RFC 9776
Internet Group Management Protocol, Version 3

Abstract

The Internet Group Management Protocol (IGMP) is the protocol used by IPv4 systems to report their IP multicast group memberships to neighboring multicast routers. Version 3 of IGMP (IGMPv3) adds support for source filtering, that is, the ability for a system to report interest in receiving packets only from specific source addresses, or from all but specific source addresses, sent to a particular multicast address. That information may be used by multicast routing protocols to avoid delivering multicast packets from specific sources to networks where there are no interested receivers.

This document specifies IGMPv3. It is a revised version of RFC 3376 that includes clarifications and fixes for errata, and it is backward compatible with RFC 3376.

This document updates RFC 2236 and obsoletes RFC 3376.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9776.

Copyright Notice

Table of Contents

1. Introduction

The Internet Group Management Protocol (IGMP) is used by IPv4 systems (hosts and routers) to report their IP multicast group memberships to any neighboring multicast routers. Note that an IP multicast router may itself be a member of one or more multicast groups, in which case it performs both the multicast router part of the protocol (to collect the membership information needed by its multicast routing protocol) and the group member part of the protocol (to inform itself and other, neighboring multicast routers of its memberships).

IGMP is also used for other IP multicast management functions, using message types other than those used for group membership reporting. This document specifies only the group membership reporting functions and messages.

This document specifies Version 3 of IGMP. Version 1, specified in [RFC1112], was the first widely deployed version and the first version to become an Internet Standard. Version 2, specified in [RFC2236], added support for low leave latency, that is, a reduction in the time it takes for a multicast router to learn that there are no longer any members of a particular group present on an attached network. Version 3 adds support for source filtering, that is, the ability for a system to report interest in receiving packets only from specific source addresses, as required to support Source-Specific Multicast (SSM) [RFC3569], or from all but specific source addresses, sent to a particular multicast address. Version 3 is designed to be interoperable with Versions 1 and 2.

This document uses "SSM-aware" to refer to systems that support SSM as defined in [RFC4607].

This document updates [RFC2236] as a proper implementation of IGMPv3 needs to implement IGMPv2 Report and Leave message handling.

This document obsoletes [RFC3376] as it provides clarifications and fixes for errata in [RFC3376]. Detailed updates for those changes are described in Appendix C.

1.1. Conventions Used in This Document

The key words "**MUST**", "**MUST NOT**", "**REQUIRED**", "**SHALL**", "**SHALL NOT**", "**SHOULD**", "**SHOULD NOT**", "**RECOMMENDED**", "**NOT RECOMMENDED**", "**MAY**", and "**OPTIONAL**" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. The Service Interface for Requesting IP Multicast Reception

Within an IP system, there is (at least conceptually) a service interface used by upper-layer protocols or application programs to ask the IP layer to enable and disable reception of packets sent to specific IP multicast addresses. In order to take full advantage of the capabilities of IGMPv3, a system's IP service interface must support the following operation:

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

where:

- "socket" is an implementation-specific parameter used to distinguish among different requesting entities (e.g., programs or processes) within the system; the socket parameter of BSD Unix system calls is a specific example.
- "interface" is a local identifier of the network interface on which reception of the specified multicast address is to be enabled or disabled. Interfaces may be physical (e.g., an Ethernet interface) or virtual (e.g., the endpoint of a Frame Relay virtual circuit or the endpoint of an IP-in-IP "tunnel"). An implementation may allow a special "unspecified" value to be passed as the interface parameter, in which case the request would apply to the "primary" or "default" interface of the system (perhaps established by system configuration). If reception of the same multicast address is desired on more than one interface, IPMulticastListen is invoked separately for each desired interface.
- "multicast-address" is the IP multicast address, or group, to which the request pertains. If reception of more than one multicast address on a given interface is desired, IPMulticastListen is invoked separately for each desired multicast address.
- "filter-mode" may be either INCLUDE or EXCLUDE. In INCLUDE mode, reception of packets sent to the specified multicast address is requested only from those IP source addresses listed in the source-list parameter. In EXCLUDE mode, reception of packets sent to the given multicast address is requested from all IP source addresses except those listed in the source-list parameter.
- "source-list" is an unordered list of zero or more IP unicast addresses from which multicast reception is desired or not desired, depending on the filter-mode. An implementation **MAY** impose a limit on the size of source-lists, but that limit **MUST NOT** be less than 64 addresses per list. When an operation causes the source-list size limit to be exceeded, the service interface **MUST** return an error.

For a given combination of socket, interface, and multicast address, only a single filter-mode and source-list can be in effect at any one time. However, either the filter-mode or the source-list, or both, may be changed by subsequent IPMulticastListen requests that specify the same socket, interface, and multicast address. Each subsequent request completely replaces any earlier request for the given socket, interface, and multicast address.

Previous versions of IGMP did not support source filters and had a simpler service interface consisting of Join and Leave operations to enable and disable reception of a given multicast address (from all sources) on a given interface. The equivalent operations in the new service interface follow:

The Join operation is equivalent to:

```
      IPMulticastListen ( socket, interface, multicast-address,
                          EXCLUDE, {} )
```

and the Leave operation is equivalent to:

```
      IPMulticastListen ( socket, interface, multicast-address,
                          INCLUDE, {} )
```

where {} is an empty source-list.

An example of an API providing the capabilities outlined in this service interface is in [RFC3678].

3. Multicast Reception State Maintained by Systems

3.1. Socket State

For each socket on which IPMulticastListen has been invoked, the system records the desired multicast reception state for that socket. That state conceptually consists of a set of records of the form:

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

The socket state evolves in response to each invocation of IPMulticastListen on the socket, as follows:

- If the requested filter-mode is INCLUDE and the requested source-list is empty, then the entry corresponding to the requested interface and multicast address is deleted if present. If no such entry is present, the request is ignored.
- If the requested filter-mode is EXCLUDE or the requested source-list is non-empty, then the entry corresponding to the requested interface and multicast address, if present, is changed to contain the requested filter-mode and source-list. If no such entry is present, a new entry is created, using the parameters specified in the request.

3.2. Interface State

In addition to the per-socket multicast reception state, a system must also maintain or compute multicast reception state for each of its interfaces. That state conceptually consists of a set of records of the form:

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

At most, one record per multicast-address exists for a given interface. This per-interface state is derived from the per-socket state, but it may differ from the per-socket state when different sockets have differing filter-modes and/or source-lists for the same multicast address and interface. For example, suppose one application or process invokes the following operation on socket s1:

```
        IPMulticastListen ( s1, i, m, INCLUDE, {a, b, c} )
```

requesting reception on interface i of packets sent to multicast address m, only if they come from source a, b, or c. Suppose another application or process invokes the following operation on socket s2:

```
        IPMulticastListen ( s2, i, m, INCLUDE, {b, c, d} )
```

requesting reception on the same interface i of packets sent to the same multicast address m, only if they come from sources b, c, or d. In order to satisfy the reception requirements of both sockets, it is necessary for interface i to receive packets sent to m from any one of the sources a, b, c, or d. Thus, in this example, the reception state of interface i for multicast address m has filter-mode INCLUDE and source-list {a, b, c, d}.

After a multicast packet has been accepted from an interface by the IP layer, its subsequent delivery to the application or process listening on a particular socket depends on the multicast reception state of that socket (and possibly also on other conditions, such as what transport-layer port the socket is bound to). So, in the above example, if a packet arrives on interface i, destined to multicast address m, with source address a, it will be delivered on socket s1 but not on socket s2. Note that IGMP Queries and Reports are not subject to source filtering and must always be processed by hosts and routers.

Filtering of packets based upon a socket's multicast reception state is a feature of this service interface. The previous service interface [RFC1112] described no filtering based upon multicast join state; rather, a join on a socket simply caused the host to join a group on the given interface, and packets destined for that group could be delivered to all sockets whether they had joined or not.

The general rules for deriving the per-interface state from the per-socket state are as follows: For each distinct (interface, multicast-address) pair that appears in any socket state, a per-interface record is created for that multicast address on that interface. Considering all socket records containing the same (interface, multicast-address) pair,

• if any such record has a filter-mode of EXCLUDE, then the filter-mode of the interface record is EXCLUDE, and the source-list of the interface record is the intersection of the source-lists of all socket records in EXCLUDE mode, minus those source addresses that appear in any socket record in INCLUDE mode. For example, if the socket records for multicast address m on interface i are:

  ◦ from socket s1: (i, m, EXCLUDE, {a, b, c, d})

◦ from socket s2: (i, m, EXCLUDE, {b, c, d, e})

◦ from socket s3: (i, m, INCLUDE, {d, e, f})

then the corresponding interface record on interface i is:

◦ (m, EXCLUDE, {b, c})

If a fourth socket is added, such as:

◦ from socket s4: (i, m, EXCLUDE, {})

then the interface record becomes:

◦ (m, EXCLUDE, {})

• if all such records have a filter-mode of INCLUDE, then the filter-mode of the interface record is INCLUDE, and the source-list of the interface record is the union of the source-lists of all the socket records. For example, if the socket records for multicast address m on interface i are:

◦ from socket s1: (i, m, INCLUDE, {a, b, c})

◦ from socket s2: (i, m, INCLUDE, {b, c, d})

◦ from socket s3: (i, m, INCLUDE, {e, f})

then the corresponding interface record on interface i is:

◦ (m, INCLUDE, {a, b, c, d, e, f})

An implementation **MUST NOT** use an EXCLUDE interface record to represent a group when all sockets for this group are in INCLUDE state. If system resource limits are reached when an interface state source-list is calculated, an error **MUST** be returned to the application that requested the operation.

The above rules for deriving the interface state are (re-)evaluated whenever an IPMulticastListen invocation modifies the socket state by adding, deleting, or modifying a per-socket state record. Note that a change of socket state does not necessarily result in a change of interface state.

4. Message Formats

IGMP messages are encapsulated in IPv4 datagrams, with an IP protocol number of 2. Every IGMP message described in this document is sent with an IP Time-to-Live of 1, IP Precedence of Internetwork Control (e.g., Type of Service 0xc0), and carries an IP Router Alert option [RFC2113] in its IP header. IGMP message types are registered per [BCP57].

There are two IGMP message types of concern to the IGMPv3 protocol described in this document:

Type Number (hex)	Message Name
0x11	Membership Query
0x22	IGMPv3 Membership Report

Table 1: New Messages Introduced by IGMPv3

An implementation of IGMPv3 **MUST** also support the following three message types, for interoperation with previous versions of IGMP (see Section 7):

Type Number (hex)	Message Name	Reference
0x12	IGMPv1 Membership Report	[RFC1112]
0x16	IGMPv2 Membership Report	[RFC2236]
0x17	IGMPv2 Leave Group	[RFC2236]

Table 2: Legacy IGMP Messages

Unrecognized message types **MUST** be silently ignored. Other message types may be used by newer versions or extensions of IGMP, by multicast routing protocols, or for other uses.

In this document, unless otherwise qualified, the capitalized words "Query" and "Report" refer to IGMP Membership Queries and IGMPv3 Membership Reports, respectively.

4.1. Membership Query Message

Membership Queries are sent by IP multicast routers to query the multicast reception state of neighboring interfaces. Queries have the following format:

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | Type = 0x11  | Max Resp Code |           Checksum            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         Group Address                         |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | Flags |S| QRV |     QQIC      |     Number of Sources (N)     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                       Source Address [1]                      |
   +-                                                             -+
   |                       Source Address [2]                      |
   +-                              .                              -+
   .                              .                              .
   .                              .                              .
   +-                                                             -+
   |                       Source Address [N]                      |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
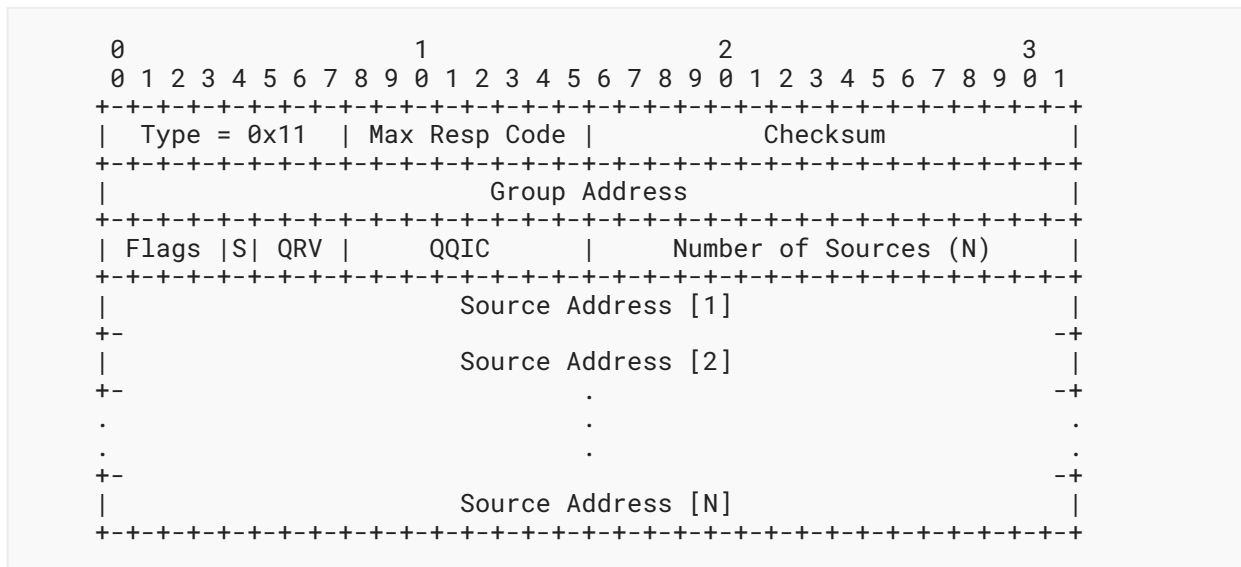
Figure 1: IGMPv3 Query Message

4.1.1. Max Response Code

The Max Resp Code field specifies the maximum time allowed before sending a responding report. The actual time allowed, called the "Max Response Time", is represented in units of 1/10 second and is derived from the Max Resp Code as follows:

  • If Max Resp Code < 128, Max Response Time = Max Resp Code

  • If Max Resp Code >= 128, Max Resp Code represents a floating-point value as follows:

```
      0 1 2 3 4 5 6 7
     +-+-+-+-+-+-+-+-+
     |1| exp | mant  |
     +-+-+-+-+-+-+-+-+

  Max Response Time = (mant | 0x10) << (exp + 3)
```
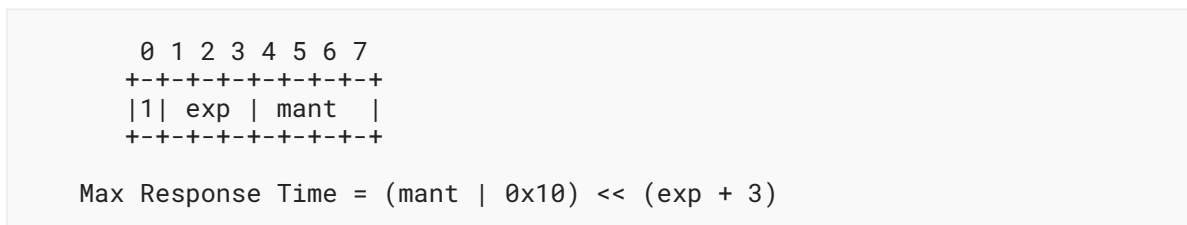
Figure 2: Max Resp Code Representation

Small values of Max Response Time allow IGMPv3 routers to tune the "leave latency" (the time between the moment the last host leaves a group and the moment the routing protocol is notified that there are no more members). Larger values, especially in the exponential range, allow tuning of the burstiness of IGMP traffic on a network.

4.1.2. Checksum

The Checksum field is the 16-bit one's complement of the one's complement sum of the whole IGMP message (the entire IP payload). For computing the checksum, the Checksum field is set to zero. When receiving packets, the checksum **MUST** be verified before processing a packet [RFC1071].

4.1.3. Group Address

The Group Address field is set to zero when sending a General Query and set to the IP multicast address being queried when sending a Group-Specific Query or Group-and-Source Specific Query (see Section 4.1.11, below).

4.1.4. Flags

The Flags field is a bitstring managed by the "IGMP Type Numbers" registry defined in [BCP57].

4.1.5. S Flag (Suppress Router-Side Processing)

When set to one, the S flag indicates to any receiving multicast routers that they are to suppress the normal timer updates they perform upon receiving a Query. It does not, however, suppress the Querier election or the normal "host-side" processing of a Query that a router may be required to perform as a consequence of itself being a group member.

4.1.6. QRV (Querier's Robustness Variable)

If non-zero, the QRV field contains the [Robustness Variable] value used by the querier, i.e., the sender of the Query. If the querier's [Robustness Variable] exceeds 7, the maximum value of the QRV field, the QRV is set to zero. Routers adopt the QRV value from the most recently received Query as their own [Robustness Variable] value, unless that most recently received QRV was zero, in which case the receivers use the default [Robustness Variable] value specified in Section 8.1 or a statically configured value.

4.1.7. QQIC (Querier's Query Interval Code)

The QQIC field specifies the [Query Interval] used by the querier. The actual interval, called the "Querier's Query Interval (QQI)", is represented in units of seconds and is derived from the QQIC as follows:

- If QQIC < 128, QQI = QQIC
- If QQIC >= 128, QQIC represents a floating-point value as follows:

```
      0 1 2 3 4 5 6 7
     +-+-+-+-+-+-+-+-+
     |1| exp | mant  |
     +-+-+-+-+-+-+-+-+

  QQI = (mant | 0x10) << (exp + 3)
```

Figure 3: QQIC Representation

Multicast routers that are not the current querier adopt the QQI value from the most recently received Query as their own [Query Interval] value, unless that most recently received QQI was zero, in which case the receiving routers use the default [Query Interval] value specified in Section 8.2.

4.1.8. Number of Sources (N)

The Number of Sources (N) field specifies how many source addresses are present in the Query. This number is zero in a General Query or a Group Specific Query and non-zero in a Group-and-Source Specific Query. This number is limited by the MTU of the network over which the Query is transmitted. For example, on an Ethernet with an MTU of 1500 octets, the IP header including the Router Alert option consumes 24 octets, and the IGMP fields up to and including the Number of Sources (N) field consume 12 octets, leaving 1464 octets for source addresses, which limits the number of source addresses to 366 (1464/4).

4.1.9. Source Address [i]

The Source Address [i] fields are a vector of n IP unicast addresses, where n is the value in the Number of Sources (N) field.

4.1.10. Additional Data

If the Packet Length field in the IP header of a received Query indicates that there are additional octets of data present, beyond the fields described here, IGMPv3 implementations **MUST** include those octets in the computation to verify the received IGMP Checksum but **MUST** otherwise ignore those additional octets. When sending a Query, an IGMPv3 implementation **MUST NOT** include additional octets beyond the fields described here.

4.1.11. Query Variants

There are three variants of the Query Message:

1. A General Query is sent by a multicast router to learn the complete multicast reception state of the neighboring interfaces (that is, the interfaces attached to the network on which the Query is transmitted). In a General Query, both the Group Address field and the Number of Sources (N) field are zero.

2. A Group Specific Query is sent by a multicast router to learn the reception state, with respect to a single multicast address, of the neighboring interfaces. In a Group Specific Query, the Group Address field contains the multicast address of interest, and the Number of Sources (N) field contains zero.

3. A Group-and-Source Specific Query is sent by a multicast router to learn if any neighboring interface desires reception of packets sent to a specified multicast address, from any of a specified list of sources. In a Group-and-Source Specific Query, the Group Address field contains the multicast address of interest, and the Source Address [i] fields contain the source address(es) of interest.

4.1.12. IP Destination Addresses for Queries

In IGMPv3, General Queries are sent with an IP destination address of 224.0.0.1, the all-systems multicast address. Group Specific and Group-and-Source Specific Queries are sent with an IP destination address equal to the multicast address of interest. However, a system **MUST** accept and process any Query whose IP Destination Address field contains any of the addresses (unicast or multicast) assigned to the interface on which the Query arrives.

4.2. IGMPv3 Membership Report Message

IGMPv3 Membership Reports are sent by IP systems to report (to neighboring routers) the current multicast reception state, or changes in the multicast reception state, of their interfaces. Reports have the following format:

```
      0                   1                   2                   3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     | Type = 0x22  |    Reserved   |            Checksum            |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |           Flags               |  Number of Group Records (M)  |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                                                               |
     .                                                               .
     .                       Group Record [1]                        .
     .                                                               .
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                                                               |
     .                                                               .
     .                       Group Record [2]                        .
     .                                                               .
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                               .                               |
     .                               .                               .
     |                               .                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                                                               |
     .                                                               .
     .                       Group Record [M]                        .
     .                                                               .
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
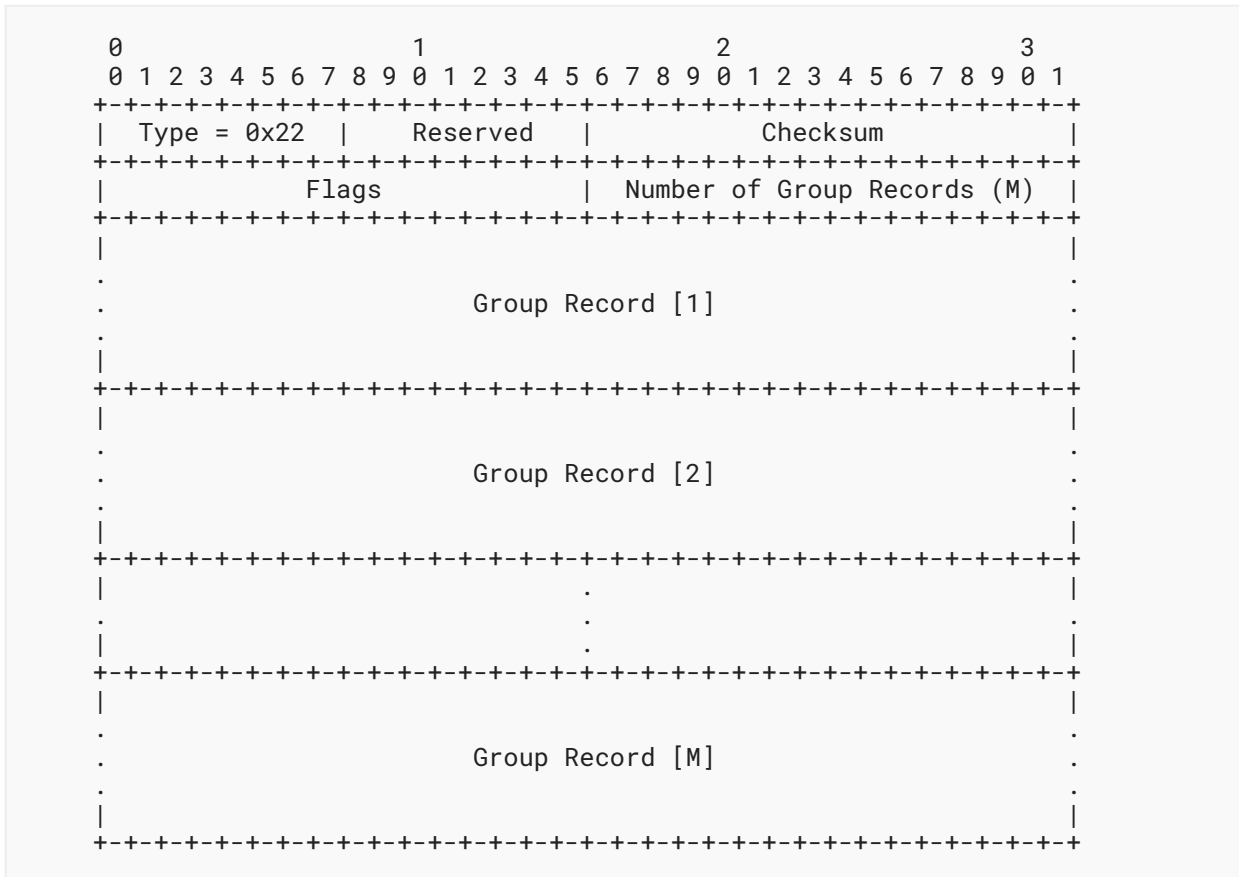
Figure 4: IGMPv3 Report Message

where each Group Record has the following internal format:

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Record Type  | Aux Data Len |     Number of Sources (N)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Multicast Address                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Source Address [1]                      |
+-                                                             -+
|                       Source Address [2]                      |
+-                                                             -+
.                              .                                .
.                              .                                .
.                              .                                .
+-                                                             -+
|                       Source Address [N]                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                              |
.                                                              .
.                          Auxiliary Data                      .
.                                                              .
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
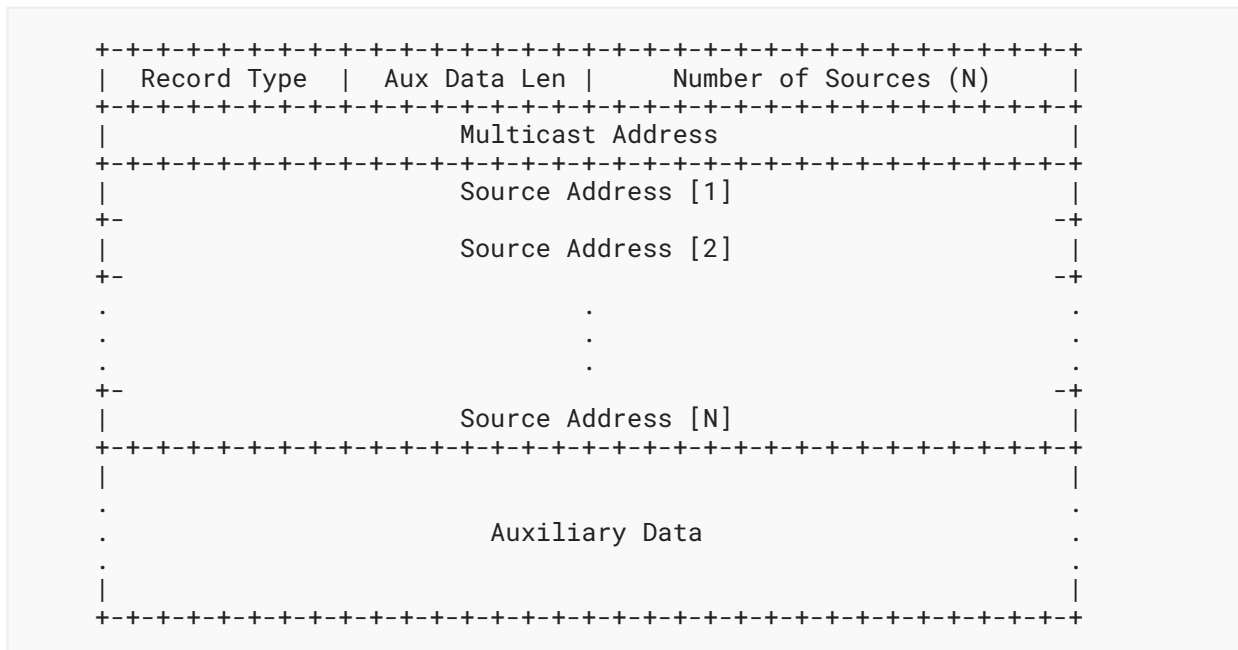
Figure 5: IGMPv3 Report Group Record

4.2.1. Reserved

The Reserved field is set to zero on transmission and ignored on reception.

4.2.2. Checksum

The Checksum field is the 16-bit one's complement of the one's complement sum of the whole IGMP message (the entire IP payload). For computing the checksum, the Checksum field is set to zero. When receiving packets, the checksum **MUST** be verified before processing a message.

4.2.3. Flags

The Flags field is a bitstring managed by the "IGMP Type Numbers" registry defined in [BCP57].

4.2.4. Number of Group Records (M)

The Number of Group Records (M) field specifies how many Group Records are present in this Report.

4.2.5. Group Record

Each Group Record is a block of fields containing information pertaining to the sender's membership in a single multicast group on the interface from which the Report is sent.

4.2.6. Record Type

See Section 4.2.13, below.

4.2.7. Aux Data Len

The Aux Data Len field contains the length of the Auxiliary Data field in this Group Record, in units of 32-bit words. It may contain zero, to indicate the absence of any auxiliary data.

4.2.8. Number of Sources (N)

The Number of Sources (N) field specifies how many source addresses are present in this Group Record.

4.2.9. Multicast Address

The Multicast Address field contains the IP multicast address to which this Group Record pertains.

4.2.10. Source Address [i]

The Source Address [i] fields are a vector of n IP unicast addresses, where n is the value in this record's Number of Sources (N) field.

4.2.11. Auxiliary Data

The Auxiliary Data field, if present, contains additional information pertaining to this Group Record. The protocol specified in this document, IGMPv3, does not define any auxiliary data. Therefore, implementations of IGMPv3 **MUST NOT** include any auxiliary data (i.e., **MUST** set the Aux Data Len field to zero) in any transmitted Group Record and **MUST** ignore any auxiliary data present in any received Group Record. The semantics and internal encoding of the Auxiliary Data field are to be defined by any future version or extension of IGMP that uses this field.

4.2.12. Additional Data

If the Packet Length field in the IP header of a received Report indicates that there are additional octets of data present, beyond the last Group Record, IGMPv3 implementations **MUST** include those octets in the computation to verify the received IGMP Checksum but **MUST** otherwise ignore those additional octets. When sending a Report, an IGMPv3 implementation **MUST NOT** include additional octets beyond the last Group Record.

4.2.13. Group Record Types

There are a number of different types of Group Records that may be included in a Report message:

- A Current-State Record is sent by a system in response to a Query received on an interface. It reports the current reception state of that interface, with respect to a single multicast address. The Record Type of a Current-State Record may be one of the following two values:

  1. MODE_IS_INCLUDE - indicates that the interface has a filter-mode of INCLUDE for the specified multicast address. The Source Address [i] fields in this Group Record contain the interface's source-list for the specified multicast address, if it is non-empty.

2. MODE_IS_EXCLUDE - indicates that the interface has a filter-mode of EXCLUDE for the specified multicast address. The Source Address [i] fields in this Group Record contain the interface's source-list for the specified multicast address, if it is non-empty. An SSM-aware host **SHOULD NOT** send a MODE_IS_EXCLUDE record type for multicast addresses that fall within the SSM address range as they will be ignored by SSM-aware routers [RFC4604].

• A Filter-Mode-Change Record is sent by a system whenever a local invocation of IPMulticastListen causes a change of the filter-mode (i.e., a change from INCLUDE to EXCLUDE, or from EXCLUDE to INCLUDE) of the interface-level state entry for a particular multicast address. The Record is included in a Report sent from the interface on which the change occurred. The Record Type of a Filter-Mode-Change Record may be one of the following two values:

3. CHANGE_TO_INCLUDE_MODE - indicates that the interface has changed to INCLUDE filter-mode for the specified multicast address. The Source Address [i] fields in this Group Record contain the interface's new source-list for the specified multicast address, if it is non-empty.

4. CHANGE_TO_EXCLUDE_MODE - indicates that the interface has changed to EXCLUDE filter-mode for the specified multicast address. The Source Address [i] fields in this Group Record contain the interface's new source-list for the specified multicast address, if it is non-empty. An SSM-aware host **SHOULD NOT** send a CHANGE_TO_EXCLUDE_MODE record type for multicast addresses that fall within the SSM address range.

• A Source-List-Change Record is sent by a system whenever a local invocation of IPMulticastListen causes a change of the source-list that is not coincident with a change of the filter-mode, of the interface-level state entry for a particular multicast address. The Record is included in a Report sent from the interface on which the change occurred. The Record Type of a Source-List-Change Record may be one of the following two values:

5. ALLOW_NEW_SOURCES - indicates that the Source Address [i] fields in this Group Record contain a list of the additional sources that the system wishes to receive, for packets sent to the specified multicast address. If the change was to an INCLUDE source-list, these are the addresses that were added to the list; if the change was to an EXCLUDE source-list, these are the addresses that were deleted from the list.

6. BLOCK_OLD_SOURCES - indicates that the Source Address [i] fields in this Group Record contain a list of the sources that the system no longer wishes to receive, for packets sent to the specified multicast address. If the change was to an INCLUDE source-list, these are the addresses that were deleted from the list; if the change was to an EXCLUDE source-list, these are the addresses that were added to the list.

If a change of source-list results in both allowing new sources and blocking old sources, then two Group Records are sent for the same multicast address, one of type ALLOW_NEW_SOURCES and one of type BLOCK_OLD_SOURCES.

We use the term "State-Change Record" to refer to either a Filter- Mode-Change Record or a Source-List-Change Record.

Unrecognized Record Type values **MUST** be silently ignored.

4.2.14. IP Source Addresses for Reports

An IGMP report is sent with a valid unicast IPv4 source address for the destination subnet. The 0.0.0.0 source address may be used by a system that has not yet acquired an IP address. Note that the 0.0.0.0 source address may simultaneously be used by multiple systems on a LAN. Routers **MUST** accept a report with a source address of 0.0.0.0.

4.2.15. IP Destination Addresses for Reports

Version 3 Reports are sent with an IP destination address of 224.0.0.22, to which all IGMPv3-capable multicast routers listen. A system that is operating in v1 or v2 compatibility modes sends v1 or v2 Reports to the multicast group specified in the Group Address field of the Report. In addition, a system **MUST** accept and process any v1 or v2 Report whose IP Destination Address field contains any of the addresses (unicast or multicast) assigned to the interface on which the Report arrives.

4.2.16. Notation for Group Records

In the rest of this document, we use the following notation to describe the contents of a Group Record pertaining to a particular multicast address:

- IS_IN (x) - Type MODE_IS_INCLUDE, source addresses x
- IS_EX (x) - Type MODE_IS_EXCLUDE, source addresses x
- TO_IN (x) - Type CHANGE_TO_INCLUDE_MODE, source addresses x
- TO_EX (x) - Type CHANGE_TO_EXCLUDE_MODE, source addresses x
- ALLOW (x) - Type ALLOW_NEW_SOURCES, source addresses x
- BLOCK (x) - Type BLOCK_OLD_SOURCES, source addresses x

where x is either:

- a capital letter (e.g., "A") to represent the set of source addresses or
- a set expression (e.g., "A+B"), where "A+B" means the union of sets A and B, "A*B" means the intersection of sets A and B, and "A-B" means the removal of all elements of set B from set A.

4.2.17. Membership Report Size

If the set of Group Records required in a report does not fit within the size limit of a single Report message (as determined by the MTU of the network on which it will be sent), the Group Records are sent in as many Report messages as needed to report the entire set.

If a single Group Record contains so many source addresses that it does not fit within the size limit of a single Report message, and if its Type is not MODE_IS_EXCLUDE or CHANGE_TO_EXCLUDE_MODE, it is split into multiple Group Records, each containing a different subset of the source addresses and each sent in a separate Report message. If its Type is MODE_IS_EXCLUDE or CHANGE_TO_EXCLUDE_MODE, a single Group Record is sent, containing

as many source addresses as can fit, and the remaining source addresses are not reported; though the choice of which sources to report is arbitrary, it is preferable to report the same set of sources in each subsequent report, rather than reporting different sources each time.

5. Description of the Protocol for Group Members

IGMP is an asymmetric protocol, specifying separate behaviors for group members -- that is, hosts or routers that wish to receive multicast packets -- and multicast routers. This section describes the part of IGMPv3 that applies to all group members. (Note that a multicast router that is also a group member performs both parts of IGMPv3, receiving and responding to its own IGMP message transmissions as well as those of its neighbors. The multicast router part of IGMPv3 is described in Section 6.)

A system performs the protocol described in this section over all interfaces on which multicast reception is supported, even if more than one of those interfaces is connected to the same network.

For interoperability with multicast routers running older versions of IGMP, systems maintain a MulticastRouterVersion variable for each interface on which multicast reception is supported. This section describes the behavior of group member systems on interfaces for which MulticastRouterVersion = 3. The algorithm for determining MulticastRouterVersion, and the behavior for versions other than 3, are described in Section 7.

The all-systems multicast address, 224.0.0.1, is handled as a special case. On all systems -- that is, all hosts and routers including multicast routers -- reception of packets destined to the all-systems multicast address, from all sources, is permanently enabled on all interfaces on which multicast reception is supported. No IGMP messages are ever sent regarding the all-systems multicast address.

There are two types of events that trigger IGMPv3 protocol actions on an interface:

- A change of the interface reception state, caused by a local invocation of IPMulticastListen.
- The reception of a Query.

(Received IGMP messages of types other than Query are silently ignored, except as required for interoperation with earlier versions of IGMP.)

The following subsections describe the actions to be taken for each of these two cases. In those descriptions, timer and counter names appear in square brackets. The default values for those timers and counters are specified in Section 8.

5.1. Action on Change of Interface State

An invocation of IPMulticastListen may cause the multicast reception state of an interface to change, according to the rules in Section 3.2. Each such change affects the per-interface entry for a single multicast address.

A change of interface state causes the system to immediately transmit a State-Change Report 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, according to Table 3. If no interface state existed for that multicast address before the change (i.e., the change consisted of creating a new per-interface record), or if no state exists after the change (i.e., the change consisted of deleting a per-interface record), then the "non-existent" state is considered to have a filter-mode of INCLUDE and an empty source-list.

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

Table 3: Transmitted Group Records for State Changes

If the computed source-list for either an ALLOW or a BLOCK State- Change Record is empty, that record is omitted from the Report message.

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]).

If more changes to the same interface state entry occur before all the retransmissions of the State-Change Report for the first change have been completed, each such additional change triggers the immediate transmission of a State-Change Report.

The contents of the new transmitted report are calculated as follows. As was done with the first report, the interface state for the affected group before and after the latest change is compared. The report records expressing the difference are built according to Table 3. However, these records are not transmitted in a message but instead are merged with the contents of the pending report to create the new State-Change report. The rules for merging the difference report resulting from the state change and the pending report are described below.

The transmission of the merged State-Change Report terminates retransmissions of the earlier State-Change Reports for the same multicast address, and becomes the first of [Robustness Variable] transmissions of State-Change Reports.

Each time a source is included in the difference report calculated above, retransmission state for that source needs to be maintained until [Robustness Variable] State-Change Reports have been sent by the host. This is done in order to ensure that a series of successive state changes do not break the protocol robustness.

If the interface reception-state change that triggers the new report is a filter-mode change, then the next [Robustness Variable] State-Change Reports will include a Filter-Mode-Change Record. This applies even if any number of source-list changes occur in that period. The host has to maintain retransmission state for the group until the [Robustness Variable] State-Change Reports have been sent. When [Robustness Variable] State-Change Reports with Filter-Mode-Change Records have been transmitted after the last filter-mode change, and if source-list changes to the interface reception have scheduled additional reports, then the next State-Change Report will include Source-List-Change Records.

Each time a State-Change Report is transmitted, the contents are determined as follows. If the report should contain a Filter-Mode-Change Record, and if the current filter-mode of the interface is INCLUDE, a TO_IN record is included in the report; otherwise, a TO_EX record is included. If instead the report should contain a Source-List-Change Record, an ALLOW and a BLOCK record are included. The contents of these records are built according to Table 4.

Record	Sources Included
TO_IN	All in the current interface state that must be forwarded
TO_EX	All in the current interface state that must be blocked
ALLOW	All with retransmission state that must be forwarded
BLOCK	All with retransmission state that must be blocked

Table 4: Change Record Construction

If the computed source-list for either an ALLOW or a BLOCK record is empty, that record is omitted from the State-Change Report.

> Note: When the first State-Change Report is sent, the non-existent pending report to merge with can be treated as a Source-Change Report with empty ALLOW and BLOCK records (no sources have retransmission state).

5.2. Action on Reception of a Query

When a system receives a Query, it does not respond immediately. Instead, it delays its response by a random amount of time, bounded by the Max Response Time value derived from the Max Resp Code in the received Query Message. A 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 as, in many cases, it can schedule a combined response. Therefore, the system 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.

When a Query with the Router Alert option arrives on an interface, provided the system has state to report, a delay for a response is randomly selected in the range (0, [Max Response Time]) where Max Response Time is derived from Max Resp Code in the received Query Message. The following rules are then used to determine if a Report needs to be scheduled and the type of Report to schedule. The rules are considered in order and only the first matching rule is applied.

1. If there is a pending response to a previous General Query scheduled sooner than the selected delay, no additional response needs to be scheduled.

2. If the received Query is a General Query, the Interface Timer is used to schedule a response to the General Query after the selected delay. Any previously pending response to a General Query is canceled.

3. If the received Query is a Group Specific Query or a Group-and- Source Specific Query and there is no pending response to a previous Query for this group, then the Group Timer is used to schedule a report. If the received Query is a Group-and-Source Specific Query, the list of queried sources is recorded to be used when generating a response.

4. If there already is a pending response to a previous Query scheduled for this group, and either the new Query is a Group Specific Query or the recorded source-list associated with the group is empty, then the group source-list is cleared and a single response is scheduled using the Group Timer. The new response is scheduled to be sent at the earliest of the remaining time for the pending report and the selected delay.

5. 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. The new response is scheduled to be sent at the earliest of the remaining time for the pending report and the selected delay.

When the timer in a pending response record expires, the system transmits, on the associated interface, one or more Report messages carrying one or more Current-State Records (see Section 4.2.13), as follows:

1. If the expired timer is the Interface Timer (i.e., it is a pending response to a General Query), then one Current-State Record is sent for each multicast address for which the specified interface has reception state, as described in Section 3.2. The Current-State Record carries the multicast address and its associated filter-mode (MODE_IS_INCLUDE or MODE_IS_EXCLUDE) and source-list. Multiple Current-State Records are packed into individual Report messages, to the extent possible.

This naive algorithm may result in bursts of packets when a system is a member of a large number of groups. Instead of using a single Interface Timer, implementations are recommended to spread transmission of such Report messages over the interval (0, [Max Response Time]). Note that any such implementation **MUST** avoid the "ack-implosion" problem, i.e., **MUST NOT** send a Report immediately on reception of a General Query.

2. If the expired timer is a Group Timer and the list of recorded sources for that group is empty (i.e., it is a pending response to a Group Specific Query), then if and only if the interface has reception state for that group address, a single Current-State Record is sent for that address. The Current-State Record carries the multicast address and its associated filter-mode (MODE_IS_INCLUDE or MODE_IS_EXCLUDE) and source-list.

3. If the expired timer is a Group Timer and the list of recorded sources for that group is non-empty (i.e., it is a pending response to a Group-and-Source Specific Query), then if and only if the interface has reception state for that group address, the contents of the responding Current-State Record is determined from the interface state and the pending response record, as specified in Table 5.

Per-Interface State	Set of Sources in the Pending Response Record	Current-State Record
INCLUDE (A)	B	IS_IN (A*B)
EXCLUDE (A)	B	IS_IN (B-A)

Table 5: Current-State Record Construction

If the resulting Current-State Record has an empty set of source addresses, then no response is sent.

Finally, after any required Report messages have been generated, the source-lists associated with any reported groups are cleared.

6. Description of the Protocol for Multicast Routers

The purpose of IGMP is to enable each multicast router to learn, for each of its directly attached networks, which multicast addresses are of interest to the systems attached to those networks. IGMPv3 adds 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 information gathered by IGMP is provided to whichever multicast routing protocol is being used by the router, in order to ensure that multicast packets are delivered to all networks where there are interested receivers.

This section describes the part of IGMPv3 that is performed by multicast routers. Multicast routers may also themselves become members of multicast groups, and therefore also perform the group member part of IGMPv3, as described in Section 5.

A multicast router performs the protocol described in this section over each of its directly attached networks. If a multicast router has more than one interface to the same network, it only needs to operate this protocol over one of those interfaces. On each interface over which this protocol is being run, the router **MUST** enable reception of multicast address 224.0.0.22 from all sources (and **MUST** perform the group member part of IGMPv3 for that address on that interface).

Multicast routers need to know only that at least one system on an attached network is interested in packets to a particular multicast address from a particular source; a multicast router is not required to keep track of the interests of each individual neighboring system. (However, see Appendix A.2, item 1 for discussion.)

IGMPv3 is backward compatible with previous versions of the IGMP protocol. In order to remain backward compatible with older IGMP systems, IGMPv3 multicast routers **MUST** also implement versions 1 and 2 of the protocol (see Section 7).

6.1. Conditions for IGMP Queries

Multicast routers send General Queries periodically to request group membership information from an attached network. These Queries are used to build and refresh the group membership state of systems on attached networks. Systems respond to these Queries by reporting their group membership state (and their desired set of sources) with Current-State Records in IGMPv3 Membership Reports.

As a member of a multicast group, a system may express interest in receiving or not receiving traffic from particular sources. As the desired reception state of a system changes, it reports these changes using Filter-Mode-Change Records or Source-List-Change Records. These records indicate an explicit state change in a group at a system in either the Group Record's source-list or its filter-mode. When a group membership is terminated at a system or traffic from a particular source is no longer desired, a multicast router must query for other members of the group or listeners of the source before deleting the group (or source) and pruning its traffic.

To enable all systems on a network to respond to changes in group membership, multicast routers send specific queries. A Group Specific Query is sent to verify there are no systems that desire reception of the specified group or to "rebuild" the desired reception state for a particular group. Group Specific Queries are sent when a router receives a State-Change Record indicating a system is leaving a group.

A Group-and-Source Specific Query is used to verify there are no systems on a network that desire receiving traffic from a set of sources. Group-and-Source Specific Queries list sources for a particular group that have been requested to no longer be forwarded. This query is sent by a multicast router to learn if any systems desire reception of packets to the specified group address from the specified source addresses. Group-and-Source Specific Queries are only sent in response to State-Change Records and never in response to Current-State Records. Section 4.1.11 describes each query in more detail.

6.2. IGMP State Maintained by Multicast Routers

Multicast routers implementing IGMPv3 keep state per group per attached network. This group state consists of a filter-mode, a list of sources, and various timers. For each attached network running IGMP, a multicast router records the desired reception state for that network. That state conceptually consists of a set of records of the form:

```
(multicast address, Group Timer, Router Filter Mode, (source records))
```

Each source record is of the form:

```
(source address, Source Timer)
```

If all sources within a given group are desired, an empty source record list is kept with filter-mode set to EXCLUDE. This means hosts on this network want all sources for this group to be forwarded. This is the IGMPv3 equivalent to an IGMPv1 or IGMPv2 group join.

6.2.1. Definition of Router Filter Mode

To reduce internal state, IGMPv3 routers keep a filter-mode per group per attached network. This filter-mode is used to condense the total desired reception state of a group to a minimum set such that all systems' memberships are satisfied. This filter-mode may change in response to the reception of particular types of Group Records or when certain timer conditions occur. In the following sections, we use the term Router Filter Mode to refer to the filter-mode of a particular group within a router. Section 6.4 describes the changes of a Router Filter Mode per Group Record received.

Conceptually, when a Group Record is received, the Router Filter Mode for that group is updated to cover all the requested sources using the least amount of state. As a rule, once a Group Record with a filter-mode of EXCLUDE is received, the Router Filter Mode for that group will be EXCLUDE.

When a Router Filter Mode for a group is EXCLUDE, the source record list contains two types of sources. The first type is the set that represents conflicts in the desired reception state; this set must be forwarded by some router on the network. The second type is the set of sources that hosts have requested to not be forwarded. Appendix A describes the reasons for keeping two different sets when in EXCLUDE mode.

When a Router Filter Mode for a group is INCLUDE, the source record list is the list of sources desired for the group. This is the total desired set of sources for that group. Each source in the source record list must be forwarded by some router on the network.

Because a reported Group Record with a filter-mode of EXCLUDE will cause a router to transition its filter-mode for that group to EXCLUDE, a mechanism for transitioning a router's filter-mode back to INCLUDE must exist. If all systems with a Group Record in EXCLUDE filter-

mode cease reporting, it is desirable for the Router Filter Mode for that group to transition back to INCLUDE mode. This transition occurs when the Group Timer expires and is explained in detail in Section 6.5.

6.2.2. Definition of Group Timers

The Group Timer is only used when a group is in EXCLUDE mode and it represents the time for the filter-mode of the group to expire and switch to INCLUDE mode. We define a Group Timer as a decrementing timer with a lower bound of zero kept per group per attached network. Group timers are updated according to the types of Group Records received.

A Group Timer expiring when a Router Filter Mode for the group is EXCLUDE means there are no listeners on the attached network in EXCLUDE mode. At this point, a router will transition to INCLUDE filter-mode. Section 6.5 describes the actions taken when a Group Timer expires while in EXCLUDE mode.

Table 6 summarizes the role of the Group Timer. Section 6.4 describes the details of setting the Group Timer per type of Group Record received.

Group Filter-Mode	Group Timer Value	Actions/Comments
INCLUDE	Timer >= 0	All members in INCLUDE mode.
EXCLUDE	Timer > 0	At least one member in EXCLUDE mode.
EXCLUDE	Timer == 0	No more listeners to group. If all source timers have expired, then delete Group Record. If there are still source record timers running, switch to INCLUDE filter-mode using those source records with running timers as the INCLUDE source record state.

Table 6: Group Timer Actions

6.2.3. Definition of Source Timers

A Source Timer is kept per source record and is a decrementing timer with a lower bound of zero. Source timers are updated according to the type and filter-mode of the Group Record received. Source timers are always updated (for a particular group) whenever the source is present in a received record for that group. Section 6.4 describes the setting of source timers per type of Group Records received.

A source record with a running timer with a Router Filter Mode for the group of INCLUDE means that there is currently one or more systems (in INCLUDE filter-mode) that desire to receive that source. If a Source Timer expires with a Router Filter Mode for the group of INCLUDE, the router concludes that traffic from this particular source is no longer desired on the attached network and deletes the associated source record.

Source timers are treated differently when a Router Filter Mode for a group is EXCLUDE. If a source record has a running timer with a Router Filter Mode for the group of EXCLUDE, it means that at least one system desires the source. It should therefore be forwarded by a router on the network. Appendix A describes the reasons for keeping state for sources that have been requested to be forwarded while in EXCLUDE state.

If a Source Timer expires with a Router Filter Mode for the group of EXCLUDE, the router informs the routing protocol that there is no longer a receiver on the network interested in traffic from this source.

When a Router Filter Mode for a group is EXCLUDE, source records are only deleted when the Group Timer expires. Section 6.3 describes the actions that should be taken dependent upon the value of a Source Timer.

6.3. IGMPv3 Source-Specific Forwarding Rules

When a multicast router receives a datagram from a source destined to a particular group, a decision has to be made whether to forward the datagram onto an attached network or not. The multicast routing protocol in use is in charge of this decision and should use the IGMPv3 information to ensure that all sources/groups desired on a subnetwork are forwarded to that subnetwork. IGMPv3 information does not override multicast routing information; for example, if the IGMPv3 filter-mode group for G is EXCLUDE, a router may still forward packets for excluded sources to a transit subnet.

To summarize, Table 7 describes the forwarding suggestions made by IGMP to the routing protocol for traffic originating from a source destined to a group. It also summarizes the actions taken upon the expiration of a Source Timer based on the Router Filter Mode of the group.

Group Filter-Mode	Group Timer Value	Action
INCLUDE	TIMER > 0	Suggest to forward traffic from source.
INCLUDE	TIMER == 0	Suggest to stop forwarding traffic from source and remove source record. If there are no more source records for the group, delete Group Record.
INCLUDE	No Source Elements	Suggest to not forward source.
EXCLUDE	TIMER > 0	Suggest to forward traffic from source.
EXCLUDE	TIMER == 0	Suggest to not forward traffic from source (DO NOT remove record).

Group Filter-Mode	Group Timer Value	Action
EXCLUDE	No Source Elements	Suggest to forward traffic from source.

Table 7: IGMP Forwarding Recommendations

6.4. Action on Reception of Reports

SSM-aware routers **SHOULD** ignore records that contain multicast addresses in the SSM address range if the record type is MODE_IS_EXCLUDE or CHANGE_TO_EXCLUDE_MODE. SSM-aware routers **SHOULD** ignore IGMPv1/IGMPv2 Report and IGMPv2 DONE messages that contain multicast addresses in the SSM address range, **SHOULD NOT** use such Reports to establish IP forwarding state, and **MAY** log an error if it receives such a message.

6.4.1. Reception of Current-State Records

When receiving Current-State Records, a router updates both its group and source timers. In some circumstances, the reception of a type of Group Record will cause the Router Filter Mode for that group to change. Table 8 describes the actions, with respect to state and timers that occur to a router's state upon reception of Current- State Records.

The following notation is used to describe the updating of source timers. The notation (A, B) will be used to represent the total number of sources for a particular group, where

- A = set of source records whose source timers > 0 (Sources that at least one host has requested to be forwarded)
- B = set of source records whose source timers = 0 (Sources that IGMP will suggest to the routing protocol not to forward)

Note that there will only be two sets when a router's filter-mode for a group is EXCLUDE. When a router's filter-mode for a group is INCLUDE, a single set is used to describe the set of sources requested to be forwarded (e.g., simply (A)).

In Tables 8 and 9, abbreviations are used for several variables (all of which are described in detail in Section 8). The variable GMI is an abbreviation for the Group Membership Interval, which is the time in which group memberships will time out. The variable LMQT is an abbreviation for the Last Member Query Time, which is the total time spent after [Last Member Query Count] retransmissions. LMQT represents the leave latency or the difference between the transmission of a membership change and the change in the information given to the routing protocol.

Within the "Actions" section of the router state tables, we use the notation 'A=J', which means that the set A of source records should have their source timers set to value J. 'Delete A' means that the set A of source records should be deleted. 'Group Timer=J' means that the Group Timer for the group should be set to value J.

Router State	Report Rec'd	New Router State	Actions
INCLUDE (A)	IS_IN (B)	INCLUDE (A+B)	(B)=GMI
INCLUDE (A)	IS_EX (B)	EXCLUDE (A*B,B-A)	(B-A)=0 Delete (A-B) Group Timer=GMI
EXCLUDE (X,Y)	IS_IN (A)	EXCLUDE (X+A,Y-A)	(A)=GMI
EXCLUDE (X,Y)	IS_EX (A)	EXCLUDE (A-Y,Y*A)	(A-X-Y)=GMI Delete (X-A) Delete (Y-A) Group Timer=GMI

Table 8: Actions on Current-State Report Reception

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

When a change in the global state of a group occurs in a system, the system sends either a Source-List-Change Record or a Filter-Mode- Change Record for that group. As with Current-State Records, routers must act upon these records and possibly change their own state to reflect the new desired membership state of the network.

Routers must query sources that are requested to be no longer forwarded to a group. When a router queries or receives a query for a specific set of sources, it lowers its source timers for those sources to a small interval of [Last Member Query Time] seconds. If Group Records are received in response to the queries which express interest in receiving traffic from the queried sources, the corresponding timers are updated.

Similarly, when a router queries a specific group, it lowers its Group Timer for that group to a small interval of [Last Member Query Time] seconds. If any Group Records expressing EXCLUDE mode interest in the group are received within the interval, the Group Timer for the group is updated and the suggestion to the routing protocol to forward the group stands without any interruption.

During a query period (i.e., [Last Member Query Time] seconds), the IGMP component in the router continues to suggest to the routing protocol that it forwards traffic from the groups or sources that it is querying. It is not until after [Last Member Query Time] seconds without receiving a record expressing interest in the queried group or sources that the router may prune the group or sources from the network.

Table 9 describes the changes in group state and the action(s) taken when receiving either Filter-Mode-Change or Source-List-Change Records. This table also describes the queries that are sent by the querier when a particular report is received.

We use the following notation for describing the queries that are sent. We use the notation 'Q(G)' to describe a Group Specific Query to G. We use the notation 'Q(G,A)' to describe a Group-and-Source Specific Query to G with source-list A. If source-list A is null as a result of the action (e.g., A*B), then no query is sent as a result of the operation.

In order to maintain protocol robustness, queries sent by actions in Table 9 need to be transmitted [Last Member Query Count] times, once every [Last Member Query Interval].

If while scheduling new queries there are already pending queries to be retransmitted for the same group, the new and pending queries have to be merged. In addition, received host reports for a group with pending queries may affect the contents of those queries. Section 6.6.3 describes the process of building and maintaining the state of pending queries.

Router State	Report Rec'd	New Router State	Actions
INCLUDE (A)	ALLOW (B)	INCLUDE (A+B)	(B)=GMI
INCLUDE (A)	BLOCK (B)	INCLUDE (A)	Send Q(G,A*B)
INCLUDE (A)	TO_EX (B)	EXCLUDE (A*B,B-A)	(B-A)=0 Delete (A-B) Send Q(G,A*B) Group Timer=GMI
INCLUDE (A)	TO_IN (B)	INCLUDE (A+B)	(B)=GMI Send Q(G,A-B)
EXCLUDE (X,Y)	ALLOW (A)	EXCLUDE (X+A,Y-A)	(A)=GMI
EXCLUDE (X,Y)	BLOCK (A)	EXCLUDE (X+(A-Y),Y)	(A-X-Y)=Group Timer Send Q(G,A-Y)
EXCLUDE (X,Y)	TO_EX (A)	EXCLUDE (A-Y,Y*A)	(A-X-Y)=Group Timer Delete (X-A) Delete (Y-A) Send Q(G,A-Y) Group Timer=GMI
EXCLUDE (X,Y)	TO_IN (A)	EXCLUDE (X+A,Y-A)	(A)=GMI Send Q(G,X-A) Send Q(G)

Table 9: Actions on Change Record Reception

6.5. Switching Router Filter Modes

The Group Timer is used as a mechanism for transitioning the Router Filter Mode from EXCLUDE to INCLUDE.

When a Group Timer expires with a Router Filter Mode of EXCLUDE, a router assumes that there are no systems with a filter-mode of EXCLUDE present on the attached network. When a router's filter-mode for a group is EXCLUDE and the Group Timer expires, the Router Filter Mode for the group transitions to INCLUDE.

A router uses source records with running source timers as its state for the switch to a filter-mode of INCLUDE. If there are any source records with source timers greater than zero (i.e., requested to be forwarded), a router switches to filter-mode of INCLUDE using those source records. Source records whose timers are zero (from the previous EXCLUDE mode) are deleted.

For example, if a router's state for a group is EXCLUDE(X,Y) and the Group Timer expires for that group, the router switches to filter- mode of INCLUDE with state INCLUDE(X).

6.6. Action on Reception of Queries

6.6.1. Timer Updates

When a router sends or receives a query with a clear Suppress Router-Side Processing flag, it must update its timers to reflect the correct timeout values for the group or sources being queried. Table 10 describes the timer actions when sending or receiving a Group Specific or Group-and-Source Specific Query with the S flag not set.

Query	Action
Q(G,A)	Source Timer for sources in A are lowered to LMQT
Q(G)	Group Timer is lowered to LMQT

Table 10: Timer Updates on Query

When a router sends or receives a query with the S flag set, it will not update its timers.

6.6.2. Querier Election

IGMPv3 elects a single querier per subnet using the same Querier election mechanism as IGMPv2, namely by IP address. When a router receives a General Query with a lower IP address, it sets the Other-Querier-Present Timer to [Other Querier Present Interval] and ceases to send general queries on the network if it was the previously elected querier. After its Other-Querier-Present Timer expires, it should begin sending General Queries.

If a router receives an older version General Query, it **MUST** use the oldest version of IGMP on the network. For a detailed description of compatibility issues between IGMP versions, see Section 7.

6.6.3. Building and Sending Specific Queries

6.6.3.1. Building and Sending Group Specific Queries

When a table action "Send Q(G)" is encountered, the Group Timer must be lowered to LMQT. The router must then immediately send a Group Specific Query as well as schedule [Last Member Query Count] - 1 query retransmission(s) to be sent every [Last Member Query Interval] over [Last Member Query Time].

When transmitting a Group Specific Query, if the Group Timer is larger than LMQT, the "Suppress Router-Side Processing" bit is set in the Query Message.

6.6.3.2. Building and Sending Group-and-Source Specific Queries

When a table action "Send Q(G,X)" is encountered by a querier in Table 9 (Section 6.4.2), the following actions must be performed for each of the sources in X of group G, with the Source Timer larger than LMQT:

- Set the number of retransmissions for each source to [Last Member Query Count].
- Lower the Source Timer to LMQT.

The router must then immediately send a Group-and-Source Specific Query as well as schedule [Last Member Query Count] - 1 query retransmission(s) to be sent every [Last Member Query Interval] over [Last Member Query Time]. The contents of these queries are calculated as follows.

When building a Group-and-source Specific Query for group G, two separate Query Messages are sent for the group. The first one has the "Suppress Router-Side Processing" bit set and contains all the sources with retransmission state and timers greater than LMQT. The second has the "Suppress Router-Side Processing" bit clear and contains all the sources with retransmission state and timers lower or equal to LMQT. If either of the two calculated messages does not contain any sources, then its transmission is suppressed.

> Note: If a Group Specific Query is scheduled to be transmitted at the same time as a Group-and-Source Specific Query for the same group, then transmission of the Group-and-Source Specific Query Message with the "Suppress Router-Side Processing" bit set may be suppressed.

7. Interoperation With Older Versions of IGMP

IGMPv3 hosts and routers interoperate with hosts and routers that have not yet been upgraded to IGMPv3. This compatibility is maintained by hosts and routers taking appropriate actions depending on the versions of IGMP operating on hosts and routers within a network.

7.1. Query Version Distinctions

The IGMP version of a Membership Query Message is determined as follows:

- IGMPv1 Query: length = 8 octets AND Max Resp Code field is zero
- IGMPv2 Query: length = 8 octets AND Max Resp Code field is non-zero
- IGMPv3 Query: length >= 12 octets

Query Messages that do not match any of the above conditions (e.g., a Query of length 10 octets) **MUST** be silently ignored.

7.2. Group Member Behavior

7.2.1. In the Presence of Older Version Queriers

In order to be compatible with older version routers, IGMPv3 hosts **MUST** operate in v1 and v2 compatibility modes. IGMPv3 hosts **MUST** keep state per local interface regarding the compatibility mode of each attached network. 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. This variable is kept per interface and is dependent on the version of General Queries received on that interface as well as the Older-Version-Querier-Present Timer for the interface.

In order to switch gracefully between versions of IGMP, hosts keep both an IGMPv1-Querier-Present Timer and an IGMPv2-Querier-Present Timer per interface. IGMPv1-Querier-Present Timer is set to [Older Version Querier Present Interval] seconds whenever an IGMPv1 Membership Query is received. IGMPv2-Querier-Present Timer is set to [Older Version Querier Present Interval] seconds whenever an IGMPv2 General Query is received.

The Host Compatibility Mode of an interface changes whenever an older version query (than the current compatibility mode) is received or when certain timer conditions occur. When the IGMPv1-Querier-Present Timer expires, a host switches to Host Compatibility Mode of IGMPv2 if it has a running IGMPv2 Querier Present timer. If it does not have a running IGMPv2 Querier Present timer, then it switches to Host Compatibility of IGMPv3. When the IGMPv2 Querier Present timer expires, a host switches to Host Compatibility Mode of IGMPv3.

The Host Compatibility Mode variable is based on whether an older version General Query was received in the last [Older Version Querier Present Interval] seconds. The Host Compatibility Mode variable value **MUST NOT** be changed by an older version Group Specific Query. The Host Compatibility Mode is set depending on the following:

Host Compatibility Mode	Timer State
IGMPv3 (default)	IGMPv2 Querier Present not running and IGMPv1 Querier Present not running

Host Compatibility Mode	Timer State
IGMPv2	IGMPv2 Querier Present running and IGMPv1 Querier Present not running
IGMPv1	IGMPv1 Querier Present running

Table 11: Host Compatibility Mode Settings

If a host receives a query that causes its Querier Present timers to be updated and correspondingly its compatibility mode, it should switch compatibility modes immediately.

When Host Compatibility Mode is IGMPv3, a host acts using the IGMPv3 protocol on that interface. When Host Compatibility Mode is IGMPv2, a host acts in IGMPv2 compatibility mode, using only the IGMPv2 protocol, on that interface. When Host Compatibility Mode is IGMPv1, a host acts in IGMPv1 compatibility mode, using only the IGMPv1 protocol on that interface.

An IGMPv1 router will send General Queries with the Max Resp Code set to 0. This **MUST** be interpreted as a value of 100 (10 seconds).

An IGMPv2 router will send General Queries with the Max Resp Code set to the desired Max Response Time, i.e., the full range of this field is linear and the exponential algorithm described in Section 4.1.1 is not used.

Whenever a host changes its compatibility mode, it cancels all its pending response and retransmission timers.

An SSM-aware host that receives an IGMPv1 Query, an IGMPv2 General Query, or an IGMPv2 Group Specific Query for a multicast address in the SSM address range **SHOULD** log an error. It is **RECOMMENDED** that implementations provide a configuration option to disable use of the Host Compatibility Mode to allow networks to operate only in SSM mode. This configuration option **SHOULD** be disabled by default.

7.2.2. In the Presence of Older Version Group Members

An IGMPv3 host may be placed on a network where there are hosts that have not yet been upgraded to IGMPv3. A host **MAY** allow its IGMPv3 Membership Record to be suppressed by either an IGMPv1 Membership Report or an IGMPv2 Membership Report. SSM-aware hosts **MUST NOT** allow its IGMPv3 Membership Record to be suppressed.

7.3. Multicast Router Behavior

7.3.1. In the Presence of Older Version Queriers

IGMPv3 routers may be placed on a network where at least one router on the network has not yet been upgraded to IGMPv3. The following requirements apply:

- If any older versions of IGMP are present on routers, the querier **MUST** use the lowest version of IGMP present on the network. This must be administratively assured; routers that desire to be compatible with IGMPv1 and IGMPv2 **MUST** have a configuration option to act in IGMPv1 or IGMPv2 compatibility modes. When in IGMPv1 mode, routers **MUST** send Periodic Queries with a Max Resp Code of 0 and truncated at the Group Address field (i.e., 8 bytes long) and **MUST** ignore Leave Group messages. They **SHOULD** also warn about receiving an IGMPv2 or IGMPv3 query, although such warnings **MUST** be rate-limited. When in IGMPv2 mode, routers **MUST** send Periodic Queries truncated at the Group Address field (i.e., 8 bytes long) and **SHOULD** also warn about receiving an IGMPv3 query (such warnings **MUST** be rate-limited). They also **MUST** fill in the Max Response Time in the Max Resp Code field, i.e., the exponential algorithm described in Section 4.1.1 is not used.

- If a router is not explicitly configured to use IGMPv1 or IGMPv2 and receives an IGMPv1 Query or IGMPv2 General Query, it **SHOULD** log a warning. These warnings **MUST** be rate-limited.

- It is **RECOMMENDED** that implementations provide a configuration option to disable use of compatibility mode to allow networks to operate only in SSM mode. This configuration option **SHOULD** be disabled by default.

7.3.2. In the Presence of Older Version Group Members

IGMPv3 routers may be placed on a network where there are hosts that have not yet been upgraded to IGMPv3. In order to be compatible with older version hosts, IGMPv3 routers **MUST** operate in v1 and v2 compatibility modes. IGMPv3 routers keep a compatibility mode per Group Record. A group's compatibility mode is determined from the Group Compatibility Mode variable, which can be in one of three states: IGMPv1, IGMPv2, or IGMPv3. This variable is kept per Group Record and is dependent on the version of Membership Reports received for that group as well as the Older-Version-Host-Present Timer for the group.

In order to switch gracefully between versions of IGMP, routers keep an IGMPv1-Host-Present Timer and an IGMPv2-Host-Present Timer per Group Record. The IGMPv1-Host-Present Timer is set to [Older Version Host Present Interval] seconds whenever an IGMPv1 Membership Report is received. The IGMPv2-Host-Present Timer is set to [Older Version Host Present Interval] seconds whenever an IGMPv2 Membership Report is received.

The Group Compatibility Mode of a Group Record changes whenever an older version report (than the current compatibility mode) is received or when certain timer conditions occur. When the IGMPv1-Host-Present Timer expires, a router switches to Group Compatibility Mode of IGMPv2 if it has a running IGMPv2 Host Present timer. If it does not have a running IGMPv2 Host Present timer, then it switches to Group Compatibility Mode of IGMPv3. When the IGMPv2-

Host-Present Timer expires and the IGMPv1-Host-Present Timer is not running, a router switches to Group Compatibility Mode of IGMPv3. Note that when a group switches back to IGMPv3 mode, it takes some time to regain source- specific state information. Source-specific information will be learned during the next General Query, but sources that should be blocked will not be blocked until [Group Membership Interval] after that.

The Group Compatibility Mode variable is based on whether an older version report was received in the last [Older Version Host Present Interval] seconds. The Group Compatibility Mode is set depending on the following:

Group Compatibility Mode	Timer State
IGMPv3 (default)	IGMPv2 Host Present not running and IGMPv1 Host Present not running
IGMPv2	IGMPv2 Host Present running and IGMPv1 Host Present not running
IGMPv1	IGMPv1 Host Present running

Table 12: Group Compatibility Mode Settings

If a router receives a report that causes its older Host Present timers to be updated and correspondingly its compatibility mode, it **SHOULD** switch compatibility modes immediately.

When Group Compatibility Mode is IGMPv3, a router acts using the IGMPv3 protocol for that group.

When Group Compatibility Mode is IGMPv2, a router internally translates the following IGMPv2 messages for that group to their IGMPv3 equivalents:

IGMPv2 Message	IGMPv3 Equivalent
Report	IS_EX({})
Leave	TO_IN({})

Table 13: IGMPv2 Compatibility Mode Message Translation

IGMPv3 BLOCK messages are ignored, as are source-lists in TO_EX() messages (i.e., any TO_EX() message is treated as TO_EX({})).

When Group Compatibility Mode is IGMPv1, a router internally translates the following IGMPv1 and IGMPv2 messages for that group to their IGMPv3 equivalents:

IGMPv2 Message	IGMPv3 Equivalent
v1 Report	IS_EX({})
v2 Report	IS_EX({})

Table 14: IGMPv1 Compatibility Mode Message Translation

In addition to ignoring IGMPv3 BLOCK messages and source-lists in TO_EX() messages as in IGMPv2 Group Compatibility Mode, IGMPv2 Leave messages and IGMPv3 TO_IN() messages are also ignored.

8. List of Timers, Counters, and Their Default Values

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

8.1. Robustness Variable

The Robustness Variable allows tuning for the expected packet loss on a network. If a network is expected to be lossy, the Robustness Variable may be increased. IGMP is robust to (Robustness Variable - 1) packet losses. The Robustness Variable **MUST NOT** be zero and **SHOULD NOT** be one. Default: 2.

8.2. 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 IGMP messages on the network; larger values cause IGMP Queries to be sent less often.

8.3. Query Response Interval

The Query Response Interval uses the Max Response Time to calculate the Max Resp Code that is inserted into the periodic General Queries. Default: 100 (10 seconds).

By varying the [Query Response Interval], an administrator may tune the burstiness of IGMP messages on the network; larger values make the traffic less bursty, as host 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].

8.4. Group Membership Interval

The Group Membership Interval is the amount of time that must pass before a multicast router decides there are no more members of a group or a particular source on a network.

This value **MUST** be ([Robustness Variable] times [Query Interval]) plus (2 * [Query Response Interval]).

8.5. Other Querier Present Interval

The Other Querier Present Interval is the length of time that must pass before a multicast router decides that there is no longer another multicast router that should be the querier. This value **MUST** be ([Robustness Variable] times [Query Interval]) plus (0.5 times [Query Response Interval]).

8.6. Startup Query Interval

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

8.7. Startup Query Count

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

8.8. Last Member Query Interval

The Last Member Query Interval (LMQI) is the Max Response Time used to calculate the Max Resp Code that is inserted into Group Specific Queries sent in response to Leave Group messages. It is also the Max Response Time used in calculating the Max Resp Code for Group-and-Source Specific Query Messages. Default: 10 (1 second).

Note that for values of LMQI greater than 12.8 seconds, a limited set of values can be represented, corresponding to sequential values of Max Resp Code. When converting a configured time to a Max Resp Code value, it is recommended to use the exact value, if possible, or the next lower value if the requested value is not exactly representable.

This value may be tuned to modify the leave latency of the network. A reduced value results in reduced time to detect the loss of the last member of a group or source.

8.9. Last Member Query Count

The Last Member Query Count is the number of Group Specific Queries sent before the router assumes there are no local members. The Last Member Query Count is also the number of Group-and-Source Specific Queries sent before the router assumes there are no listeners for a particular source. Default: [Robustness Variable].

8.10. Last Member Query Time

The Last Member Query Time is the time value represented by [Last Member Query Interval] times [Last Member Query Count]. It is not a tunable value, but it may be tuned by changing its components.

8.11. Unsolicited Report Interval

The Unsolicited Report Interval is the time between repetitions of a host's initial report of membership in a group. Default: 1 second.

8.12. Older Version Querier Present Interval

The Older Version Querier Present Interval is the timeout for transitioning a host back to IGMPv3 mode once an older version query is received. When an older version query is received, hosts set their Older-Version-Querier-Present Timer to [Older Version Querier Present Interval].

It is **RECOMMENDED** to use the default values for calculating the interval value as hosts do not know the values configured on the querying routers. This value **SHOULD** be [Robustness Variable] times [Query Interval] plus (10 times the Max Response Time in the last received Query Message).

8.13. Older Host Present Interval

The Older Host Present Interval is the timeout for transitioning a group back to IGMPv3 mode once an older version report is sent for that group. When an older version report is received, routers set their Older-Host-Present Timer to [Older Host Present Interval].

This value **MUST** be ([Robustness Variable] times [Query Interval]) plus [Query Response Interval].

8.14. Configuring Timers

This section is meant to provide advice to network administrators on how to tune these settings to their network. Ambitious router implementations might tune these settings dynamically based upon changing characteristics of the network.

8.14.1. Robustness Variable

The Robustness Variable tunes IGMP to expected losses on a link. IGMPv3 is robust to ([Robustness Variable] - 1) packet losses, e.g., if the Robustness Variable is set to the default value of 2, IGMPv3 is robust to a single packet loss but may operate imperfectly if more losses occur. On lossy subnetworks, the Robustness Variable should be increased to allow for the expected level of packet loss. However, increasing the Robustness Variable increases the leave latency of the subnetwork. (The leave latency is the time between when the last member stops listening to a source or group and when the traffic stops flowing.)

8.14.2. Query Interval

The overall level of periodic IGMP traffic is inversely proportional to the Query Interval. A longer Query Interval results in a lower overall level of IGMP traffic. The Query Interval **MUST** be equal to or longer than the Max Response Time inserted in General Query Messages.

8.14.3. Max Response Time

The burstiness of IGMP traffic is inversely proportional to the Max Response Time. A longer Max Response Time will spread Report messages over a longer interval. However, a longer Max Response Time in Group Specific and Source-and-Group Specific Queries extends the leave latency. (The leave latency is the time between when the last member stops listening to a source or group and when the traffic stops flowing.) The expected rate of Report messages can be calculated by dividing the expected number of Reporters by the Max Response Time. The Max Response Time may be dynamically calculated per Query by using the expected number of Reporters for that Query as follows:

Query Type	Expected Number of Reporters
General Query	All systems on the subnetwork
Group Specific Query	All systems that had expressed interest in the group on the subnetwork
Source-and-Group Specific Query	All systems on the subnetwork that had expressed interest in the source and group

Table 15: Expected Number of IGMP Reporters

A router is not required to calculate these populations or tune the Max Response Time dynamically; these are simply guidelines.

9. Security Considerations

IGMP provides any form of confidentiality. This means any device on a link can passively receive any IGMP message on the link. Such access can lead to privacy concerns around potentially sensitive multicast groups or the ability to identify/map the devices on a link.

We consider the ramifications of a forged message of each type and describe the usage of an IPsec Authentication Header (AH) to authenticate messages if desired.

9.1. 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 Leave messages, traffic might flow to groups with no members for up to [Group Membership Interval].

A Denial-of-Service (DoS) attack on a host could be staged through forged Group-and- Source Specific Queries. The attacker can find out about membership of a specific host with a General Query. After that, it could send a large number of Group-and-Source Specific Queries, each with a large source-list and the Maximum Response Time set to a large value. The host will have to

store and maintain the sources specified in all of those Queries for as long as it takes to send the delayed response. This would consume both memory and CPU cycles in order to augment the recorded sources with the source-lists included in the successive Queries.

To protect against such a DoS attack, a host stack implementation could restrict the number of Group-and-Source Specific Queries per group membership within this interval and/or record only a limited number of sources.

Forged Query Messages from the local network can be easily traced. There are three measures necessary to defend against externally forged Queries:

- Routers **SHOULD NOT** forward Queries. This is easier for a router to accomplish if the Query carries the Router Alert option.
- Hosts **SHOULD** ignore v2 or v3 Queries without the Router Alert option.
- Hosts **SHOULD** ignore v1, v2, or v3 General Queries sent to a multicast address other than 224.0.0.1, the all-systems address.

9.2. Current-State Report Messages

A forged Report message may cause multicast routers to think there are members of a group on a network when there are not. Forged Report messages from the local network are meaningless, as joining a group on a host is generally an unprivileged operation, so a local user may trivially gain the same result without forging any messages. Forged Report messages from external sources are more troublesome; there are two defenses against externally forged Reports:

1. Ignore the Report if you cannot identify the source address of the packet as belonging to a network assigned to the interface on which the packet was received. This solution means that Reports sent by mobile hosts without addresses on the local network will be ignored. Report messages with a source address of 0.0.0.0 **SHOULD** be accepted on any interface.
2. Ignore Report messages without Router Alert options [RFC2113] and require routers to not forward Report messages. (The requirement is not a requirement of generalized filtering in the forwarding path, as the packets already have Router Alert options in them.) This solution breaks backwards compatibility with implementations of IGMPv1 or earlier versions of IGMPv2 that did not require a Router Alert.

A forged v1 Report message may put a router into "v1 members present" state for a particular group, meaning that the router will ignore Leave messages. This can cause traffic to flow to groups with no members for up to [Group Membership Interval]. This can be solved by providing routers with a configuration switch to ignore v1 messages completely. This breaks automatic compatibility with v1 hosts, so it should only be used in situations where "fast leave" is critical.

A forged v2 Report message may put a router into "v2 members present" state for a particular group, meaning that the router will ignore IGMPv3 source-specific state messages. This can cause traffic to flow from unwanted sources for up to [Group Membership Interval]. This can be

solved by providing routers with a configuration switch to ignore v2 messages completely. This breaks automatic compatibility with v2 hosts, so it should only be used in situations where source include and exclude is critical.

9.3. State-Change Report Messages

A forged State-Change Report message will cause the Querier to send out Group Specific or Source-and-Group Specific Queries for the group in question. This causes extra processing on each router and on each member of the group, but it cannot cause loss of desired traffic. There are two defenses against externally forged State-Change Report messages:

1. Ignore the State-Change Report message if you cannot identify the source address of the packet as belonging to a subnet assigned to the interface on which the packet was received. This solution means that State-Change Report messages sent by mobile hosts without addresses on the local subnet will be ignored. State- Change Report messages with a source address of 0.0.0.0 **SHOULD** be accepted on any interface.

2. Ignore State-Change Report messages without Router Alert options [RFC2113] and require routers to not forward State-Change Report messages. (The requirement is not a requirement of generalized filtering in the forwarding path, as the packets already have Router Alert options in them.)

9.4. IPsec Usage

In addition to these measures, IPsec in AH mode [RFC4302] may be used to protect against remote attacks by ensuring that IGMPv3 messages came from a system on the LAN (or, more specifically, from a system with the proper key). When using IPsec, the messages sent to 224.0.0.1 and 224.0.0.22 should be authenticated using AH. When keying, there are two possibilities:

1. Use a symmetric signature algorithm with a single key for the LAN (or a key for each group). This allows validation that a packet was sent by a system with the key. This has the limitation that any system with the key can forge a message; it is not possible to authenticate the individual sender precisely. It also requires disabling IPsec's Replay Protection.

2. When appropriate key management standards have been developed, use an asymmetric signature algorithm. All systems need to know the public key of all routers, and all routers need to know the public key of all systems. This requires a large amount of key management but has the advantage that senders can be authenticated individually so, e.g., a host cannot forge a message that only routers should be allowed to send.

This solution only directly applies to Query and Leave messages in IGMPv1 and IGMPv2 as Reports are sent to the group being reported, and it is not feasible to agree on a key for host-to-router communication for arbitrary multicast groups.

10. IANA Considerations

All IGMP types described in this document are managed via [BCP57].

IANA has replaced each reference to [RFC3376] with a reference to this document in both the "IGMP Type Numbers" and "IPFIX Information Elements" registries.

11. References

11.1. Normative References

[BCP57]   Best Current Practice 57, <https://www.rfc-editor.org/info/bcp57>.
            At the time of writing, this BCP comprises the following:

            Haberman, B., Ed., "IANA Considerations for Internet Group Management
            Protocols", BCP 57, RFC 9778, DOI 10.17487/RFC9778, March 2025, <https://
            www.rfc-editor.org/info/rfc9778>.

[RFC1112]   Deering, S., "Host extensions for IP multicasting", STD 5, RFC 1112, DOI 10.17487/
            RFC1112, August 1989, <https://www.rfc-editor.org/info/rfc1112>.

[RFC2113]   Katz, D., "IP Router Alert Option", RFC 2113, DOI 10.17487/RFC2113, February
            1997, <https://www.rfc-editor.org/info/rfc2113>.

[RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14,
            RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/
            rfc2119>.

[RFC2236]   Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, DOI
            10.17487/RFC2236, November 1997, <https://www.rfc-editor.org/info/rfc2236>.

[RFC4302]   Kent, S., "IP Authentication Header", RFC 4302, DOI 10.17487/RFC4302, December
            2005, <https://www.rfc-editor.org/info/rfc4302>.

[RFC4604]   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, DOI 10.17487/RFC4604,
            August 2006, <https://www.rfc-editor.org/info/rfc4604>.

[RFC4607]   Holbrook, H. and B. Cain, "Source-Specific Multicast for IP", RFC 4607, DOI
            10.17487/RFC4607, August 2006, <https://www.rfc-editor.org/info/rfc4607>.

[RFC8174]   Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP
            14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/
            rfc8174>.

11.2. Informative References

[RFC1071]   Braden, R., Borman, D., and C. Partridge, "Computing the Internet checksum",
            RFC 1071, DOI 10.17487/RFC1071, September 1988, <https://www.rfc-editor.org/
            info/rfc1071>.

[RFC3376]   Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet
            Group Management Protocol, Version 3", RFC 3376, DOI 10.17487/RFC3376,
            October 2002, <https://www.rfc-editor.org/info/rfc3376>.

[RFC3569]   Bhattacharyya, S., Ed., "An Overview of Source-Specific Multicast (SSM)", RFC
            3569, DOI 10.17487/RFC3569, July 2003, <https://www.rfc-editor.org/info/
            rfc3569>.

[RFC3678]   Thaler, D., Fenner, B., and B. Quinn, "Socket Interface Extensions for Multicast
            Source Filters", RFC 3678, DOI 10.17487/RFC3678, January 2004, <https://
            www.rfc-editor.org/info/rfc3678>.

Appendix A.   Design Rationale

A.1. The Need for State-Change Messages

IGMPv3 specifies two types of Membership Reports: Current-State and State-Change. This section describes the rationale for needing both types of Reports.

Routers need to distinguish Membership Reports that were sent in response to Queries from those that were sent as a result of a change in interface state. Membership reports that are sent in response to Membership Queries are used mainly to refresh the existing state at the router; they typically do not cause transitions in state at the router. Membership Reports that are sent in response to changes in interface state require the router to take some action in response to the received report (see Section 6.4).

The inability to distinguish between the two types of reports would force a router to treat all Membership Reports as potential changes in state, and it could result in increased processing at the router as well as an increase in IGMP traffic on the network.

A.2. Host Suppression

In IGMPv1 and IGMPv2, a host would cancel sending pending Membership Reports if a similar report was observed from another member on the network. In IGMPv3, this suppression of host Membership Reports has been removed. The following points explain the reasons behind this decision.

1. Routers may want to track per-host membership status on an interface. This allows routers to implement fast leaves (e.g., for layered multicast congestion control schemes) as well as track membership status for possible accounting purposes.
2. Membership Report suppression does not work well on bridged LANs. Many bridges and Layer 2 / Layer 3 switches that implement IGMP snooping do not forward IGMP messages across LAN segments in order to prevent Membership Report suppression. Removing Membership Report suppression eases the job of these IGMP snooping devices.
3. By eliminating Membership Report suppression, hosts have fewer messages to process; this leads to a simpler state machine implementation.

4. In IGMPv3, a single Membership Report now bundles multiple multicast Group Records to decrease the number of packets sent. In comparison, the previous versions of IGMP required that each multicast group be reported in a separate message.

A.3. Switching Router Filter Modes from EXCLUDE to INCLUDE

If hosts exist in both EXCLUDE and INCLUDE modes for a single multicast group in a network, the router must be in EXCLUDE mode as well (see Section 6.2.1). In EXCLUDE mode, a router forwards traffic from all sources unless that source exists in the exclusion source-list. If all hosts in EXCLUDE mode cease to exist, it would be desirable for the router to switch back to INCLUDE mode seamlessly without interrupting the flow of traffic to existing receivers.

One of the ways to accomplish this is for routers to keep track of all sources desired by hosts that are in INCLUDE mode even though the router itself is in EXCLUDE mode. If the Group Timer now expires in EXCLUDE mode, it implies that there are no hosts in EXCLUDE mode on the network (otherwise, a Membership Report from that host would have refreshed the Group Timer). The router can then switch to INCLUDE mode seamlessly with the list of sources currently being forwarded in its source-list.

Appendix B.   Summary of Changes from IGMPv2

While the main additional feature of IGMPv3 is the addition of source filtering, the following is a summary of other changes from [RFC2236].

- State is maintained as Group + List-of-Sources, not simply Group as in IGMPv2.
- Interoperability with IGMPv1 and IGMPv2 systems is defined as operations on the IGMPv3 state.
- The IP service interface has changed, to allow specification of source-lists.
- The Querier includes its Robustness Variable and Query Interval in Query packets to allow synchronization of these variables on non- Queriers.
- The Max Response Time in Query Messages has an exponential range, changing the maximum from 25.5 seconds to about 53 minutes, for use on links with a huge number of systems.
- Hosts retransmit state-change messages for increased robustness.
- Additional data sections are defined, to allow later extensions.
- Report packets are sent to 224.0.0.22, to assist Layer 2 switches in snooping.
- Report packets can contain multiple Group Records, to allow reporting of full current state using fewer packets.
- Hosts no longer perform suppression, to simplify implementations and permit explicit membership tracking.
- A new S flag in Query Messages fixes robustness issues, which were also present in IGMPv2.

Appendix C.   Summary of Changes from RFC 3376

The following is a list of changes made since [RFC3376] was published.

- Modified the definition of Older Version Querier Present Interval to address Erratum 4375.
- Modified the metadata to fix the Obsoletes vs. Updates relationship with [RFC2236] per Erratum 1501.
- Updated the introductory text to describe the Updates relationship with [RFC2236] per Erratum 7339.
- Updated the definition of Group Membership Interval to address Erratum 6725.
- Updated the text relating to the Router Filter Mode to address Erratum 5562.
- Clarified the use of General Queries in the Querier election process.

Acknowledgments

Contributors

Brad Cain, Steve Deering, Isidor Kouvelas, Bill Fenner, and Ajit Thyagarajan are the authors of [RFC3376], which forms the bulk of the content contained herein.

Anuj Budhiraja, Toerless Eckert, Olufemi Komolafe, and Tim Winters have contributed valuable content to this specification.

Author's Address

Brian Haberman (EDITOR)
Johns Hopkins University Applied Physics Lab
Email: brian@innovationslab.net