Internet Engineering Task Force
Internet Draft
Expiration Date: May 2003

P. Savola
CSC/FUNET

November 2002

IPv6 Multicast Deployment Issues

draft-savola-v6ops-multicast-issues-01.txt

Status of this Memo

Abstract

There are many issues concerning the deployment and implementation, and to a lesser degree, specification of IPv6 multicast. This memo describes known problems, trying to raise awareness. Currently, global IPv6 interdomain multicast is completely impossible except using SSM: there is no way to convey information about multicast sources between PIM RPs. Site-scoped multicast is also problematic when used alongside to global multicast because of that. A few possible solutions are outlined or referred to. In addition, an issue regarding link-local multicast-blocking Ethernet switches is brought up. Finally, a feature request for MLD snooping switches is noted.

Internet Draft draft-savola-v6ops-multicast-issues-01.txt November 2002

Table of Contents

1. Introduction

There are many issues concerning the deployment and implementation, and to a lesser degree, specification of IPv6 multicast. This memo describes known problems, trying to raise awareness.

Currently, global IPv6 interdomain multicast is completely impossible except using SSM: there is no way to convey information about multicast sources between PIM RPs. Site-scoped multicast is also problematic when used alongside to global multicast because of that. A few possible solutions are outlined or referred to. These are discussed in section 2.

In addition, an issue regarding link-local multicast -blocking Ethernet switches is brought up. Finally, a feature request for MLD snooping switches is noted. These are discussed in sections 3 and 4, respectively.

[MULTIGAP] analyses the more generic set of issues with multicast; this memo focuses on critical issues regarding IPv6.

2. Issues with Multiple PIM Domains and Any Source Multicast

For both administrative and technical reasons, there must be multiple Protocol-Independent Multicast (PIM) [PIM] domains in the Internet, which means there will be multiple PIM Rendezvous Points (RPs) -- and communication mechanisms between these RPs will become critical.

These issues only come up with classical Any Source Multicast; Source-Specific Multicast [SSM] does not require RPs and is not affected, as the multicast "channel" is identified by the combination <source address, group address> and can be communicated out-of-band.

In IPv4, notification of multicast sources between these PIM RPs is done with Multicast Source Discovery Protocol (MSDP) [MSDP]. Many consider this a sub-optimal, but unfortunately necessary, solution; when it was specified, it was only meant as a "stop-gap" measure.

Below, some issues and solutions/work-arounds are described.

2.1. Changing the Multicast Usage Model

As "Any Source Multicast" -model has been found to be complex and a bit problematic, there may be an incentive to move to SSM for good (at least for most cases). Below two paragraphs are adapted from [PIMSO]:

The most serious criticism of the SSM architecture is that it does not support shared trees which may be useful for supporting many-to-many applications. In the short-term this is not a serious concern since the multicast application space is likely to be dominated by one-to-many applications. Some other classes of multicast applications that are likely to emerge in the future are few-to-few (e.g. private chat rooms, whiteboards), few-to-many (e.g., video conferencing, distance learning) and many-to-many (e.g., large chat rooms, multi-user games). The first two classes can be easily handled using a few one-to-many source-based trees.

The issue of many-to-many multicasting service on top of a SSM architecture is an open issue at this point. However, some feel that even many-to-many applications should be handled with multiple one-to-many instead of shared trees.

In any case, even though SSM would avoid mentioned problems, it is far from being generally implemented, much less deployed, yet.

Nonetheless, few seem to realize that SSM is currently the only way to get global interdomain multicast in IPv6.

2.2. Implementing MSDP for IPv6

One could argue that currently, the easiest stop-gap solution (to a stop-gap solution) would be to specify IPv6 TLV's for MSDP. This would be fairly straightforward, and existing implementations would probably be relatively easily modifiable.

There has been some resistance to this, as MSDP was not supposed to last this long in the first place, though. Whether this is a "good" or "bad" decision is a matter of opinion.

2.3. Implementing Another Multicast Routing Protocol

One possibility might be to specify and/or implement a different multicast routing protocol. In fact, Border Gateway Multicast Protocol (BGMP) [BGMP] has been specified for a few years; however, it seems quite complex and there have been no implementations. Lacking deployment experience and specification analysis, it is difficult to say which problems it might solve (and possibly, which new ones to introduce).

In conclusion, looking for a solution in BGMP may not be realistic in this time frame.

2.4. Embedding the Address of the RP in Multicast Address

One way to work around these problems would be to allocate and assign multicast addresses in such a fashion that the address of the RP could be automatically calculated from the multicast address.

At the first glance, this appears to be an impossible problem: the address of the RP, as well as the multicast address, are both 128 bits long; in the general case, embedding one in the other is impossible.

However, making some assumptions about multicast addressing, this is can be done -- a proposed solution is presented in a different memo [V6RPADDR]. Some minor changes in existing PIM specifications would have to be done to take advantage of this, though (but non-modified implementations would be no worse than today).

One should note that MSDP is also used in "Anycast RP" [ANYCASTRP] -technique, for sharing the state information between different RP's in one PIM domain; without further specification, anycast-RP technique could not be used with "embedded RP address" mechanism.

However, a "cold failover" variant of anycast-RP (for redundancy only) would still be possible. In this mechanism, multiple routers

would be configured with the RP address, but only one would be active at the time: if the RP goes down, another takes its place. The multicast state stored in the RP would be lost, though, unless it is copied by some out-of-band mechanism (e.g. placing the backup RP absolutely on-the-path and have it snoop all the relevant packets).

2.5. Site-local Group Scoping

Site-local groups must be their own PIM domains to prevent site-local data leaking to other sites. A more complex possibility would be to implement something resembling "BSR border" feature which would filter out all site-local components in PIM packets: if this is not done very carefully, site-local information will leak to the global network. This is operationally difficult, and PIM working group has come to consensus that a scope boundary MUST also be a a site boundary for certain critical PIM messages (e.g. C-RP and Bootstrap).

Especially if site-local multicast is used (and the site also wants to engage in global multicast), there will be a huge number of domains and communication required between them. This will increase the need for a global multicast solution.

3. Neighbor Discovery Using Multicast

Neighbor Discovery [NDISC] uses link-local multicast in the most common link-layer media, not broadcast as does ARP with IPv4. This may cause some operational problems with some equipment.

The author has seen one brand of managed Ethernet switches, and heard reports of a few unmanaged switches, that do not forward IPv6 link-layer multicast packets to other ports at all. In essence, native IPv6 is impossible with this equipment. Investigation is still going on whether these issues can be worked around.

However, it seems likely this may be a problem with some switches that build multicast forwarding state based on Layer 3 information (and do not support IPv6); state using Layer 2 information would work just fine [MLDSNOOP].

For the deployment of IPv6, it would be important to find out how this can be fixed (e.g. how exactly this breaks specifications) and how one can identify which equipment could case problems like these (and whether there are workarounds).

One workaround might be to implement a toggle in the nodes that would use link-layer broadcast instead of multicast as a fallback solution. However, this would have to be used in all the systems in the same segment one wishes to communicate with.

4. MLD Snooping Ethernet Switches

Especially if multicast traffic is relatively heavy (e.g. video streaming), it becomes particularly important to have some feature like Multicast Listener Discovery (MLD) snooping implemented in some equipment, most importantly Ethernet switches [MLDSNOOP].

In addition, there have been some misunderstandings wrt. which multicast addresses (in particular, link-locals) MLD reports (utilized in the snooping) should be generated for. If all implementations do not generate enough MLD reports, the introduction of MLD snooping could cause them being blocked out. Clarifications and analysis on what MLD snooping switches can reasonably expect would be very important.

One could also argue that MLD snooping might make the devices too complex, requiring the processing of datagrams above the link-layer, and should be discouraged [MULTIGAP]: the whole idea of L2-only devices having be able to peek into L3 datagrams seems like a severe layering violation -- and often the devices aren't upgradeable in any way. Better mechanisms could be having routers tell switches which multicasts to forward where (e.g. [CGMP]) or by using some other mechanisms [GARP].

5. Security Considerations

Only deployment/implementation issues are considered, and these do not have any particular security considerations; security considerations for each technology are covered in the respective specification.

One fairly obvious issue raised in this memo is that if there is no adminsitrative PIM domain border between site-local multicast domains, the site-local traffic could very easily flow into other sites and the Internet.

6. Acknowledgements

Early discussions with Stig Venaas, Jerome Durand, Tim Chown et al. led to the writing of this draft. Brian Haberman offered extensive comments along the way. "Itojun" Hagino brought up the need for MLD snooping in a presentation. Bill Nickless pointed out issues in the gap analysis and provided a pointer to GARP/GMRP; H[illegible]vard Eidnes made a case for a protocol like CGMP. Leonard Giuliano pointed out a more complete analysis of SSM with different kind of applications.

7. References

7.1. Normative References

[NDISC] Narten, T., Nordmark, E., Simpson W., "Neighbor Discovery for IP Version 6 (IPv6)", RFC2461, December 1998.

7.2. Informative References

[ANYCASTRP] Dorian, K. et al, q(Anycast RP mechanism using PIM and MSDP", work-in-progress, draft-ietf-mboned-anycast-rp-08.txt, May 2001.

[BGMP] Thaler, D., "Border Gateway Multicast Protocol (BGMP)", work-in-progress, draft-ietf-bgmp-spec-03.txt. June 2002.

[CGMP] Cisco, "Cisco Group Management Protocol", e.g. http://www.cisco.com/en/US/tech/tk648/tk363/tk105/tech_protocol_home.html

[GARP] Tobagi, F., et al, "Study of IEEE 802.1p GARP/GMRP Timer Values", (for introduction to GARP/GMRP, see section 2), Sep 1997.

[MLDSNOOP] Christensen, M., Solensky, F., "IGMP and MLD snooping switches", work-in-progress, draft-ietf-magma-snoop-02.txt, June 2002.

[MSDP] Farinacci, D. et al, "Multicast Source Discovery Protocol (MSDP)", work-in-progress, draft-ietf-msdp-spec-13.txt (expired), 2002.

[MULTIGAP] Meyer, D., Nickless, B., "Internet Multicast Gap Analysis [...]", work-in-progress, draft-ietf-mboned-iesg-gap-analysis-00.txt, July 2002.

[PIM] Fenner, B. et al, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", work-in-progress, draft-ietf-pim-sm-v2-new-05.txt, March 2002.

[PIMSO] Bhattacharyya, S. et al, "Deployment of PIM-SO at Sprint (PIM-SO)", work-in-progress, draft-bhattach-diot-pimso-00.txt (expired), March 2000.

[SSM] Holbrook, H. et al, "Source-Specific Multicast for IP", work-in-progress, draft-ietf-ssm-arch-00.txt, November 2001.

[V6RPADDR] Savola, P., Haberman, B., "Embedding the Address of RP

in IPv6 Multicast Address", work-in-progress, draft-savola-mboned-mcast-rpaddr-00.txt, October 2002.

Author's Address

Pekka Savola
CSC/FUNET
Espoo, Finland
EMail: psavola@funet.fi