Internet Engineering Task Force
Internet-Draft
Obsoletes: 3913,2189,2201,1584,1585 (if approved)
Intended status: Best Current Practice
Expires: May 25, 2007

P. Savola
CSC/FUNET

November 21, 2006

Overview of the Internet Multicast Routing Architecture
draft-ietf-mboned-routingarch-07.txt

Status of this Memo

Copyright Notice

Abstract

The lack of up-to-date documentation on IP multicast routing protocols and procedures has caused a great deal of confusion. To clarify the situation, this memo describes the routing protocols and techniques currently (as of this writing) in use. This memo also

Obsoletes and reclassifies to Historic a number of older multicast protocols.

Table of Contents

1. Introduction

Good, up-to-date documentation of IP multicast is close to non-existent. This issue is severely felt with multicast routing protocols and techniques. The consequence is that those who wish to learn of IP multicast and how the routing works in the real world do not know where to begin. Multicast addressing is described in a companion document [I-D.ietf-mboned-addrarch].

The aim of this document is to provide a brief overview of multicast routing protocols and techniques.

This memo deals with:

- setting up multicast forwarding state (Section 2.1),
- distributing multicast topology information (Section 2.2),
- learning active sources (Section 2.3),
- configuring and distributing the PIM RP information (Section 2.4),
- mechanisms for enhanced redundancy (Section 2.5),
- interacting with hosts (Section 2.6), and
- restricting the multicast flooding in the link layer (Section 2.7).

Section 2 starts by describing a simplistic example how these classes of mechanisms fit together. Some multicast data transport issues are also introduced in Appendix A.

This memo obsoletes and re-classifies to Historic [RFC2026] Border Gateway Multicast Protocol (BGMP), Core Based Trees (CBT), Multicast OSPF (MOSPF) RFCs: [RFC3913], [RFC2189], [RFC2201], [RFC1584], and [RFC1585]. The purpose of the re-classification is to give the readers (both implementors and deployers) an idea what the status of a protocol is; there may be legacy deployments of some of these protocols, which are not affected by this reclassification. See Section 2.1 for more on each protocol.

Further historical perspective may be found in, for example, [RFC1458], [IMRP-ISSUES], and [IM-GAPS].

1.1. Multicast-related Abbreviations

ASM	Any Source Multicast
BGMP	Border Gateway Multicast Protocol
BSR	Bootstrap Router
CBT	Core Based Trees
CGMP	Cisco Group Management Protocol
DR	Designated Router
DVMRP	Distance Vector Multicast Routing Protocol
GARP	(IEEE 802.1D-2004) Generic Attribute Reg. Protocol
GMRP	GARP Multicast Registration Protocol
IGMP	Internet Group Management Protocol
MBGP	Multi-protocol BGP (*not* "Multicast BGP")
MLD	Multicast Listener Discovery
MRP	(IEEE 802.1ak) Multiple Registration Protocol
MMRP	(IEEE 802.1ak) Multicast Multiple Registration Proto.
MOSPF	Multicast OSPF
MSDP	Multicast Source Discovery Protocol
PGM	Pragmatic General Multicast
PIM	Protocol Independent Multicast
PIM-DM	PIM - Dense Mode
PIM-SM	PIM - Sparse Mode
PIM-SSM	PIM - Source-Specific Multicast
RGMP	(Cisco's) Router Group Management Protocol
RP	Rendezvous Point
SSM	Source-specific Multicast

2. Multicast Routing

In order to give a simplified summary how each of these class of mechanisms fits together, consider the following multicast receiver scenario.

When a host wants to receive a transmission, it first needs to find out the multicast group address (and with SSM, source address) using unspecified means. Then it will signal its interest to its router using IGMP or MLD (Section 2.6). To deliver a multicast transmission, the router will need to know how to build the distribution tree which includes all the sources (Section 2.3) and/or to locate the RP (Section 2.4) or one of RPs (Section 2.5). In scenarios where multicast is routed via different topology than unicast, a means to distribute topology information is required (Section 2.2). Nonetheless, using whatever topology information is available, the first-hop router initiates setting up hop-by-hop multicast forwarding state (Section 2.1). When multicast transmission arrives at the receiver's LAN, it is flooded to every port unless flooding reduction such as IGMP snooping is employed

(Section 2.7).

2.1. Setting up Multicast Forwarding State

The most important part of multicast routing is setting up the multicast forwarding state. This section describes the protocols commonly used for this purpose.

2.1.1. PIM-SM

By far, the most common multicast routing protocol is PIM-SM [RFC4601]. The PIM-SM protocol includes both Any Source Multicast (ASM) and Source-Specific Multicast (SSM) functionality; PIM-SSM is a subset of PIM-SM. Most current routing platforms support PIM-SM.

2.1.2. PIM-DM

Whereas PIM-SM has been designed to avoid unnecessary flooding of multicast data, PIM-DM [RFC3973] assumed that almost every subnet at a site had at least one receiver for a group. PIM-DM floods multicast transmissions throughout the network ("flood and prune") unless the leaf parts of the network periodically indicate that they are not interested in that particular group.

PIM-DM may be an acceptable fit in small and/or simple networks, where setting up an RP would be unnecessary, and possibly in cases where a large percentage of users is expected to want to receive the transmission so that the amount of state the network has to keep is minimal.

PIM-DM was used as a first step in transitioning away from DVMRP. It also became apparent that most networks would not have receivers for most groups, and to avoid the bandwidth and state overhead, the flooding paradigm was gradually abandoned. Transitioning from PIM-DM to PIM-SM was easy as PIM-SM was designed to use compatible packet formats and dense-mode operation could also be satisfied by a sparse protocol. PIM-DM is no longer in widespread use.

Many implementations also support so-called "sparse-dense" configuration, where Sparse mode is used by default, but Dense is used for configured multicast group ranges (such as Auto-RP in Section 2.4.3) only. Lately, many networks have transitioned away from sparse-dense to only sparse mode.

2.1.3. Bi-directional PIM

Bi-directional PIM [I-D.ietf-pim-bidir] is a multicast forwarding protocol that establishes a common shared-path for all sources with a

single root. It can be used as an alternative to PIM-SM inside a single domain. It doesn't have data-driven events or data-encapsulation. As it doesn't keep source-specific state, it may be a lucrative approach especially in sites with a large number of sources.

As of this writing, there is no inter-domain solution to configure a group range to use bi-directional PIM.

2.1.4. DVMRP

Distance Vector Multicast Routing Protocol (DVMRP) [RFC1075] [I-D.ietf-idmr-dvmrp-v3] [I-D.ietf-idmr-dvmrp-v3-as] was the first protocol designed for multicasting, and to get around initial deployment hurdles. It also included tunneling capabilities which were part of its multicast topology functions.

Currently, DVMRP is used only very rarely in operator networks, having been replaced with PIM-SM. The most typical deployment of DVMRP is at a leaf network, to run from a legacy firewall only supporting DVMRP to the internal network. However, GRE tunneling [RFC2784] seems to have overtaken DVMRP in this functionality, and there is relatively little use for DVMRP except in legacy deployments.

2.1.5. MOSPF

MOSPF [RFC1584] was implemented by several vendors and has seen some deployment in intra-domain networks. However, since it is based on intra-domain OSPF it does not scale to the inter-domain case, operators have found it is easier to deploy a single protocol for use in both intra-domain and inter-domain networks and so it is no longer being actively deployed.

2.1.6. BGMP

BGMP [RFC3913] did not get sufficient support within the service provider community to get adopted and moved forward in the IETF standards process. There were no reported production implementations and no production deployments.

2.1.7. CBT

CBT [RFC2201][RFC2189] was an academic project that provided the basis for PIM sparse mode shared trees. Once the shared tree functionality was incorporated into PIM implementations, there was no longer a need for a production CBT implementation. Therefore, CBT never saw production deployment.

2.1.8. Interactions and Summary

It is worth noting that it is possible to run different protocols with different multicast group ranges. For example, treat some groups as dense or bi-dir in an otherwise PIM-SM network; this typically requires manual configuration of the groups or a mechanism like BSR (Section 2.4.3). It is also possible to interact between different protocols, for example use DVMRP in the leaf network, but PIM-SM upstream. The basics for interactions among different protocols have been outlined in [RFC2715].

The following figure gives a concise summary of the deployment status of different protocols as of this writing.

	Interdomain	Intradomain	Status
PIM-SM	Yes	Yes	Active
PIM-DM	Not anymore	Not anymore	Little use
Bi-dir PIM	No	Yes	Some uptake
DVMRP	Not anymore	Stub only	Going out
MOSPF	No	Not anymore	Inactive
CBT	No	No	Never deployed
BGMP	No	No	Never deployed

From this table, it is clear that PIM-Sparse Mode is the only multicast routing protocol that is deployed inter-domain and, therefore, is most frequently used within multicast domains as well. This is partially result of not working on inter-domain RP/group configuration mechanisms since PIM-SM and MSDP (Section 2.3.2).

2.2. Distributing Topology Information

PIM has become the de-facto multicast forwarding protocol, but as its name implies, it is independent of the underlying unicast routing protocol. When unicast and multicast topologies are the same ("congruent"), i.e., use the same routing tables (routing information base, RIB), it has been considered sufficient just to distribute one set of reachability information to be used in conjunction with a protocol that sets up multicast forwarding state (e.g., PIM-SM).

However, when PIM which by default built multicast topology based on the unicast topology gained popularity, it became apparent that it would be necessary to be able to distribute also non-congruent multicast reachability information in the regular unicast protocols. This was previously not an issue, because DVMRP built its own reachability information.

The topology information is needed to perform efficient distribution of multicast transmissions and to prevent transmission loops by applying it to the Reverse Path Forwarding (RPF) check.

This subsection introduces these protocols.

2.2.1. Multi-protocol BGP

Multiprotocol Extensions for BGP-4 [I-D.ietf-idr-rfc2858bis] (often referred to as "MBGP"; however, it is worth noting that "MBGP" does *not* stand for "Multicast BGP") specifies a mechanism by which BGP can be used to distribute different reachability information for unicast (SAFI=1) and multicast traffic (SAFI=2). Multiprotocol BGP has been widely deployed for years, and is also needed to route IPv6. Note that SAFI=3 was originally specified for "both unicast and multicast" but has since then been deprecated.

These extensions are in widespread use wherever BGP is used to distribute unicast topology information. Multicast-enabled networks that use BGP should use Multiprotocol BGP to distribute multicast reachability information explicitly even if the topologies are congruent to make an explicit statement about multicast reachability. A number of significant multicast transit providers even require this, by doing the RPF lookups solely based on explicitly advertised multicast address family.

2.2.2. OSPF/IS-IS Multi-topology Extensions

Similar to BGP, some IGPs also provide the capability for signalling a differing topologies, for example IS-IS multi-topology extensions [I-D.ietf-isis-wg-multi-topology]. These can be used for a multicast topology that differs from unicast. Similar but not so widely implemented work exists for OSPF [I-D.ietf-ospf-mt].

It is worth noting that interdomain incongruence and intradomain incongruence are orthogonal, so one doesn't require the other. Specifically, interdomain incongruence is quite common, while intradomain incongruence isn't, so you see much more deployment of MBGP than MT-ISIS/OSPF. Commonly deployed networks have managed well without protocols handling intradomain incongruence. However, the availability of multi-topology mechanisms may in part replace the typically used workarounds such as tunnels.

2.2.3. Issue: Overlapping Unicast/multicast Topology

An interesting case occurs when some routers do not distribute multicast topology information explicitly while others do. In particular, this happens when some multicast sites in the Internet

are using plain BGP while some use MBGP.

Different implementations deal with this in different ways. Sometimes, multicast RPF mechanisms first look up the multicast routing table, or M-RIB ("topology database") with a longest prefix match algorithm, and if they find any entry (including a default route), that is used; if no match is found, the unicast routing table is used instead.

An alternative approach is to use longest prefix match on the union of multicast and unicast routing tables; an implementation technique here is to copy the whole unicast routing table over to the multicast routing table. The important point to remember here, though, is to not override the multicast-only routes; if the longest prefix match would find both a (copied) unicast route and a multicast-only route, the latter should be treated as preferable.

Another implemented approach is to just look up the information in the unicast routing table, and provide the user capabilities to change that as appropriate, using for example copying functions discussed above.

2.2.4. Summary

The following table summarizes the topology distribution approaches described in this Section. In particular, it is recommended that if interdomain routing uses BGP, multicast-enabled sites should use MP-BGP SAFI=2 for multicast and SAFI=1 for unicast even if the topology was congruent.

	Interdomain	Intradomain
Congruent topology	Yes	Yes
BGP without SAFI	Not recomm.	Yes
MP-BGP SAFI=1 only	Not recomm.	Not recomm.
MP-BGP SAFI=2	Recommended	Yes
MP-BGP SAFI=3	Doesn't work	Doesn't work
IS-IS multi-topology	No	Yes
OSPF multi-topology	No	Few implem.

2.3. Learning (Active) Sources

Typically, multicast routing protocols must either assume that the receivers know the IP addresses of the (active) sources for a group in advance, possibly using an out-of-band mechanism (SSM), or the transmissions are forwarded to the receivers automatically (ASM).

Learning active sources is a relatively straightforward process with a single PIM-SM domain and with a single RP, but having a single PIM-SM domain for the whole Internet is a completely unscalable model for many reasons. Therefore it is required to be able to split up the multicast routing infrastructures to smaller domains, and there must be a way to share information about active sources using some mechanism if the ASM model is to be supported.

This section discusses the options.

2.3.1. SSM

Source-specific Multicast [RFC4607] (sometimes also referred to as "single-source Multicast") does not count on learning active sources in the network. Recipients need to know the source IP addresses using an out of band mechanism which are used to subscribe to the (source, group) channel. The multicast routing uses the source address to set up the state and no further source discovery is needed.

As of this writing, there are attempts to analyze and/or define out-of-band source discovery functions which would help SSM in particular [I-D.lehtonen-mboned-dynssm-req].

2.3.2. MSDP

Multicast Source Discovery Protocol [RFC3618] was invented as a stop-gap mechanism, when it became apparent that multiple PIM-SM domains (and RPs) were needed in the network, and information about the active sources needed to be propagated between the PIM-SM domains using some other protocol.

MSDP is also used to share the state about sources between multiple RPs in a single domain for, e.g., redundancy purposes [RFC3446]. The same can be achieved using PIM extensions [RFC4610]. See Section 2.5 for more information.

There is no intent to define MSDP for IPv6, but instead use only SSM and Embedded-RP instead [I-D.ietf-mboned-ipv6-multicast-issues].

2.3.3. Embedded-RP

Embedded-RP [RFC3956] is an IPv6-only technique to map the address of the RP to the multicast group address. Using this method, it is possible to avoid the use of MSDP while still allowing multiple multicast domains (in the traditional sense).

The model works by defining a single RP address for a particular

group for all of the Internet, so there is no need to share state about that with any other RPs. If necessary, RP redundancy can still be achieved with Anycast-RP using PIM.

2.3.4. Summary

The following table summarizes the source discovery approaches and their status.

	IPv4	IPv6	Status
Bi-dir single domain	Yes	Yes	OK but for intra-domain only
PIM-SM single domain	Yes	Yes	OK
PIM-SM with MSDP	Yes	No	De-facto v4 inter-domain ASM
PIM-SM w/ Embedded-RP	No	Yes	Best inter-domain ASM option
SSM	Yes	Yes	No major uptake yet

2.4. Configuring and Distributing PIM RP Information

PIM-SM and Bi-dir PIM configuration mechanisms exist which are used to configure the RP addresses and which groups are to use those RPs in the routers. This section outlines the approaches.

2.4.1. Manual RP Configuration

It is often easiest just to manually configure the RP information on the routers when PIM-SM is used.

Originally, static RP mapping was considered suboptimal since it required explicit configuration changes every time the RP address changed. However, with the advent of anycast RP addressing, the RP address is unlikely to ever change. Therefore, the administrative burden is generally limited to initial configuration. Since there is usually a fair amount of multicast configuration required on all routers anyway (e.g., PIM on all interfaces), adding the RP address statically isn't really an issue. Further, static anycast RP mapping provides the benefits of RP load sharing and redundancy (see Section 2.5) without the complexity found in dynamic mechanisms like Auto-RP and Bootstrap Router (BSR).

With such design, an anycast RP uses an address that is configured on a loopback interfaces of the routers currently acting as RPs, and state is distributed using PIM [RFC4610] or MSDP [RFC3446].

Using this technique, each router might only need to be configured with one, portable RP address.

2.4.2. Embedded-RP

Embedded-RP provides the information about the RP's address in the group addresses which are delegated to those who use the RP, so unless no other ASM than Embedded-RP is used, the network administrator only needs to configure the RP routers.

While Embedded-RP in many cases is sufficient for IPv6, other methods of RP configuration are needed if one needs to provide ASM service for other than Embedded-RP group addresses. In particular, service discovery type of applications may need hard-coded addresses that are not dependent on local RP addresses.

As the RP's address is exposed to the users and applications, it is very important to ensure it does not change often, e.g., by using manual configuration of an anycast address.

2.4.3. BSR and Auto-RP

BSR [I-D.ietf-pim-sm-bsr] is a mechanism for configuring the RP address for groups. It may no longer be in as wide use with IPv4 as it was earlier, and for IPv6, Embedded-RP will in many cases be sufficient.

Cisco's Auto-RP is an older, proprietary method for distributing group to RP mappings, similar to BSR. Auto-RP has little use today.

Both Auto-RP and BSR require some form of control at the routers to ensure that only valid routers are able to advertise themselves as RPs. Further, flooding of BSR and Auto-RP messages must be prevented at PIM borders. Additionally, routers require monitoring that they are actually using the RP(s) the administrators think they should be using, for example if a router (maybe in customer's control) is advertising itself inappropriately. All in all, while BSR and Auto-RP provide easy configuration, they also provide very significant configuration and management complexity.

It is worth noting that both Auto-RP and BSR were deployed before the use of a manually configured anycast-RP address became relatively commonplace, and there is actually relatively little need for them today unless there is a need to configure different properties (e.g., sparse, dense, bi-dir) in a dynamic fashion.

2.4.4. Summary

The following table summarizes the RP discovery mechanisms and their status. With the exception of Embedded-RP, each mechanism operates within a PIM domain.

	IPv4	IPv6	Deployment
Static RP	Yes	Yes	Especially in ISPs
Auto-RP	Yes	No	Legacy deployment
BSR	Yes	Yes	Some, anycast simpler
Embedded-RP	No	Yes	Growing

2.5. Mechanisms for Enhanced Redundancy

A couple of mechanisms, already described in this document, have been used as a means to enhance redundancy, resilience against failures, and to recover from failures quickly. This section summarizes these techniques explicitly.

2.5.1. Anycast RP

As mentioned in Section 2.3.2, MSDP is also used to share the state about sources between multiple RPs in a single domain for, e.g., redundancy purposes [RFC3446]. The purpose of MSDP in this context is to share the same state information on multiple RPs for the same groups to enhance the robustness of the service.

Recent PIM extensions [RFC4610] also provide this functionality. In contrast to MSDP, this approach works for both IPv4 and IPv6.

2.5.2. Stateless RP Failover

It is also possible to use some mechanisms for smaller amount of redundancy as Anycast RP, without sharing state between the RPs. A traditional mechanism has been to use Auto-RP or BSR (see Section 2.4.3) to select another RP when the active one failed. However, the same functionality could be achieved using a shared-unicast RP address ("anycast RP without state sharing") without the complexity of a dynamic mechanism. Further, Anycast RP offers a significantly more extensive failure mitigation strategy, so today there is actually very little need to use stateless failover mechanisms, especially dynamic ones, for redundancy purposes.

2.5.3. Bi-directional PIM

Because bi-directional PIM (see Section 2.1.3) does not switch to shortest path tree (SPT), the final multicast tree is may be established faster. On the other hand, PIM-SM or SSM may converge more quickly especially in scenarios (e.g., unicast routing change) where bi-directional needs to re-do the Designated Forwarder election.

2.5.4. Summary

The following table summarizes the techniques for enhanced redundancy.

	IPv4	IPv6	Deployment
Anycast RP w/ MSDP	Yes	No	De-facto approach
Anycast RP w/ PIM	Yes	Yes	New, simpler than MSDP
Stateless RP fail.	Yes	Yes	Causes disturbance
Bi-dir PIM	Yes	Yes	Deployed at some sites

2.6. Interactions with Hosts

Previous sections have dealt with the components required by routers to be able to do multicast routing. Obviously, the real users of multicast are the hosts: either sending or receiving multicast. This section describes the required interactions with hosts.

2.6.1. Hosts Sending Multicast

After choosing a multicast group through a variety of means, hosts just send the packets to the link-layer multicast address, and the designated router will receive all the multicast packets and start forwarding them as appropriate.

In intra-domain or Embedded-RP scenarios, ASM senders may move to a new IP address without significant impact on the delivery of their transmission. SSM senders cannot change the IP address unless receivers join the new channel or the sender uses an IP mobility technique that is transparent to the receivers.

2.6.2. Hosts Receiving Multicast

Hosts signal their interest in receiving a multicast group or channel by the use of IGMP [RFC3376] and MLD [RFC3810]. IGMPv2 and MLDv1 are still commonplace, but are also often used in new deployments. Some vendors also support SSM mapping techniques for receivers which use an older IGMP/MLD version where the router maps the join request to an SSM channel based on various, usually complex means of configuration.

2.6.3. Summary

The following table summarizes the techniques host interaction.

	IPv4	IPv6	Notes
Host sending	Yes	Yes	No support needed
Host receiving ASM	IGMP	MLD	Any IGMP/MLD version
Host receiving SSM	IGMPv3	MLDv2	Also SSM-mapping

2.7. Restricting Multicast Flooding in the Link Layer

Multicast transmission in the link layer, for example Ethernet, typically includes some form of flooding the packets through a LAN. This causes unnecessary bandwidth usage and discarding unwanted frames on those nodes which did not want to receive the multicast transmission.

Therefore a number of techniques have been developed, to be used in Ethernet switches between routers, or between routers and hosts, to limit the flooding.

Some mechanisms operate with IP addresses, others with MAC addresses. If filtering is done based on MAC addresses, hosts may receive unnecessary multicast traffic (filtered out in the hosts' IP layer) if more than one IP multicast group addresses maps into the same MAC address, or if IGMPv3/MLDv2 source filters are used. Filtering based on IP destination addresses, or destination and sources addresses, will help avoid these but requires parsing of the Ethernet frame payload.

These options are discussed in this section.

2.7.1. Router-to-Router Flooding Reduction

A proprietary solution, Cisco's RGMP [RFC3488] has been developed to reduce the amount of flooding between routers in a switched networks. This is typically only considered a problem in some Ethernet-based Internet Exchange points or VPNs.

There have been proposals to observe and possibly react ("snoop") PIM messages [I-D.ietf-l2vpn-vpls-pim-snooping].

2.7.2. Host/Router Flooding Reduction

There are a number of techniques to help reduce flooding both from a router to hosts, and from a host to the routers (and other hosts).

Cisco's proprietary CGMP [CGMP] provides a solution where the routers notify the switches, but also allows the switches to snoop IGMP

packets to enable faster notification of hosts no longer wishing to receive a group. Fast leave behaviour support for IGMPv3 hasn't been implemented. Due to IGMP report suppression in IGMPv1 and IGMPv2, multicast is still flooded to ports which were once members of a group as long as there is at least one receiver on the link. Flooding restrictions are done based on multicast MAC addresses. IPv6 is not supported.

IEEE 802.1D-2004 specification describes Generic Attribute Registration Protocol (GARP), and GARP Multicast Registration Protocol (GMRP) [GMRP] is a link-layer multicast group application of GARP that notifies switches about MAC multicast group memberships. If GMRP is used in conjunction with IP multicast, then the GMRP registration function would become associated with an IGMP "join." However, this GMRP-IGMP association is beyond the scope of GMRP. GMRP requires support at the host stack and it has not been widely implemented. Further, IEEE 802.1 considers GARP and GMRP obsolete being replaced by Multiple Registration Protocol (MRP) and Multicast Multiple Registration Protocol (MMRP) that are being specified in IEEE 802.1ak [802.1ak]. MMRP is expected to be mainly used between bridges. Some further information about GARP/GMRP is also available in Appendix B of [RFC3488].

IGMP snooping [RFC4541] appears to be the most widely implemented technique. IGMP snooping requires that the switches implement a significant amount of IP-level packet inspection; this appears to be something that is difficult to get right, and often the upgrades are also a challenge.

Snooping switches also need to identify the ports where routers reside and therefore where to flood the packets. This can be accomplished using Multicast Router Discovery protocol [RFC4286], looking at certain IGMP queries [RFC4541], looking at PIM Hello and possibly other messages, or by manual configuration. An issue with PIM snooping at LANs is that PIM messages can't be turned off or encrypted, leading to security issues [I-D.ietf-pim-lasthop-threats].

IGMP proxying [RFC4605] is sometimes used either as a replacement of a multicast routing protocol on a small router, or to aggregate IGMP/ MLD reports when used with IGMP snooping.

2.7.3. Summary

The following table summarizes the techniques for multicast flooding reduction inside a single link for router-to-router and last-hop LANs.

	R-to-R	LAN	Notes
Cisco's RGMP	Yes	No	Replaced by PIM snooping
PIM snooping	Yes	No	Security issues in LANs
IGMP/MLD snooping	No	Yes	Common, IGMPv3 or MLD bad
Multicast Router Disc	No	Yes	Few if any implem. yet
IEEE GMRP and MMRP	No	No	No host/router deployment
Cisco's CGMP	No	Yes	Replaced by other snooping

3. Acknowledgements

Tutoring a couple multicast-related papers, the latest by Kaarle Ritvanen [RITVANEN] convinced the author that up-to-date multicast routing and address assignment/allocation documentation is necessary.

Leonard Giuliano, James Lingard, Jean-Jacques Pansiot, Dave Meyer, Stig Venaas, Tom Pusateri, Marshall Eubanks, Dino Farinacci, Bharat Joshi, Albert Manfredi, Jean-Jacques Pansiot, Spencer Dawkins, Sharon Chisholm, John Zwiebel, Dan Romascanu, and Thomas Morin provided good comments, helping in improving this document.

4. IANA Considerations

This memo includes no request to IANA.

5. Security Considerations

This memo only describes different approaches to multicast routing, and this has no security considerations; the security analysis of the mentioned protocols is out of scope of this memo.

However, there has been analysis of the security of multicast routing infrastructures [RFC4609], IGMP/MLD [I-D.daley-magma-smld-prob], and PIM last-hop issues [I-D.ietf-pim-lasthop-threats].

6. References

6.1. Normative References

[I-D.ietf-idr-rfc2858bis]
Bates, T., "Multiprotocol Extensions for BGP-4", draft-ietf-idr-rfc2858bis-10 (work in progress),

March 2006.

[I-D.ietf-isis-wg-multi-topology]
Przygienda, T., "M-ISIS: Multi Topology (MT) Routing in IS-IS", draft-ietf-isis-wg-multi-topology-11 (work in progress), October 2005.

[I-D.ietf-mboned-addrarch]
Savola, P., "Overview of the Internet Multicast Addressing Architecture", draft-ietf-mboned-addrarch-05 (work in progress), October 2006.

[I-D.ietf-ospf-mt]
Psenak, P., "Multi-Topology (MT) Routing in OSPF", draft-ietf-ospf-mt-06 (work in progress), February 2006.

[I-D.ietf-pim-bidir]
Handley, M., "Bi-directional Protocol Independent Multicast (BIDIR-PIM)", draft-ietf-pim-bidir-08 (work in progress), October 2005.

[RFC2026] Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

[RFC3376] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.

[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.

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

[RFC3973] Adams, A., Nicholas, J., and W. Siadak, "Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)", RFC 3973, January 2005.

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006.

[RFC4607] Holbrook, H. and B. Cain, "Source-Specific Multicast for IP", RFC 4607, August 2006.

6.2. Informative References

[802.1ak] "IEEE 802.1ak - Multiple Registration Protocol", <http://www.ieee802.org/1/pages/802.1ak.html>.

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

[GMRP] "GARP Multicast Registration Protocol", <http://www.javvin.com/protocolGMRP.html>.

[I-D.daley-magma-smld-prob]
Daley, G. and G. Kurup, "Trust Models and Security in Multicast Listener Discovery", draft-daley-magma-smld-prob-00 (work in progress), July 2004.

[I-D.ietf-idmr-dvmrp-v3]
Pusateri, T., "Distance Vector Multicast Routing Protocol", draft-ietf-idmr-dvmrp-v3-11 (work in progress), December 2003.

[I-D.ietf-idmr-dvmrp-v3-as]
Pusateri, T., "Distance Vector Multicast Routing Protocol Applicability Statement", draft-ietf-idmr-dvmrp-v3-as-01 (work in progress), May 2004.

[I-D.ietf-l2vpn-vpls-pim-snooping]
Hemige, V., "PIM Snooping over VPLS", draft-ietf-l2vpn-vpls-pim-snooping-00 (work in progress), August 2006.

[I-D.ietf-mboned-ipv6-multicast-issues]
Savola, P., "IPv6 Multicast Deployment Issues", draft-ietf-mboned-ipv6-multicast-issues-02 (work in progress), February 2005.

[I-D.ietf-pim-lasthop-threats]
Savola, P. and J. Lingard, "Last-hop Threats to Protocol Independent Multicast (PIM)", draft-ietf-pim-lasthop-threats-00 (work in progress), October 2006.

[I-D.ietf-pim-sm-bsr]
Bhaskar, N., "Bootstrap Router (BSR) Mechanism for PIM", draft-ietf-pim-sm-bsr-09 (work in progress), June 2006.

[I-D.lehtonen-mboned-dynssm-req]

Lehtonen, R., "Requirements for discovery of dynamic SSM sources", draft-lehtonen-mboned-dynssm-req-00 (work in progress), February 2005.

[IM-GAPS] Meyer, D. and B. Nickless, "Internet Multicast Gap Analysis from the MBONED Working Group for the IESG [Expired]", draft-ietf-mboned-iesg-gap-analysis-00 (work in progress), July 2002.

[IMRP-ISSUES] Meyer, D., "Some Issues for an Inter-domain Multicast Routing Protocol [Expired]", draft-ietf-mboned-imrp-some-issues-01 (work in progress), September 1997.

[RFC1075] Waitzman, D., Partridge, C., and S. Deering, "Distance Vector Multicast Routing Protocol", RFC 1075, November 1988.

[RFC1458] Braudes, B. and S. Zabele, "Requirements for Multicast Protocols", RFC 1458, May 1993.

[RFC1584] Moy, J., "Multicast Extensions to OSPF", RFC 1584, March 1994.

[RFC1585] Moy, J., "MOSPF: Analysis and Experience", RFC 1585, March 1994.

[RFC2189] Ballardie, T., "Core Based Trees (CBT version 2) Multicast Routing -- Protocol Specification --", RFC 2189, September 1997.

[RFC2201] Ballardie, T., "Core Based Trees (CBT) Multicast Routing Architecture", RFC 2201, September 1997.

[RFC2715] Thaler, D., "Interoperability Rules for Multicast Routing Protocols", RFC 2715, October 1999.

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, March 2000.

[RFC3208] Speakman, T., Crowcroft, J., Gemmell, J., Farinacci, D., Lin, S., Leshchiner, D., Luby, M., Montgomery, T., Rizzo, L., Tweedly, A., Bhaskar, N., Edmonstone, R., Sumanasekera, R., and L. Vicisano, "PGM Reliable Transport Protocol Specification", RFC 3208, December 2001.

[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.

[RFC3488] Wu, I. and T. Eckert, "Cisco Systems Router-port Group Management Protocol (RGMP)", RFC 3488, February 2003.

[RFC3740] Hardjono, T. and B. Weis, "The Multicast Group Security Architecture", RFC 3740, March 2004.

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

[RFC4286] Haberman, B. and J. Martin, "Multicast Router Discovery", RFC 4286, December 2005.

[RFC4541] Christensen, M., Kimball, K., and F. Solensky, "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches", RFC 4541, May 2006.

[RFC4605] Fenner, B., He, H., Haberman, B., and H. Sandick, "Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")", RFC 4605, August 2006.

[RFC4609] Savola, P., Lehtonen, R., and D. Meyer, "Protocol Independent Multicast - Sparse Mode (PIM-SM) Multicast Routing Security Issues and Enhancements", RFC 4609, October 2006.

[RFC4610] Farinacci, D. and Y. Cai, "Anycast-RP Using Protocol Independent Multicast (PIM)", RFC 4610, August 2006.

[RITVANEN] Ritvanen, K., "Multicast Routing and Addressing", HUT Report, Seminar on Internetworking, May 2004, <http://www.tml.hut.fi/Studies/T-110.551/2004/papers/>.

Appendix A. Multicast Payload Transport Extensions

A couple of mechanisms have been, and are being specified, to improve the characteristics of the data that can be transported over multicast.

These go beyond the scope of multicast routing, but as reliable

multicast has some relevance, these are briefly mentioned.

A.1. Reliable Multicast

Reliable Multicast Working Group has been working on experimental specifications so that applications requiring reliable delivery characteristics, instead of simple unreliable UDP, could use multicast as a distribution mechanism.

One such mechanism is Pragmatic Generic Multicast (PGM) [RFC3208]. This does not require support from the routers, bur PGM-aware routers may act in router assistance role in the initial delivery and potential retransmission of missing data.

A.2. Multicast Group Security

Multicast Security Working Group has been working on methods how the integrity, confidentiality, and authentication of data sent to multicast groups can be ensured using cryptographic techniques [RFC3740].

Author's Address

Pekka Savola
CSC - Scientific Computing Ltd.
Espoo
Finland

Email: psavola@funet.fi

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).