IPv6 Multicast Deployment Issues
draft-ietf-mboned-ipv6-multicast-issues-02.txt

Status of this Memo

Copyright Notice

Abstract

This memo describes known issues with IPv6 multicast, and provides
historical reference of how some earlier problems have been resolved.

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 to raise awareness, and documents how
previous problems have been resolved.

Section 2 describes the justifications for providing an inter-domain
multicast solution using Any Source Multicast (ASM) with IPv6.
Section 3 in turn describes which options were considered for filling
the requirements for the IPv6 inter-domain multicast solutions.
These sections are provided for historical reference of the
discussion and consensus in the IETF MBONED working group.

Section 4 lists issues that have come up with IPv6 multicast but have
not yet been at least fully resolved, and may require raised
awareness.

1.1 Multicast-related Abbreviations

```
ASM      Any Source Multicast
BSR      Bootstrap Router
CGMP     Cisco Group Management Protocol
DR       Designated Router
IGMP     Internet Group Management Protocol
MLD      Multicast Listener Discovery
MSDP     Multicast Source Discovery Protocol
PIM      Protocol Independent Multicast
PIM-SM   Protocol Independent Multicast - Sparse Mode
RP       Rendezvous Point
SSM      Source-specific Multicast
```

2. Justification for IPv6 Inter-domain ASM

This section documents the reasons and the discussion which led to
the agreement that solution to IPv6 inter-domain ASM was necessary.

The main reason was that SSM [RFC3569] was not considered to solve
all the relevant problems (e.g., many-to-many applications, source
discovery), and that SSM was not sufficiently widely deployed and
used. As these issues are more generic than just IPv6, they are
described in Appendix A.

2.1 Groups of Different Non-global Scopes

Many ASM applications are used with a smaller scope than global; some
of these have a wider scope than others. However, groups of smaller

scope typically need to be in their own PIM-SM domains to prevent
inappropriate data leakage.

Therefore if a site has groups of different scopes, it is important
to have multiple PIM domain borders. However, this need can be
obviated by using globally-scoped multicast addresses instead. It is
easier to set scoping using globally scoped addresses, rather than
having to configure (nesting) local multicast scopes.

In consequence there will be a need for inter-domain multicast
solutions, as a means to simplify and obviate the need for
operational hassles with local scoping. As many applications are
relying on ASM characteristics, this further increases the need for
an inter-domain ASM solution.

3. Different Solutions to Inter-domain Multicast

When ASM is used, the Internet must be divided into multiple PIM-SM
domains for both administrative and technical reasons, which means
there will be multiple PIM-SM RPs which need to share the source IP
addresses between themselves.

On the other hand, SSM does not require RPs and also works in the
inter-domain without such communication. Section 2 describes the
justification why Inter-domain ASM was still considered to be
required. This section describes different solutions which were
discussed to providing inter-domain multicast for IPv6.

For inter-domain multicast, MBONED WG came to consensus to continue
using SSM, and also use Embedded-RP for ASM as appropriate.

This section provides historical reference of the discussion and
decisions.

3.1 Changing the Multicast Usage Model

As ASM model has been found to be complex and a bit problematic, some
felt that this is a good incentive to move to SSM for good (at least
for most cases). Below two paragraphs are adapted from
[I-D.bhattach-diot-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 be preferable in many cases, SSM
was not sufficiently widely available to completely replace ASM (see
Appendix A), and that the IETF should not try to force the
application writers to change their multicast usage models.

3.2 Implementing MSDP for IPv6

In IPv4, notification of multicast sources between these PIM-SM RPs
is done with Multicast Source Discovery Protocol (MSDP) [RFC3618].
The protocol is widely considered a sub-optimal solution and even
dangerous to deploy; when it was specified, it was only meant as a
"stop-gap" measure.

The easiest stop-gap solution (to a stop-gap solution) would have
been to specify IPv6 TLV's for MSDP. This would be fairly
straightforward, and existing implementations would probably be
relatively easy to modify.

There is and has been resistance to this, as MSDP was not supposed to
last this long in the first place; there is clear consensus that
there should be no further work on it [I-D.ietf-mboned-msdp-deploy].

3.3 Implementing Another Multicast Routing Protocol

One possibility might have been to specify and/or implement a
different multicast routing protocol.

In fact, Border Gateway Multicast Protocol (BGMP) [RFC3913] has been
specified; however, it is quite complex and there have been no
implementations nor desire to write any. Lacking deployment
experience and specification analysis, it is difficult to say which
problems BGMP might solve (and possibly, which new ones BGMP might
introduce). One probable reason why BGMP failed to attract
continuing interest was it's dependance on similarly heavy-weight
multicast address allocation/assignment protocols.

As of this writing, no other inter-domain protocols have been
specified, and BGMP is not considered a realistic option.

3.4 Embedding the RP Address in an IPv6 Multicast Address

One way to work around these problems was to allocate and assign
multicast addresses in such a fashion that the address of the RP
could be automatically calculated from the IPv6 multicast address.

Making some assumptions about how the RPs would configure Interface
Identifiers, this is can achieved as described in [RFC3956]; PIM-SM
implementations need to implement the Embedded RP group-to-RP mapping
mechanism which processes this encoding.

To completely replace the need for MSDP for IPv6, a different way to
implement "Anycast RP" [RFC3446] is also needed. One such approach
is described in [I-D.ietf-pim-anycast-rp].

4. Issues with IPv6 Multicast

This section describes issues that have come up with IPv6 multicast
but have not yet been at least fully resolved.

4.1 Issues with Embedded RP

4.1.1 RP Failover with Embedded RP

Embedded RP provides a means for ASM multicast without inter-domain
MSDP. However, to continue providing failover mechanisms for RPs, a
form of state sharing, Anycast-RP, should still be supported.
Instead of MSDP, this can be achieved using a PIM-SM extension
[I-D.ietf-pim-anycast-rp].

One should note that as Embedded RP does not require MSDP peerings
between the RPs, it's possible to deploy more RPs in a PIM domain.
For example, the scalability and redundancy could be achieved by
co-locating RP functionality in the DRs: each major source, which
"owns" a group, could have its own DR act as the RP. This has about
the same redundancy characteristics as using SSM -- so there may not
be an actually very urgent need for Anycast-RP if operational methods
to include fate-sharing of the groups is followed.

In any case, "cold failover" redundancy without state sharing is
still an option. This does not offer any load-balancing of RPs or
shared trees, but provides only long-term redundancy. In this
mechanism, multiple routers would be configured with the RP address
(with appropriate unicast metrics), but only one of them would be
active at any time: if the main RP goes down, another takes its
place. However, the multicast state stored in the RP would be lost,
unless it is synchronized by some out-of-band mechanism.

4.1.2 Embedded RP and Control Mechanisms

With ASM and MSDP deployment, the ISPs can better control who is using their RPs.

With Embedded RP, anyone could use a third-party RP to host their groups unless some mechanisms, for example access-lists, are in place to control the use of the RP [RFC3956].

Such abuse is of questionable benefit, though, as anyone with a /64 could form an RP of its own.

Whether this is a sufficiently serious problem worth designing a (potentially complex) solution for is still under debate, as of this writing.

4.2 Neighbor Discovery Using Multicast

Neighbor Discovery [RFC2461] uses link-local multicast in Ethernet media, not broadcast as ARP does with IPv4. This has been seen to cause operational problems with some equipment. This section documents these as "lessons (hopefully) learned" so that other vendors could better avoid them.

There are equipment which do not forward (IPv6) multicast frames appropriately; these could be considered "bugs", but are sufficiently commonplace so that the behaviour is worth mentioning.

In particular, many WLAN IEEE 802.11b access points, working in the bridged mode, do not forward IPv6 Ethernet multicast frames across the bridge. When procuring WLAN equipment, it is probably a good idea to check out this functionality explicitly.

In some Ethernet switches, IPv6 frames are likewise not forwarded. The problem has likely been with building multicast forwarding state based on Layer 3 information (which the vendor does support with IPv6); state using Layer 2 information would work just fine [I-D.ietf-magma-snoop]. Therefore the snooping swich developers should be aware of the tradeoff of using Layer 2 vs Layer 3 information on multicast data forwarding, especially if IPv6 snooping is not supported.

There are no good workarounds for these problems, except disseminating information about them (e.g., at http://www.v6fix.net) and complaining to the vendor.

4.3 Functionality Like MLD Snooping

On Ethernet, multicast frames are forwarded to every port, even
without subscribers (or IPv6 support).

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, to
reduce the amount of flooding [I-D.ietf-magma-snoop].

Looking at the actual problem from a higher view, it is not clear
that MLD snooping is the right long-term solution. It makes the
switches complex, requires the processing of datagrams above the
link-layer, and should be discouraged
[I-D.ietf-mboned-iesg-gap-analysis]: the whole idea of L2-only
devices having to peek into L3 datagrams seems like a severe layering
violation -- and often the devices aren't upgradeable (if there are
bugs or missing features, which could be fixed later) 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 and implementation issues are considered, and these
do not have any particular security considerations; security
considerations for each technology are covered in the respective
specifications.

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; Havard 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

[I-D.ietf-mboned-msdp-deploy]
          McBride, M., "Multicast Source Discovery Protocol (MSDP)
          Deployment Scenarios",
          Internet-Draft draft-ietf-mboned-msdp-deploy-06, March
          2004.

   [I-D.ietf-pim-anycast-rp]
             Farinacci, D., "Anycast-RP using PIM",
             Internet-Draft draft-ietf-pim-anycast-rp-02, June 2004.

   [I-D.ietf-pim-sm-v2-new]
             Fenner, B., Handley, M., Holbrook, H. and I. Kouvelas,
             "Protocol Independent Multicast - Sparse Mode PIM-SM):
             Protocol Specification (Revised)",
             Internet-Draft draft-ietf-pim-sm-v2-new-11, October 2004.

   [RFC2461] Narten, T., Nordmark, E. and W. Simpson, "Neighbor
             Discovery for IP Version 6 (IPv6)", RFC 2461, December
             1998.

   [RFC3446] Kim, D., Meyer, D., Kilmer, H. and D. Farinacci, "Anycast
             Rendevous Point (RP) mechanism using Protocol Independent
             Multicast (PIM) and Multicast Source Discovery Protocol
             (MSDP)", RFC 3446, January 2003.

   [RFC3569] Bhattacharyya, S., "An Overview of Source-Specific
             Multicast (SSM)", RFC 3569, July 2003.

   [RFC3618] Fenner, B. and D. Meyer, "Multicast Source Discovery
             Protocol (MSDP)", RFC 3618, October 2003.

   [RFC3810] Vida, R. and L. Costa, "Multicast Listener Discovery
             Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.

   [RFC3913] Thaler, D., "Border Gateway Multicast Protocol (BGMP):
             Protocol Specification", RFC 3913, September 2004.

   [RFC3956] Savola, P. and B. Haberman, "Embedding the Rendezvous
             Point (RP) Address in an IPv6 Multicast Address",
             RFC 3956, November 2004.

7.2 Informative References

   [CGMP]   "Cisco Group Management Protocol",
             <http://www.javvin.com/protocolCGMP.html>.

   [GARP]   Tobagi, F., Molinero-Fernandez, P. and M. Karam, "Study of
             IEEE 802.1p GARP/GMRP Timer Values", 1997.

   [I-D.bhattach-diot-pimso]
             Bhattacharyya, S., Diot, C., Giuliano, L. and R. Rockell,
             "Deployment of PIM-SO at Sprint (PIM-SO)", March 2000.

   [I-D.ietf-magma-snoop]

              Christensen, M., Kimball, K. and F. Solensky,
              "Considerations for IGMP and MLD Snooping Switches",
              Internet-Draft draft-ietf-magma-snoop-11, May 2004.

   [I-D.ietf-mboned-iesg-gap-analysis]
              Meyer, D. and B. Nickless, "Internet Multicast Gap
              Analysis from the MBONED Working Group for the IESG",
              Internet-Draft draft-ietf-mboned-iesg-gap-analysis-00,
              July 2002.

   [I-D.ietf-pim-sm-bsr]
              Fenner, B., "Bootstrap Router (BSR) Mechanism for PIM",
              Internet-Draft draft-ietf-pim-sm-bsr-04, July 2004.

   [RFC3678] Thaler, D., Fenner, B. and B. Quinn, "Socket Interface
              Extensions for Multicast Source Filters", RFC 3678,
              January 2004.

   [SSMSNOOP]
              "Operational Problems with IGMP snooping switches", March
              2003, <http://www.ietf.org/proceedings/03mar/148.htm>.

Author's Address

   Pekka Savola
   CSC/FUNET
   Espoo
   Finland

   Email: psavola@funet.fi

Appendix A. SSM Deployment Issues

   To be deployed, SSM requires changes to:

   1. routers

   2. IGMP/MLD-snooping Ethernet switches

   3. hosts

   4. application programming interfaces (APIs)

   5. multicast usage models

   Introducing SSM support in the routers has been straightforward as
   PIM-SSM is a subset of PIM-SM [I-D.ietf-pim-sm-v2-new].

IGMP-snooping Ethernet switches have been a more difficult issue
[SSMSNOOP]; some which perform IGMPv2 snooping discard IGMPv3 reports
or queries, or multicast transmissions associated to them. If MLDv1
snooping had been implemented (or is implemented in a similar
manner), this would likely have affected that as well.

Host systems require MLDv2 [RFC3810] support. The situation has
improved with respect to MLDv2 support for end systems, and
interoperability has increased after the publication of the RFC due
to the stabilization of the ICMP types used.

The multicast source filtering API specification has also been
completed [RFC3678]; its deployment is likely roughly equal (or
slightly worse) than MLDv2. The API is required for creating
(cross-platform) SSM applications.

The most difficult issue, multicast usage models, remains a problem
as of this writing as described below. SSM is an excellent fit for
one-to-many distribution topologies, and porting such applications to
use SSM would likely be rather simple. However, a significant number
of current applications are many-to-many (e.g., conferencing
applications) which cannot be converted to SSM without significant
effort, including, for example, out-of-band source discovery. For
such applications to be usable for IPv6 at least in a short to medium
term, ASM -like techniques seem to be required.

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment