lower IPv6 Source Address. To be valid, the Query message MUST come from a link-local IPv6 Source Address, be at least 24 octets long, and have a correct MLD checksum.
- "other querier present timer expired" occurs when the timer set to note the presence of another querier with a lower IP address on the link expires. This event is significant only when in the Non-Querier state.

There are three actions that may be taken in response to the above events:

- "start general query timer" for the attached link to [Query Interval].
- "start other querier present timer" for the attached link to [Other Querier Present Interval].
- "send general query" on the attached link. The General Query is sent to the link-scope all-nodes address (FF02::1), and has a Maximum Response Delay of [Query Response Interval].



A router starts in the Initial state on all attached links, and immediately transitions to Querier state.

In addition, to keep track of which multicast addresses have listeners, a router may be in one of three possible states with respect to any single IPv6 multicast address on any single attached link:

- "No Listeners Present" state, when there are no nodes on the link that have sent a Report for this multicast address. This is the initial state for all multicast addresses on the router; it requires no storage in the router.
- "Listeners Present" state, when there is a node on the link that has sent a Report for this multicast address.

- "Checking Listeners" state, when the router has received a Done message but has not yet heard a Report for the identified address.

There are five significant events that can cause router state transitions:

- "report received" occurs when the router receives a Report for the address from the link. To be valid, the Report message MUST come from a link-local IPv6 Source Address, be at least 24 octets long, and have a correct MLD checksum.
- "done received" occurs when the router receives a Done message for the address from the link. To be valid, the Done message MUST come from a link-local IPv6 Source Address, be at least 24 octets long, and have a correct MLD checksum. This event is significant only in the "Listeners Present" state and when the router is a Querier.
- "multicast-address-specific query received" occurs when a router receives a Multicast-Address-Specific Query for the address from the link. To be valid, the Query message MUST come from a link-local IPv6 Source Address, be at least 24 octets long, and have a correct MLD checksum. This event is significant only in the "Listeners Present" state and when the router is a Non-Querier.
- "timer expired" occurs when the timer set for a multicast address expires. This event is significant only in the "Listeners Present" or "Checking Listeners" state.
- "retransmit timer expired" occurs when the timer set to retransmit a Multicast-Address-Specific Query expires. This event is significant only in the "Checking Listeners" state.

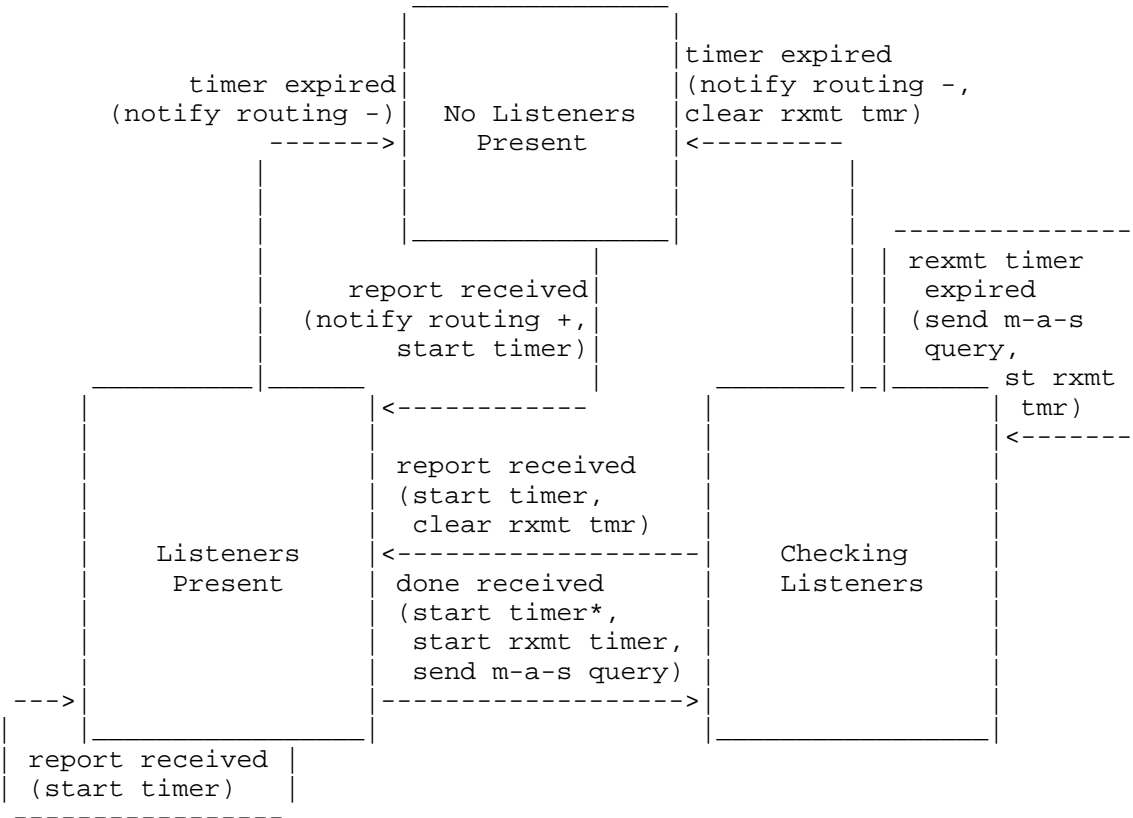
There are seven possible actions that may be taken in response to the above events:

- "start timer" for the address on the link - also resets the timer to its initial value [Multicast Listener Interval] if the timer is currently running.
- "start timer*" for the address on the link - this alternate action sets the timer to the minimum of its current value and either [Last Listener Query Interval] * [Last Listener Query Count] if this router is a Querier, or the Maximum Response Delay in the Query message * [Last Listener Query Count] if this router is a non-Querier.

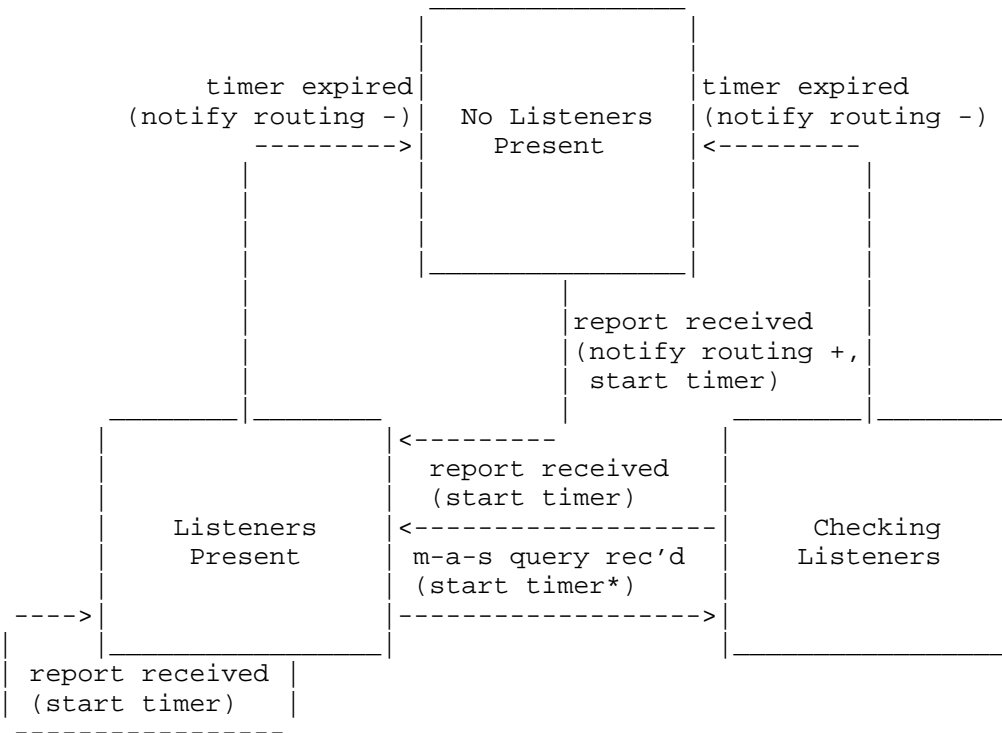
- "start retransmit timer" for the address on the link [Last Listener Query Interval].
- "clear retransmit timer" for the address on the link.
- "send multicast-address-specific query" for the address on the link. The Multicast-Address-Specific Query is sent to the address being queried, and has a Maximum Response Delay of [Last Listener Query Interval].
- "notify routing +" internally notify the multicast routing protocol that there are listeners to this address on this link.
- "notify routing -" internally notify the multicast routing protocol that there are no longer any listeners to this address on this link.

The following state diagrams apply per group per link. There are two diagrams; one for routers in Querier state and one for routers in Non-Querier state. The transition between Querier and Non-Querier state on a link is handled specially. All groups on that link in "No Listeners Present" or "Listeners Present" states switch state transition diagrams when the Querier/Non-Querier state transition occurs. However, any groups in "Checking Listeners" state continue with the same state transition diagram until the "Checking Listeners" state is exited. E.g. a router that starts as a Querier, receives a Done message for a group and then receives a Query from a router with a lower address (causing a transition to the Non-Querier state) continues to send multicast-address-specific queries for the group in question until it either receives a Report or its timer expires, at which time it starts performing the actions of a Non-Querier for this group.

The state transition diagram for a router in Querier state follows:



The state transition diagram for a router in Non-Querier state is similar, but non-Queriers do not send any messages and are only driven by message reception.



7. List of timers and default values

Most of these timers are configurable. If non-default settings are used, they MUST be consistent among all routers on a single link. Note that parentheses are used to group expressions to make the algebra clear.

7.1. Robustness Variable

The Robustness Variable allows tuning for the expected packet loss on a link. If a link is expected to be lossy, the Robustness Variable may be increased. MLD is robust to $(\text{Robustness Variable} - 1)$ packet losses. The Robustness Variable MUST NOT be zero, and SHOULD NOT be one. Default: 2

7.2. Query Interval

The Query Interval is the interval between General Queries sent by the Querier. Default: 125 seconds.

By varying the [Query Interval], an administrator may tune the number of MLD messages on the link; larger values cause MLD Queries to be sent less often.

7.3. Query Response Interval

The Maximum Response Delay inserted into the periodic General Queries. Default: 10000 (10 seconds)

By varying the [Query Response Interval], an administrator may tune the burstiness of MLD messages on the link; larger values make the traffic less bursty, as node responses are spread out over a larger interval. The number of seconds represented by the [Query Response Interval] must be less than the [Query Interval].

7.4. Multicast Listener Interval

The Multicast Listener Interval is the amount of time that must pass before a router decides there are no more listeners for an address on a link. This value MUST be ((the Robustness Variable) times (the Query Interval)) plus (one Query Response Interval).

7.5. Other Querier Present Interval

The Other Querier Present Interval is the length of time that must pass before a router decides that there is no longer another router which should be the querier on a link. This value MUST be ((the Robustness Variable) times (the Query Interval)) plus (one half of one Query Response Interval).

7.6. Startup Query Interval

The Startup Query Interval is the interval between General Queries sent by a Querier on startup. Default: 1/4 the Query Interval.

7.7. Startup Query Count

The Startup Query Count is the number of Queries sent out on startup, separated by the Startup Query Interval. Default: the Robustness Variable.

7.8. Last Listener Query Interval

The Last Listener Query Interval is the Maximum Response Delay inserted into Multicast-Address-Specific Queries sent in response to Done messages, and is also the amount of time between Multicast-Address-Specific Query messages. Default: 1000 (1 second)

This value may be tuned to modify the "leave latency" of the link. A reduced value results in reduced time to detect the departure of the last listener for an address.

7.9. Last Listener Query Count

The Last Listener Query Count is the number of Multicast-Address-Specific Queries sent before the router assumes there are no remaining listeners for an address on a link. Default: the Robustness Variable.

7.10. Unsolicited Report Interval

The Unsolicited Report Interval is the time between repetitions of a node's initial report of interest in a multicast address. Default: 10 seconds.

8. Message Destinations

This information is provided elsewhere in the document, but is summarized here for convenience.

Message Type	IPv6 Destination Address
-----	-----
General Query	link-scope all-nodes (FF02::1)
Multicast-Address-Specific Query	the multicast address being queried
Report	the multicast address being reported
Done	link-scope all-routers (FF02::2)

9. Security Considerations

We consider the ramifications of a forged message of each type. Note that the requirement that nodes verify that the IPv6 Source Address of all received MLD messages is a link-local address defends them from acting on forged MLD messages originated off-link, so we discuss only the effects of on-link forgery.

Query message:

A forged Query message from a machine with a lower IP address than the current Querier will cause Querier duties to be assigned to the forger. If the forger then sends no more Query messages, other routers' Other Querier Present timer will time out and one will resume the role of Querier. During this time, if the forger ignores Done messages, traffic might flow to addresses with no listeners for up to [Multicast Listener Interval].

A forged Query message sent to an address with listeners will cause one or more nodes that are listeners to that address to send a Report. This causes a small amount of extra traffic on the link, but causes no protocol problems.

Report message:

A forged Report message may cause routers to think there are listeners for an address present on a link when there are not. However, since listening to a multicast address is generally an unprivileged operation, a local user may trivially gain the same result without forging any messages.

Done message:

A forged Done message will cause the Querier to send out Multicast-Address-Specific Queries for the address in question. This causes extra processing on each router and on each of the address's listeners, and extra packets on the link, but cannot cause loss of desired traffic.

10. Acknowledgments

MLD was derived from IGMPv2 [IGMPv2], which was designed by Rosen Sharma and Steve Deering and documented by Bill Fenner.

11. References

- [ADDR-ARCH] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 2373, July 1998.
- [ICMPv6] Conta, A. and S. Deering, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 2463, December 1998.
- [IGMPv2] Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, November 1997.
- [IPv6] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.
- [IPv6-ETHER] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", RFC 2464, December, 1998.
- [KEYWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RTR-ALERT] Partridge, C. and A. Jackson, "IPv6 Router Alert Option", RFC 2711, October 1999.
- [STD-PROC] Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

12. Authors' Addresses

Stephen E. Deering
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA

Phone: +1 408 527 8213
EMail: deering@cisco.com

William C. Fenner
AT&T Research
75 Willow Road
Menlo Park, CA 94025
USA

Phone: +1 650 867 6073
EMail: fenner@research.att.com

Brian Haberman
IBM Corporation
800 Park Office Drive
Research Triangle Park, NC 27709
USA

Phone: +1 919 254 2673
EMail: haberman@raleigh.ibm.com

13. Full Copyright Statement

Copyright (C) The Internet Society (1999). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgement

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

