Internet Engineering Task Force
Internet Draft
Expiration Date: April 2003

P. Savola
CSC/FUNET

B. Haberman
Caspian Networks

October 2002

Embedding the Address of RP in IPv6 Multicast Address

draft-savola-mboned-mcast-rpaddr-00.txt

Status of this Memo

Abstract

As has been noticed, there is exists a huge deployment problem with global, interdomain IPv6 multicast: PIM RPs have no way of communicating the information about multicast sources to other multicast domains, as there is no MSDP, and the whole interdomain Any Source Multicast model is rendered unusable; SSM avoids these problems. This memo outlines a way to embed the address of the RP in the multicast address, solving the interdomain multicast problem. The problem is three-fold: specify an address format, adjust the operational procedures and configuration if necessary, and modify receiver-side PIM implementations. In consequence, there would be no need for interdomain MSDP.

Internet Draft draft-savola-mboned-mcast-rpaddr-00.txt October 2002

Table of Contents

1. Introduction 2
2. Unicast-Prefix-based Address Format 3
3. Modified Unicast-Prefix-based Address Format 3
4. Embedding the Address of the RP in the Multicast Address ... 4
5. Examples 5
5.1. Example 1 5
5.2. Example 2 5
5.3. Example 3 5
5.4. Example 4 6
6. Operational Requirements 6
6.1. Anycast-RP 6
6.2. Guidelines for Assigning IPv6 Addresses to RPs 6
7. Required PIM Modifications 6
8. Scalability/Usability Analysis 7
9. Acknowledgements 8
10. Security Considerations 8
11. References 8
11.1. Normative References 8
11.2. Informative References 9
Authors' Addresses 9
A. Open Issues/Discussion 9

1. Introduction

As has been noticed [V6MISSUES], there is exists a huge deployment problem with global, interdomain IPv6 multicast: PIM [PIM] RPs have no way of communicating the information about multicast sources to other multicast domains, as there is no MSDP [MSDP], and the whole interdomain Any Source Multicast model is rendered unusable; SSM [SSM] avoids there problems.

This memo outlines a way to embed the address of the RP in the multicast address, solving the interdomain multicast problem. The problem is three-fold: specify an address format, adjust the operational procedures and configuration if necessary, and modify receiver-side PIM implementations. In consequence, there would be no need for interdomain MSDP.

The solution is founded upon unicast-prefix-based IPv6 multicast addressing [UNIPRFXM] and making some assumptions about IPv6 address assignment for the RPs in the PIM domain.

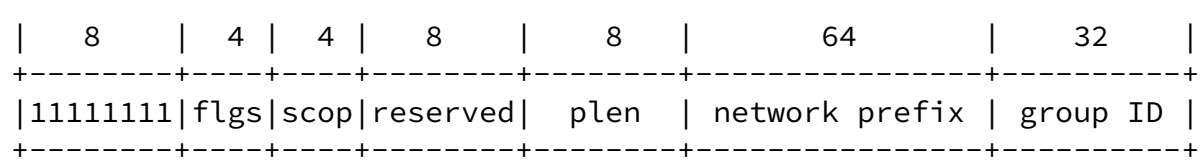
It is self-evident that one can't embed, in the general case, two 128-bit addresses in one 128-bit address. In this memo, some assumptions on how this could be done are made. If these assumptions can't be followed, either operational procedures and configuration must be slightly changed or this mechanism not be used.

The assignment of multicast addresses is outside the scope of this document.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2. Unicast-Prefix-based Address Format

As described in [UNIPRFXM], the multicast address format is as follows:



Where flgs are "0011". (The first two bits are yet undefined and thus zero.)

3. Modified Unicast-Prefix-based Address Format

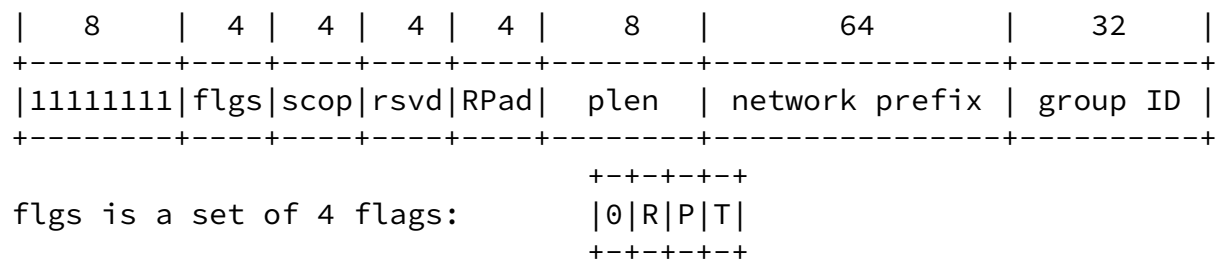
This memo proposes a modification to the unicast-prefix-based address format:

1. If the second high-order bit in "flgs" is set to 1, the address of the RP is embedded in the multicast address, as described in this memo.

2. If the second high-order bit in "flgs" was set to 1, interpret

the last low-order 4 bits of "reserved" field as signifying the RP interface ID, as described in this memo.

In consequence, the address format becomes:



R = 1 indicates a multicast address that embeds the address of the PIM RP. Then P MUST BE set to 1, and consequently T MUST be set to 1, as specified in [UNIPRFXM].

In the case that R = 1, the last 4 bits of previously reserved field ("RPad") are interpreted as embedding the interface ID of the RP, as specified in this memo.

R = 0 indicates a multicast address that does not embed the address of the PIM RP and follows the semantics defined in [ADDRARCH] and [UNIPRFXM]. In this context, the value of "RPad" has no meaning.

4. Embedding the Address of the RP in the Multicast Address

The address of the RP can only be embedded in unicast-prefix -based addresses, but the scheme could be extended to other forms of multicast addresses as well. Further, the mechanism cannot be combined with SSM.

To identify whether an address is a multicast address as specified in this memo and to be processed any further, it must satisfy all of the bullets:

- it MUST be part of the prefix FF7::/12
- "plen" MUST NOT be 0 (ie. not SSM)
- "plen" MUST NOT be greater than 96

The address of the RP can be obtained from a multicast address by taking the following steps:

1. take the last 96 bits of the multicast address add 32 zero bits at the end,

2. zero the last 128-"plen" bits, and

3. replace the last 4 bits with the contents of "RPad".

One should note that there are several operational scenarios when [UNIPRFXM] statement "All non-significant bits of the network prefix field SHOULD be zero" is ignored. This is to allow multicast address assignments to third parties which still use your RP; see example 2 below.

"Plen" higher than 64 SHOULD NOT be used as that would overlap with the upper bits of multicast group-id.

The implementation MUST perform at least the same address validity checks to the calculated RP address as to one received via other means (like MSDP), to avoid e.g. the address being "::" or "::1".

5. Examples

5.1. Example 1

The network administrator of 3FFE:FFFF::/32 wants to set up an RP for the network and all of his customers. He chooses network prefix=3FFE:FFFF and plen=32, and wants to use this addressing mechanism. The multicast addresses he will be able to use are of the form:

FF7x:y20:3FFE:FFFF:zzzz:zzzz:<group-id>

Where "x" is the multicast scope, "y" the interface ID of the RP address, and "zzzz:zzzz" will be freely assignable within the PIM domain. In this case, the address of the PIM RP would be:

3FFE:FFFF::y

(and "y" could be anything from 0 to F); the address 3FFE:FFFF::y/128 is added as a Loopback address and injected to the routing system.

5.2. Example 2

As above, the network administrator can also allocate multicast addresses like "FF7x:y20:3FFE:FFFF:DEAD::/80" to some of his customers within the PIM domain. In this case the RP address would still be "3FFE:FFFF::y" (note the prefix length rule: "plen" does not need to have anything to do with real unicast/multicast address prefix lengths).

5.3. Example 3

In the above network, the network admin sets up addresses as above, but an organization wants to have their own PIM domain; that's reasonable. The organization can pick multicast addresses like "FF7x:y30:3FFE:FFFF:BEEF::/80", and then their RP address would be "3FFE:FFFF:BEEF::y".

5.4. Example 4

In the above networks, if the admin wants to specify the RP to be in a non-zero /64 subnet, he could always use something like "FF7x:y40:3FFE:FFFF:BEEF:FEED::/96", and then their RP address would be "3FFE:FFFF:BEEF:FEED::y". There are still 32 bits of multicast group-id's to assign to customers and self.

6. Operational Requirements

6.1. Anycast-RP

One should note that MSDP is also used, in addition to interdomain connections between RPs, in anycast-RP [ANYCASTRP] -technique, for sharing the state information between different RPs in one PIM

domain.

Anycast-RP mechanism is incompatible with this addressing method unless unless MSDP is specified and implemented. Alternatively, another method for sharing state information could be defined.

Anycast-RP and other possible RP failover mechanisms are outside of the scope of this memo.

6.2. Guidelines for Assigning IPv6 Addresses to RPs

With this mechanism, the RP can be given basically any network prefix up to /64 (and even beyond, by using the upper bits of multicast group-id). The interface identifier will have to be manually configured.

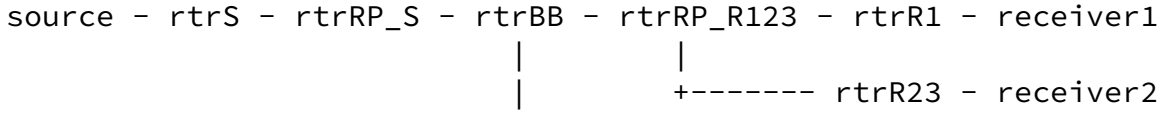
If an administrator wishes to use an RP address that does not conform to the addressing topology, that address can be injected into the routing system via a host route. This RP address SHOULD be assigned out of the network's prefix in order to ensure aggregation at the border.

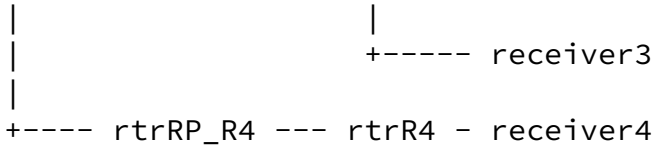
7. Required PIM Modifications

The use of multicast addresses with embedded RP addresses requires additional PIM processing. Namely, a PIM router will need to be able to recognize the encoding and derive the RP address from the address using the rules in section 4.

The two key places where these modifications are used are the Designated Routers (DRs) on the receiver networks and the RPs in the receiving domain (see figure below). For the DR's (rtrR1, rtrR23, and rtrR4), this would be similar to the RPT -> SPT switchover

scenario. For the RPs (rtrRP_R123 and rtrRP_R4) the scenario would be the same as building an SPT to a foreign source based on MSDP information. In particular, there is no need to have all routers on the path modified: this is a major benefit for quick deployment.





In addition, the administration of the PIM domain will require a policy decision on where the SPT towards the encoded RP should be built.

The extraction of the RP information from the multicast address should be done during forwarding state creation. That is, if no state exists for the multicast address, PIM must take the embedded RP information into account when creating forwarding state. Depending on administrative policy, this could result in a receiver's DR initiating an SPT towards the foreign RP, or the local RP initiating an SPT towards the foreign RP.

It should be noted that this approach removes the need to run inter-domain MSDP. Multicast distribution trees in foreign networks can be joined by issuing an SPT towards the RP address encoded in the multicast address.

8. Scalability/Usability Analysis

Interdomain MSDP model for connecting PIM domains is mostly hierarchical. The "embedded RP address" changes this to a mostly flat, full-mesh virtual topology.

This may or may not cause some effects; it may or may not be desirable. At the very least, it makes many things much more robust as the number of third parties is minimized. A good scalability analysis is needed.

In some cases (especially if e.g. every home user is employing site-local multicast), some degree of hierarchy would be highly desirable, for scalability (e.g. take the advantage of shared multicast state) and administrative point-of-view.

9. Acknowledgements

Jerome Durand commented on an early draft of this memo. Marshall Eubanks noted an issue regarding short plen values.

10. Security Considerations

The address of the PIM RP is embedded in the multicast address. RPs may be a good target for Denial of Service attacks, and in this way, the target would be clearly visible. However, it could be argued that if interdomain multicast was to be made work e.g. with MSDP, the address would have to be visible anyway (through via other channels, which may be more easily securable).

RPs may become a bit more single points of failure as anycast-RP mechanism is not (at least immediately) available. This can be partially mitigated by the fact that some other forms of failover are still possible, and there should be less need to store state as with MSDP.

The implementation MUST perform at least the same address validity checks to the embedded RP address as to one received via other means (like MSDP), to avoid the address being e.g. "::" or "::1".

TBD: the implications (if any) with regard to embedding the RP address in the packets (e.g. packet laundering and DoS do not seem possible due to the way multicast works, but more analysis is needed).

11. References

11.1. Normative References

[ADDRARCH] Hinden, R., Deering, S., "IP Version 6 Addressing Architecture", RFC2373, July 1998.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[UNIPRFXM] Haberman, B., Thaler, D., "Unicast-Prefix-based IPv6 Multicast Addresses", RFC3306, August 2002.

11.2. Informative References

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

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

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

[V6MISSUES] Savola, P., "IPv6 Multicast Deployment Issues", work-in-progress, draft-savola-v6ops-multicast-issues-00.txt, October 2002.

Authors' Addresses

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

Brian Haberman
Caspian Networks
One Park Drive
Suite 400
Research Triangle Park, NC 27709
EMail: bkhabs@nc.rr.com
Phone: +1-919-949-4828

A. Open Issues/Discussion

One could argue that there can be more RPs than the 4-bit "RPad" allows for, especially if anycast-RP cannot be used. In that light, extending "RPad" to take full advantage of whole 8 bits would seem reasonable. However, this would use up all of the reserved bits, and leave no room for future flexibility. In case of large number of RPs, an operational workaround could be to split the PIM domain: for

example, using two /33's instead of one /32 would gain another 16 RP addresses.

Some hierarchy (e.g. two-level, "ISP/customer") for RPs could possibly be added if necessary, but that would be torturing one 128 bits even more.