MAGMA Working Group
Internet Draft
January 2002
Expiration Date: July 2002

M. Christensen
jCAPS
F. Solensky

# IGMP and MLD snooping switches
<draft-ietf-magma-snoop-01.txt>

## Status of this Memo

This document is the successor of draft-ietf-magma-snoop-00 which was presented at the 52'nd IETF in Salt Lake City. The main differences between this and the previous version is a restructuring of the draft to introduce the main result as early as possible. Also the draft has been trimmed to a smaller size. No new results are presented, as the draft is expected to published as Informational within the next four months.

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026 [RFC2026].

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

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

## Abstract

This memo describes the requirements for IGMP and MLD snooping switches. The requirements for IGMPv2 snooping switches are based on best current practices. IGMPv3 and MLDv2 snooping are also covered in this draft although we are not aware of any such implementations at the time of writing.

The memo also describes the interoperability problems and issues that can arise when a mixed deployment of IGMPv3 and IGMPv2 capable hosts and routers are interconnected by a switch with IGMP snooping capabilities.

Areas which are of relevance to IGMP and MLD snooping switches, such as link layer topology changes and Ethernet specific encapsulation issues are also considered.

It is intended as an accompanying document to the IGMPv3 and MLDv2 specifications.

## 1. Introduction

In recent years, a number of commercial vendors have introduced products described as "IGMP snooping switches" to the market. These devices do not adhere to the conceptual model that provides the strict separation of functionality between different communications layers in the ISO model and instead utilizes information in the upper level protocol headers as factors to be considered in the processing at the lower levels.

This is analogous to the manner in which a router can act as a firewall by looking into the transport protocol's header before allowing a packet to be forwarded to its destination address.

In the case of multicast traffic, an IGMP snooping switch provides the benefit of conserving bandwidth on those segments of the network where no node has expressed interest in receiving packets addressed to the group address. This is in contrast to normal switch behaviour where multicast traffic is typically forwarded on all interfaces.

Many switch datasheets state support for IGMP snooping, but no requirements for this exist today. It is the authors hope that the information presented in this draft will supply this information.

The requirements presented here is based on the following information sources: The IGMP specifications [RFC112][RFC2236][IGMPv3], vendor supplied technical documents [CISCO], bug reports [MSOFT], discussions with people invovled in design of IGMP snooping switches, MAGMA mailinglist discussions, and on replies by switch vendors to an

implementation questionnaire.

The discussions in this document are based on IGMP which applies to IPv4 only. For IPv6 we must use MLD instead. Because MLD is based on IGMP we do not repeat the whole discussion and requirements for MLD





snooping switches. Instead we point out the few cases where there is a difference compared to IGMP.

## 2. IGMP Snooping Requirements

The following sections list the requirements for an IGMP snooping switch. The requirement is stated and is supplemented by a discussion. All implementation discussions are examples only and there may well be other ways to achieve the same functionality.

### 2.1. Forwarding rules

The IGMP snooping functionality is here separated in a control section (IGMP forwarding) and data section (Data forwarding).

#### 2.1.1. IGMP Forwarding Rules

1) A snooping switch MUST only forward IGMP Membership Reports on ports where multicast routers are attached. Alternatively stated: A snooping switch MUST NOT forward IGMP Membership Reports to ports on which only hosts are attached.

This is the main IGMP snooping functionality. Sending membership reports to other hosts can result (For IGMPv2 and IGMPv1) in the unwanted prevention of a host wishing to join a specific multicast group. This is not a problem in a IGMPv3 only network because there is no cancellation of IGMP Membership reports.

The switch supporting IGMP snooping MUST maintain a list of multicast routers and the ports on which they are attached. This list SHOULD be built using IGMP Multicast Router Discovery [MRDISC]. IGMP snooping

switches MAY alternatively build this list based on

a) The arrival port for IGMP Queries (sent by multicast routers) or

b) List of ports configured (by management) as having multicast routers attached.

Implementation example: IGMP snooping can be achieved by redirecting all IGMP packets (IP packets with IP-PROTO = 2) to the CPU (or similar higher layer entity) for IGMP message processing and table management.

2) IGMP snooping switches MAY implement "proxy-reporting" in which





reports received from downstream hosts are summarized and used to build internal membership states. An IGMP proxy-reporting switch would then report it's own state in response to upstream queriers. If the switch does not have an IP address it SHOULD use the address 0.0.0.0 as source in these reports.

An IGMP proxy-reporting switch may act as Querier for the downstream hosts while proxy reporting to the 'real' upstream queriers.

It should be noted that there may be multiple IGMP proxy-reporting switches in the network all using the 0.0.0.0 source IP address. In this case the switches can be identified by their, link layer source MAC address.

IGMP membership reports should not be rejected because of a source IP address of 0.0.0.0.

3) The switch that supports IGMP snooping MUST forward all unrecognized IGMP messages and MUST NOT attempt to make use of any information beyond the end of the network layer header. In particular, messages where any reserved fields are non-zero MUST NOT be subject to "normal" snooping since this could indicate an incompatible change to the message format.

4) An IGMP snooping switch SHOULD be aware of link layer topology changes. Following a topology change the switch SHOULD initiate the transmission of a General Query on all ports in order to reduce network convergence time.

5) An IGMP snooping switch MUST discard from the IGMP snooping process IGMP packets where IP or IGMP headers have checksum errors.

### 2.1.2. Data Forwarding Rules

6) Packets with a destination IP (DIP) address in the 224.0.0.X range which are *not* IGMP MUST be forwarded on all ports.

This requirement is based on fact that many hosts exist today, which does not Join IP multicast addresses in this range before sending or listening to IP multicast. Furthermore since the 224.0.0.X address range is defined as link local (not to be routed) it seems unnecessary to keep state for each address in this range.

7) Packets with a destination IP address outside 224.0.0.X which are *not* IGMP SHOULD be forwarded according to group based port membership





tables and MUST also be forwarded on router ports.

This is the core IGMP snooping requirement for the data path.

Discussion: An implementation could maintain separate membership and multicast router tables in software and then "merge" these tables into a current forwarding cache.

8) If a switch receives a *non* IGMP multicast packet without having first processed Membership Reports for the group address, it MUST forward the packet on all ports. In other words, the switch must allow for the possibility that connected hosts and routers have been upgraded to support future versions or extensions of IGMP that the switch does not yet recognize. A switch MAY have a configuration option that suppresses this operation, but default behavior MUST be to allow flooding of unregistered packets.

9) IGMP snooping switches MAY maintain forwarding tables based on either MAC addresses or IP addresses. If a switch supports both types of forwarding tables then the default behavior MUST be to use IP addresses.

Discussion: Forwarding based on MAC addresses is subject to the problem associated with the 32-fold IP address to 1 MAC address mapping.

10) Switches which rely on information in the IP header SHOULD verify that the IP header checksum is correct. If the checksum fails the packet SHOULD be silently discarded.

## 2.2. IGMP snooping related problems

A special problem arise in the network consisting of IGMPv3 routers as well as IGMPv2 and IGMPv3 hosts interconnected by a IGMPv2 snooping switch. IGMPv3 has a mechanism to fall back to IGMPv2 when receiving IGMPv2 membership reports. This means that the network will converged on IGMPv2 eventually. However, the convergence time will lead to supression of v3 Hosts for several minutes.

Therefore it is recommended that in such a network, the multicast router is configured to use IGMPv2.

# 3. IPv6 Considerations

In order to avoid confusion, the previous discussions have been based on IGMPv3 functionality which only applies to IPv4 multicast. In the case of IPv6 most of the above discussions are still valid with a few





exceptions which we will describe here.

In IPv6 the protocol for multicast group maintenance is called Multicast Listener Discovery (MLDv2). IPv6 is not widely deployed today and neither is IPv6 multicast. However, it is anticipated that at some time IPv6 switches capable of MLD snooping will appear.

The three main differences between IGMPv3 and MLDv2 are

- MLDv2 uses ICMPv6 message types instead of IGMP message types.
- The ethernet encapsulation is a mapping of 32bits of the 128bit DIP addresses into 48bit DMAC addresses [IPENCAPS].
- Multicast router discovery is done using Neighbor Discovery Pro-

tocol (NDP) for IPv6. NDP uses ICMPv6 message types.

A minor difference which applies to the requirements section is that in IPv6 there is no checksum in the IP header. This is the reason that the checksum validation requirement is listed as a MAY.

The fact that MLDv2 is using ICMPv6 adds new requirements to a snooping switch because ICMPv6 has multiple uses aside from MLD. This means that it is no longer sufficient to detect that the next-header field of the IP header is ICMPv6 in order to redirect packets to the CPU. If this was the case the CPU queue assigned for MLD would potentially be filled with non-MLD related packets. Furthermore ICMPv6 packets destined for other hosts would not reach their destination. A solution is either to require that the snooping switch looks further into the packets or to be able to detect a multicast DMAC address in conjunction with ICMPv6. The first solution is desirable only if it is configurable which message types should trigger a CPU redirect and which should not. The reason is that a hardcoding of message types is unflexible for the introduction of new message types. The second solution introduces the risk of new protocols, which are not related to MLD but uses ICMPv6 and multicast DMAC addresses wrongly being identified as MLD. It is suggested that solution one is the preferred if the switch is capable of triggering CPU redirects on individual ICMPv6 message types. If this is not the case then use solution two.

The mapping from IP multicast addresses to multicast DMAC addresses introduces a potentially enormous overlap. The structure of an IPv6 multicast address is shown in figure 5. Theoretically 2**80, two to the power of 80 (128 - 8 - 4 - 4 - 32) unique DIP addresses could map to one DMAC address. This should be compared to 2**5 in the case of IPv4.

Initial allocation of IPv6 multicast addresses, however, uses only the





lower 32 bits of group ID. This eliminates the address ambiguity for the time being but it should be noted that the allocation policy may change in the future.

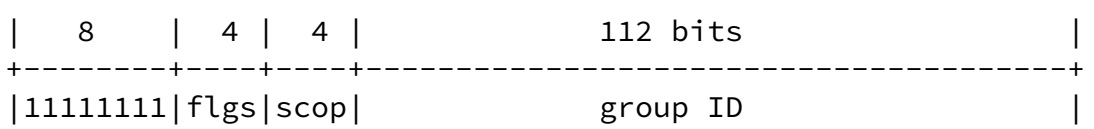

Figure 5

In the case of IPv6 forwarding can be made on the basis of DMAC addresses in the forseable future.

Finally, we mention the reserved address range FF0X:0:0:0:0:0:X:X where X is any value. This range is similar to 224.0.0.X for IPv4 and is reserved to routing protocols and resource discovery [RFC2375]. In the case of IPv6 it is suggested that packets in this range are forwarded on all ports if they are not MLD packets.

## 4. Security Considerations

Security considerations for IGMPv3 are accounted for in [IGMPv3]. The introduction of IGMP snooping switches adds the following considerations with regard to IP multicast.

The exclude source failure which could cause traffic from sources that are 'black listed' to reach hosts that have requested otherwise. This can also occur in certain network topologies without IGMP snooping.

It is possible to generate packets which make the switch wrongly believe that there is a multicast router on the segment on which the source is attached. This will potentially lead to excessive flooding on that segment. The authentication methods discussed in [IGMPv3] will also provide protection in this case.

IGMP snooping switches which rely on the IP header of a packet for their operation and which do not validate the header checksum potentially will forward packets on the wrong ports. Even though the IP headers are protected by the ethernet checksum this is a potential vulnerability.

Generally though, it is worth to stress that IP multicast must so far be considered insecure until the work of for example the suggested Multicast Security (MSEC) working group or similar is completed or at least has matured.





## 5. IGMP Questionnaire

As part of this work the following questions were asked both on the MAGMA discussion list and sent to known switch vendors implementing IGMP snooping. The individual contributions have been anonymized upon request.

The questions were:

Q1 Does your switches perform IGMP Join aggregation? In other words, are IGMP joins intercepted, absorbed by the hardware/software so that only one Join is forwarded to the querier?

Q2 Is multicast forwarding based on MAC addresses? Would datagrams addressed to multicast IP addresses 224.1.2.3 and 239.129.2.3 be forwarded on the same ports-groups?

Q3 Is it possible to forward multicast datagrams based on IP addresses (not routed). In other words, could 224.1.2.3 and 239.129.2.3 be forwarded on different port-groups with unaltered TTL?

Q4 Are multicast datagrams within the range 224.0.0.1 to 224.0.0.255 forwarded on all ports whether or not IGMP Joins have been sent?

Q5 Are multicast frames within the MAC address range 01:00:5E:00:00:01 to 01:00:5E:00:00:FF forwarded on all ports whether or not IGMP joins have been sent?

Q6 Does your switch support forwarding to ports on which IP multicast routers are attached in addition to the ports where IGMP Joins have been received?

Q7 Is your IGMP snooping functionality fully implemented in hardware?

Q8 Is your IGMP snooping functionality partly software implemented?

Q9 Can topology changes (for example spanning tree configuration changes) be detected by the IGMP snooping functionality so that for example new queries can be sent or tables can be updated to ensure robustness?

The answers were:

| | Switch Vendor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Q1 Join aggregation | x | x | x | |
| Q2 Layer-2 forwarding | x | x | x | x |
| Q3 Layer-3 forwarding | (1) | | (1) | |
| Q4 224.0.0.X aware | (1) | x | (1) | (2) |
| Q5 1:00:5e:0:0:X aware | x | x | x | (2) |
| Q6 Mcast router list | x | x | x | x |
| Q7 Hardware implemented | | | | |
| Q8 Software assisted | x | x | x | x |
| Q9 Topology change aware | x | x | x | x |

x Means that the answer was Yes.
(1) In some products (typically high-end) Yes, in others No.
(2) Currently no, but will be real soon.

## 6. IPR Issues

It appears that a number of patents have been filed which may apply to this draft or parts thereof. None of these patents, listed below, have been filed by the authors of this draft or companies in which they are or have been employed.

We, the authors, have not tried to evaluate whether these patents are violated by implementing IGMP snooping according to this document. The IETF/IESG have been notified about the patents.

International patent: WO 96/34474, Oct. 31 1996, Title: "Broadcast transmission in a data network"

US patent: Number 5,608,726, Mar. 4, 1997 (filed 1995) Title: "Networking Bridge with Multicast forwarding table"

US patent: Number 5,898,686, Apr. 27, 1999 (filed 1996) Title: "Networking Bridge with Multicast forwarding table"

Australian patent: Application No. 65440/98 Title: "System, Device, and Method for Managing Multicast Group Memberships in a Multicast Network"

## 7. References

[BRIDGE] IEEE 802.1D, "Media Access Control (MAC) Bridges"

[CISCO] Cisco Tech Notes, "Multicast In a Campus Network: CGMP and IGMP snooping", http://www.cisco.com/warp/public/473/22.html

[IANA] Internet Assigned Numbers Authority, "Internet Multicast Addresses", http://www.isi.edu/in-notes/iana/assignments/multicast-addresses

[IGMPv3] Cain, B., "Internet Group Management Protocol, Version 3", draft-ietf-idmr-igmp-v3-06.txt, November 2000

[IPENCAPS] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", RFC2464, December 1998.

[MLDv2] Vida, R., "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", draft-vida-mld-v2-00.txt, February 2001.

[MRDISC] Biswas, S. "IGMP Multicast Router Discovery", draft-ietf-idmr-igmp-mrdisc-06.txt, May 2001.

[MSOFT] Microsoft support article Q223136, "Some LAN Switches with IGMP Snooping Stop Forwarding Multicast Packets on RRAS Startup", http://support.microsoft.com/support/kb/articles/Q223/1/36.ASP

[RFC1112] Deering, S., "Host Extensions for IP Multicasting", RFC 1112, August 1989.

[RFC2026] Bradner, S. "The Internet Standards Process -- Revision 3", RFC2026, October 1996.





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

[RFC2375] Hinden, R. "IPv6 Multicast Address Assignments", RFC2375, July 1998.

## 8. Acknowledgements

We would like to thank (in no identifiable order) Hugh Holbrook, Bill Fenner, Yiqun Cai, Edward Hilquist, Toerless Eckert, Kevin Humphries, Karen Kimball and Martin Bak for comments and suggestions on this document. Furthermore, the following companies are acknowledged for their contributions: Vitesse Semiconductor Corporation, Cisco Systems, Hewlett-Packard, Enterasys Networks.

## 9. Author's Addresses:

Morten Jagd Christensen
jCAPS
Begoniavej 13
2820 Gentofte
DENMARK
email: morten@jcaps.com

Frank Solensky

email: solenskyf@acm.org