

PIM working group
Internet Draft
Intended status: Proposed standard
Expires: May 2016

S. Sami
November 24, 2015

Protocol independent multicast- Next Generation (PIM-NG):
Protocol Specifications

draft-sami-pim-ng-08.txt

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

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/lid-abstracts.txt>

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

This Internet-Draft will expire on May 24, 2016.

Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

This document specifies protocol independent multicast- Next generation or simply called PIM-NG. PIM-NG uses the underlying unicast routing information, but unlike PIM-SM it doesn't necessarily need to build unidirectional shared trees rooted at a Rendezvous Point (RP) per group, due to the fact that the processes through which a source registers with the Rendezvous Point and a host finds the source of the multicast destination groups it needs are done in a whole new way and hence, the source of multicast group (G) is discovered faster. So at some points PIM-NG works faster than its predecessor. Also the new Domain concept, unique to PIM-NG, along the RPF check method used in PIM-NG specifications provides many features along a robust and flexible control and administration over multicast networks.

Table of Contents

| | |
|--|----|
| 1. Introduction..... | 9 |
| 2. Terminology..... | 9 |
| 2.1. Definitions..... | 10 |
| 3. IP Address considerations..... | 14 |
| 4. PIM-NG processes..... | 15 |
| 4.1. New packet Header..... | 15 |
| 4.2. Source Discovery..... | 16 |
| 4.2.1. Source register process with RP..... | 17 |
| 4.2.2. Communication between client and the source..... | 21 |
| 4.2.3. Source specific multicast..... | 31 |
| 4.2.4. Source Discovery Delay Prevention..... | 33 |
| 4.2.4.1. Delay prevention in Default-Mode..... | 34 |
| 4.2.4.2. Delay prevention in Admin-Mode..... | 36 |
| 4.3. Communication between PIM-NG routers..... | 39 |
| 4.4. RP discovery..... | 44 |
| 4.4.1. Static RP discovery..... | 45 |
| 4.4.2. Dynamic RP discovery..... | 46 |
| 4.4.2.1. Multicast IP addresses used..... | 47 |
| 4.4.2.2. Dynamic RP discovery TYPE-1..... | 47 |
| 4.4.2.2.1. RP introduction process..... | 49 |
| 4.4.2.2.2. Back up C-RP considerations..... | 51 |
| 4.4.2.2.3. PIM-NG clients processes in Dynamic-RP..... | 58 |
| 4.4.2.2.3.1. New PIM-NG router/client..... | 59 |
| 4.4.2.3. Dynamic RP discovery type 2..... | 61 |
| 4.4.2.3.1. Client related concepts..... | 63 |
| 4.4.2.3.2. C-MAPPER concepts..... | 63 |
| 4.4.2.3.2.1. Value of the A-BIT..... | 66 |
| 4.4.2.3.2.2. C-MAPPER preparation..... | 66 |
| 4.4.2.3.3. C-RP concepts..... | 69 |
| 4.4.2.3.3.1. C-RP redundancy..... | 69 |
| 4.4.2.3.3.2. C-RP processes..... | 71 |
| 4.4.2.3.3.3. Redundant C-RPs..... | 72 |
| 4.4.2.3.3.4. PIM-NG Anycast RP..... | 78 |
| 4.4.2.3.3.5. PIM-NG C-RP Mesh-Group..... | 79 |
| 4.4.2.3.3.6. Backup C-RP considerations..... | 81 |
| 4.5. C-MAPPER Redundancy in PIM-NG..... | 81 |
| 4.5.1. SC-MAPPER considerations..... | 91 |
| 4.5.2. C-MAPPER Mesh-Group..... | 91 |
| 4.5.2.1. Active C-MAPPER Election..... | 94 |
| 4.5.3. ANYCAST RP rules..... | 95 |
| 4.5.4. Configuration process of Redundant C-MAPPERS..... | 95 |
| 4.5.5. C-RP and Redundant C-MAPPERS..... | 96 |
| 4.6. Multiple multicast domains..... | 97 |
| 4.6.1. Definitions and concepts..... | 99 |

| | |
|---|-----|
| 4.6.1.1. Domain..... | 99 |
| 4.6.1.2. PIM-EDGE-ROUTER (PER/PPER)..... | 107 |
| 4.6.1.3. Tree Root (TR)..... | 114 |
| 4.6.1.4. Domain-set and RPF check..... | 117 |
| 4.6.2. Inter-domain connectivity concepts..... | 121 |
| 4.6.2.1. Public Domains..... | 121 |
| 4.6.2.1.1. Intra-domain concepts..... | 122 |
| 4.6.2.1.2. Inter-domain concepts..... | 124 |
| 4.6.2.2. Private domains..... | 128 |
| 4.6.2.2.1. Intra-Domain processes..... | 129 |
| 4.6.2.2.2. Inter-domain concepts..... | 135 |
| 4.6.3. Core-Domain implementation..... | 138 |
| 4.6.4. Multiple multicast domains scenario..... | 143 |
| 4.6.5. PIM-NG Sub-Domain..... | 149 |
| 4.6.6. PIM-NG Stub-Domain..... | 161 |
| 4.7. PIM-NG Bidirectional logic..... | 164 |
| 4.7.1. Requirements..... | 164 |
| 4.7.2. Bidirectional Multicast Group Discovery..... | 165 |
| 4.7.2.1. Manual Mode..... | 165 |
| 4.7.2.2. Autosense mechanism..... | 167 |
| 4.7.3. Redundant TRs..... | 170 |
| 4.7.3.1. ANYCAST Approach..... | 170 |
| 4.7.3.2. TR Grouping..... | 171 |
| 4.7.4. Bidirectional Tree formation..... | 173 |
| 4.7.4.1. Tree formation between TRs..... | 173 |
| 4.7.4.2. Completing the Tree formation..... | 175 |
| 4.7.5. Inter-Domain Bidirectional connectivity rules..... | 177 |
| 4.8. PIM-SM compatibility..... | 181 |
| 4.8.1. PIM-SM compatibility and SA messages..... | 181 |
| 4.8.2. PIM-SM compatibility and join/prune messages..... | 184 |
| 4.9. Loop prevention..... | 186 |
| 4.10. DR election and PIM Assert Message..... | 187 |
| 5. Security Considerations..... | 187 |
| 5.1. Attacks based on forged messages..... | 188 |
| 5.1.1. Unicast forged messages..... | 188 |
| 5.1.2. Forged link local messages..... | 188 |
| 5.1.3. Forged multicast messages..... | 189 |
| 5.2. Non-cryptographic authentication mechanisms..... | 190 |
| 5.3. Authentication..... | 191 |
| 5.3.1. Protecting Multicast Introduction Message..... | 192 |
| 5.4. Denial-Of-Service attacks..... | 192 |
| 6. IANA Considerations..... | 193 |
| 6.1. PIM-NG multicast destination group addresses..... | 193 |
| 6.2. PIM-NG packets and values of type field..... | 193 |
| 6.3. PIM-NG Domain numbers..... | 193 |
| 7. Conclusions..... | 194 |
| 8. References..... | 194 |

8.1. Normative References.....194
8.2. Informative References.....195
9. Acknowledgments.....196

Table of figures

| | |
|---|----|
| Figure 1 the new packet Header..... | 15 |
| Figure 2 source register request packet..... | 18 |
| Figure 3 multicast mapping table..... | 18 |
| Figure 4 C-RP acknowledge message..... | 19 |
| Figure 5 source communication with the RP..... | 20 |
| Figure 6 Request for source packet..... | 23 |
| Figure 7 Request for Source packet format..... | 24 |
| Figure 8 C-RP Acknowledge packet format..... | 25 |
| Figure 9 C-RP answer to the host in case that source doesn't exist... | 26 |
| Figure 10 Join/prune message format..... | 29 |
| Figure 11 Encoded Source Address Format..... | 30 |
| Figure 12 Communication between Host and Source..... | 32 |
| Figure 13 Client Request Table (CRT)..... | 38 |
| Figure 14 New-Activated-Source-Notification message format..... | 38 |
| Figure 15 original PIM-SM hello packet compared with the new hello packet format..... | 40 |
| Figure 16 PIM domain topology table (PDTT)..... | 42 |
| Figure 17 C-MAPPER-INTRODUCTION message sent to 239.0.1.190..... | 50 |
| Figure 18 SC-MAPPER/RP answers to a host REQUEST-FOR-C-MAPPER message | 53 |
| Figure 19 SC-MAPPER/RP multicast introductions to all PIM-NG-CLIENTS. | 54 |
| Figure 20 PIM-NG network with backup RP (SC-RP)..... | 55 |
| Figure 21 RP introduction message format..... | 55 |
| Figure 22 C-MAPPER/RP unicast HELLO/KEEP-ALIVE message to back up RPs | 57 |

| | |
|--|-----|
| Figure 23 request- for-C-MAPPER packet sent to the SC-MAPPER..... | 59 |
| Figure 24 New client added..... | 61 |
| Figure 25 network design with one multicast domain..... | 62 |
| Figure 26 C-MAPPER introduction message formats..... | 65 |
| Figure 27 multicast domain with one C-MAPPER..... | 69 |
| Figure 28 network designs with one multicast domain..... | 70 |
| Figure 29 C-RP introduction message format..... | 71 |
| Figure 30 PEER-MAPPING table and PEER-LIST table..... | 74 |
| Figure 31 AMMT compared with MMT..... | 76 |
| Figure 32 AMMT Format..... | 77 |
| Figure 33 network with single multicast domain and redundant C-MAPPERS | 82 |
| Figure 34 PEER-MAPPING Table..... | 83 |
| Figure 35 introduction message format sent to ALL PIM-NG MAPPERS 239.0.1.188..... | 86 |
| Figure 36 A- Multicast MAPPING TABLE..... | 87 |
| Figure 37 sample illustration of the above process..... | 90 |
| Figure 38 network with 2 PIM-NG multicast domains..... | 99 |
| Figure 39 PER introduction message..... | 113 |
| Figure 40 Internal Multicast Source Table..... | 113 |
| Figure 41 TR introduction message..... | 116 |
| Figure 42 Joined-Groups table used by both TR'(s) and Clients..... | 117 |
| Figure 43 Inter-domain connectivity and RPF check..... | 119 |
| Figure 44 Intra-Domain connectivity and RPF check..... | 120 |
| Figure 45 network with 2 multicast domains..... | 125 |

| | |
|---|-----|
| Figure 46 Multicast network with private PIM-NG domains..... | 129 |
| Figure 47 core domain implementation..... | 141 |
| Figure 48 join/prune message being sent towards the source..... | 143 |
| Figure 49 network with multiple multicast domains..... | 144 |
| Figure 50 multicast network and multicast advertisements..... | 146 |
| Figure 51 Mesh-Groups and C-MAPPER peering..... | 157 |
| Figure 52 PPER and PER communication with C-MAPPER..... | 159 |
| Figure 53 Sample PIM-NG domain Divided to Sub-Domains..... | 160 |
| Figure 54 Stub-Domain as the transitory multicast domain..... | 163 |
| Figure 55 Bidirectional Groups Table (BGT)..... | 165 |
| Figure 56 Bidirectional Tree formation with ANYCAST TR Address..... | 176 |
| Figure 57 Bidirectional Tree formation with TR grouping..... | 177 |

1. Introduction

This document specifies a protocol for efficiently routing multicast groups that may span wide-area (and inter-domain) internets. This protocol is called protocol independent multicast-Next Generation or simply put PIM-NG. The name is chosen due to the fact that this new protocol has some behaviors similar to PIM-SM, in parts related to using the underlying unicast routing information base to find the best path to reach a source and not being limited to any specific routing protocol.

But as the overall process of Multicast source discovery is a whole new story it is called PIM-NG or the next generation of PIM-SM[7].

PIM-NG provides many new features and concepts related to multicast routing which together open new doors and possibilities such as removing the need for Shared path trees to form, being able to easily consider transit multicast domains, use as much as 254*255 RPs within a single multicast domain and using redundant Tree Roots per bidirectional tree and eventually redundant Bidirectional Trees per Bidirectional Group. Also it provides the possibility of considering transit multicast domains easily. In addition to the above this protocol is a lot more secure than current implementations because of its unique multicast domain concept introduced for the first time.

2. Terminology

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", "ONLY" and "OPTIONAL" are to be interpreted as described in RFC 2119 [1] and indicate requirement levels for compliant PIM-NG implementations.

Commands used in this document, MUST NOT be interpreted as the solid commands, and are only used as an example to simplify the explanation of the processes used by PIM-NG.

2.1. Definitions

The following terms have special significance for PIM-NG:

Candidate Rendezvous Point (C-RP):

C-RP is a router that has been configured to take the role of a multicast information station in a PIM-NG domain. Unlike PIM-SM specifications in RFC4601 [7], it is not used as the root of the non-source-specific distribution tree for a multicast group. Any multicast source will inform the RP about its existence by sending a register message (S,G) to the RP, and RP will save the unicast address of the source for further use in a special table. And any host looking for the source of any desired multicast group will send a request of (*,G) or (S,G) to the RP to receive the unicast address of the desired multicast source.

Candidate MAPPER (C-MAPPER):

C-MAPPER is a router in charge of introducing the existing components of a PIM-NG multicast domain to all PIM-NG population within a multicast domain and also is responsible for the exchange of A-Multicast Mapping Table between different PIM-NG multicast domains. It acts like a BSR [9] in parts related to introducing the C-RPs to all PIM-NG routers. The difference between a C-MAPPER and a BSR is that the C-MAPPER doesn't do the group to RP mapping. It will only introduce the existing components such as C-RP'(s),

Client:

Client is a router that either wants to register a source, or is looking for a source. To be more specific any none C-RP or C-MAPPER router can be called a client. And each C-RP and C-MAPPER can act as a client too, which is not recommended.

Tree Root (TR):

A PIM-NG-AWARE router that after being configured, MUST be introduced to the multicast domain to become the root of each (S,G) tree. Any client in need of receiving multicast traffic from a source will send its join/prune messages towards the existing TR in the domain, or if no TR is considered in the domain, directly to the existing TRs in the core-domain.

EDGE-Client:

It is a PIM-NG client within a PIM-NG multicast domain that can act as a boundary between downstream clients and upstream clients, to limit the propagation of multicast introduction messages sent by C-MAPPER'(s) or C-RP'(s). an EDGE-CLIENT can be placed at the edge of multi-access network'(s), which are part of one unified PIM-NG multicast domain.

PIM-EDGE-ROUTER (PER):

It is a PIM-NG aware router that connects 2 or more separate PIM-NG multicast domains and Sub-Domains by using the underlying IGP protocol used inside the network or a routing protocol capable of transferring multicast traffic like MBGP.

PRIVATE-PIM-EDGE-ROUTER (PPER):

It is PIM-NG PER that is placed at the edge of a private PIM-NG multicast domain. A PPER is also in charge of NAT operations in the domain. As soon as the PPER is configured it MUST introduce itself to the closest C-MAPPER in the domain.

BORDER-PIM-ROUTER (BPR):

It is a PER or PPER which is placed between a PIM-NG multicast Domain and a PIM-SM multicast domain, and has one or more internal interfaces connected to the PIM-NG multicast domain and one or more interfaces connected to the PIM-SM multicast domain. In case of using PER as the BPR, the PER MUST introduce itself to the closest C-MAPPER.

Second candidate Rendezvous Point (SC-RP):

SC-RP is a router configured or elected to act as the backup C-RP if needed. And if C-RP goes offline will immediately take its place.

Second Candidate MAPPER (SC-MAPPER):

SC-MAPPER is a router configured or chosen to act as the backup C-MAPPER if needed. And if C-MAPPER goes offline will immediately take its place

Internal interface (II):

All interfaces of a PIM-NG router are internal interfaces by default, and are assumed to be connected to either internal domain from a PIM-

NG-PER/PPER or PIM-NG-BPR point of view or a multi-access-network from a PIM-NG-CLIENT point of view.

External interface (EI):

It is an interface configured as an external interface. External interface is assumed to be connected to an external domain from a PIM-NG-PER point of view or connected to the domain a PIM-NG-CLIENT is a member of and thus, multicast introductions received on an external interface won't be forwarded to internal interfaces. Also a PIM-NG-BPR can have an external interface which is by default the interface connected to a PIM-SM domain.

PIM-NG DOMAIN:

A domain is actually a public or private PIM-NG multicast network including its own set of C-MAPPERS, C-RPs and clients isolated from other domains. Clients and C-RPs inside one domain do not react to C-MAPPER introduction messages that might be received from other Domains. The only points of connection between 2 different domains are the C-MAPPERS and if used PIM-EDGE-ROUTERS. Each DOMAIN can be connected to either one or more PIM-NG-DOMAIN and if needed PIM-SM domains or a single PIM-NG-CORE-DOMAIN.

PIM-NG CORE-DOMAIN:

A special domain implementation of PIM-NG, which if applied gives a hierarchical design approach, alongside a good and sound multicast traffic flow control to the administrators of different CORE-DOMAINS. A CORE-DOMAIN can be connected to one or more CORE-DOMAINS and one or more PIM-NG-DOMAINS. PIM-NG-DOMAINS MUST BE connected to the outside world or World Wide Web through a CORE-DOMAIN if they need to advertise their multicast sources globally.

Multicast mapping table (MMT):

After a router is configured as a C-RP, a table called multicast-mapping table is created on it. This table will then hold the information needed to be used by clients to find a source. After a C-RP receives a register message from a source it will put an entry for that source in this table which indicates the unicast address of the source plus the multicast destination group it is representing in the format (S,G) alongside the unicast address of the client which is sending the register message. It is done to bring compatibility with the needs of SSM.

Abbreviated multicast mapping table (A-multicast mapping table) (AMMT):

Is an abbreviated form of Multicast mapping table which only holds the information regarding each source which exists in the domain and the unicast address of that source. This table is created on C-MAPPERS and C-RPs and is used in related processes.

PIM domain topology table (PDTT):

This table is used to store the information needed to find C-RP/RPs, C-MAPPER, TRs and other components that may exist in a PIM-NG multicast Domain. It holds the unicast address of such components alongside other information needed.

Core topology table (CTT):

It is created on C-MAPPERS inside the core-domain. It only holds the address of any existing TR'(s) inside the core domain. And is the only PIM-NG topology table that CAN be exchanged between the PIM-NG-CORE-DOMAIN and any existing PIM-NG-DOMAIN connected to the core domain. Also this table is the ONLY PIM-NG topology table that CAN be exchanged between peer C-MAPPERS in different PIM-NG-DOMAINS so that all PIM-NG-DOMAINS will know about the TR'(s) inside the core domain. This table MUST NOT be exchanged between PIM-NG-CORE-DOMAINS.

Peer mapping table:

It is created on C-RPs and C-MAPPERS as soon as a C-RP or C-MAPPER becomes peer with new C-RP or C-MAPPER.

Internal multicast source table (IMST):

This table is created on PPER'(s) and in case of private network'(s) within a unified PIM-NG multicast domain connected to other parts of the Network using NAT, on EDGE-CLIENTS, and helps the PPER or EDGE-CLIENT to act on behalf of a host in search of a source by putting an entry inside the table to keep track of the behavior of the domain. It holds the address of the requesting client and the multicast group requested.

Client request table (CRT):

This table is created on C-RP'(s) and holds the unicast address of clients to which the C-RP sends a NULL-ACK in response to the clients request to find the source for multicast group(G)along the multicast group (G)which the client needs to find the generating source.

Bidirectional Groups Table (BGT):

This table is created on C-MAPPERS and C-RPs and hold the information of active or to be active Bidirectional multicast groups.

Bidirectional Translator TR (TTR):

A TR part of all available TR groups which is used to communicate with TRs in other multicast domains with regards to exchanging the information of active Bidirectional multicast groups.

3. IP Address considerations

PIM-NG processes need to use 3 new multicast destination addresses from Internetwork Control Block [2]. These addresses are going to be used in PIM-NG processes and are needed to be assigned. For the simplicity of explanations in this document, multicast address 239.0.1.188, 239.0.1.189 and 239.0.1.190 from the Scoped Multicast Ranges are used as advised by IANA. But PIM-NG needs the use of addresses from the "internetwork control block" and the use of the 3 addresses from the scoped multicast range is only for the sake of simplicity in the process of explanation.

Also it MUST BE noted that any IP addresses, whether multicast or unicast, used in this document from this point forward ARE ONLY used as an example to simplify the explanation process. For example addresses 228.8.8.8 and 229.9.9.9 are used only to simplify the explanation process.

4. PIM-NG processes

PIM-NG processes are related tightly to the new packet formats defined for it. So in each section related packet formats are going to be explained or shown.

4.1. New packet Header

The new packet'(s) format is designed and defined in the way, so it can support the needs of PIM-NG.

The new packet header format specification for PIM-NG is as follows:

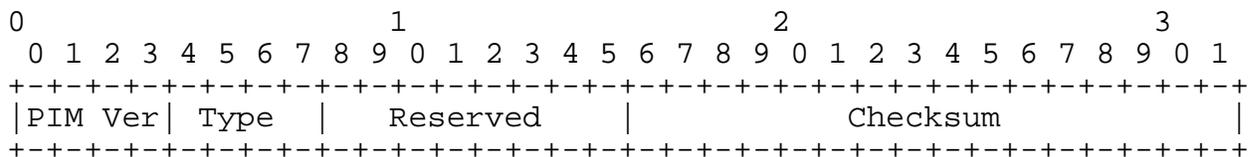


Figure 1 the new packet Header

PIM-ver : is 3

As the new packet format will be used in different situations, the new type field definitions must be explained. The new definitions are as follows:

5 bit TYPE field to support up to 32 different functions:

| Message type | destination |
|-----------------------------|---|
| 0- Hello | Multicast to All-PIM routers |
| 1- Register | Unicast to RP and EDGE-CLIENT |
| 2- Keep alive to RP | Unicast from source to RP |
| 3- Join/prune | Multicast to ALL-PIM-ROUTERS |
| 4- Request For Source | Unicast to RP |
| 5- Ack to Client/source | Unicast from RP or EDGE-Client To source |
| 6- Assert | Multicast to ALL-PIM routers |
| 7- Host request to C-MAPPER | Unicast to C-MAPPER |
| 8- RP introduction | Multicast to ALL-RPs |
| 9- RP introduction | Unicast to C-MAPPER |
| 10- C-MAPPER introduction1 | Multicast to ALL-PIM-NG routers |
| 11- C-MAPPER introduction2 | Multicast to ALL-MAPPERS |

| | |
|--------------------------|---|
| 12- Request-for-C-MAPPER | Unicast to SC-MAPPER |
| 13- C-MAPPER ack | Unicast to Client |
| 14- Edge | Unicast to C-MAPPER |
| 15- BPR | Unicast to C-MAPPER |
| 16- TR | Unicast to C-MAPPER and Peer-TR |
| 17- NASN | Unicast from C-RP to C-MAPPER And Multicast from C-MAPPER to ALL-PIM-NG routers |

The above definitions will be explained, as we proceed through out different sections of PIM-NG specifications.

4.2. Source Discovery

In the original PIM-SM specifications of the communication between a SOURCE and RP we see that the source sends Register messages to the RP and the RP will only accept the Register message, if any host or hosts asked for that particular multicast group. Otherwise the RP will send a register-stop message back to the source and the source starts a register-suppression timer of 60 seconds. And 5 seconds before the suppression timer expires the source sends a register-message with its null-register bit set. Now if the RP doesn't know any hosts asking for that specific group it will send another register-stop. The process goes on until a host asks the RP about that group.

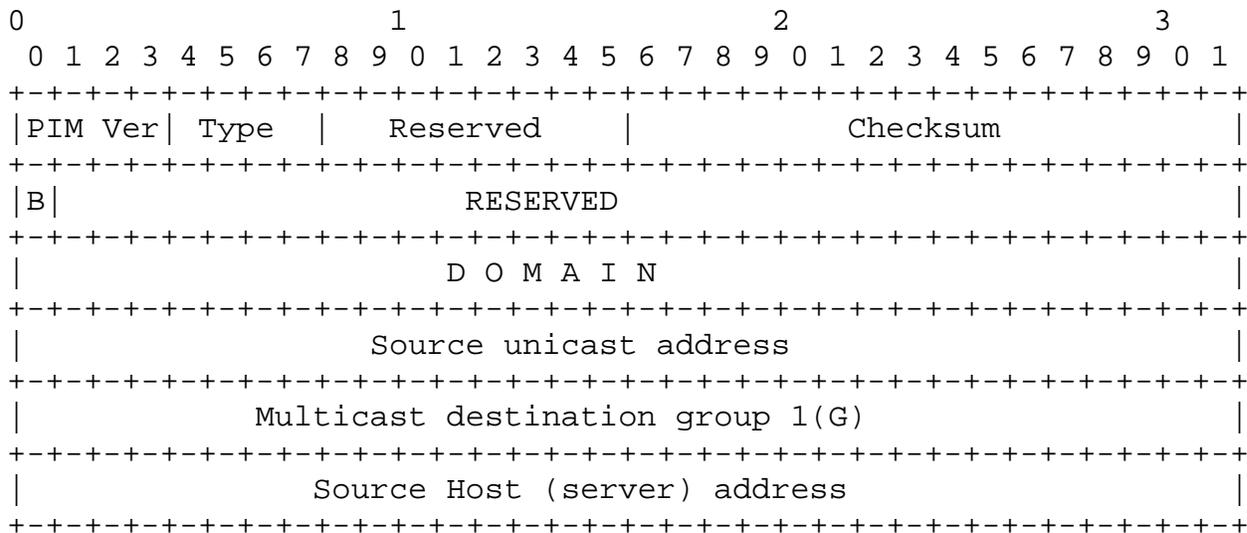
Now what happens if right after the suppression timer starts by the source a host comes up and asks for that specific source? As it is explained in the original PIM-SM specifications, the host will have to send its join messages to the RP until the RP hears again from the source and this time, due to the fact that there is a host asking for the group the RP won't send a register stop and will forward the packet down the RPT .This process didn't seem so efficient, so some changes has been made to the way a SOURCE communicates with the RP alongside the new packet definitions

4.2.1. Source register process with RP

In PIM-NG the process to initially deliver the multicast traffic to a host asking for it, is somehow different from that of PIM-SM. It initially starts by:

- 1- Source introducing itself to a router called RP(rendezvous point)
- 2- RP keeps and enters the information related to the source in to a table for further use.
- 3- Host asks the RP about the source of an specific group
- 4- Host sends a join request to the source directly

For the sake of simplicity let's consider that all the routers know the address of C-RP. The source of the multicast traffic starts its process by introducing itself to the C-RP, by sending Unicast-Encapsulated register messages to it. The source does the introduction process completely different from the original PIM-SM specifications. In PIM-SM source introduces itself by sending a packet containing the multicast data to be sent. But the introduction process in PIM-NG is done through sending a register-request packet containing only the address of the source along the multicast group address the source represents (0). The C-RP receives the message and by looking in its type field, it knows that it is a Source register request message sent from a source. From here the C-RP will act differently from what is defined in the PIM-SM specifications.



Inside the register-message, each source sends a keep-alive timer value defaulting to 30sec to the C-RP, which later will be used by the C-RP in the process.

Then the C-RP sends a unicast acknowledge-message(0) back to the source, using the unicast address of the source .after receiving the acknowledge the source proceeds with sending periodic keep-alive messages to C-RP , to tell the C-RP that it is alive and so the C-RP wont delete the entry related to the source from its multicast-mapping table. For each keep-alive C-RP receives it will send an acknowledgement. If the C-RP doesn't hear the first keep alive, it will start to count down the expiry timer for the source which is by default 3*keep-alive timer .if the C-RP doesn't hear from the source in the expiry time duration it will delete the entry for that source.

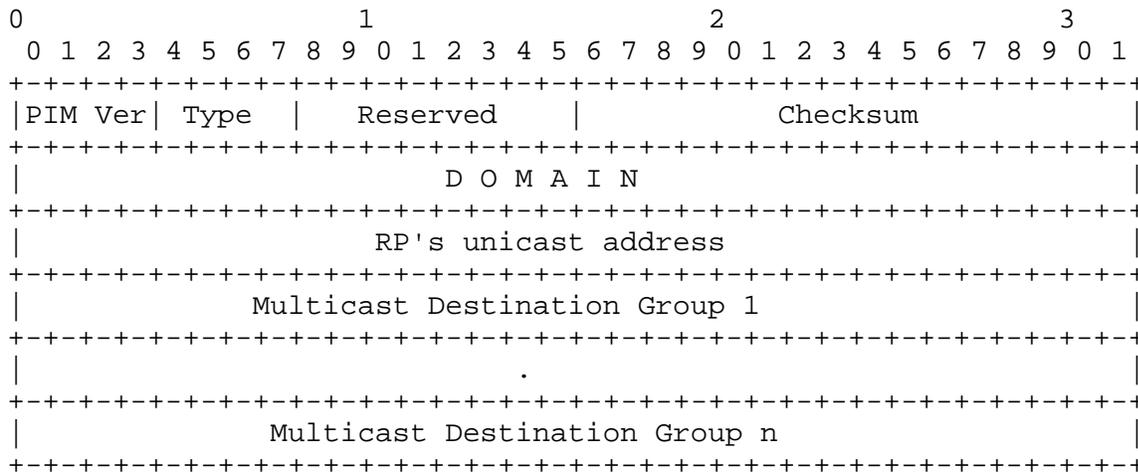


Figure 4 C-RP acknowledge message

- o Type : RP acknowledge to source

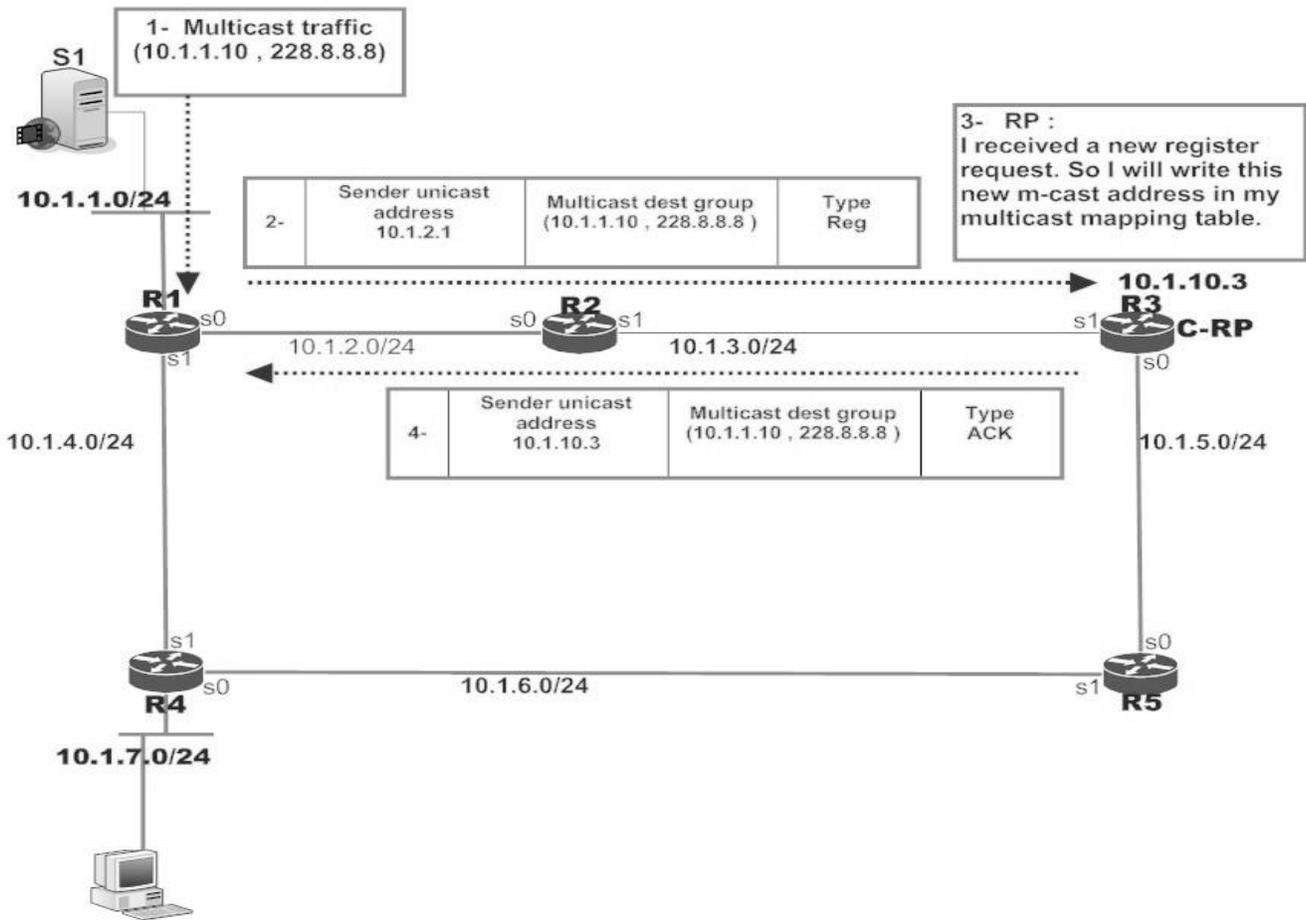


Figure 5 source communication with the RP

Now let's start by explaining the process related to the communication of the source and RP as shown in 0:

- 1- The server behind R1 (called S1), starts sending multicast to 228.8.8.8 and R1 receives those multicasts on its connected LAN interface.
- 2- R1 will try to contact the C-RP and introduce itself and the multicast group it represents, by sending the Unicast-Encapsulated register-request message to the address of the C-RP. C-RP receives a PIM packet and looks at the TYPE-FIELD.

- 3- After looking at the TYPE_FIELD C-RP finds out that it is a register-request which is sent from a source. So the C-RP looks in to the source unicast address and the Multicast group destination address fields and writes the unicast address of the source alongside the multicast group address it represents alongside the keep-alive timer and the expiry timer as a new entry in its MMT.
- 4- RP sends acknowledge back to the SOURCE.
- 5- After receiving the acknowledge the source starts its keep-alive timer ,and will send periodic keep-alive messages as long as it wants to send traffic for the multicast group address it represents. These keep-alive messages can be simply a register message.

So this was the process related to the communication between a source and the C-RP. In the next section the process related to the communication between a host and a source, or how a host sends join request for a multicast group address to the source will be explained.

4.2.2. Communication between client and the source

The process through which, a host finds the source and communicates with it in order to receive the multicast traffic, is in parts different from that of the original PIM-SM specifications. So let's have a flash back at the process of the original PIM-SM:

1. Host sends a join message of (*, G) to the RP, and joins the RPT for the (*, G).
2. RP sends the join request to the source with (S,G), indicating that a host is asking for the traffic or the RP simply forwards the

multicast packets received from the source down RPT towards the host.

3. Source sends the multicast packet to the RP ,and RP forwards it to the host
4. After receiving the first couple of packets , the host sends a join request directly to the source
5. So the host leaves the RPT and switches to SPT for (S,G)
6. Host sends a prune to the upstream router, so the RP will stop forwarding that traffic.

This process works fine, but not as efficient and as fast as possible. Waiting for the first couple of packets or even first packet to be received and then switch to SPT for the (S,G) and after that sending a prune to the upstream router ,doesn't seem so efficient. What happens when a host sends a join for a particular group (*, G) or (S, G) right after the RP sent a register-stop to the source? The process seems to waste some valuable time.

So PIM-NG uses a new process to reach the source, alongside new messages. We are going to consider that the host/hosts know about the C-RP for now.

Let's start by assuming that a HOST behind R4(0), asks for 228.8.8.8 traffic using IGMPV2-3 .the request arrives at R4, and now it is up to R4 to find the source and send Join message to the source. But in order to find the source R4 knows that it will have to ask C-RP as the information station of the Multicast domain about the address of the source.

So R4 sends a unicast-encapsulated PIM-NG request message to the C-RP. R4 sets the TYPE-FIELD to " Request For Source" and puts its own unicast IP address in the "SOURCE UNICAST ADDRESS" field and puts the multicast group address 228.8.8.8 in the format of (* , 228.8.8.8) in

the " Multicast group destination address " field, And sends the packet to the destination address of C-RP.

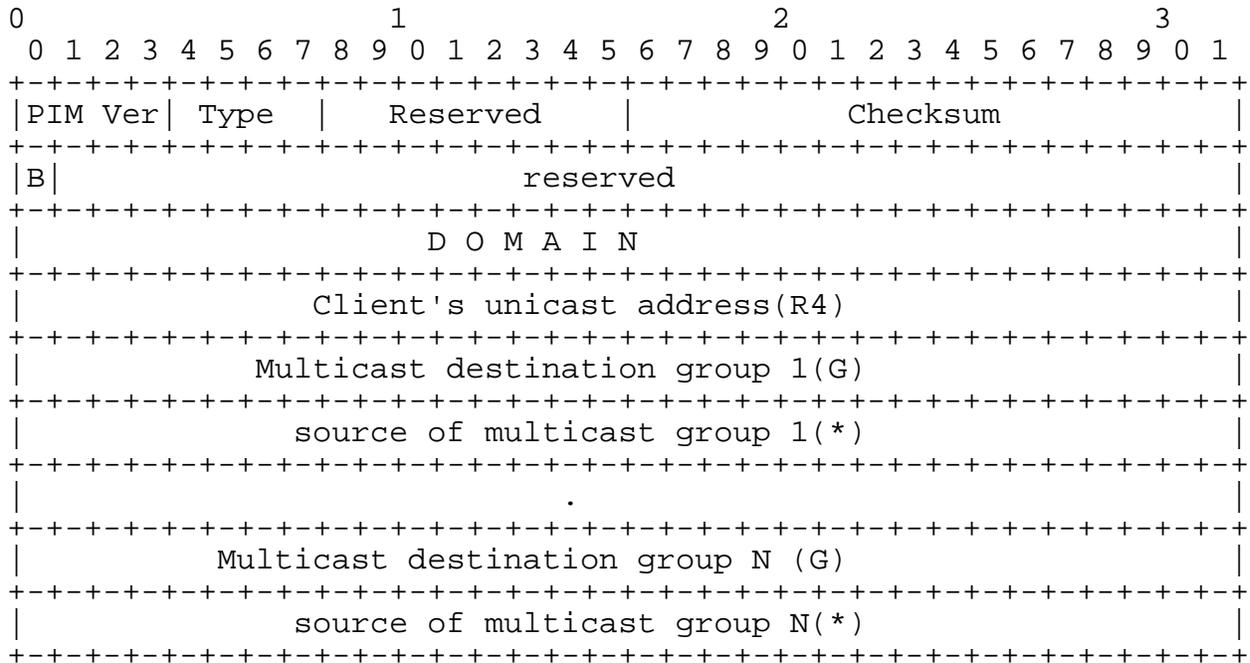


Figure 6 Request for source packet

- o Type: Request For Source

There might be a case that there are two source HOSTs sending traffic for the 228.8.8.8 group and the host behind R4 , asks through IGMP to receive the traffic generated by a specific source (i.e.10.1.1.10) .then the format of the request packet it sends will be as shown in 0.this is different from SSM multicast. As it is considered that the CLIENT/HOST knows the address of the server or host originating the traffic destined for group (G), but not the address of CLIENT who is responsible for forwarding the traffic on behalf of the server and acting as the DR.

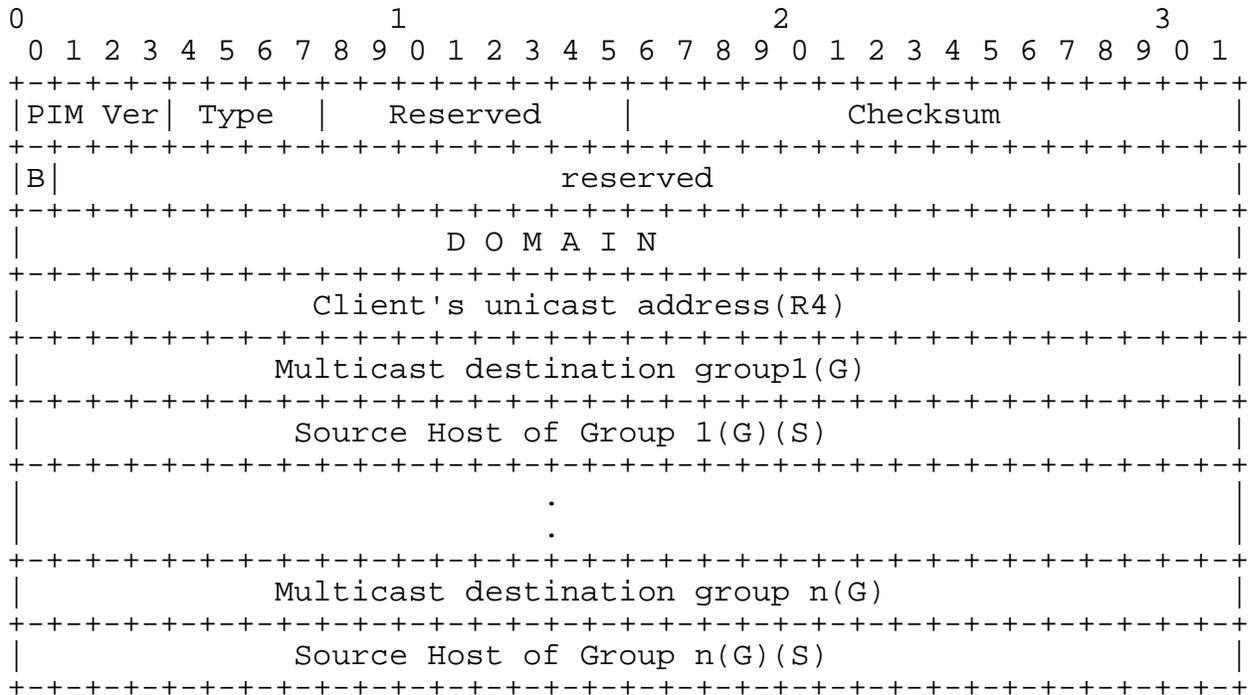


Figure 7 Request for Source packet format

- o Type: Request For Source

C-RP receives the packet, and checks the type-field.the type-field indicates that it is a host request to find the unicast address of a source. So C-RP checks the Multicast group destination address field and finds out that the host is either:

- 1- Looking for the unicast address of a source representing 228.8.8.8 multicast traffic (*, G).
- 2- Looking for the unicast address of a specific source representing 228.8.8.8 multicast traffic (*, S, G). Different from PIM-SSM.

Then the C-RP looks in to its MMT or AMMT if the network is consisted of separate PIM-NG domains, and acts in two different ways:

- 1- C-RP finds an entry in its MMT which indicates that the source exists:

In this case, instead of forwarding the traffic towards the source and join the SPT for (S,G), C-RP sends back the address of the source directly toward R4 using R4's unicast address found in the "source unicast address " field with a packet shown in 0 :

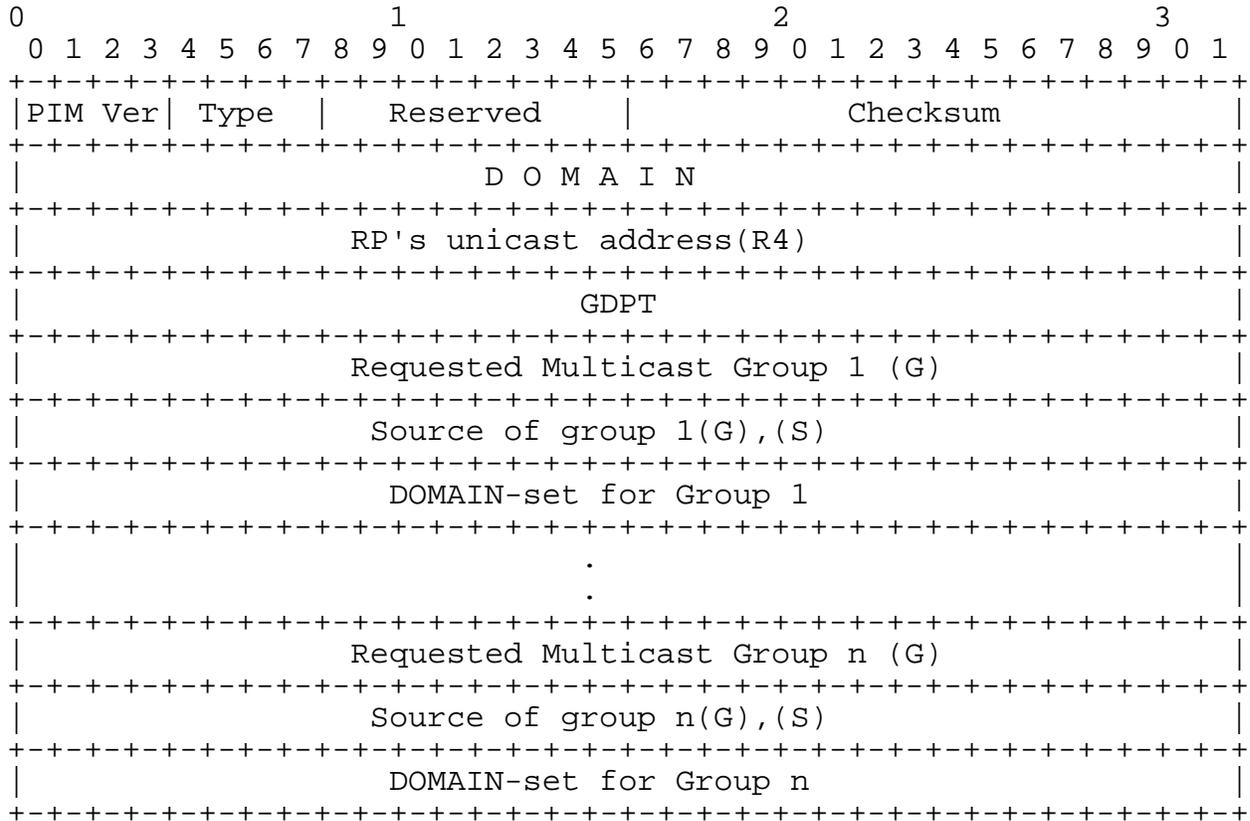


Figure 8 C-RP Acknowledge packet format

- o Type: RP acknowledge to host
- o DOMAIN-SET: is part of the AMMT and its use will be explained later. It shows the domain in which a source is registered. And also the path towards the source, so that a client can decide on generating the join/prune message better.
- o GDPT Filed: Holds the global Delay prevention timer set on C-RP and is used to automatically synchronize this timer on all the clients receiving a Null-ACK from C-RP.

C-RP sets the type field to "RP acknowledge to host" and puts its own unicast IP address in the source unicast address field and fills out

the "Multicast group destination address" field in the format of (S, G).

One other thing that the C-RP sends back to the client is the Domain set, which shows the domain in which the source is resided and also the path through which the join/prune message can reach the source. In case that the source resides in a separate domain and to be more specific the PIM-NG network is consisted of different and separate domains, clients will use the domain set to decide whether a join/prune message MUST path a core domain first or MUST NOT which the related concepts will be explained later. A Domain-Set with the Null value in it indicates that a source is resided within a Multi-Access Network which the related processes will be discussed in the appropriate sections.

2- C-RP doesn't find an entry in its multicast-mapping table:

In this case C-RP answers to the host, with a packet which indicates that the source doesn't exist now and the host must try again later. 0 shows the packet sent by RP to the host:

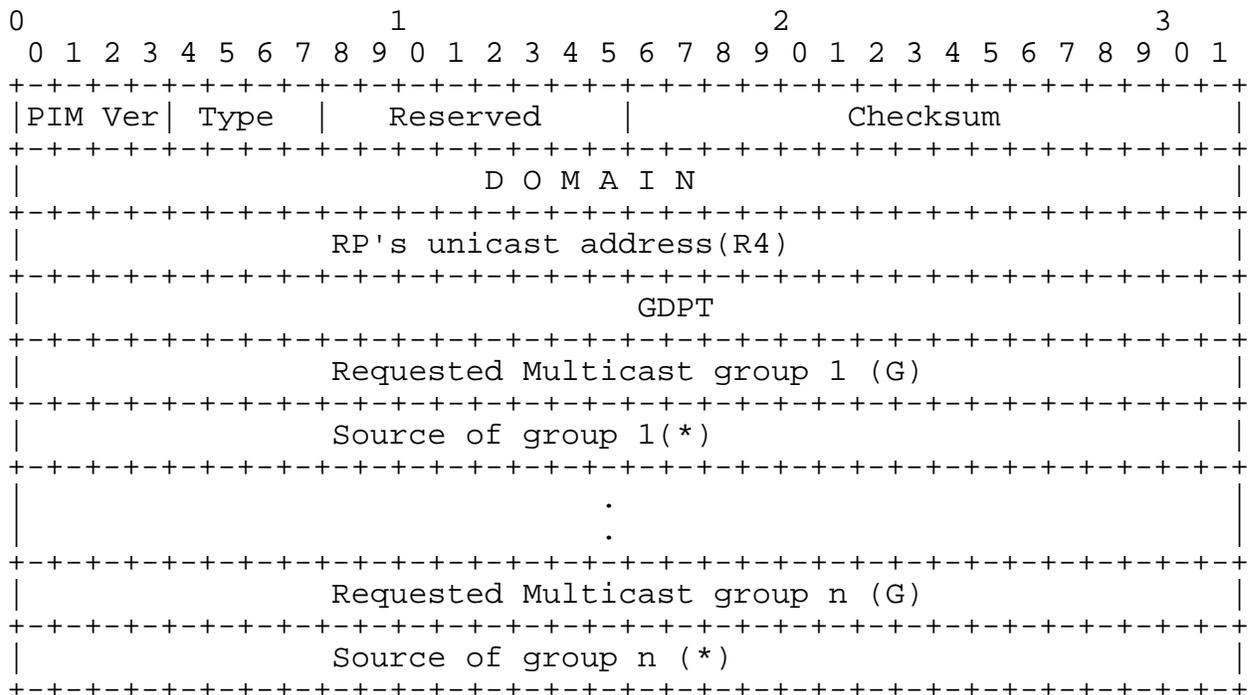


Figure 9 C-RP answer to the host in case that source doesn't exist

- o Type: RP acknowledge

R4 receives a unicast-encapsulated PIM_NG packet, and looks into the TYPE-FIELD. The type field indicates that it is the C-RP answer to the host request for finding the unicast address of a source. So R4 will check the source unicast address field and makes sure it was sent from the C-RP and then looks into the Multicast group destination address field.

R4 will act in 2 different ways, depending which one of the 2 above conditions is met:

- 1- If the entry inside the " Source of group field" is in the format of (*, G), then R4 understands that the source doesn't exist for now and it has to try again later. The time of resending the request will be equal to the HELLO time or 30 seconds.
- 2- If the entry inside the "Source of group field" is in the format of (S, G) then R4 will take S, and looks into its unicast routing table to find the best way for reaching to S which is in our example the unicast address of R1. After finding the best path to R1, R4 sends a join/prune message to the upstream neighbor in the best path towards the source so the join/prune message goes hop-by-hop until it reaches the desired destination. And depending on whether a TR exists in the domain or not and whether any core-domain is considered in the PIM-NG overall network, client will :
 - o Join the (S,G)tree in case there are no residing TRs inside the domain
 - o Join the (S, G, RPT) which is assumed by PIM-NG specifications the shortest path tree rooted at TR in case a TR exists in the domain. Joining the (S,G,RPT) tree inside each single domain is strongly advised due to the fact that it will reduce the unnecessary join/prune messages being sent towards a source in case that new clients come up later which need to receive the same traffic.
 - o Join the (S, G, RPT) which is considered by PIM-NG specifications the shortest path tree towards a source rooted at core-domain-TR. For this tree to be formed 2 conditions MUST exist :

1. Core-domain MUST do exist, and any existing TR inside the core-domain MUST be known to the PIM-NG population.
2. A source MUST be reachable by passing through the core-domain. Which the decision is up to the client generating the join/prune message, by checking the DOMAIN-SET associated with a source.

This process MUST be done if the above conditions are met, due to the fact that, if a PIM-NG-Core-domain exists there will be one or more PIM-NG-DOMAINS connected to it and if one or more new clients in other domains come up later which are in need of receiving the same traffic, it will be waste of time and network resources to send the new join/prune messages towards the source. Related concepts will be explained later.

R4 MUST save the state for the join message being sent. By saving the incoming interface for (*, G) or (S, G, TR) if it is a SSM join, and the outgoing interface for the (S, G) or (S, G, TR).and use this for RPF check later when receiving packets destined for (G). The related processes are done the way described in RFC4601 [7].

The packet that is being sent by R4 to R1 as a join/prune message is shown in 0 along the new encoded source address format

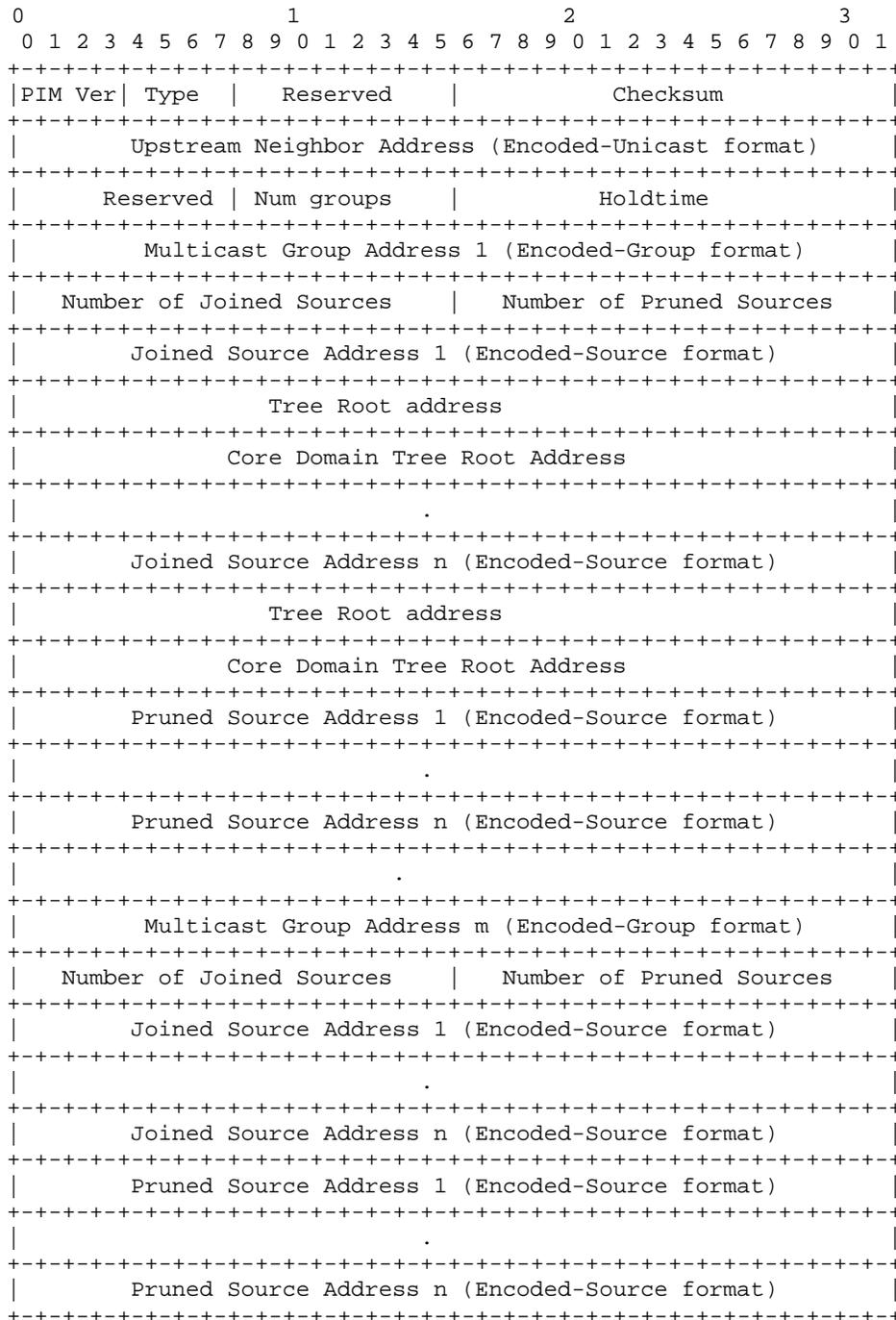


Figure 10 Join/prune message format

- o Type : join/prune

- o Tree-Root address: if any TR exists in the Domain, this field holds the address of the TR to towards which the message MUST be forwarded first, before being forwarded towards the source. The concepts will be discussed later.
- o Core Domain Tree-Root Address: holds the unicast address of the TR resided in the core domain, and is used only in topologies that contain Core-domain implementations, and incase the source is not an internal source and MUST be reached by passing a core domain. The related concepts will be discussed later.

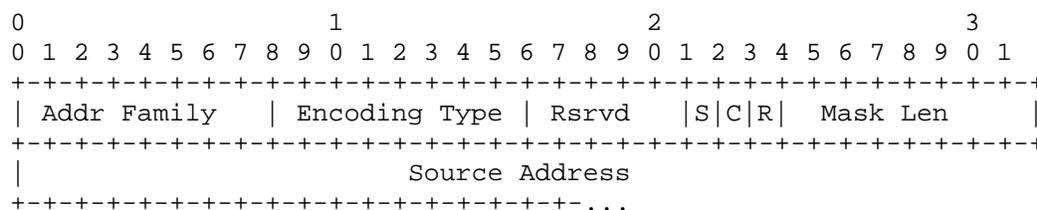


Figure 11 Encoded Source Address Format

- o S-BIT: only for compatibility with PIM-V1.
- o R-BIT: Tree Root bit is set by the client who is sending the join/prune message for the first time for a source in a topology with one or more TR'(s) to join the (S, G, TR) tree. a value of 1 is set by the creator of the join/prune message to tell upstream routers that the packet MUST be forwarded first towards the existing TR in the domain which the address of it can be found in the appropriate field .and a value of 0 can be set by the TR after receiving the join/prune message and before forwarding it towards the source. A value of 0 dictates to upstream routers that the join/prune had reached an existing TR In the domain and from now on MUST be forwarded towards either the source, if the source is an internal source or there are no core domain implementations, or the TR inside the core domain if the source is an external source that can be reached by core domain. This bit can also be set by a PER or by a PER/BPR in case a join/prune message is received from another PIM-NG domain or a PIM-SM domain and must be forwarded towards any existing TR in the domain to either reach a source inside the domain or towards a source in other domains.
- o C-BIT: Core-Domain-BIT can be set by either the client originating a join/prune message or the TR inside the core domain. If a client is originating a join/prune message for a source which is not an internal source and can be reached by passing through an existing core domain, it will set this bit to 1. A value of 1 means that the

packet MUST be forwarded towards the TR inside the core domain first with the unicast address that can be found in the appropriate field. And a value of 0 can be set by the TR inside the core domain, which dictates to upstream routers that the join/prune message MUST be forwarded hop-by-hop towards the source from now on. This bit can also be set by a BPR in case it receives a join/prune message from a PIM-SM neighbor router. The related processes and concepts will be discussed later.

Source (R1) receives the join/prune message. And joins the shortest path tree for (G) and starts sending the multicast data destined for (G) down the shortest path tree towards the receiver.

In order to receive packets from R1, R4 will have to send periodic keep-alive or join messages every 30 seconds to R1. R1 receives the join-request and will know that R4 still needs the traffic. If R1 (source) doesn't receive 2 keep-alive messages from the host (R4) it will assume that the host doesn't exist anymore and will automatically stop forwarding multicast traffic to the host (R4).

If R4 no longer needs to receive the traffic from R1, it will inform R1 by sending a prune message to the upstream PIM-NG router forwarding the traffic .this is done due to the fact that there might be other hosts connected to the upstream routers in need of receiving the traffic.

4.2.3. Source specific multicast

Source specific multicast (SSM) [6] can be implemented under the following rules:

- o PIM-NG routers MUST NOT send the unicast-encapsulated Request for Source message for SSM addresses.
- o PIM-NG routers MUST NOT send a register message for any packet that is destined to an SSM address.
- o If any TR exists in a PIM-NG domain, a PIM-NG router MUST join the (S, G, RPT) rooted at the TR which is considered the shortest path tree towards a source by PIM-NG specifications.
- o ONLY incase that in a PIM-NG multicast Domain TR is implemented and the TR and C-RP are the same components, a Client is allowed to create the join/prune message for the SSM address and send it towards the existing C-RP, which is also the TR.

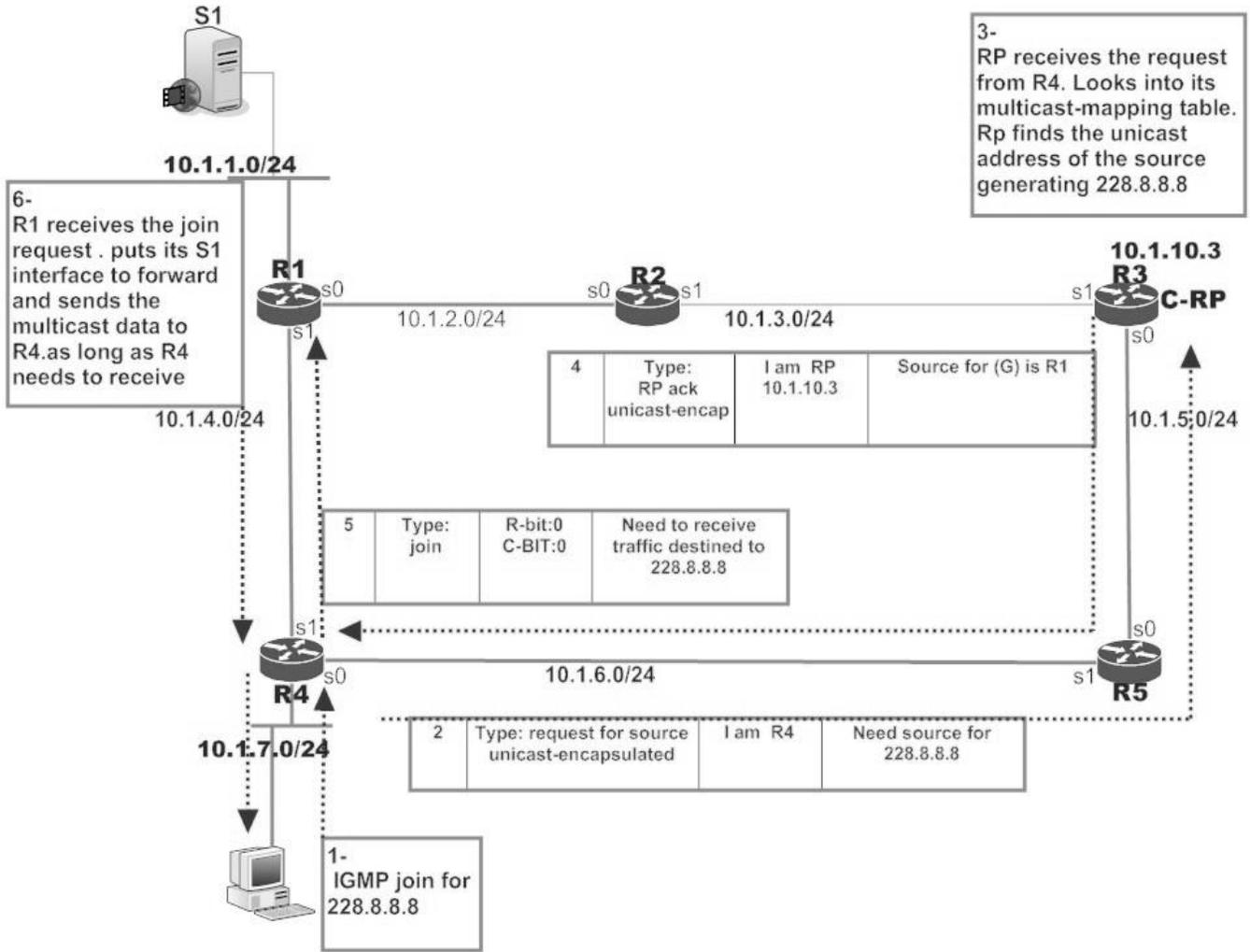


Figure 12 Communication between Host and Source

4.2.4. Source Discovery Delay Prevention

One of the important factors in a multicast domain is the time through which a client discovers or finds the source of a multicast group (G).

Up to this point the processes through which a source registers with C-RP and finally a client discovers the (S) for (G) by sending a request for source message to the C-RP have been explained. Throughout the explanations we saw that if a source is registered, since the information regarding the source is saved in a special table called MMT as soon as a client sends a request for source to C-RP the C-RP will be able to send back the unicast address of (S) for (G) in an acknowledge message to the client.

Now what happens if a client sends its request for source message to find the (S) for (G) to the C-RP and the source (S) for (G) is not registered yet?

As it has been explained by PIM-NG specifications, in such a case the C-RP will send a NULL-ACK for the requested (G) back to the client which is in the form of (*, G) and the client will have to start a 30 second timer after receiving the NULL-ACK from C-RP and send periodical request for source messages until the source becomes active and the C-RP responds with an acknowledge containing (S) for (G).

Now what will happen if the (S) for (G) becomes active right after the C-RP sends the NULL-ACK to the client?

As explained the client will have to try again in 30 seconds and send its periodical request again to the C-RP. In this case and considering that the (S) for (G) becomes active 1 second after the C-RP responds to client with a NULL-ACK, the worst case scenario will be that an unwanted 30 second delay will occur until the client sends its request again and finds the (S) for (G).

To eliminate this delay in source discovery, PIM-NG introduces a mechanism or process through which the C-RP will:

- o Save the state of the clients to which it sends a NULL-ACK for the requested (G).
- o As soon as the (S) for (G) becomes active and registers the C-RP will inform the clients who had been responded with a NULL-ACK for that specific source (S).

Through the above the C-RP will be able to inform the client as soon as the source becomes active and the client will be able to find (S) for (G) as fast as possible.

To do so PIM-NG specifications introduces its Source Discovery Delay Prevention method which is divided in to 2 modes:

- o Default-Mode.
- o Admin-Mode.

4.2.4.1. Delay prevention in Default-Mode

This mode is activated by default as soon as a PIM-NG-Aware Router is configured to become a C-RP and is suitable for all types of multicast network topologies form small sized networks to large sized networks. But it is strongly advised by PIM-NG specifications to use this mode in small sized networks in which the number of clients are low.

Bellow rules and specifications apply to Default-Mode:

- o It is done automatically by existing C-RPs.
- o It is suitable for small sized network topologies but applicable to medium and large sized networks too.
- o In this mode C-RP will save the state of clients to which a NULL-ACK is sent in a special table called Client Request Table (CRT). And by state PIM-NG means the unicast address of the client alongside the multicast group (G) which the client needs to find the associated source(S) for it and finally a timer which defaults to 33sec (30+3).
- o C-RP MUST only put an entry in CRT for the clients to which it sends a NULL-ACK.
- o For each (G) that the C-RP sends a NULL-ACK an entry MUST be put in CRT.
- o The timer which defaults to 33sec(30sec for the default periodical client request + 3sec) is a countdown timer and resets to 33sec ONLY when both of the bellow conditions are true:
 - 1- The source is not active and registered within 33seconds.

- 2- C-RP receives clients periodical Request For source within 33sec which shows that the client is still interested in finding (S) for (G).
- o An entry is deleted from CRT when any of the bellow conditions is true:
 - 1- The source for (G) becomes active within the time period of 25 seconds or since the timer in CRT is a countdown timer 33 to 8 seconds which results in C-RP sending an Acknowledge containing (S) for (G) to the client.
 - 2- The source becomes active in the time period between 8 to 0 seconds and C-RP receives the periodical Request for Source form the client and responds to client's request by an Acknowledge of (S, G).
 - 3- The source is not activated within the 33sec time interval and C-RP doesn't receive the periodical Request of client for that specific (G) in 33sec which indicates that the client is not interested anymore in finding (S) for (G).
 - 4- The source is activated and registered in the time interval between 8 to 0 seconds but the C-RP doesn't receive the periodical Request for Source from the Client which shows that there client is not interested anymore in finding (S) for (G).
 - o As long as there is an entry in CRT, for each new source (S) that becomes active C-RP MUST check the contents of CRT and check the (G) represented by newly activated source(S) against the entries inside CRT and if there is a match then depending on the time that is shown by the timer C-RP MUST act in one of the bellow ways:
 - 1- If the timer shows a time between 33 to 8 seconds, the C-RP MUST immediately send its unicast-encapsulated Acknowledge containing (S) for (G) to the client and after the timer goes off and becomes equal to 0 delete that entry from the CRT.
 - 2- If the timer shows a time between 8 to 0 seconds, the C-RP MUST NOT send an Acknowledge back to the client and MUST wait to receive the client's periodical Request for Source and then send an Acknowledge in response to the Client's Request and delete the entry from CRT. Note that each 30 seconds the client who has received a NULL-ACK will send periodical Request to C-RP and since when the timer shows 8sec we will

have 3 seconds to the next periodical Request it will help to reduce C-RP's work to wait for the next periodical Request, also in case that the client is not interested anymore in finding the (S) for (G) then at this point it will be a waste of C-RP resources to send an Acknowledge after the timer passes the 8sec.

- o For Default-Mode to work without problem the timers on all routers MUST be the same. Thus the C-RP sends the value of the Timer set on it in the Null-ACK packet in the GDPT field so that the synchronization is done automatically.
- o If the default timer is changed by an administrator, the time range in which the C-RP is to send a notification MUST be between the higher limit and the 8 seconds. so if for instance the time is set to 60 then C-RP MUST only send notifications between 60 to 8 seconds.

The above being explained and since the involved processes will keep the C-RP a little busier it is suggested to use this mode in topologies with a low number of clients although this mode works perfectly under any topology. Also it must be noted and considered that due to the unique Domain concept(4.6.1.1.) introduced by PIM-NG which allows implementing the Sub-Domain concept(section 4.6.5) it is possible to divide any large multicast domain to a number of smaller domains which then makes it possible to easily use the Default-Mode under any network topology.

4.2.4.2. Delay prevention in Admin-Mode

This mode is a suitable choice for large network topologies with many clients which will reduce the work overhead of C-RP'(s). and is suggested to be used in scenarios in which C-MAPPER'(s) are considered and also the Sub-Domain concept is not considered.

Bellow rules and specifications apply to Admin-Mode:

- o The Admin-Mode as the name implies needs to be activated on all existing C-RPs by an administrator.
- o For each Client's Request for Source to find the (S) for (G) that the C-RP sends a NULL-ACK of (*, G) an entry MUST be put in CRT.
- o The entry in CRT is as explained in Default-Mode.

- o As long as there is an entry in CRT, for each new source (S) that becomes active C-RP MUST check the contents of CRT and check the (G) represented by newly activated source(S) against the entries inside CRT and if there is a match then depending on the time that is shown by the timer C-RP MUST act in one of the bellow ways:
 - 1- If the timer shows a time between 33 to 8 seconds and Dynamic RP discovery type 1(4.4.2.2.) is in use, the C-RP which is also acting as the C-MAPPER MUST send a Multicast Notification called New-Activated-Source-Notification(NASN) to the entire population of PIM-NG-AWARE routers using the destination address 239.0.1.190 which is used by this draft for the simplicity in explaining the processes as the address used by C-MAPPER'(s) when sending multicast introductions or notification within the multicast domain. This notification contains either only all the (G) for which there is an entry in CRT or both the (G) and the unicast address of the source (S) representing (G). both methods can be used although the later is faster approach but less secure because it is not a good approach to let all the clients know about the unicast address of all the sources that are active in the domain.
 - 2- If the timer shows a time between 33 to 8 seconds and Dynamic RP discovery type 2(4.4.2.3.) is in use, then the C-RP MUST send a unicast-encapsulated New-Activated-Source-Notification to the closest Active C-MAPPER the way it sends its introductions. This notification contains either only all the (G) for which there is an entry in CRT or both the (G) and the unicast address of the source (S) representing (G). both methods can be used although the later is faster approach but less secure because it is not a good approach to let all the clients know about the unicast address of all the sources that are active in the domain. After C-MAPPER receives the C-RP notification it will send out a multicast notification to the entire population of PIM-NG-AWARE routers within the domain using the destination address 239.0.1.190.
 - 3- If the timer shows a time between 8 to 0 seconds, the C-RP MUST NOT send the notification to the C-MAPPER and MUST wait to receive the client's periodical Request for Source and then send an Acknowledge in response to the Client's Request and delete the entry from CRT. Note that each 30 seconds the client who has received a NULL-ACK will send periodical Request to C-RP and since when the timer shows 8sec we will have 3 seconds to the next periodical Request it will help to reduce C-RP's work to wait for the next periodical Request,

also in case that the client is not interested anymore in finding the (S) for (G) then at this point it will be a waste of C-RP resources to send an Acknowledge after the timer passes the 8sec.

- 4- In all the above conditions, after the C-RP receives the client's request for source representing (G) for each (G) there is an entry in CRT, the C-RP MUST delete the entry for that (G) from the CRT after the timer goes off and is equal to 0.
- o All the clients receive the notification and those that have received a NULL-ACK for (G) will send their queries to C-RP again only if there is an entry for (G) they had received a NULL-ACK for. Since considering that if the clients send their queries immediately RP might get too busy, then each client MUST activate a 3 second timer and after the timer goes off will send its query. Or if the notification contains the unicast address of the source too then clients will be able to join the SPT for (S, G) immediately. it is possible to make modifications so that this special notification is only readable by clients that have received a NULL-ACK

```

+-----+
|Client unicast addr| Requested Group (G) |Timer      |
+-----+
|                   |                       |           |
+-----+
    
```

Figure 13 Client Request Table (CRT)

```

0          1          2          3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|PIM Ver| Type  |   Reserved   |                               |Checksum|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |D O M A I N|                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |C-MAPPER's unicast address|                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |Activated Group 1 (G-1)|                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |. . .|                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               |Activated Group N (G-N)|                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
    
```

Figure 14 New-Activated-Source-Notification message format

- o Type: NASN (New-Activated-Source-Notification)

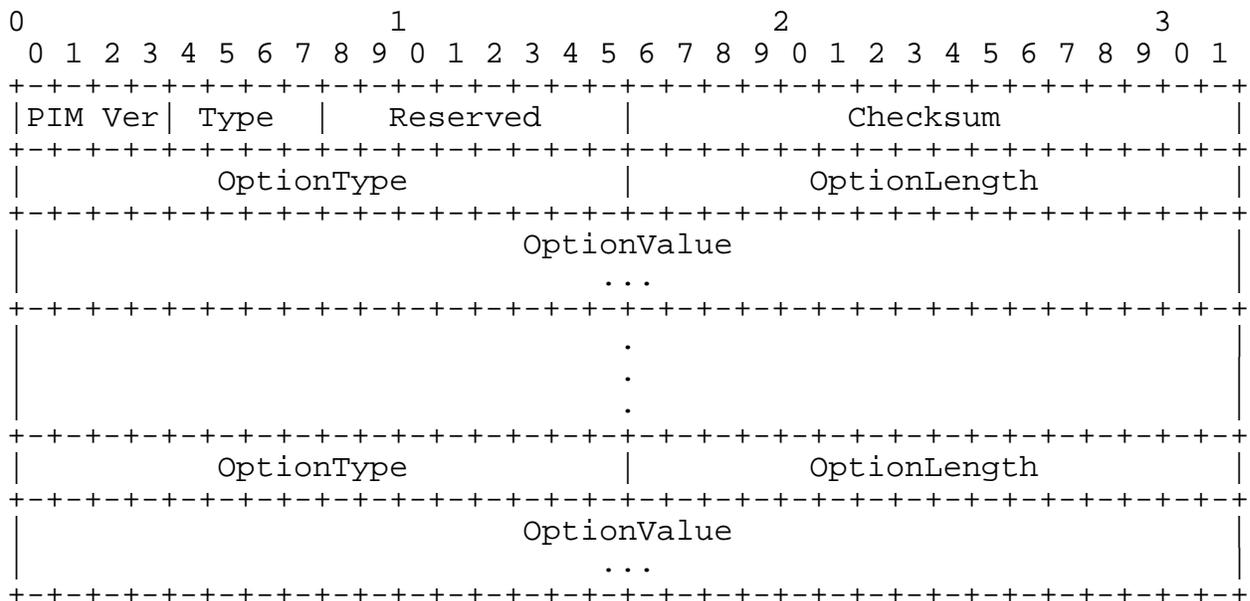
4.3. Communication between PIM-NG routers

It is important, that connected PIM-NG neighbor routers maintain a steady state of connection. In order to do this PIM-NG routers will send periodic HELLO messages or simply called Keep-Alive messages to their neighbors.

This HELLO messages will be sent to the address 224.0.0.13, all PIM routers.

What is changed is not the process of sending, rather the contents of the hello message. A new hello packet is defined and compared to the original PIM-SM specifications, 0.

Original PIM-SM HELLO packet



PIM-NG hello packet

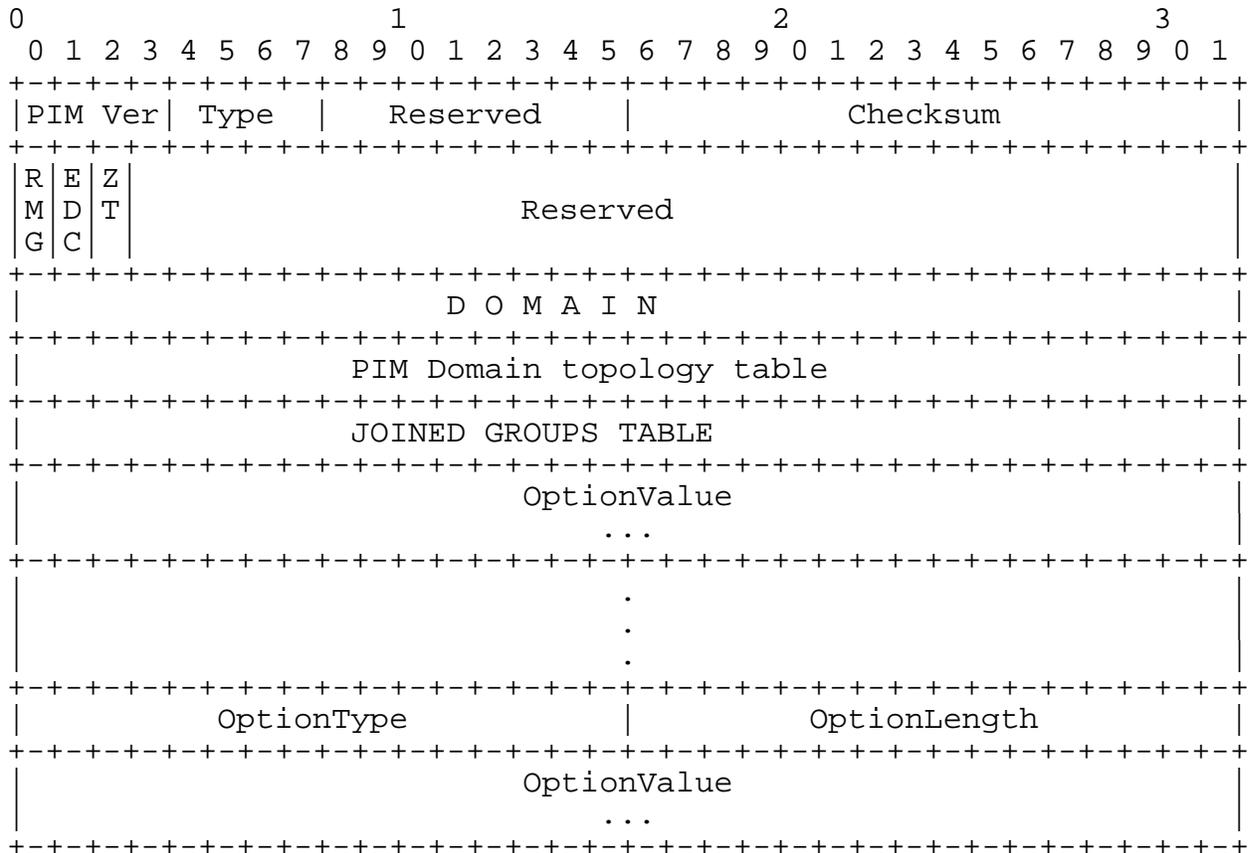


Figure 15 original PIM-SM hello packet compared with the new hello packet format

Type: HELLO

New fields are added to the new packet which in later sections their usage will be covered:

- 1- RM-BIT: can be set to 1 by the sending neighbor to inform the other neighbors about the existence of the C-MAPPER incase dynamic methods are being used. A value of 0 indicates that the neighbor doesn't know about any RP or C-MAPPER.
- 2- EDG-BIT: if set indicates to all downstream CLIENTS that a PIM EDGE-Client exists in the network to limit the propagation of multicast introductions sent by C-MAPPERS and C-RPs. And to

register a multicast source the clients MUST send the register message directly to the EDGE-CLIENT which is responsible of NAT operations. The EDGE-CLIENT can be considered as the DR of a Multi-Access Network which is part of one PIM-NG Multicast Domain. Downstream routers need to look into PIM DOMAIN TOPOLOGY TABLE to find the address of EDGE-CLIENT, if the EDGE-BIT is set.

- 3- ZTC-BIT: if set indicates that there has been a change in the multicast groups the CLIENT has joined the SPT for. a PIM-NG-CLIENT will set this bit in case there has been a change in its JOINED GROUP TABLE and will send the table in the hello message.
- 4- Domain field: a 32 bit field used to announce the Domain a PIM-NG-ROUTER is resided in to the neighbors. Only PIM-NG-ROUTERS with matching Domains can become neighbor.
- 5- PIM Domain topology table: this table as described before in the definition section, holds the information related to the C-RP'(s), C-MAPPER'(s) and other components. What is included in this table is actually the unicast addresses of each of the components mentioned.
- 6- JOINED GROUPS TABLE: holds the list of all the multicast groups a PIM-NG-CLIENT has joined the SPT for and is receiving the related traffic, in the format of (S, G) or (S, S, G) which the first S indicates the unicast address of the PIM-NG-CLIENT which is sending the traffic to the domain and the second S indicates the address of the host or server inside the LAN behind the client. this table will help clients in the process of sending join messages to the source of a multicast group in case their neighbors are already receiving that traffic :
 - o the client in need of receiving multicast traffic for (*,G)or(S,G) first looks inside this table to see if the neighbor PIM-NG-CLIENTS has already joined the group and receiving the traffic ,before asking the C-RP for the unicast address of the source.
 - o If there is an entry in this table for the multicast group it needs it will send a join/prune message to the neighbor for that group.

- o If there is no entry for the multicast group it needs in this table, it will then ask the C-RP.
- o A CLINET will send the table only incase that there has been a change in its internal JOINED GROUP TABLE , like joining the SPT for a new group or leaving/pruning from a group .in such cases client must inform the neighbors by setting the ZTC-BIT in its hello message.

```

+-----+
|Source unicast addr| Role | priority |Domain|TR Group|
+-----+
|                   |     |           |      |         |
+-----+
|                   |     |           |      |         |
+-----+

```

Figure 16 PIM domain topology table (PDTT)

- o Source unicast address field: holds the unicast address, related with either one of C-MAPPER, C-RP, SC-MAPPER, BPR or EDGE-CLIENT.
- o Role: indicates that the address entry inside the source unicast address field belongs to what component of the PIM-NG domain:
 1. C-MAPPER(C-M) and SC-MAPPER(SC-M)with binary codes 1 and 2
 2. C-RP and SC-RP with binary codes 3 and 4
 3. A ROLE of EDGE with binary code 9, can be seen in designs, that to reduce the propagation of multicast introduction messages, a Client is chosen to act as the EDGE-CLIENT and will not pass any multicast introductions received on external interfaces to internal interfaces. In such a design the EDGE-CLIENT IS responsible of introducing only the existing TR'(S) to its downstream clients by sending the PIM-DOMAIN TOPOLOGY TABLE to downstream neighbors. An EDGE-CLIENT MUST BE ONLY used in parts of a PIM-NG multicast network where no C-MAPPER, C-RP or TR exists. And the downstream routers MUST BE only clients. Also in such networks a multicast source can use private IP addresses, and such a source MUST send register messages to the existing EDGE-CLIENT'(s). Related concepts will be explained in the appropriate section'(S).

4. A role of PPER (PRIVATE-PIM-EDGE-ROUTER) with binary code 8 can be seen when a private PIM-NG domain is connected to another domain through the use of PIM-EDGE-ROUTERS which are responsible for the process of NETWORK ADDRESS TRANSLATION (NAT).
 5. A role of BPR with binary code 10 can be seen in designs where a PIM-NG multicast domain is connected to a PIM-SM multicast domain. The information related to BPR is locally significant to C-MAPPER and won't be sent to all PIM-NG routers.
 6. A role of STC-MAPPER (STANDBY-C-MAPPER) with binary code 5 can be seen in network designs with PEER-C-MAPPERS inside the same Domain. In the case of existing Peer C-MAPPERS, C-RPs will use the closest C-MAPPER to propagate the information associated with existing sources inside the domain.
 7. A role of TR with binary code 6 can be seen in case any TR implementations are considered inside the domain. Please do note that a C-MAPPER or C-RP can take the role of TR too or the TR can be a separate and unique PIM-NG-AWARE-ROUTER inside the domain.
 8. A role of C-TR (core-TR) with binary code 7 can be seen in PIM-NG-DOMAINS. It indicates the existence of a core-domain and the unicast address of any existing TR inside the core domain for further use by clients inside each domain.
- o Priority: this field holds the priority related to an EDGE-CLIENT and is used only in hello packets sent to private networks by an edge-client. Related concepts will be discussed later.
 - o DOMAIN: in case separate PIM-NG multicast domains are going to be connected to each other through the use of an EDGE router, the Domain to which a PIM-EDGE-ROUTER is connected will be informed inside this field. This field is usable by C-MAPPERS inside each domain and thus won't be sent to ALL-PIM-NG-CLIENTS by C-MAPPER when sending introductions to 239.0.1.190.
 - o TR Group: this field indicates the TR group each TR belongs too and is set to 0 by default. It is ONLY set by C-MAPPER or the Active C-MAPPER and is used in Bidirectional PIM-NG processes.

When a PIM-NG-CLIENT receives a join message from its neighbor for a multicast group, that it is joined to and is receiving the traffic for, it will then forward the traffic on the interface that the neighbor is connected to and towards that neighbor.

The definitions given by the original PIM-SM [7] are applicable to the other fields.

4.4. RP discovery

RP discovery in general, points to the processes involved in, how a Client finds the C-RP in the first place to send a request to and reach the destination it wants.

In PIM-NG RP discovery is done in 3 different conditions or better to say network topologies. Choosing the best solution depends on the specifications of the network and will be decided by network administrators.

Bellow the 3 different RP discovery solutions are briefly defined, and will be explained later:

1- Static RP discovery :

The unicast address of C-RP is statically given to each PIM-NG router through special commands.

2- Dynamic RP discovery type 1 :

Type 1 is used in networks not so big to need redundant RPs. and also used in order to make the process of RP discovery easier and more scalable. This type uses its own set of commands for the sake of simplicity in explaining the processes .in this type of RP-DISCOVERY a router is given the role of both C-MAPPER and C-RP

3- Dynamic RP discovery type 2 :

Type 2 is used in large networks with many routers, which will need to have C-RP redundancy. Also this type is suitable for the future needs of World Wide Web (internet). This type uses its own set of commands for the sake of simplicity in explaining the processes. In this type C-MAPPER and C-RP'(s) are different routers.

In the following sections the above three RP discovery methods will be discussed.

Also please note that any command line written in this document from this point forward is just to provide a good sense of understanding, and to provide simplicity in explaining the processes involved and shall not be taken as the solid command line to be used. Also to make the explanation process and the understanding easier some of the processes and specifications are explained through scenarios and examples.

4.4.1. Static RP discovery

As the name indicates, the unicast address of the C-RP which is the IP address of a loop back interface created on RP, is set on each PIM router so the PIM routers will be able to communicate with RP.

The process is as follows:

- 1- A loopback interface, typically loopback0 is created on RP and an IP address is assigned to it.
- 2- The candidate RP router will know that it is the RP and will have to use its loopback 0 by initiating the command :

```
<# IP PIM-NG SET-RP SOURC-LO "x" INTERFACE X, INTERFACE Y, INTERFACE Z.>
```

by initiating this command on the router, the router understands that it is the RP and it will have to use its loopback 0 unicast address to communicate, or better to say it will use its loopback 0 in the " source unicast address field " and also if it sees its loopback 0 unicast IP address as the destination of a PIM-NG packet it will have to answer to that packet. And also it will have to bring its interface "X, Y, Z" in to the PIM game.

- 3- On all the other PIM routers the bellow command is initiated :

```
<# IP PIM-NG CLIENT RP-ADDRESS "X.Y.Z.W" INTERFACE "X", INTERFACE "Y">
```

This command tells the router that it is a client of the RP with address of "X.Y.Z.W" .and to find any source or to be able to send multicast traffic to any host looking for that traffic, it will have to contact the RP through the processes explained before. And also it will have to bring its interface "X, Y" in to the PIM game.

Since the C-MAPPER plays a significant role when it comes to connecting different PIM-NG multicast domains and for exchanging the information related to existing multicast sources in each domain the C-RP takes the role of a C-MAPPER too in the static mode. But no C-MAPPER introduction messages will be generated by the C-RP within the multicast domain and the additional role of C-MAPPER is only used when exchanging information related to existing multicast sources with other domains.

A PIM-NG specification advises that a mechanism be implemented through which any existing components such as PPER'(s) and BPR'(s) MUST learn about the C-MAPPER or better to say the existing C-RP through static configuration of the C-RP unicast address.

In case any TR exists within the multicast domain or the domain is connected to a PIM-NG core domain, PIM-NG strongly advises that the unicast address of existing TR'(S) be introduced to each Client through static configuration since there will be no C-MAPPER introduction in the static RP discovery mode.

At the end PIM-NG specifications strongly advises to implement the static RP discovery method in design plans where the multicast domain is of small size and the domain is not connected to other domains. And if the multicast domain whether big or small is supposed to be connected to another domain it is strongly advised by PIM-NG specifications to use one of the Dynamic RP discovery methods described in the following sections.

Although in the final sections of PIM-NG specifications a new concept will be explained which is related to connection of a Multi-Access network which can also be considered a small sized multicast domain to another domain or a service provider multicast domain without the need for existing of any C-MAPPER, C-RP or TR in that Multi-Access network which is unique to PIM-NG.

4.4.2. Dynamic RP discovery

As stated before PIM-NG uses 2 types of dynamic RP discovery methods depending on the needs and the size of the network:

- 1- Dynamic discovery type 1
- 2- Dynamic discovery type 2

4.4.2.1. Multicast IP addresses used

Before explaining the processes involved in PIM-NG DYNAMIC-RP-DISCOVERY-TYPE1, there are some multicast destination group addresses that must be defined which are used in both DYNAMIC-RP-DISCOVERY methods:

1- Multicast group address 239.0.1.190 :

This address is reserved to be used by C-MAPPER at the time of introducing itself to other PIM-NG routers. PIM-NG routers/clients will listen to this multicast group address to find the C-MAPPER and the existing C-RPS. So a new PIM-NG C-MAPPER will send its introduction to this destination address.

2- Multicast group address 239.0.1.189: this address is required to be used as the reserved range for the communication of RPs in order to hold an election between RPs incase backup RPs are needed. Also this address is used to find peers automatically in case redundant C-RP is needed in the network.

3- Multicast group address 239.0.1.188 : this address is required to be used as the reserved range for the communication of MAPPERS in order to hold an election between MAPPERS incase backup MAPPERS are needed. The usage of this range will be covered later.

Now that the addresses are defined, it is time to explain the process through which the PIM-NG dynamic RP discovery is done and implemented.

4.4.2.2. Dynamic RP discovery TYPE-1

There might be some network designs in which, the presence of only one dedicated router as the RP is adequate but for administration purposes, administrators prefer to use an automatic mechanism, so that every PIM router will know the RP. An example of such networks can be an enterprise or a company which uses multicast applications such as voice and video conference often and not always. In other words the network doesn't need to have multiple redundant RPs and also they don't use multicast applications all the time, but for the sake of easy administration and maintenance administrators prefer to use the automatic mechanism.

In such a network a router will be used as the designated RP or candidate RP (C-RP), and will introduce itself to other PIM

routers/clients as both C-MAPPER and C-RP .and as soon as all the clients understand about the presence of the C-MAPPER and C-RP the rest of the process is as before.

The process is as follows:

1- A router is chosen as the C-RP. Then like the specifications of the static RP a loopback interface is configured on it. The loopback is better to be the loopback 0 interface.

2- The commands :

```
<#IP PIM-NG DYNAMIC-RP1 SOURCE-LO "X" INTERFACE [TYPE] "X" ,  
INTERFACE [TYPE] "Y"> and
```

```
<#IP PIM-NG DOMAIN [X]> are initiated on the RP.
```

Above commands tells the router that it is both C-MAPPER and C-RP in the network and the Dynamic discovery method used is type1 and it should use its interface loopback "X" as the C-MAPPER and C-RP unicast address in introductions. And it should bring its interface "x, y,.." into the PIM-NG game.

So router sees that it is the C-RP and the discovery protocol used is Dynamic and the discovery method is type1, and knows that:

- o It is the only RP in the network and also it must introduce itself as the C-MAPPER.
- o It is resided in domain X
- o It should Send multicast introduction messages to multicast address 239.0.1.190, out its interface "x, y" and any other interfaces configured to be in the PIM-NG game.
- o It should put the interfaces defined in the command in to forwarding for the multicast address destinations 239.0.1.188 and 239.0.1.189.
- o It should Send periodic C-MAPPER-introduction messages every 60 seconds.
- o If by any means it is reloaded as soon as it comes back up, it should send a multicast introduction to 239.0.1.190 ASAP.

3- All the other none-RP PIM-NG routers are configured as the clients of C-MAPPER/RP.

4- on PIM-NG none-RP routers(clients) commands :

```
<#IP PIM-NG DYNAMIC-RP CLIENT INTERFACE [TYPE] "X", INTERFACE  
[TYPE] "Y">
```

```
<#IP PIM-NG DOMAIN [X]> are initiated.
```

above commands tells the router that it is in a PIM-NG network and the protocol by which it can find the RP is Dynamic RP discovery .also it understands that it should bring its interface "x ,y" and any other interfaces dictated, in PIM-NG game.

After entering the above commands in a none-RP or a client router'(s) the router knows that:

- o it is in PIM-NG domain X
- o It should use dynamic methods to find the RP .so it will wait to receive a C-MAPPER introduction message.
- o it should listen to multicast address 239.0.1.190 to hear the C-MAPPER-introduction message to learn about the C-RP/RPs
- o it should bring the interfaces mentioned in the command in PIM-NG game
- o It should put the interfaces mentioned in the command in forward for the address 239.0.1.190 so others will hear the C-MAPPER introduction message. Also it should put those interfaces in to forwarding state for group addresses 239.0.1.189 and 239.0.1.188.
- o If it doesn't receive the address of the C-MAPPER in the first hello-acknowledge received from the neighbor, it should wait to hear from the C-MAPPER.

Now that the overall process is covered, in the next section the way RP and clients communicate will be explained and the new packet formats will be shown.

4.4.2.2.1. RP introduction process

As soon as the router designated as the C-RP and C-MAPPER, is configured and knows that it is the only RP in the network, it will start the process of introducing itself as the C-MAPPER to the network by sending the C-MAPPER-INTRODUCTION message to the multicast destination address 239.0.1.190 (0).

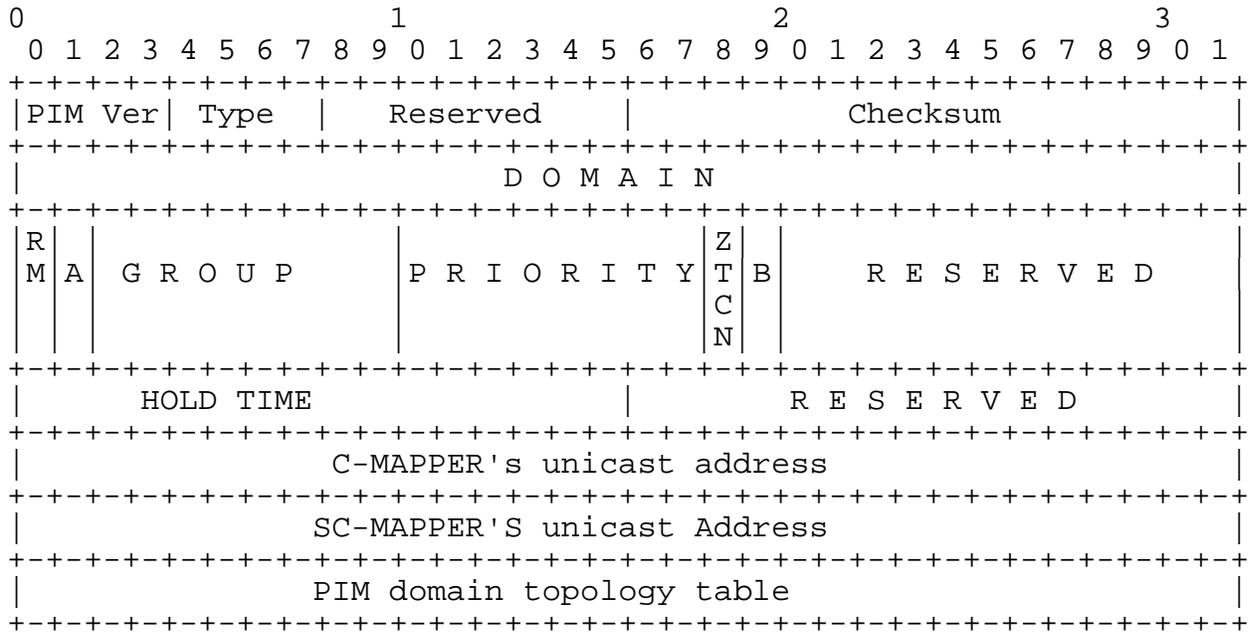


Figure 17 C-MAPPER-INTRODUCTION message sent to 239.0.1.190

- o Type: C-MAPPER-INTRODUCTION
- o RM-BIT field : if set to 1 in a C-MAPPER introduction message sent to 239.0.1.190 , tells to all PIM-NG CLIENTS that only 1 RP exists in the domain .so the address of the C-MAPPER is the unicast address of C-RP . If this bit is not set to 1 then indicates to all PIM-NG CLIENTS that in order to find existing C-RPs they should read the contents of PDTT.
- o A-BIT field: is set to 0 in this type. Its usage will be discussed later.
- o GROUP and PRIORITY fields: 8 bit fields, and are set to all zeros when sending introduction to 239.0.1.190 .and clients won't read the contents of these fields.
- o ZTCN-BIT field : if a C-MAPPER needs to inform ALL-PIM-NG-CLIENTS about a change it will set this bit, also if a C-MAPPER needs to inform a change in the domain to SC-MAPPER or PEER C-MAPPERS ,it will set this bit .use will be discussed later.
- o B-BIT: Bidirectional Bit. If set to 1 in such introduction message indicates that the domain uses the manual mode(4.7.2.1.) and for

any multicast group found in Bidirectional Groups Table sources and receivers MUST immediately join the SPT rooted at closest TR.

- o DOMAIN: indicates the domain in which a C-MAPPER resides. It can be used as either a security factor or as an scoping mechanism so that clients and C-RPs inside one domain wont react to the introductions sent from a C-MAPPER in another domain.
- o PIM Domain topology table (PDTT): when the RM-BIT is set the C-MAPPER won't send this table in its introductions sent to 239.0.1.190, unless a TR is considered in the domain.

The Type field is set to C-MAPPER-INTRODUCTION1 so the clients receiving the packet will know that it is an introduction message sent from the C-MAPPER and will look directly into the "SOURCE UNICAST ADDRESS" field. The address field contains the unicast address of the interface loopback "X" of the C-MAPPER.and the next field contains the address of the SC-MAPPER if any exists.

After sending the first introduction message , C-MAPPER/RP sets a timer of 60 seconds and starts counting down to 0 ,and resends the introduction message so all the PIM-NG routers will know that the C-MAPPER still exists.

4.4.2.2.2. Back up C-RP considerations

If any backup C-RP is needed to be considered in such networks explained above for the sake of high availability a mechanism for the negotiation between the RPs and election of the candidate RP(C-RP) is needed.

The process is as follows:

- 1- The addition of **<GROUP [0-255]>** to the command tells the router its group number and the **PEER RP GROUP[X]** tells the router that there is another RP which it must become peer with but since the group number of the peer is the same as its own value an election is needed to elect the C-RP and SC-RP. The values in front of the GROUP means that all the RPs belong to one unified group. The usage of the GROUP will later be more clarified. and the **<PRIORITY [VALUE BETWEEN 0-255]>** at the end of the command will define each C-RP's priority in the election process :

```
<#IP PIM-NG DYNAMIC-RP1 SOURCE-LO "X" INTERFACE [TYPE] "X" ,  
INTERFACE [TYPE] "Y">
```

<#IP PIM-NG GROUP [X] PRIORITY[VALUE]>

<#IP PIM-NG PEER RP GROUP[X] ADDRESS>

- 2- If the address of the peer is not used, RPs will send multicast introduction messages to the reserved address 239.0.1.189, so the other RP/RPS will find each other.
- 3- An election based on the highest priority configured on each C-RP , or highest C-RP unicast address will be held between the C-RPs
- 4- The candidate RP will take the responsibility of both C-MAPPER and C-RP as explained above
- 5- RPs will send unicast keep-alive messages to each other every 30 second.
- 6- The candidate RP, C-RP, will send a copy of its multicast-mapping table only incase that any changes occur in the domain. And C-RP periodical introduction messages to SC-RP wont contain any MULTICAST MAPPING table.
- 7- In case any changes occur C-RP will set the Z-BIT inside its introduction message to SC-RP which indicates that a change has occurred.
- 8- If more than 2 RPs are considered in a network design, an election between none candidate RPS will be held to choose the second best choice (SC-RP) to be the RP if the candidate RP is dead.
- 9- The candidate C-RP which has the role of C-MAPPER too, will send the unicast address of the second best candidate as SC-MAPPER in its introduction message sent to all PIM-NG routers every 60 second for further use by PIM-NG clients.
- 10- PIM-NG clients will write the address of the C-MAPPER, SC-MAPPER and C-RP in a special table called PIM Domain Topology Table (PDTT) for further use.
- 11- If SC-MAPPER/RP doesn't receive any INTRODUCTION/KEEPALIVE-MESSAGE from the C-MAPPER/RP in 2*INTRODUCTION/KEEPALIVE-MESSAGE plus 5 seconds timer (2*30+5) it will immediately take the place of the C-RP and send a multicast introduction to 239.0.1.190.

12- If the SC-MAPPER/RP receives a REQUEST-FOR-C-MAPPER message from a client asking about the existence of the C-MAPPER in 70 seconds it will react in 2 ways :

- o It will immediately query the C-MAPPER/RP by sending an introduction message and if it receives a KEEP-ALIVE message from the C-MAPPER/RP which means the C-MAPPER/RP is still alive then it assumes that the client which is looking for C-MAPPER/RP is having problems .so it will return the address of the current C-MAPPER/RP back to the host in a unicast C-MAPPER-INTRODUCTION (Figure 17)

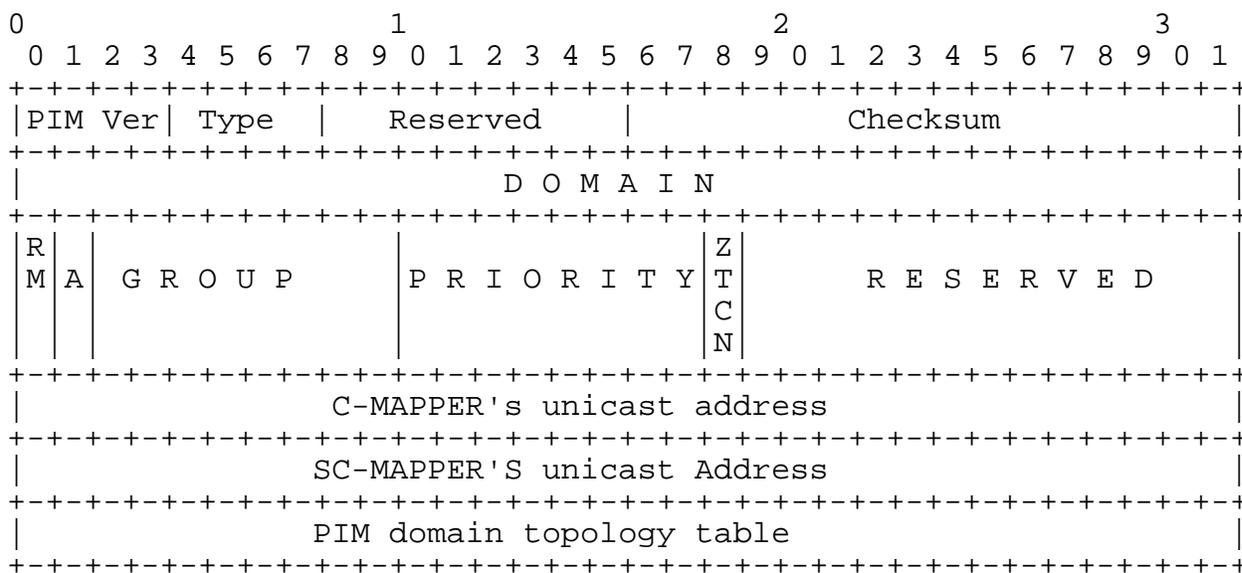
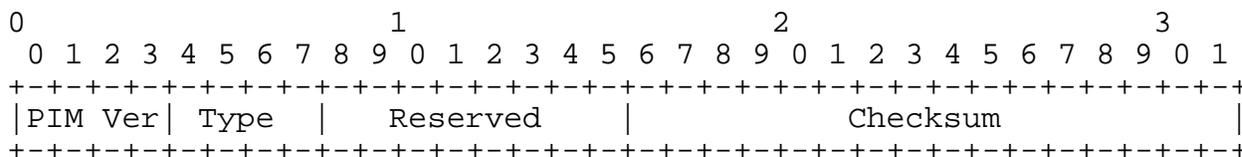


Figure 18 SC-MAPPER/RP answers to a host REQUEST-FOR-C-MAPPER message

- o Type: C-MAPPER acknowledge
- o If SC-MAPPER/RP hasn't received KEEP_ALIVE messages for the past 2*30 seconds , then it should not receive such a message, as by now SC-MAPPER/RP must had taken the place of C-MAPPER/RP and sent out a C-MAPPER introduction to introduce itself. (Figure 18).



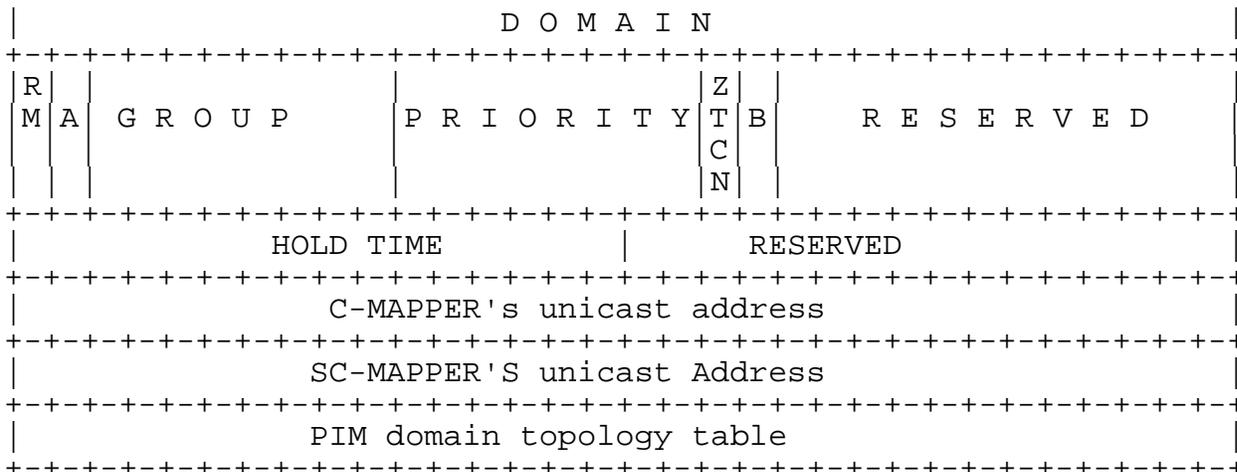


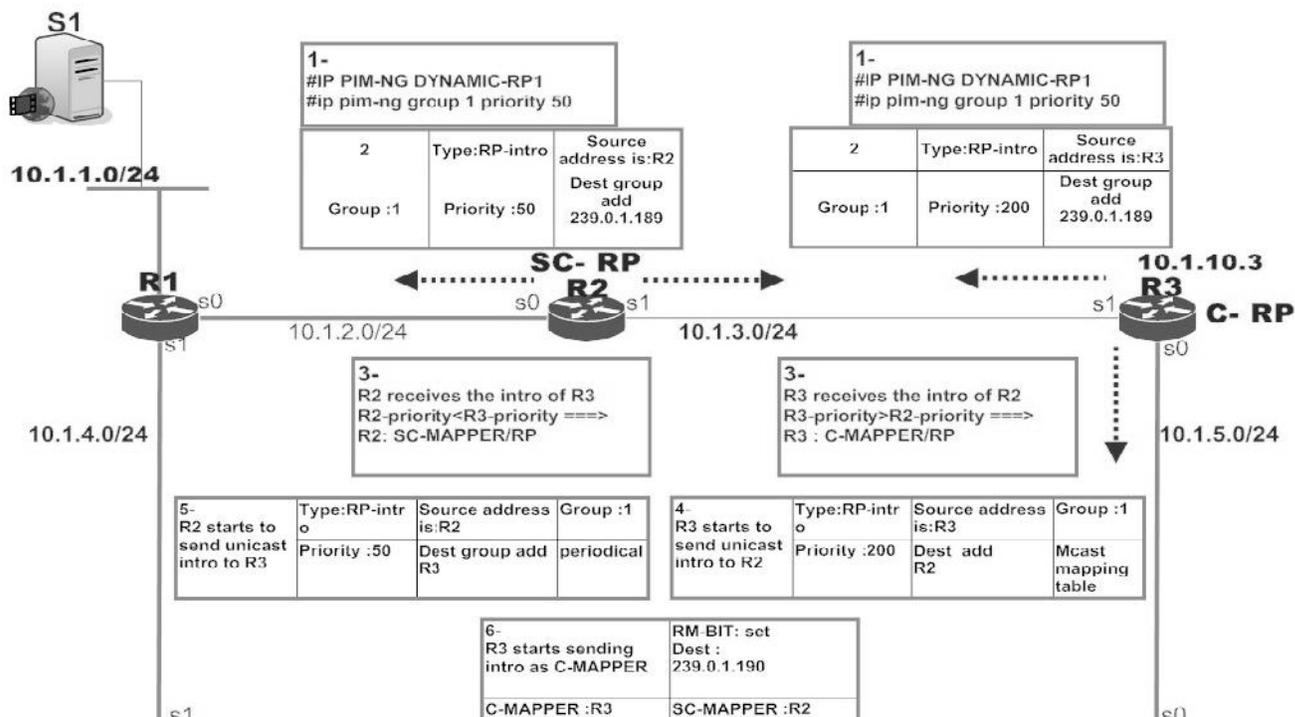
Figure 19 SC-MAPPER/RP multicast introductions to all PIM-NG-CLIENTS

Type: C-MAPPER INTRODUCTION1

If clients see the address of the SC-MAPPER in the C-MAPPER unicast address field they will assume that the current C-MAPPER is dead and will update their tables with the new one.

13- If the previous C-MAPPER/RP gets back again then another election will be held between RPS and the C-RP will introduce itself.

14- All PIM-NG routers will put their interfaces inside the game of PIM-NG into forwarding for the group address 239.0.1.189 and 239.0.1.188 and won't listen to these packets. Only will forward them.



Sami

Figure 20 PIM-NG network with backup RP (SC-RP)

RP introduction packet formats are shown below:

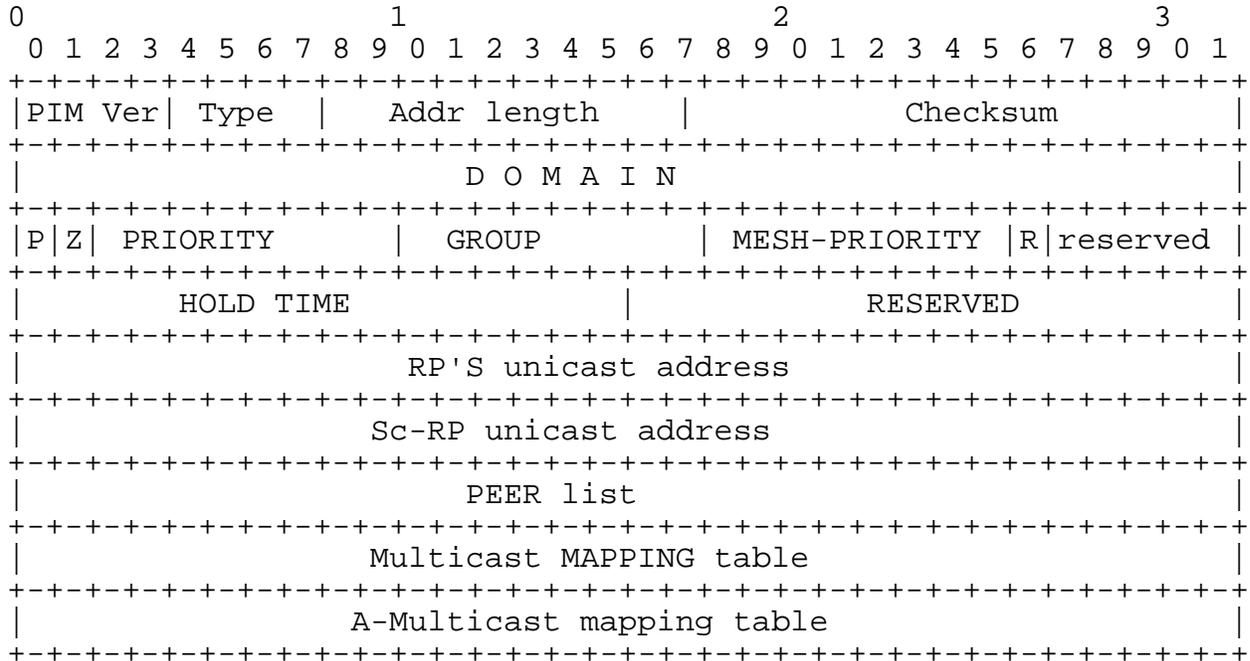


Figure 21 RP introduction message format

- o Type: RP introduction
- o P-BIT field (PEER): is used when a C-RP wants to become peer with another C-RP. C-RP sets this bit to 1 and sends an introduction to 239.0.1.189. By seeing this bit set only C-RPs configured to become peer with another C-RP from a different Group will react to the message contents. Others will only forward.
- o Z-BIT field (ZONE TOPOLOGY CHANGE NOTIFICATION): this bit is set to 1 in case a C-RP needs to inform the SC-MAPPER C-MAPPER or peering C-RP about any changes occurred. when this bit is set ,the C-RP will then sends out one of the tables shown in Figure-34 depending on the fact that to whom this introduction is being sent and in what type of domain it is resided.
- o Group field: 8 bit field containing the group that the C-RP is a member of. Helps the C-RPs to see if the packet is sent from a C-RP they are configured to become peer with or in cases that backup C-RP is needed. This field is set to all ZEROs when sending to a C-MAPPER.
- o Priority field: is only used when finding and communicating with SC-RP. Otherwise set to all ZEROs.
- o Mesh-priority: will be used in the process of electing the active C-RP in a MESH-Group.
- o R-BIT: if set to 1 by a C-RP, indicates to the C-MAPPER that the RP has the role of TR too.
- o SC-RP unicast address: will be sent only when introducing to PEER-C-RPs or C-MAPPER.and not to a back up RP.
- o DOMAIN: indicates the domain the C-RP is a member of. helps the C-RP to differ between the packets received from C-RPs inside other domain in an special design of multiple PIM-NG multicast domains.
- o Multicast MAPPING table (MMT): it is sent only when introducing to SC-RP.and is sent in case any changes occur.
- o A-Multicast MAPPING table (AMMT): this table is sent by a C-RP to C-MAPPER in case it realizes that the domain it is resided in is connected to another domain, or there are multiple C-MAPPERS resided in the domain. When the C-RP receives a C-MAPPER introduction with the A-BIT set, it starts to send this table to the C-MAPPER. Also this table is sent by C-RP to PEER-C-RPs if any PEER-C-RP exists.

- o Hold time: is used only when introducing to C-MAPPER and indicates the amount of time C-MAPPER will expect to hear C-RP keep-alive message.

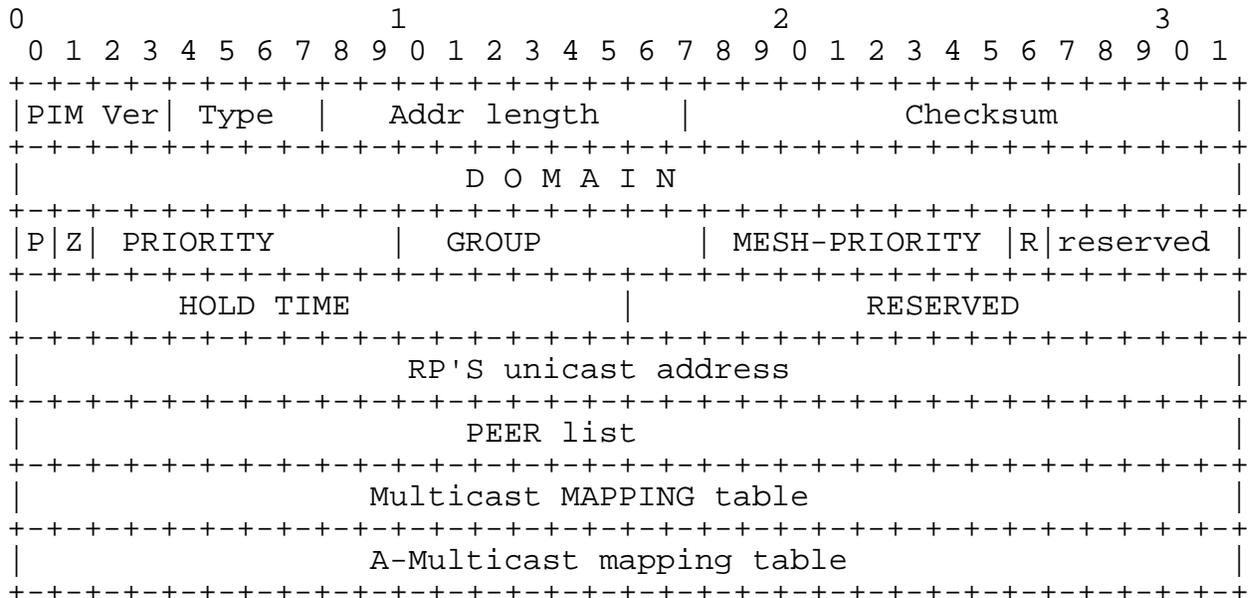


Figure 22 C-MAPPER/RP unicast HELLO/KEEP-ALIVE message to back up RPs

- o Type: RP introduction

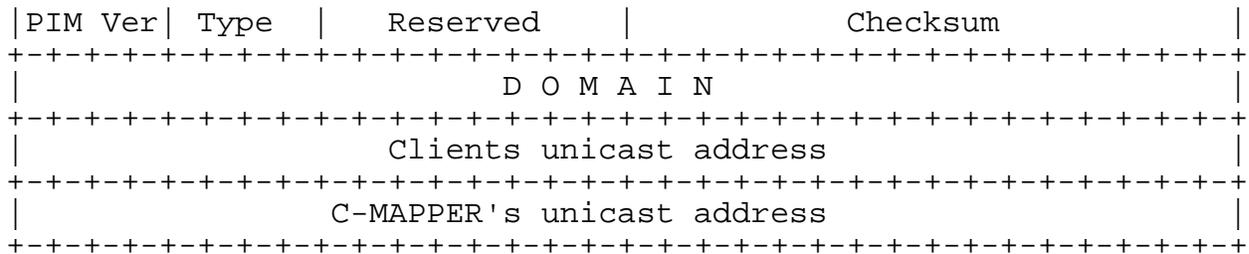


Figure 23 request- for-C-MAPPER packet sent to the SC-MAPPER

- o Type: REQUEST-FOR-C-MAPPER
- o The Client (i.e. R4) puts its own address in the client unicast address field and the unicast address of the SC-MAPPER/RP in the C-MAPPER unicast address field.
- o If no SC-MAPPER/RP exists then it will query the neighboring PIM-NG routers by sending a HELLO packet with the RM-BIT set to 0, indicating that it doesn't know any C-MAPPER.
- o If it doesn't receive the C-MAPPER's unicast address from the neighbor in the first hello inside the Pim-domain-topology-table or if the C-MAPPER address received from the neighbor PIM-NG router is the same as its own current entry for the C-MAPPER then it should wait to hear from the C-MAPPER.

4.4.2.2.3.1. New PIM-NG router/client

In this section, actions taken by a new PIM-NG router added to the network using DYNAMIC-DISCOVERY without any knowledge about the C-RP is explained:

- 1- Puts its interfaces defined by the related commands in to forwarding for the multicast address groups 239.0.1.189/40/188
- 2- Sends HELLO message to all PIM routers address of 224.0.0.13 out of the interfaces in the game of PIM-NG, to introduce itself to the neighbor PIM-NG routers.
- 3- If in the first received HELLO message from the neighbor it sees the unicast address of the C-MAPPER it will use it and updates its PMTT.

- 4- If no C-MAPPER address is seen in the HELLO packet received from the upstream neighbor, it will wait to hear an introduction from the C-MAPPER or receive it inside the next hello messages sent from the neighbor.
- 5- If in the Hello message received the EDGE-BIT is set, it assumes that it is inside a private network, and it may have to act differently in order to register a source with the C-RP. The related concepts will be discussed later in a separate section.

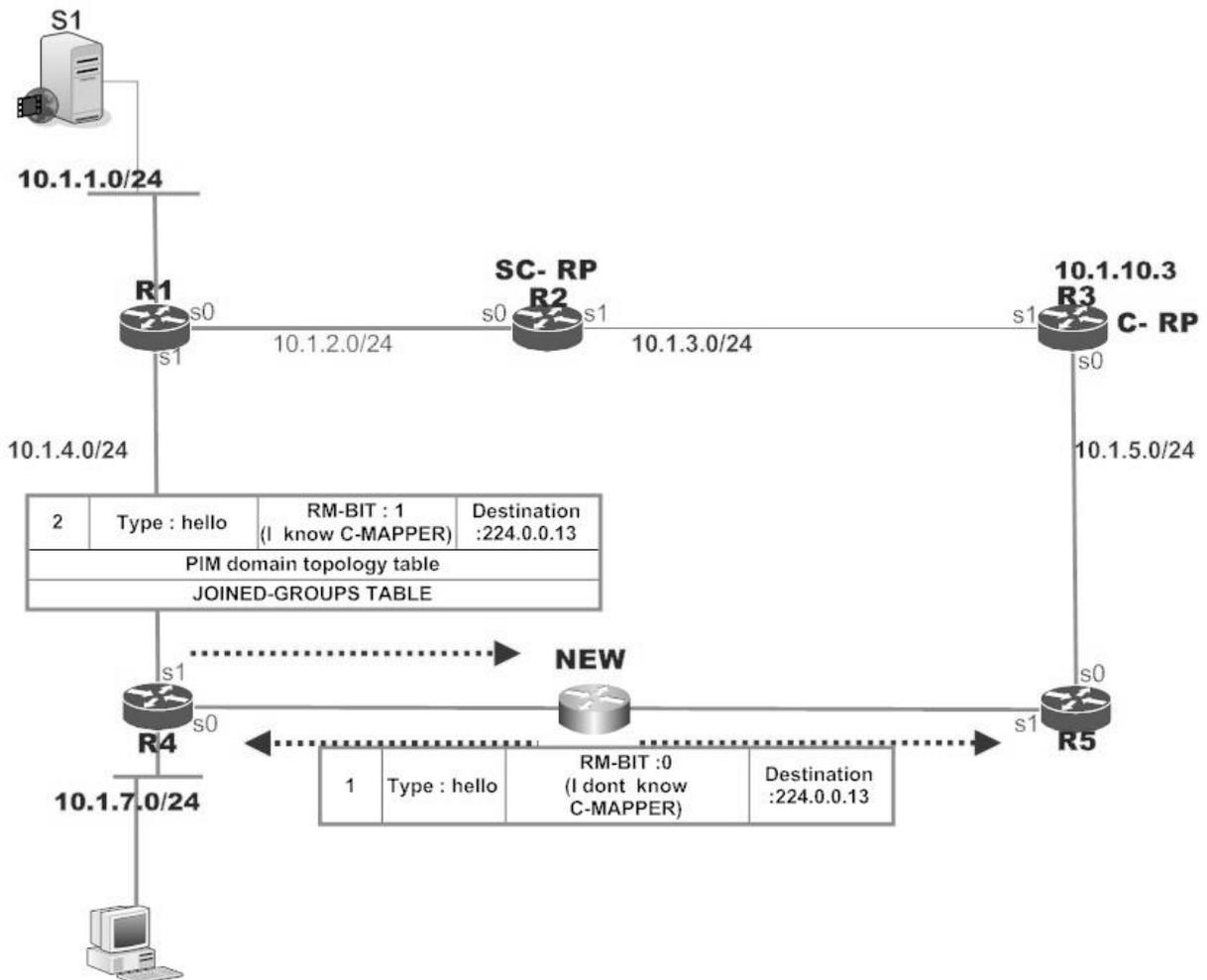


Figure 24 New client added

4.4.2.3. Dynamic RP discovery type 2

In previous sections the processes related to the dynamic RP discovery type 1 in which there are no needs for redundant RPs and networks seems to be of small to medium size has been described. Now let us consider larger networks. In larger networks (i.e. internet or enterprises) with many multicast actions and sessions in process, the high availability of the RPs and also load balance between RPs through using redundant RPs becomes an important factor.

In fact hosts in different parts of a large network must be able to find the desired source as fast as possible, if the source exists. So as described by the original PIM-SM specifications, the network will need redundant RPs alongside a new feature called MAPPING-AGENT or simply PIM-NG-C-MAPPER.

C-MAPPER is in charge of introducing C-RPs to all PIM-NG routers. And the process is in parts like the original PIM-SM RFC 4601[7], and in parts different. Bellow the steps involved in the process of PIM-NG is briefly described and later will be explained in details:

- 1- C-MAPPER introduces itself to all PIM-NG routers, so that everyone knows about its address.
- 2- C-RP'(s) introduce themselves to the C-MAPPER, by sending unicast RP introductions to the C-MAPPER's address. So they wait to receive a C-MAPPER introduction.
- 3- If any TR exists in the multicast domain, it MUST introduce itself to the closest C-MAPPER.
- 4- C-MAPPER will send discovery messages or introduction messages to PIM-NG routers to inform them about the existence of the RPs so sources can use the RP address to introduce themselves to the nearest RP.

- 5- Clients or simply put all none PIM-NG C-MAPPER, RP and TR routers will use the C-MAPPER to find RP'(s) or TR'(s).
 - 6- Clients or simply put ALL none PIM-NG C-MAPPER, RP and TR routers will either only-forward specific multicast address groups or listen to them .
- 0 will be used as the basic topology to explain the process related to an enterprise network with one multicast domain.

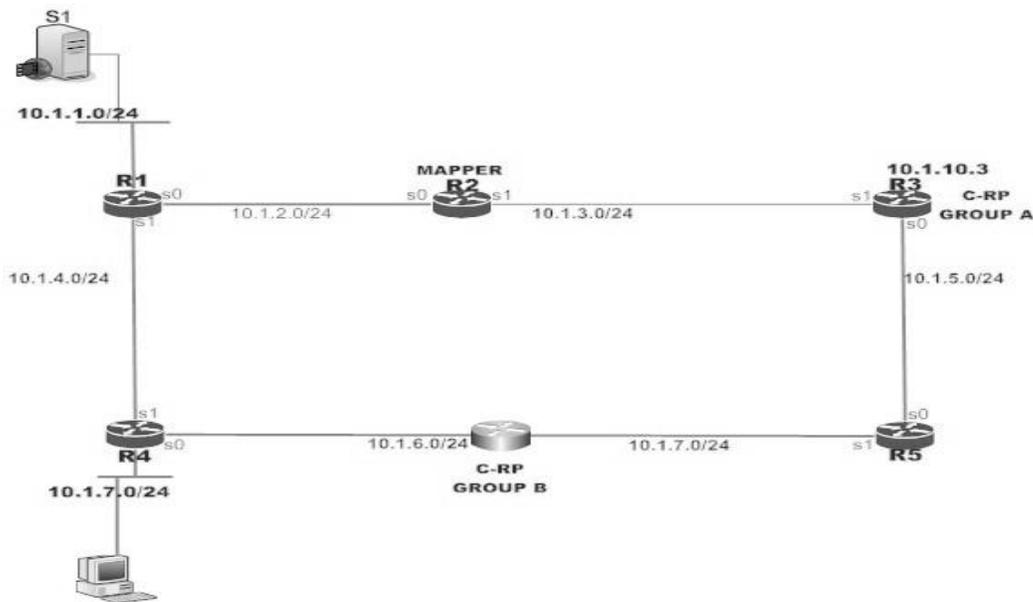


Figure 25 network design with one multicast domain

In the following sections processes related to different components of a PIM-NG domain will be explained.

4.4.2.3.1. Client related concepts

The processes related to a client are as described in section 4.4.2.2.3. .

Clients will choose the closest C-RP or C-MAPPER according to their unicast routing information. And if there is a tie, they will choose based on the highest C-RP or C-MAPPER IP address that can be found inside the PMTT information received with the C-MAPPER introduction message.

4.4.2.3.2. C-MAPPER concepts

C-MAPPER is in charge of discovering C-RP'(s) and introducing them to the population of PIM-NG routers through sending INTRODUCTION-MESSAGES to the destination address 239.0.1.190. Sending these periodic introductions causes any existing C-RP'(s) or TR'(s) and PIM-NG clients to learn the unicast address of the C-MAPPER and SC-MAPPER in case any BACKUP MAPPER is considered in the network.

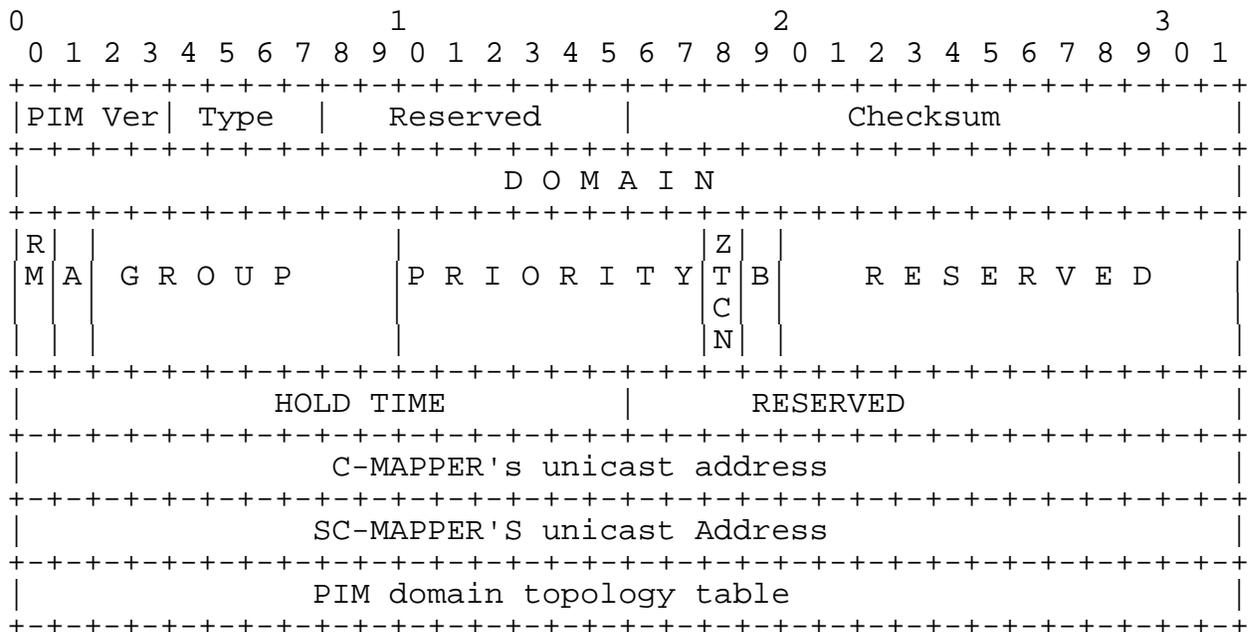
In the introduction message C-MAPPER will send its own unicast address alongside the unicast address of the C-RP/RPs inside the PMTT.

The introduction messages are sent under these circumstances:

1. Periodically as keep-alive messages so all the population knows about the existence of the C-MAPPPER, and new routers can learn about the C-MAPPER and the C-RP/RPs. sent to 239.0.1.190
2. Whenever it receives an introduction from a new C-RP. Sent to 239.0.1.190
3. Whenever it receives an introduction from a C-RP indicating a change in the group addresses the C-RP represents. Sent to 239.0.1.190

4. Whenever it receives a request from a client, The C-MAPPER will send a unicast introduction or acknowledgment in this case to the address of the client.
5. If other MAPPERS exist in the network as backup C-MAPPERS the MAPPERS will send introduction messages to the address group 239.0.1.188 to find each other and hold elections to elect the C-MAPPER and SC-MAPPER (backup MAPPER) or to become PEERS.

Message format sent to 239.0.1.190(ALL-PIM-NG-CLIENTS)



- o Type: MAPPER introduction1
- o Group: is not used and will be set to all 0's.
- o Priority: is not used and will be set to all 0's.

Message format sent to 239.0.1.188(ALL-PIM-NG-MAPPERS) and PEER-MAPPERS

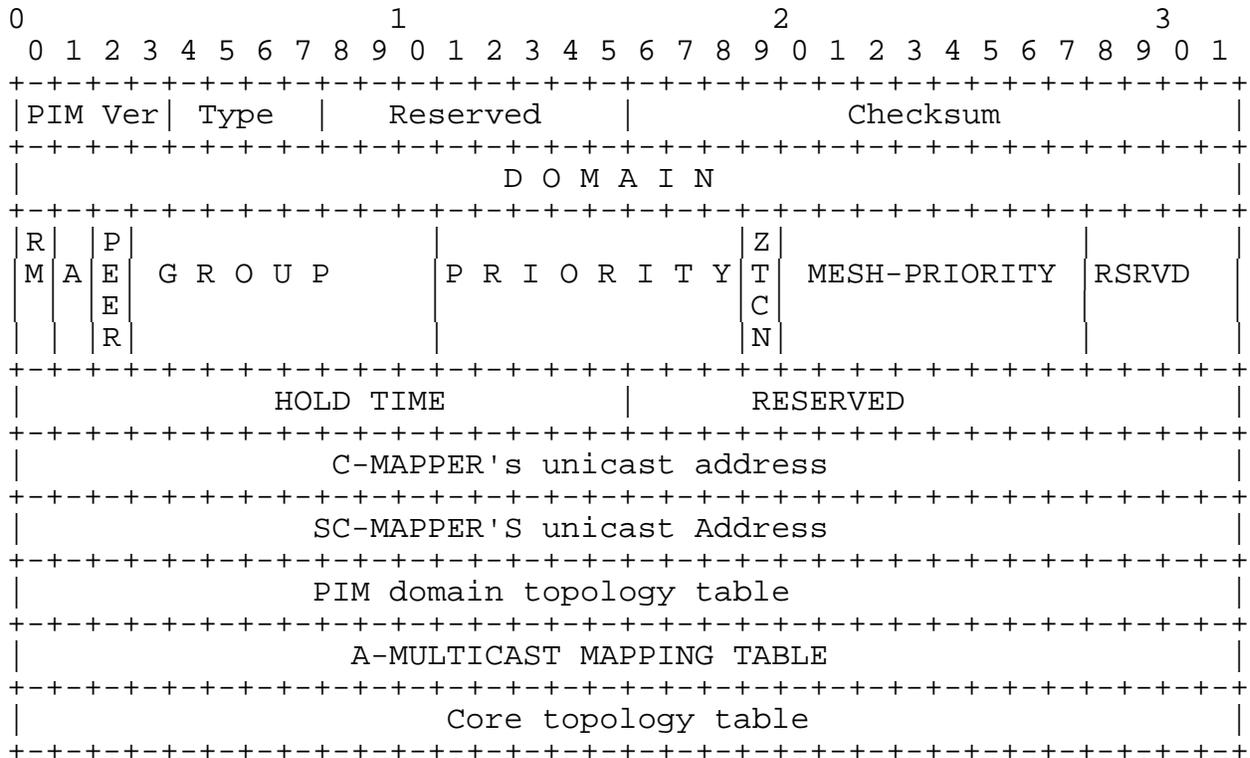


Figure 26 C-MAPPER introduction message formats

- o PEER-BIT field: if set indicates that the introduction message is meant for a PEER C-MAPPER and not for SC-MAPPERS and will be discussed later.
- o A-BIT: won't be set in this message type.
- o Hold time: will be set to all 0's, when sending the introduction to 239.0.1.188. when sending to found peers this field holds the hold time timer that must be used for the peer.

C-MAPPER will set the type field to MAPPER introduction 1 or 2 depending on the destination group it is sending the message to. And will put its own unicast address in the C-MAPPER unicast address field. And also will use the SC-MAPPER unicast address field to introduce the SC-MAPPER if any exists.

After introducing itself it will introduce any existing C-RP'(s) or TR'(s) in the domain, by sending PMTT.

As this periodical introductions sent by C-MAPPER are used by all PIM-NG routers to confirm the existence of the C-MAPPER and also will be considered by all PIM-NG none-RP routers (clients) as C-RP'(s) or any existing TR KEEP-ALIVE message , the C-MAPPER sends periodical unicast introductions every 30 seconds to the C-RP/RPs and will expect to hear from C-RP/RPs every 30 seconds by receiving a unicast RP-introduction.

C-MAPPER will set the A-BIT (4.4.2.3.2.1.) as soon as it becomes PEER with another C-MAPPER with either different domain value or the same domain value and different group.

4.4.2.3.2.1. Value of the A-BIT

This bit is set by the C-MAPPER as soon as it becomes PEER with another C-MAPPER. It indicates to the C-RPs that:

- o They are in a network with more than 1 C-MAPPER which became peers or in a network with different and separate Multicast domains, with each domain connected to the other one through peer C-MAPPERS.
- o C-RP/RPs will have to send A-MULTICAST-MAPPING TABLE to the C-MAPPER.

4.4.2.3.2.2. C-MAPPER preparation

The process begins by choosing a router to act as the C-MAPPER in the PIM-NG network. For the sake of simplicity of understanding and explaining the processes , first the processes related to a single C-MAPPER is explained and later processes related to redundant C-MAPPERS will be explained(4.5.).

In a single multicast domain (0) with one C-MAPPER, the process starts by choosing one router to act as C-MAPPER .when it is chosen the rest is as follows:

- o commands :

```
<#IP PIM-NG DYNAMIC-RP2>
```

```
<#IP PIM-NG DOMAIN [X]>
```

```
<#IP PIM-NG MAPPER>
```

```
<#IP PIM-NG SOURCE-LO"X">
```

```
<#IP PIM-NG INTERFACE "X", INTERFACE"Y">
```

Are initiated on the router. This command tells the router that:

- 1- It is in PIM-NG DOMAIN X that uses DYNAMIC-RP-DISCOVERY TYPE2
- 2- It is the C-MAPPER in the network
- 3- It should bring its interfaces "x, y" and any other interfaces configured as a PIM-NG interface in to the PIM-NG game.
- 4- As no other commands are entered, it is the only MAPPER and it is in a single multicast domain. So it MUST NOT set the A-BIT when sending introduction messages out of its interfaces in the game of PIM-NG.
- 5- It should send its introduction to multicast destination address of 239.0.1.190 as the C-MAPPER, so all PIM-NG routers and especially C-RP/RPs will learn its address.
- 6- As a PIM-NG router it should maintain connectivity with its neighboring PIM-NG routers through sending HELLO messages out of the interfaces defined by the command.
- 7- As a PIM-NG C-MAPPER it should send introduction/keep-alive messages every 60 seconds to 239.0.1.190, so every router in the PIM-NG network knows about its existence and the new ones will learn its address and the C-RP/RPs address. This periodic introduction/keep-alive messages will also act as the introduction/keep-alive on behalf of the C-RP/RPs.
- 8- As soon as it receives a C-RP introduction, the C-MAPPER will check the DOMAIN value to see if it matches its own value and if it does the C-MAPPER will put an entry in its PMTT alongside the multicast groups and the associated

unicast addresses it may receives from each C-RP inside the AMMT.

- 9- The C-MAPPER will maintain connectivity with each C-RP by sending periodic unicast introduction messages to the address of each C-RP. And will expect to receive an introduction from each C-RP in return.
- 10- If it doesn't receive an introduction from each C-RP for 60+10sec it will delete the entry for the RP and announces the loss in its next introduction.
- 11- In case any SC-RP exists , if C-MAPPER doesn't receive a keep alive from the C-RP for (60sec+10=70sec) , it will automatically send an intro to the address of Sc-MAPPER.and after receiving the first intro from Sc-MAPPER it will inform all PIM-NG clients in an introduction message.

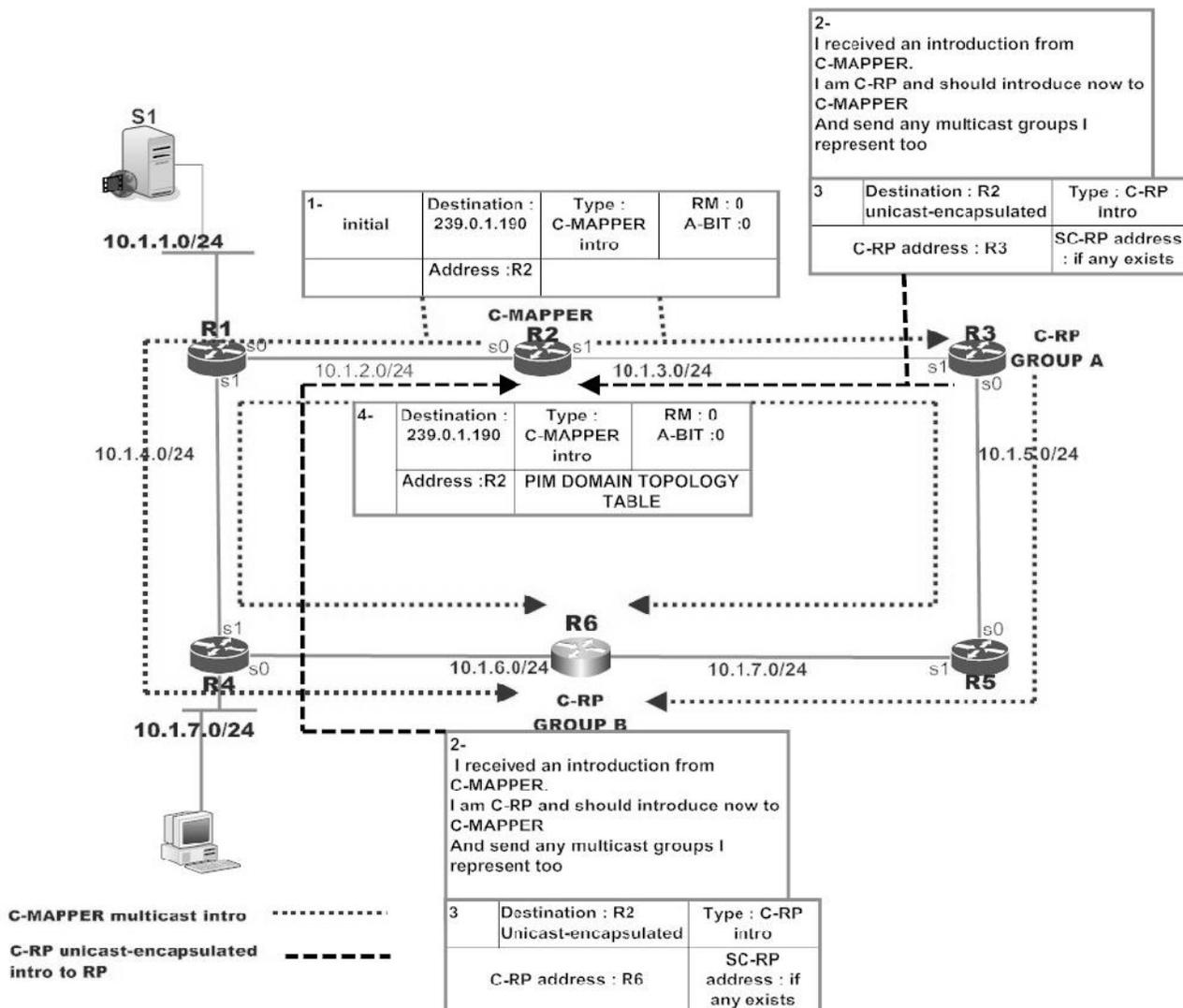


Figure 27 multicast domain with one C-MAPPER

4.4.2.3.3. C-RP concepts

As explained before, the role of a C-RP in general is like an information registry station. Any client that needs to register a source will communicate with the C-RP and any client in need to find a source for a multicast group (G) will have to ask the C-RP. These parts had been covered in the first sections and we are not going to talk about the related processes again.

What had been explained earlier about the C-RP, was mainly related to its processes in a domain with a small or medium size that the presence of only one C-RP and if needed, a backup C-RP (SC-RP) could support the needs of the network. But in larger networks beside the high availability factor, redundancy becomes vital.

In this case there may be the need for presence of more than 1 C-RP all working together to bring both high availability and redundancy to the network.

4.4.2.3.3.1. C-RP redundancy

In a single domain (0) with redundant C-RPs, the main task is for the C-RPs to find each other and also introduce themselves to all PIM-NG clients.

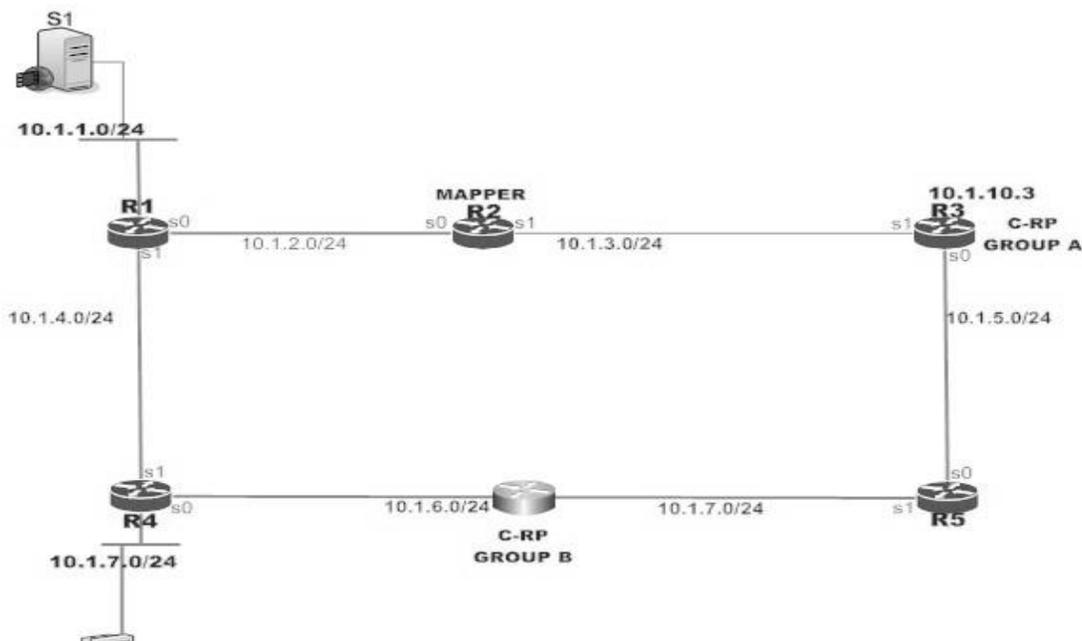
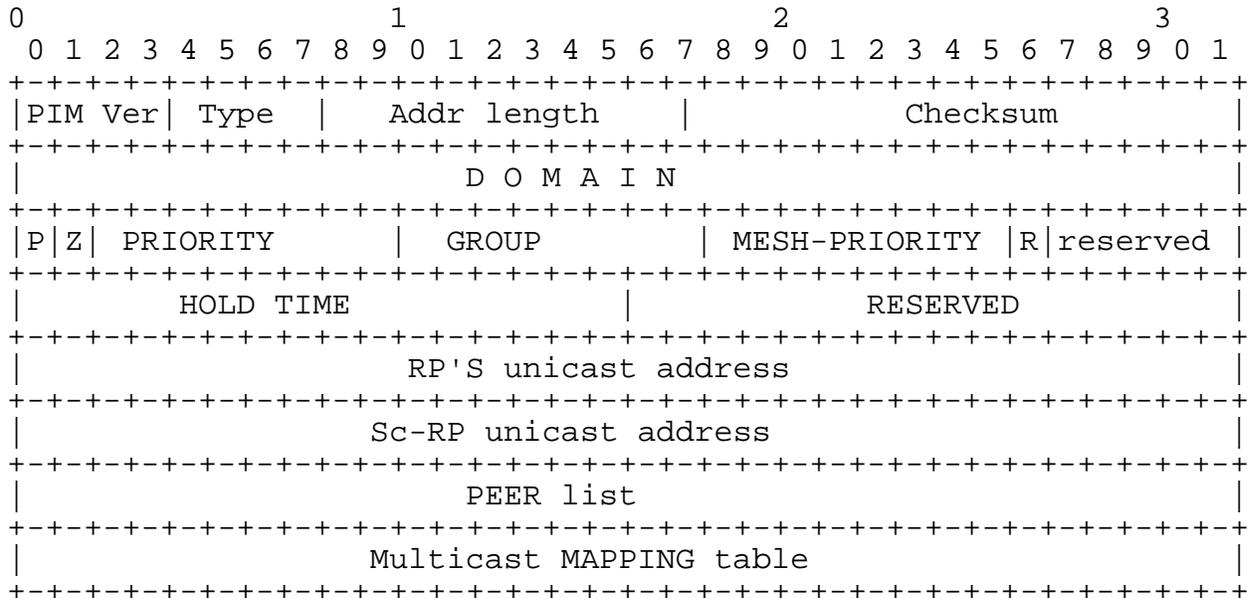


Figure 28 network designs with one multicast domain

After a PIM-NG router is configured to become a C-RP and is told that Dynamic-Discovery type 2 is in use. It needs to introduce itself to all PIM-NG clients. To do so a PIM-NG C-RP needs to wait to learn about the existence of a C-MAPPER in the domain. As explained in the previous sections this is done when a C-RP receives a C-MAPPER introduction sent to 239.0.1.190, containing the unicast address of the C-MAPPER. Then it will introduce itself as the C-RP to the C-MAPPER by sending a unicast C-RP introduction message to the address of C-MAPPER. Format of this introduction and related definitions can be seen bellow.



```

|                                     A-Multicast mapping table                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
Figure 29 C-RP introduction message format

```

- o Type : RP introduction dynamic
- o Priority : set to 0
- o Group : set to 0
- o Hold time: Hold time is the amount of time a receiver must keep the neighbor reachable, in seconds. If the Hold time is set to '0xffff', the receiver of this message never times out the neighbor. This may be used with dial-on-demand links, to avoid keeping the link up with periodic Hello messages.
- o P-BIT : set to 0
- o PEER-LIST : won't be sent
- o Multicast Mapping Table : won't be sent to C-MAPPER and only to Sc-RP
- o A- Multicast Mapping Table : will be sent to C-MAPPER ,SC-RP and peer C-RPs

4.4.2.3.3.2. C-RP processes

The processes involved are as follows:

- 1- A router is chosen to become C-RP and also told that it is in a network that a C-MAPPER exists as a separate component by initiating a set of commands :

```
<#IP PIM-NG DYNAMIC-RP2>
```

```
<#IP PIM-NG DOMAIN [X]>
```

```
<#IP PIM-NG RP>
```

```
<#IP PIM-NG SOURCE-LO"X">
```

```
<#IP PIM-NG INTERFACE "X", INTERFACE"Y">
```

- 2- The above set of commands tells the router that it is a RP in PIM-NG-domain(X). And it should use its LOOP-BACK "X" interface as the RP source address when introducing itself. Also it has to bring the interfaces "X, Y" inside the PIM-NG game and put them into forward for 239.0.1.190, 239.0.1.189 and 239.0.1.188.
- 3- Because the RP discovery method used is of type2, a PIM-NG RP waits until it hears the C-MAPPER introduction.

- 4- After receiving C-MAPPER introduction message first it checks the Domain field inside the introduction message to see if the domain value matches its own domain value it puts the source address of the C-MAPPER into its local PIM domain topology table.
- 5- The C-RP'(s) MUST check the A-BIT, to understand how should it contact with the C-MAPPER.
- 6- RP sends a unicast introduction to the unicast address of C-MAPPER and waits to receive a unicast introduction from C-MAPPER which is interpreted as an acknowledgment from the C-MAPPER.
- 7- After the acknowledgment from the C-MAPPER is received, RP starts to sending periodical unicast introductions as keep alive every 60 seconds to C-MAPPER in response to C-MAPPER periodical unicast introductions.
- 8- If any SC-RP exists, the C-RP will also send the address of SC-RP for further use by C-MAPPER.
- 9- If any changes occur in the domain, like the registration or deletion of a multicast group source (G) the RP will only update its own tables since the A-BIT is set to 0.

4.4.2.3.3.3. Redundant C-RPs

Now that the processes and concepts related to each C-RP are explained it is time to bring redundancy to the network and make C-RPs redundant.

Considering the design shown in 0, we have 2 C-RPs in our network ,one belongs to GROUP-A and the other GROUP-B.and each one of the C-RPs are configured up to this point to act as a C-RP and introduce itself to the C-MAPPER. Also it should be noted that both C-RPs MUST be inside the same DOMAIN.

In order to bring redundancy in to the network, each C-RP must be told to become peer with a C-RP in another group, so they will be able to exchange their entire AMMT. This way we will cover both needs of the network which are high availability (backup C-RP) and redundancy at the same time. But it MUST be noted that a C-RP cannot and MUST NOT become peer with a C-RP in another domain.

The processes involved are as follows:

- 1- Each C-RP must be told that it belongs to a GROUP, and also that it is peer with a C-RP from another Group. This is done by initiating a set of commands on each RP .consider the following set of commands are initiated on both C-RPs (0) :

```
<#IP PIM-NG GROUP [A]>
```

```
<#IP PIM-NG PEER RP GROUP [B]|PEER ADDRESS>
```

- 2- After the above commands are initiated on C-RP-GROUP-A (C-RP-A for simplicity) it starts a series of processes.
- 3- First of all it puts an entry for the new peer in a table called PEER-LIST.
- 4- Sends out a multicast C-RP introduction message to all PIM-NG C-RPs with the destination address of 239.0.1.189.
- 5- In the introduction message, C-RP-A sets the P-BIT to 1 which tells to existing C-RPs that the introduction message is meant for a C-RP that is looking for its peer. This is done duo to the fact that there might be SC-RPs in the network, and this bit will make them to ignore the packet.
- 6- Sets the value of GROUP field to its own GROUP value.
- 7- Alongside its unicast address and any existing SC-RP, C-RP will send keep-alive timer value to C-RP-B. it shows the keep-alive timer value that C-RP-B has to set for C-RP-A.
- 8- One other table that is sent in this initial introduction message is the entire AMMT of C-RP-A.
- 9- After receiving an introduction message from the other C-RP which in our case is C-RP-B, C-RP-A will update its PEER MAPPING table with information needed. And starts sending periodic unicast introduction messages as keep-alive messages to the unicast address of C-RP-B every 60 seconds. These periodic introductions do not contain the MMT.
- 10- After the initial phase of finding C-RP-B is done, C-RP-A will send the entire AMMT table just in case there is a change in the domain. This change can be the registration of a new source with C-RP-A or deletion of a source from the MMT. In such a case C-RP-A will set the Z-BIT in its unicast introduction message.

11- In a case, that C-RP-A receives an introduction from its peer with the Z-BIT set and containing the AMMT, it will only update its own AMMT and won't inform the C-MAPPER about the change if the A-BIT inside the received C-MAPPER introduction messages is set to 1. This is due to this fact that the change had been informed to the C-MAPPER before by C-RP-B, and if C-RP-A is going to inform again it will be a double introduction process and will waste bandwidth and C-MAPPERS resources and is not a necessary task nor recommended.

12- If C-RP-A doesn't receive an intro from the C-RP-B for 60 sec it will send a final unicast intro to the unicast address of C-RP-B to see if it is alive or not. If after this last unicast introduction C-RP-A doesn't receive any introductions from C-RP-B and there is any SC-RP-B, C-RP-A will immediately send a unicast introduction to the address of SC-RP-B.

13- If no SC-RP-B is considered in the domain, then C-RP-A will start sending periodic multicast introduction messages to the address 239.0.1.189 to find its peer every 60 seconds. Until it hears again from the C-RP-B or other C-RPs it might have become peer with.

The same set of actions will be done by the other C-RP (C-RP-B).

Now that the related processes are explained let's take a look at the format of PEER-MAPPING table and PEER-LIST table which is created on both C-MAPPER'(s) and C-RP'(s) in Figure 30.

PEER-MAPPING table

```

+-----+
| PEER ADDR | Domain | GROUP | Mesh-Group | Priority | status | keep-alive | expiry |
+-----+
|           |       |       |           |         |        |           |       |
+-----+
|           |       |       |           |         |        |           |       |
+-----+

```

Figure 30 PEER-MAPPING table and PEER-LIST table

- o Peer unicast address: unicast address of the peer which will be found later.
- o Status: gets a value of either A (active) or S (standby) and is used by C-MAPPERS in case more than one C-MAPPER exists.

Multicast mapping table

```

+-----+
|Source addr(S)|Dest group (G)|Source HOST| keep alive |expiry time |
+-----+
|              |              |          |           |           |
+-----+
|              |              |          |           |           |
+-----+

```

A-Multicast mapping table

```

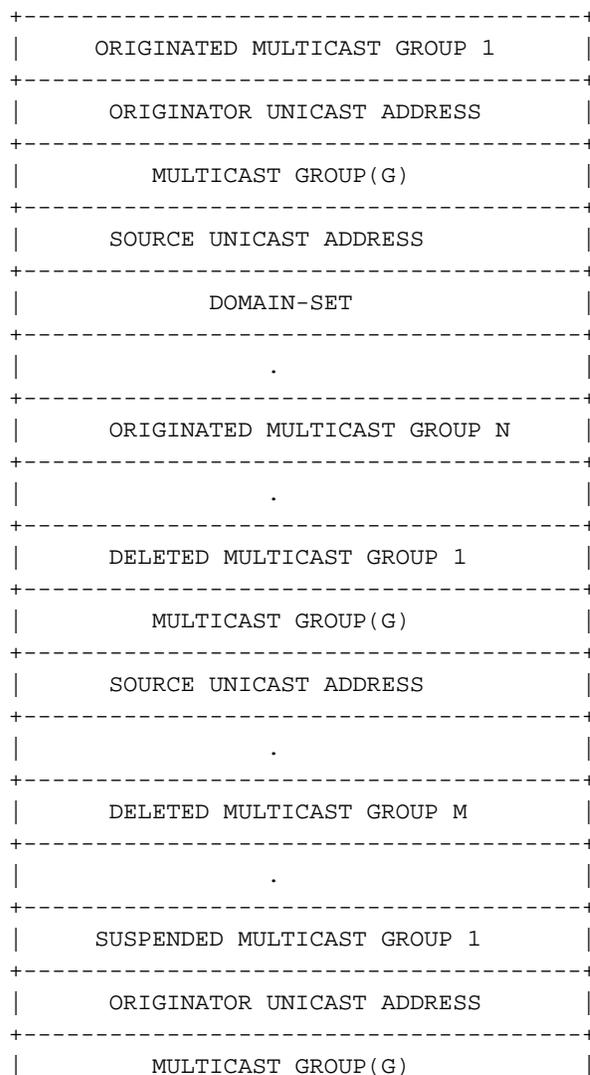
+-----+
|Source addr|Dest group (G)|Soruce Host|Originator|DOMAIN-set|status|
+-----+
|           |              |          |         |          |     |
+-----+
|           |              |          |         |          |     |
+-----+

```

- o Status field: takes the values equal to suspend or active. If an entry inside the AMMT has the status of suspend, it means that it cannot be used until either it is removed or activated again but its usage is not meant for this process and will be discussed later. So if any entry inside this table has the status of suspend C-RPs won't give the source address to any clients that may be looking for the source address of a group (G) which matches that entry.
- o Domain-set: shows the domain in which a source is generated and registered. It is called domain-set due to the fact that when separate and different PIM-NG multicast domains are connected to each other, a C-MAPPER inside one domain will add its domain number to the domain value before sending the AMMT to a peer C-MAPPER in another domain. So at the end there will be a string of domain

numbers like, 2-1-core2-core1-3-2, which will be used later. But a C-RP only adds its domain value for any sources that is registered with it or in a singly multicast domain design with one C-MAPPER and multiple PEER C-RP'(s), for any source that is received from a PEER-C-RP before sending it to another PEER-C-RP. This way the domain-set itself can be used to RPF check.

Figure 31 AMMT compared with MMT



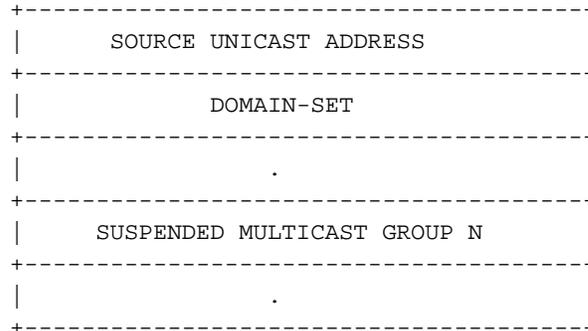
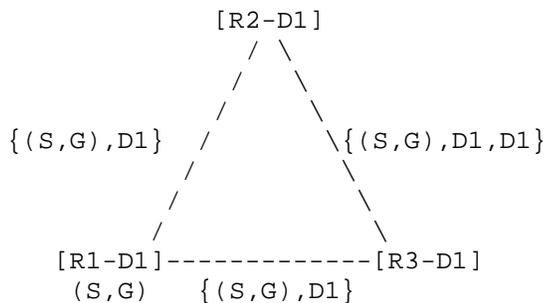


Figure 32 AMMT Format

The above(0) is the real format of the A-MULTICAST MAPPING TABLE (AMMT) being exchanged between either peering C-RP'(s) or a C-RP and a C-MAPPER in case there are multiple PIM-NG domains or peer C-MAPPERS in the domain ,with the bellow definitions and specifications :

- o In a single multicast domain network design which only one C-MAPPER exists ,and multiple PEER-C-RPs are considered for redundancy , each C-RP MUST add its domain value to the domain-set of the received sources from a peer C-RP before sending them to another peer C-RP. The domain-set will be used by peer C-RPs to RPF check the received sources inside the domain. Consider the bellow network design including 3 C-RPs all part of multicast domain(1),in which a source registers with R1 and R1 starts to advertise the new originated source to its peers R2 and R3.R3 receives an update for (S,G) with the domain-set of D1 from R1 and another update for (S,G) with domain-set [D1,D1] from R2.so at this point the received update from R2 fails the RPF check because of the longer domain-set string and R3 will only accept the received update from R1.



- o Suspended multicast group: a C-RP MUST NOT put an entry for a suspended multicast group. This field MUST only be filled out by a C-MAPPER or PPER which loses connectivity with its PEER-C-MAPPER (PPER) from another domain.

Now that the processes related to bringing redundancy to the domain by using redundant C-RPs in a single multicast domain network and also the C-RP related processes in such a network has been covered we are going to explain the reactions of a C-RP in a multiple domain network and the related concepts.

4.4.2.3.3.4. PIM-NG Anycast RP

Up to this point we have discussed the concepts related to redundant RPs of up to 2 C-RPs in a domain. In such a multicast domain it won't be a problem for the C-MAPPER to introduce all existing C-RPs one by one inside the PDTT. But if the network size becomes bigger with the need for many C-RPs to bring redundancy and high availability to the network ,the process of introducing all existing C-RPs in the domain by C-MAPPER to ALL-PIM-NG-ROUTERS including CLIENTS and any other C-MAPPERS will be a waste of network resources.

So PIM-NG specifications suggest the use of ANYCAST-RP concept [11] with bellow specifications and rules:

- o Each C-RP MUST use 2 different unicast addresses. One will be used in the processes related to introducing to existing C-MAPPER'(s) and PEER-C-RPs. And one unicast address will be used as the ANYCAST address. So it is better to put it this way, each C-RP MUST use 2 interface loopbacks for its processes .one interface with the ANYCAST address and one with the address used for introductions.
- o Clients MUST communicate with the closest C-RP.
- o Existing C-MAPPERS MUST become aware of ANYCAST concept usage in the domain alongside the ANYCAST ADDRESS in use ,by initiating a command such as :

<#IP PIM-NG ANYCAST ADDRESS [A.B.C.D]>

On any existing C-MAPPER'(s). So that C-MAPPERS will use the configured ANYCAST address when introducing C-RPs to the domain which will only be one entry inside the PIM DOMAIN TOPOLOGY TABLE compared to up to 256 different C-RPs.

- o If only one C-MAPPER exists in the domain or the domain is not connected to other domains, C-RP'(s) won't need to introduce themselves to the C-MAPPER. But C-RP'(s) MUST become peer with each other in order to exchange the contents of AMMT with each other to bring redundancy.
- o If the A-BIT in the received C-MAPPER introduction message is set then each C-RP will have to introduce itself to the closest C-MAPPER and start exchanging AMMT individually with the closest C-MAPPER. and in such case it is advised not to peer C-RPs with each other. This is due to the fact that C-MAPPERS will inform C-RPs about any new sources that are registered with a C-RP.

4.4.2.3.3.5. PIM-NG C-RP Mesh-Group

Up to this point of explaining PIM-NG specifications we, have explained the process of bringing redundancy to a PIM-NG domain by peering C-RPs and in case the network is a large sized network ,the usage of ANYCAST RP concept with its specifications have been explained.

In some multicast domain designs, the needs of the network may dictate the use of many C-RP'(s) in the domain to both bring redundancy and high availability. ANYCAST RP concept solves a problem regarding the introduction of C-RPs to the domain in a way that C-MAPPER'(s) will only need to introduce one unicast address as the address of the C-RP. But one problem remains unsolved which is related to the fact that with ANYCAST RP, each C-RP still has to introduce itself directly to the C-MAPPER and exchange the contents of A-MULTICAST MAPPING TABLE. Now if the number of existing C-RPs goes high, the amount of data that is going to be transferred between large number of C-RPs and C-MAPPER will waste both network and C-MAPPER resources.

To solve the above issue PIM-NG specifications introduces the concept of C-RP MESH-GROUP with bellow specifications:

- o A PIM-NG domain can contain up to 25 C-RP Mesh-Groups. With each group containing up to 10 C-RPs.
- o A Mesh-Group priority value between [0-255] MUST be defined on each C-RP that is going to be a member of a Mesh-Group. This priority is different from the priority that had been used at the time of choosing backup or SC-RP and defaults to 100. The higher the better.

- o The C-RP with the highest priority in the Mesh-Group will become the active C-RP and better to say the speaker on behalf of the Mesh-Group. to do so an election MUST be held between members of the Mesh-Group :
 1. Each C-RP introduces itself to its peers by sending a multicast introduction to the destination address 239.0.1.189, which contains its GROUP number, and the priority that will be used in the election process.
 2. Each C-RP receiving such introduction message will compare the priority in the received message with its own priority and puts an entry in its PEER-MAPPING-TABLE for that peer with the status of either active or standby.
 3. If the priority associated with a peer in a received introduction is lower than the priority set on the C-RP, then that peer will get the status of standby, and if the associated priority of a peer is higher, that peer gets an active status.
 4. At the end of this election process, each C-RP knows the active C-RP.
- o It is advised by PIM-NG specifications that a mechanism be implemented and used through configuration, that each member of a Mesh-Group becomes aware about the number of members within a Mesh-Group.
- o The active C-RP in a Mesh-Group is responsible for interacting with the closest C-MAPPER and exchanging the AMMT which contains any sources that registers with any members of the C-RP Mesh-Group.
- o If the current active C-RP dies and doesn't have any SC-RP, the next C-RP with the highest priority takes its place immediately.
- o Each member of the Mesh-Group MUST inform all the other members about any changes that occurs. So for instance, if we have 10 members in a Mesh-Group and C-RP1 receives information regarding a new source from C-RP2, it MUST NOT forward it to other members due to the fact that it has been informed to other members by C-RP2.

- o A C-RP can be a member of different Mesh-Groups, but such an implementation is not advised.
- o The above can be done by initiating command lines such as the bellow command lines on each C-RP :

```
<#IP PIM-NG RP-MESH GROUP [1-25]>
```

```
<#IP PIM-NG RP-MESH GROUP MEM-COUNT [1-10]>
```

```
<#IP PIM-NG MESH-PEER GROUP [1-255]>
```

```
<#IP PIM-NG MESH-PRIORITY [0-255]>
```

PIM-NG specifications dictates that, such a command MUST indicate the use of C-RP Mesh-Group, total number of members in the Mesh-Group, Mesh-Group that a C-RP is a member of, Group number of all the members of Mesh-Group that the C-RP is supposed to become peer with and finally the priority of the Current C-RP in the Mesh-Group.

4.4.2.3.3.6. Backup C-RP considerations

In a single PIM-NG multicast network design with redundant C-RPs there seems that no SC-RP is needed to be considered .as each C-RP can be used as the backup for the other one.

But in larger networks, the existence of backup RP seems necessary and of high importance due to the fact that the high availability of each C-RP becomes an important factor.

If any SC-RP should be considered, the processes are as explained before with one additional table being exchanged between C-RP and SC-RP, which is the A-MULTICAST-MAPPIN-TABLE.

4.5. C-MAPPER Redundancy in PIM-NG

Peering C-MAPPERS will bring redundancy to existing C-MAPPERS .in large networks and special designs the redundancy between C-MAPPERS inside the same domain might become handy.

The main use of peer C-MAPPERS is going to be felt when connecting 2 or more separate PIM-NG domains, which will be discussed later.

So for the sake of simplicity in explaining the processes involved we are going to explain the processes in a PIM-NG network with one multicast domain. See 0.

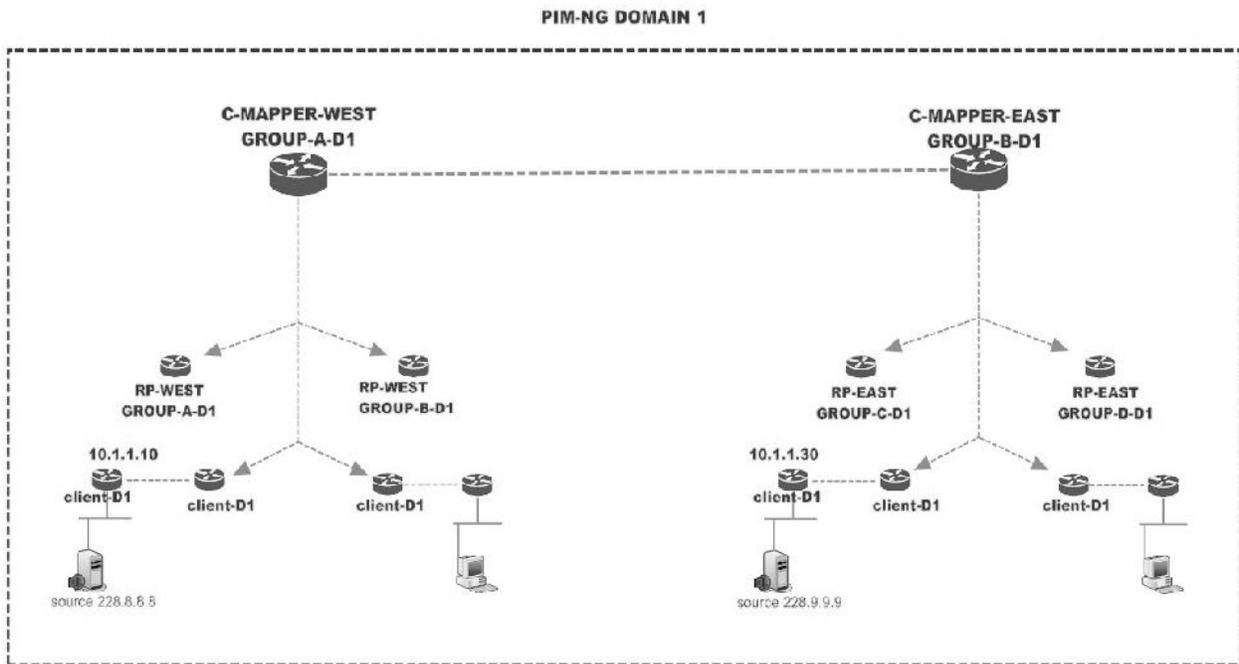


Figure 33 network with single multicast domain and redundant C-MAPPERS

As you see in 0 we have a single domain and redundant MAPPERS. If you look at the picture each C-MAPPER is assigned to a unique GROUP.

PIM-NG specifications dictate that each C-MAPPER can only communicate with MAPPERS inside the same GROUP and domain. In other words each MAPPER can hear the introduction of MAPPERS in other groups sent to 239.0.1.188 but won't show any reaction to it and will only forward it, unless it is told that it can listen to the introduction of C-MAPPER's in other groups and use the information inside the introduction message.

The process begins by telling each C-MAPPER that it is peer with a C-MAPPER from other groups. Consider the sample network illustrated in 0. C-MAPPER-WEST(called west for simplicity) which is in domain 1 and is assigned to GROUP-A is told that it is peer with a C-MAPPER in domain 1 and GROUP-B which is in this case C-MAPPER-EAST(called east for simplicity).this must be done on both sides. As soon as west is told that it is peer with east a series of actions occurs:

- 1- West puts an entry for east or simply GROUP-B in a special table called PEER-MAPPING.and also an entry in the Hold Time field of the introduction message which shows the keep-alive timer each peer should set for the C-MAPPER-west .

```

+-----+
| PEER ADDR|Domain|GROUP|Mesh-Group|Priority|status|keep-alive|expiry |
+-----+
|          |      |     |          |      |      |          |      |
+-----+
|          |      |     |          |      |      |          |      |
+-----+

```

Figure 34 PEER-MAPPING Table

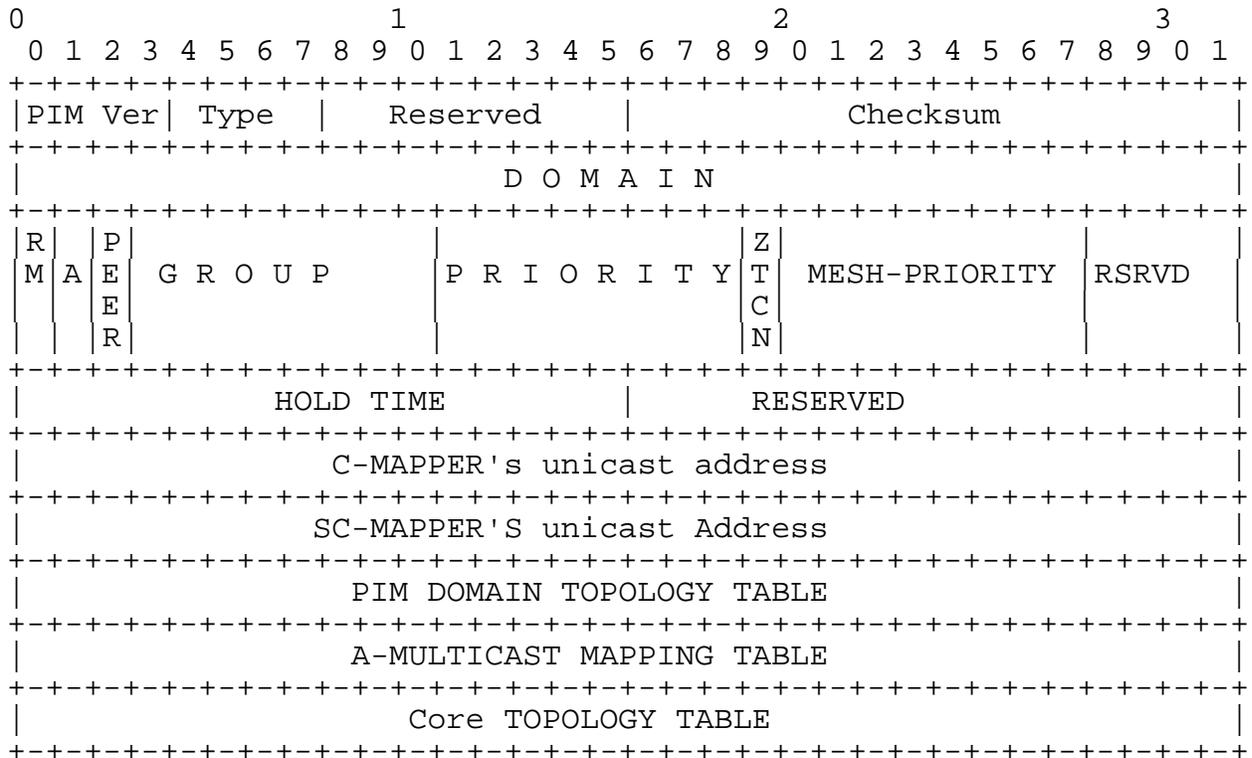
- o Peer unicast address: unicast address of the peer which will be later found.
- o Status : gets a value of either A(active) or S(standby)

- 2- West sends a multicast introduction to the destination address of 239.0.1.188 which will be heard by all the MAPPERS .in this introduction message the west introduces its unicast address plus the Domain number assigned to it and the GROUP it represents and sends its AMMT along with the PEER-LIST table. This introduction

messages will be read only by those C-MAPPERS that are supposed to become peers with the sender, and in this case east.

- 3- When West becomes peer with a C-MAPPER it sets the A-BIT in its introduction messages sent to ALL-PIM-NG-CLIENTS and All-C-RPs, so that C-RPs start sending AMMT to C-MAPPERS.
- 4- If the C-MAPPERS inside the domain are forming C-MAPPER MESH-GROUPS (4.5.2.) then, in its introduction message sent to 239.0.1.188 to find peers from the same domain, west will send its MESH-PRIORITY value. This priority value is going to be used in an election process between peer C-MAPPERS forming a MESH-GROUP inside the same domain which will be discussed later.
- 5- If the C-MAPPERS are not forming a MESH-GROUP then no priority needs to be sent.
- 6- West will keep sending the introduction every 30 seconds until it receives an acknowledge or simply put introduction from east ,which can be either a unicast introduction to west sent from east or a multicast introduction sent from east in order to find its peers.
- 7- After the introduction from east is heard, and the Domain number and group numbers are checked and verified, west keeps sending periodical introduction messages every 60 seconds by default or every X second equal to the time set in the KEEP-ALIVE-TIMER inside PEER-LIST.
- 8- If the Mesh-Priority value of EAST is lower than its own Mesh-Priority it means that , it is the ACTIVE-C-MAPPER and responsible of introducing east or any other STANDBY-C-MAPPER(STC-MAPPER) existing inside the domain to ALL-PIM-NG ROUTERS so that C-RPs will learn the address of STANBY-C-MAPPERS to use the closest C-MAPPER.
- 9- Introduction messages will be sent by west under these conditions :
 - o Triggered multicast introductions whenever west is told to become peer with a new C-MAPPER.
 - o Periodical unicast introductions, which will act as keep-alive messages after finding the peer. In this type of introduction no AMMT will be exchanged.
 - o Triggered unicast introduction to the address of east to inform a change in the network, like a new entry in the AMMT by setting a

flag called ZONE-TOPOLOGY-CHANGE-NOTIFICATION (ZTCN) and sending the entire new AMMT.



| |
|-----------------------------|
| DELETED MULTICAST GROUP M |
| . |
| SUSPENDED MULTICAST GROUP 1 |
| ORIGINATOR UNICAST ADDRESS |
| MULTICAST GROUP(G) |
| SOURCE UNICAST ADDRESS |
| DOMAIN-SET |
| . |
| SUSPENDED MULTICAST GROUP N |
| . |

Figure 36 A- Multicast MAPPING TABLE

The above table (0) is the real format of A-MULTICAST MAPPING TABLE that is exchanged between peer C-MAPPERS with bellow definitions:

- o Originator unicast address: holds the address of the C-MAPPER that, a source is originated in its domain and it has received the information regarding the source from a connected C-RP. It is used by PEER-C-MAPPERS inside the same domain as a means for RPF check of a received update regarding a source that is inside their domain. Also the Originator address will be used when a PIM-NG domain is going to be connected to a PIM-SM domain to bring compatibility.
- o A C-MAPPER MUST change the contents of originator unicast address field with its own address, for sources that are originated inside its own domain and received from a downstream C-RP.
- o A C-MAPPER MUST NOT change the contents of originator unicast address field of sources that are received from a peer C-MAPPER or PPER.
- o Each C-MAPPER MUST add its domain number value to the domain-set of a received source when sending the AMMT to a peer C-MAPPER either inside the same domain or in another domain. This domain-set is used as a means of RPF check which will be discussed later.

- o If a C-MAPPER or PPER wants to send an update to a peer in another domain, it MUST first remove any additional domain values equal to its domain value from the string, and then add its domain value to the string and send it to other domains. The additional domain values are those added by C-MAPPERS inside the domain to be used for RPF check.
- o Suspended multicast group: holds the information regarding any multicast sources that MUST be treated as suspended .when a C-RP receives this information it will not answer to requests from clients for that source until the sources are again active or deleted.

Now that west has done its part of the work, it is time to see what are the processes related to its peer C-MAPPER-EAST.

As soon as east is told that it is peer with Group-A, it will do as described above .also it will wait to hear a multicast introduction destined for 239.0.1.188. When east receives such packets it will do the following:

- 1- First of all it will check the PEER field to see if it is set. If this field is set, it means that it is sent from a C-MAPPER that is looking for a peer or wants to introduce itself to a peer. In our example east receives a packet destined for 239.0.1.188 .checks the PEER field .and sees that it is set by sender.
- 2- East checks the Domain field to see if it is sent from inside the domain or not.
- 3- Then east will check the GROUP field to see if the packet belongs to a group it is looking for and wants to become peer with. In this case east will see A in the GROUP field.
- 4- It checks its PEER-MAPPING table and sees an entry for GROUP-A .so it understands that the packet is sent from its peer.
- 5- If the Mesh-Group concept is implemented, Then east checks the Mesh-Priority value to see if the sender has a higher value or not. And without the Mesh-Group concept EAST MUST start sending introductions to 239.0.1.190 introducing itself to PIM-NG population. And use the information regarding WEST to only exchange information regarding multicast sources.
- 6- If the Mesh-Priority value is higher it means that it is going to be in standby mode and only connect with existing C-RPs

through unicast introduction messages it will receive from C-RPs. And won't send multicast introductions to 239.0.1.190 as long as the ACTIVE C-MAPPER exists.

- 7- East will then look in the SOURCE-UNICAST-ADDRESS field to find the address of the sender. And in our example east finds the address of west, and puts the west's unicast address in its PEER-MAPPING table alongside the associated status.
- 8- Now it is time for east to check the PEER-LIST table inside the introduction message. It finds an entry for GROUP-A alongside the associated KEEP-ALIVE timer. East then puts the associated keep-alive timer in the right field in its PEER-MAPPING table.
- 9- It will read the content of the AMMT sent from west and enters the new mappings in its own AMMT for further use.
- 10- It will send a unicast introduction to the unicast address of west as an acknowledgement and also for west to learn east's address.
- 11- If any C-RP introduces itself to east, then east will have to send the unicast address of C-RPs to west inside the PDTT so that west will be able to introduce the existing C-RPs to ALL-PIM-NG-CLIENTS by sending multicast introduction to 239.0.1.190.
- 12- East will inform C-RPs connected to it, about any changes received inside the A-MULTICAST MAPPING table by sending unicast introductions containing the A-MULTICAST MAPPING table to C-RPs it knows.

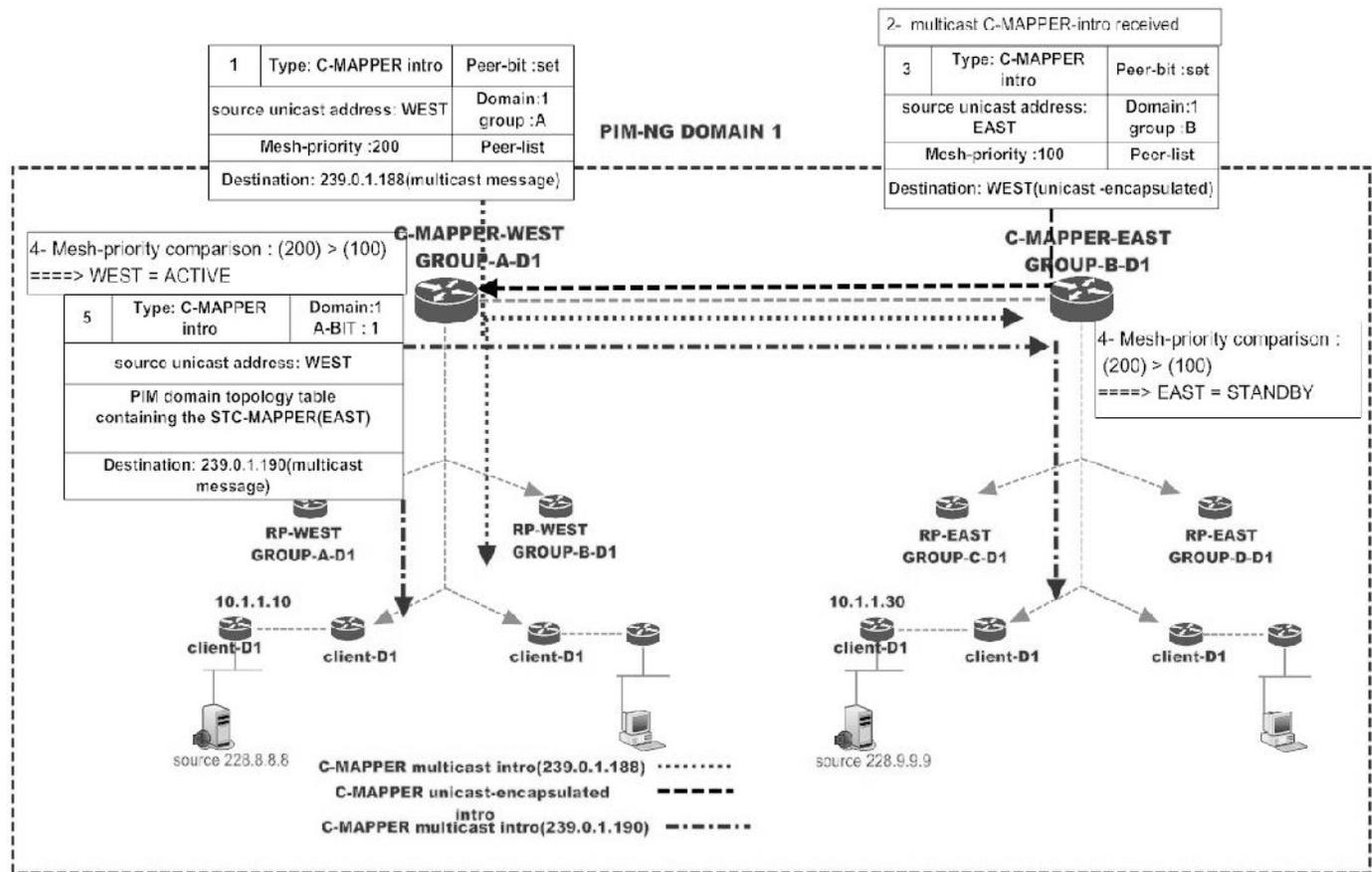


Figure 37 sample illustration of the above process

The above concepts are mostly related to a design in which C-MAPPER Mesh-Group concept is considered to reduce the amount of C-MAPPER introduction message sent to 239.0.1.190.so without the Mesh-Group concept:

- o each C-MAPPER starts sending introduction messages to 239.0.1.190 individually and since the C-RP'(s), RC'(s) and clients residing in a PIM-NG multicast domain MUST communicate with the closest C-MAPPER, then peer C-MAPPERS(WEST,EAST) MUST ONLY exchange AMMT and do not need to exchange the information regarding C-RP'(s) or TR'(s).
- o In the case of peer C-MAPPERS without the Mesh-Group concept, if the needs of multicasting in a domain dictates that the information regarding existing C-RP'(s) or TR'(s) MUST be

exchanged between the peer C-MAPPERS, PIM-NG specification strongly advises that this MUST be done through command initiation as an optional feature.

- o Duo to the fact that, in designs with more than one existing C-MAPPER, the amount of C-MAPPER introductions destined for 239.0.1.190 will go high and will consume network resources; it is STRONGLY advised to use the Mesh-Group concept.

4.5.1. SC-MAPPER considerations

Actions taken by C-MAPPERS and SC-MAPPER are as follows:

- 1- In the case of existing backup MAPPER'(s), C-MAPPER will send the learned unicast address of the peers in its PDTT along the new multicast groups it may learn in its introduction messages to the SC-MAPPER so in case that the C-MAPPER crashes and dies the SC-MAPPER will be able to immediately take its place and introduces itself on behalf of its GROUP to the peering C-MAPPER. C-MAPPER will inform the SC-MAPPER about any changes in the domain by sending AMMT.
- 2- As explained in previous sections, if the SC-MAPPER doesn't receive a C-MAPPER introduction for 2*30 seconds it will immediately take its place and send a C-MAPPER introduction to the domain. And in this case as the SC-MAPPER is aware of the existence of PEER C-MAPPERS it will send unicast introductions to the unicast address of peers.
- 3- If C-MAPPER-EAST (0) doesn't receive an introduction from its peer for 70 seconds (default timer+10sec) it will send an introduction to the address of SC-MAPPER-WEST .although if everything works accordingly by this time C-MAPPER-EAST must have received the introduction message of SC-MAPPER-WEST.

4.5.2. C-MAPPER Mesh-Group

In a single PIM-NG multicast domain with needs for redundant C-MAPPERS or PEER-C-MAPPERS ,a mechanism MUST be used to prevent each single one of existing C-MAPPERS from sending multicast introduction

messages to the destination group 239.0.1.190 which will be heard by ALL-PIM-NG-CLIENTS and PIM-NG-C-RPs. this prevention mechanism MUST be implemented due to the fact that ,if each C-MAPPER is going to send the mentioned multicast introduction message , the multicast traffic created will consume lots of bandwidth and network resources.

To do so PIM-NG specifications define the concept of C-MAPPER MESH-GROUP and ACTIVE-C-MAPPER and STANDBY-C-MAPPER (STC-MAPPER) in a way that:

- o Each Domain can have up to 25 C-MAPPER MESH-GROUPs, with each group ONLY including up to 10 C-MAPPERS.
- o Through an election between the existing C-MAPPERS, one C-MAPPER becomes the ACTIVE C-MAPPER and all the other C-MAPPERS become STANDBY C-MAPPERS.
- o The ACTIVE C-MAPPER is in charge of introducing other C-MAPPERS or STC-MAPPERS to ALL-PIM-NG-CLIENTS and C-RPs inside the domain it is resided in, in case ANYCAST RP is not considered, by means of sending multicast introduction messages to 239.0.1.190 and sending the unicast address and role of other C-MAPPERS inside the PIM domain topology table. So that C-RPs in different parts of the domain will find the closest C-MAPPER and introduce to that C-MAPPER, whether it is an ACTIVE or STANBY C-MAPPER.
- o STC-MAPPERS will maintain connectivity with the active C-MAPPER, which the related processes had been explained.
- o Each STC-MAPPER will send the information related to new C-RPs it finds to the ACTIVE-C-MAPPER by sending unicast introduction messages and sending the PIM domain topology table inside it which contains the information regarding the newly found C-RPs.
- o Active C-MAPPER then introduces the existing C-RPs and TR'(s) to ALL-PIM-NG-CLIENTS by sending multicast introduction messages to 239.0.1.190, which include PDDT. Through the above process each CLIENT will be able to contact the closest C-RP in case it has a source to register or it needs to find a source.
- o Each C-MAPPER in a MESH-GROUP (STC-MAPPER and ACTIVE-C-MAPPER) MUST inform other C-MAPPERS about any changes regarding the existing multicast sources in the domain by updating its AMMT with the information it receives from its connected C-RPs. This way if, for instance C-MAPPER1 receives an update from C-MAPPER2, it will not need to send it to C-MAPPER3 as it has been done by C-MAPPER2.

- C-MAPPERS, both ACTIVE and PASSIVE, MUST then inform their connected or downstream C-RPs about any changes regarding the existing multicast sources by sending the AMMT. This way ALL-PIM-NG C-RPs in different parts of the network will have enough information about the domain to answer the requests of clients in need of finding a source. This process will eliminate the need for PEER C-RPs.
- If the design of multicast domain, dictates to form more than one C-MAPPER Mesh-Group, PIM-NG specifications STRONGLY advises to place a boundary between different Mesh-Group areas by simply putting PER'(s) where ever needed. This will limit the propagation of C-MAPPER introduction messages sent by ACTIVE-C-MAPPER in each Mesh-Group area to other areas and will reduce the network resources consumption by unwanted multicast traffic.
- If no boundary is considered between different areas, components of the PIM-NG domain such as C-RP, Client, and ETC, MUST react to the introduction message of the closest ACTIVE C-MAPPER according to the contents of the SOURCE UNICAST ADDRESS field of the introduction message and the information inside the unicast routing table. The only issue within the domain MUST BE ONLY the consumption of network resources by unnecessary multicast traffic related to multiple ACTIVE C-MAPPER introduction messages.
- In the case of implementing more than one Mesh-Group, each area separated from other areas by PER'(s) will act normally, and existing components such as C-RP'(s), Clients and other components will respond to the introduction messages received from the active C-MAPPER in the area.
- If a PIM-NG domain is consisted of more than one C-MAPPER Mesh-Group and separated in to different areas, PIM-NG specification STRONGLY advises to peer 2 C-MAPPERS from each Mesh-Group with each other using the static methods so that the information related to multicast sources in each area reaches other areas. Using the static method of connecting 2 different Mesh-Group will eliminate the need for the C-MAPPER introduction messages sent to 239.0.1.188 to find other C-MAPPERS, to be forwarded by PER'(s) between areas.
- If a C-MAPPER in a Mesh-Group is becoming peer with a C-MAPPER in another Group, depending on the needs of multicasting in the domain, it may need to advertise the information regarding the existing C-RP'(s) and TR'(s) too, so in case anything happens to C-RP'(s) or TR'(s) in one area, the network stays stable and

converges automatically. PIM-NG specifications advise to do this through command initiation on C-MAPPERS as an optional feature.

4.5.2.1. Active C-MAPPER Election

The election process starts as soon as a c-MAPPER inside a domain becomes aware that it is part of a MESH-GROUP and is peer with a C-MAPPER inside the same domain but with a different group. The related processes are explained bellow:

- 1- As soon as each C-MAPPER becomes aware that it is peer with another C-MAPPER in a MESH-GROUP in the same domain but with a different group, it starts the process of sending multicast introduction messages to ALL-PIM-NG C-MAPPERS destination address of 239.0.1.188.
- 2- Each C-MAPPER will set the PEER-BIT, so that incase there are any SC-MAPPERS inside the network, they won't read the contents of the message.
- 3- Each C-MAPPER sends its domain number, group number, alongside the priority to be used in the process of election.
- 4- As soon as each C-MAPPER receives an introduction from its peer C-MAPPER, it will compare the Mesh-Priority value inside the introduction message with its own Mesh-Priority value.
- 5- If the received priority value is higher than its own priority value then, it will put an entry for the peering C-MAPPER inside its PEER-MAPPING table with the status of ACTIVE. And by this time the C-MAPPER understands that it is the STC-MAPPER and must act as one.
- 6- If the received priority value is lower than its own priority value then an entry will be put in the PEER-MAPPING table with the status of STANDBY. And by this time the C-MAPPER knows that it is the active C-MAPPER and should act as an ACTIVE-C-MAPPER.
- 7- In case there is more than 2 C-MAPPERS inside the domain, and thus each C-MAPPER is configured to become peer with more than 1 C-MAPPER from its own domain, then each C-MAPPER must do the above processes until it finds all of its peers and comes to a conclusion that who is the ACTIVE-C-MAPPER inside MESH-GROUP in the domain.

- 8- It is strongly advised by PIM-NG specification that a mechanism be implemented through configuration that informs each C-MAPPER about the total number of C-MAPPERS that are members of a Mesh-Group.
- 9- By this time each C-MAPPER knows that which C-MAPPER is the active one .so without doing much more process, each C-MAPPER will silently accept its role and starts the related process.

Through the above processes existing C-MAPPERS in a domain can decide who is the ACTIVE-C-MAPPER and the rest of them will be STC-MAPPERS.

4.5.3. ANYCAST RP rules

When ANYCAST RP is in use, and more than one C-MAPPER is implemented in the domain, a C-MAPPER MUST NOT send the information regarding any new C-RP it finds through the introduction processes, to its peers. This is due to the fact that from the CLIENTS point of view ALL-C-RPs has the same unicast address.

Each C-MAPPER MUST maintains connectivity with each C-RP it finds.

4.5.4. Configuration process of Redundant C-MAPPERS

With explanations in previous sections the process is straight forward also refer to 0 if needed:

- 1- Each component residing inside a domain is supposed to have a domain number. So it is assumed that up to this point ALL-PIM-NG routers inside a domain has been configured with a command which indicates the domain a CLIENT , RP and MAPPER resides in :

```
<#IP PIM-NG DOMAIN [X]>
```

The concepts related to DOMAIN are explained in section 4.6.1.1. .

- 2- Each C-MAPPER needs to be assigned a GROUP number :

```
<#IP PIM-NG MAPPER GROUP [1-255]>
```

The above commands tells the C-MAPPER that it belongs to a specific GROUP, which in our illustrated example will be GROUP-A(west) and GROUP-B(east). please keep in mind that the group value is a number between 1 to 255

- 3- Each C-MAPPER needs to know its peer GROUP and if needed the address of the peer:

```
<#IP PIM-NG PEER MAPPER DOMAIN[X] [GROUP [0-255]] [MAPPER address]]>
```

Or in case of a mesh group

```
<#IP PIM-NG MAPPER-MESH-GROUP [1-25]>
```

```
<#IP PIM-NG MAPPER-MESH-GROUP MEM-COUNT [1-25]>
```

```
<#IP PIM-NG MESH-PEER MAPPER DOMAIN[X] [GROUP [0-255]] [MAPPER address]]>
```

```
<#IP PIM-NG MESH-PRIORITY [0-255(default 100)]>
```

THE above command lines dictates to the C-MAPPER that it must become peer with a C-MAPPER from DOMAIN X and GROUP[0-255] ,which in our example will be GROUP-B for west and GROUP-A from east. And also it should use the configured priority value to be used in ACTIVE-C-MAPPER election.

Or in case of implementing the Mesh-Group concept dictates the Mesh-Group Number, total number of Mesh-Group members, group number of the peer or the unicast address of the peer C-MAPPER and finally the priority used in the Active C-MAPPER election.

- 4- In case the unicast address of the peering C-MAPPER is used the router will send unicast introduction messages to its peer. Although not advised because of the dynamic nature of this process, and especially when there are backup peers in each group, but it is a way to do it.

4.5.5. C-RP and Redundant C-MAPPERS

In such network designs with redundant C-MAPPERS, Active C-MAPPER informs the C-RPs about the existing C-MAPPERS inside the domain by

sending multicast introduction messages to ALL-PIM-NG-CLIENTS destination address 239.0.1.190.

Also as described before as soon as a C-MAPPER becomes peer with another C-MAPPER it will set the A-BIT when sending introduction messages to either ALL-PIM-NG-CLIENTS (239.0.1.190) or unicast address of a C-RP.

The above processes dictate to the existing C-RPs that they MUST:

- o Find the closest C-MAPPER based on the unicast address of the C-MAPPER and the information inside the unicast routing information base.
- o Introduce to the closest C-MAPPER.
- o If any changes occur inside the domain like, the registration of a new source or deletion of a new source, each C-RP must inform the C-MAPPER by sending the AMMT to the C-MAPPER. This way other C-RPs inside the domain will receive the information regarding a new source registration inside the network from their closest C-MAPPER which they had introduced themselves to it.
- o If a C-RP receives a request from a CLIENT for a source, C-RP MUST send the Domain-set associated with a source to the client, so that the client knows that where the source is generated and use the DOMAIN-SET later when sending a join request to the source.

4.6. Multiple multicast domains

Throughout the specifications of PIM-NG up to this point the concepts and processes that occur in a single PIM-NG multicast domain have been explained.

But in large sized networks such as an enterprise network or the World Wide Web (WWW) itself, the need for using multiple multicast domains instead of a single multicast domain is needed to be considered.

Examples of such cases can be:

- o Enterprise networks ,which the needs of the network dictates to divide the multicast domain in to 2 or more separate domains to have a better control over the propagation of the multicast traffic.
- o Enterprises or companies, with buildings and networks in different places. With each network or building connected by means of GRE-TUNNELS or MP-BGP to carry multicast traffic.
- o The internet as the biggest network of all, which is divided to different multicast domains, each connected to the other ones by different methods like MP-BGP as the carrier protocol of multicast traffic.

connecting 2 or more PIM-NG multicast domains to exchange the information regarding the multicast sources originated in each domain, is done through making 2 or more C-MAPPERS from one domain peer with one or more C-MAPPER'(s) in the other domain. The C-MAPPERS then will start to exchange their AMMT'(s) with each other through sending unicast-encapsulated introduction messages to each other. And because of the bellow facts:

- o The propagation of multicast introduction messages of one domain in to other domains is not desired.
- o A PIM-NG domain may be 1 or more Autonomous Systems away.
- o There may be a transitory AS between 2 PIM-NG domains, which only provides the connectivity in terms of forwarding multicast traffic destined for group (G) or join/prune messages. Such a scenario can be seen in designs where 2 PIM-NG domains are either connected through an ISP or a PIM-SM network.
- o There might be one or more AS'(s) or networks between 2 PIM-NG domains, that must not receive the information regarding some of the sources that are being originated, or must not receive them at all. But their network infrastructure is needed to forward join/prune messages and multicast traffics.

C-MAPPERS MUST NOT use the automated mechanisms described up to this point to become peers and MUST use the unicast address of their desired peers in other domains.

Also PIM-NG uses a unique method to RPF check the received information regarding multicast sources which provides the existence of transitory Ass unlike the RPF check method used in PIM-SM and MSDP which has some limitations in this part.

So PIM-NG specifications uses some rules, processes and definitions specific to PIM-NG to connect different PIM-NG domains or a network of PIM-NG domains with a network of PIM-SM domains, which are explained in the following sections.

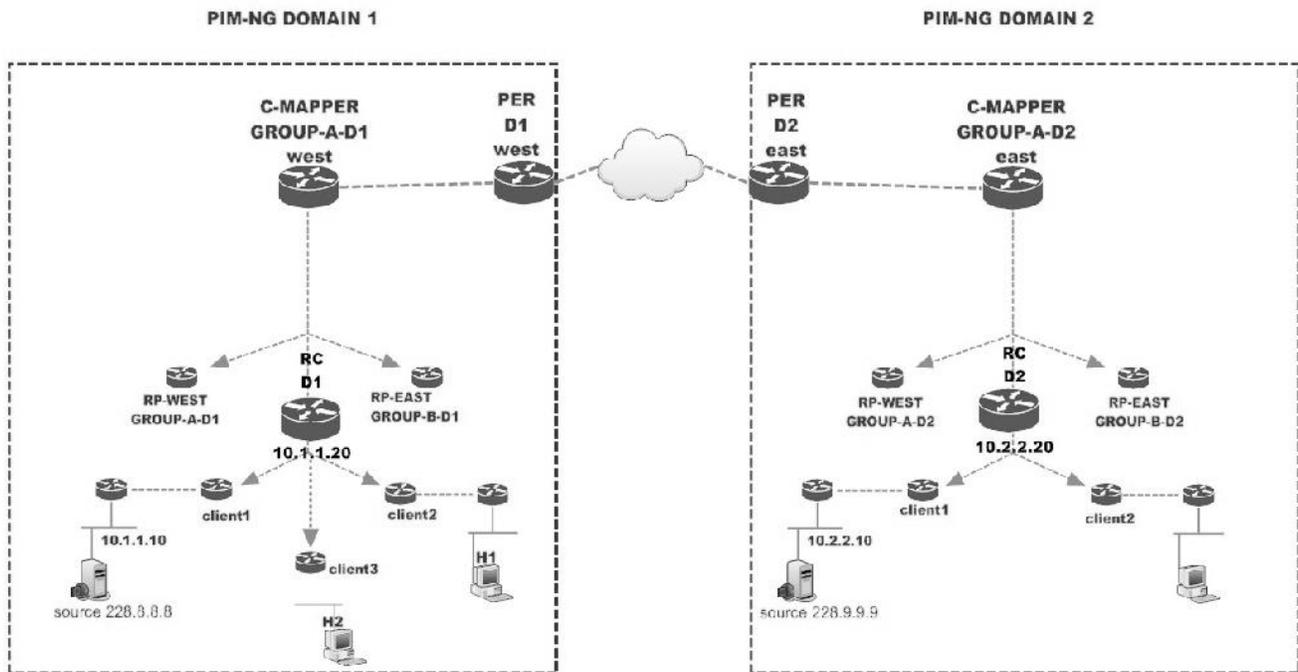


Figure 38 network with 2 PIM-NG multicast domains

4.6.1. Definitions and concepts

In the following sections different features and rules that are used when connecting multiple PIM-NG multicast domains, and related definitions and concepts are explained.

4.6.1.1. Domain

PIM-NG introduces the concept of domain in a way that each domain is distinguished from the other domains by the domain number or value. Each domain has its separate set of CLIENTS, C-RPs, TRs and C-MAPPERS.

PIM-NG uses a 32 bit domain field to support future needs of multicasting and classifies the range in to 4 groups which are explained bellow. But PIM-NG specifications suggest to different methods of domain number assignment which both of them classifies the range in to 4 groups:

- 1- This method is unique to PIM-NG specifications and we STRONGLY advise the use of this method which will need the domain numbers to be either globally unique and assigned by IANA or locally unique and assigned by either the RIRs or an entity or enterprise with a unique PIM-NG-CORE-DOMAIN number:
 - o Domain range [1-9900] as PIM-NG-DOMAIN.
 - o Domain range [9901-9999] as PIM-NG-DOMAIN for private, experimental and documentation use.
 - o DOMAIN range [10000-4294000000] as PIM-NG-Core-DOMAIN.
 - o Domain range [4294000001-4294967293] as PIM-NG-Core-DOMAIN for private, experimental and documentation use.
 - o Domain numbers 0 and 4294967294 are reserved.
- 2- This method won't need any assignment by IANA or any RIR, since a domain number will be the same as the globally unique autonomous system number (AS number) assigned by IANA to be used as PIM-NG-CORE-DOMAIN numbers, and PIM-NG-DOMAIN numbers will be chosen from the private ranges. But this method is only introduced to make the domain number assignment easier and is not the advised method:
 - o Domain range [1-64511] and [64536-4200000000] as PIM-NG-CORE-DOMAIN numbers which due to the fact the numbers are assigned by IANA as AS numbers, they are globally unique.
 - o Domain range [64512-65512] as PIM-NG-DOMAIN numbers. This range needs to be locally unique from the RIR-CORE-DOMAIN point of view or the CORE-DOMAIN of an entity or enterprise.
 - o DOMAIN range [65513-65535] as PIM-NG-DOMAIN numbers for private, experimental and documentation use.
 - o Domain range [4200000001-4294967293] as PIM-NG-Core-DOMAIN for private, experimental and documentation use.
 - o Domain numbers 0 and 4294967294 are reserved.

Now that the domain number classifications are explained, we are going to explain the rules and specifications that apply to the above classifications.

The bellow specifications and rules apply to ALL-PIM-NG domains:

- 1- All clients inside a domain MUST have the same domain number.
- 2- All C-RPs inside a domain MUST have the same domain number, as the clients.
- 3- All C-MAPPERS inside a domain MUST have the same domain number as the CLIENTS.
- 4- All TRs inside a domain MUST have the same domain number as the C-MAPPER.
- 5- Clients MUST only respond to or accept HELLO messages from a neighbor with matching domain number.
- 6- Clients MUST only accept and respond to C-MAPPER introduction messages sent to 239.0.1.190, which has a matching domain number.
- 7- C-RPs MUST only accept and respond to C-MAPPER introduction messages sent to 239.0.1.190 or unicast C-MAPPER introduction messages with matching domain number.
- 8- C-RPs can only become peer with C-RPs inside the same domain, so a C-RP MUST NOT accept or respond to the C-RP introduction message sent to 239.0.1.189 from other domains.
- 9- C-MAPPERS can become peer with C-MAPPERS from the same domain or from other domains.
- 10- C-MAPPERS MUST BE the only means to exchange the information regarding the existing sources in each domain by exchanging A-MULTICAST MAPPING TABLES.
- 11- IF a DOMAIN is needed to be connected to another domain it MUST be connected through one or more PIM-EDGE-ROUTERS (PER/PPER) to prevent the propagation of unnecessary multicast introduction traffic in each domain. In case an enterprise needs to divide its multicast domain in to 2 or more sub-domains, it is strongly advised to use PERs to prevent the fellow of unnecessary multicast introduction traffic of one domain to other domains.

- 12- C-MAPPER'(s) MUST introduce all existing C-RPs and TRs they find inside the domain to ALL-PIM-NG-ROUTERS inside the domain (230.0.1.190).
- 13- C-MAPPERS MUST remove any private DOMAIN, private CORE-DOMAIN or sub-domain number before advertising any multicast source globally or to be more specific to a peer C-MAPPER in another domain.

The bellow specifications and rules apply to PIM-NG-DOMAINS:

- 1- PIM-NG-DOMAIN number values MUST be unique from the connected PIM-NG-CORE-DOMAIN point of view .so the values MUST be assigned by a local organization or Regional Internet Registry who has a globally unique PIM-NG-CORE-DOMAIN number, in case the enterprise doesn't have a PIM-NG-CORE-DOMAIN number which is globally unique.
- 2- A PIM-NG-DOMAIN can be either a public network or a private network which is connected to the public network by using the NETWORK ADDRESS TRANSLATION. And by private PIM-NG specification means that all the components of the PIM-NG domain such as C-MAPPERS and even sources can use private unicast addresses.
- 3- An enterprise can use any PIM-NG-DOMAIN number from the classified range freely as long as it doesn't need to advertise its multicast sources globally and by advertise from now on we mean sending any existing multicast sources to a peer C-MAPPER inside a PIM-NG-CORE-DOMAIN with a globally unique domain number by sending A-MULTICAST-MAPPING-TABLE inside the unicast-encapsulated C-MAPPER introduction message.
- 4- If an enterprise needs to advertise its multicast sources globally but does not wish to or doesn't need to apply to get a globally unique PIM-NG-CORE-DOMAIN number , it MUST apply to get a unique PIM-NG-DOMAIN number from the Regional Internet Registry and advertise its multicast sources through the Regional Internet Registry's PIM-NG-CORE-DOMAIN.
- 5- If an enterprise with more than one PIM-NG-DOMAINS all connected through an Internet Service Provider(ISP) backbone ,doesn't need to advertise its multicast sources globally and only needs to use them internally , it doesn't not need to become connected to the ISP by making one or more C-MAPPERS peer with a C-MAPPER inside the ISP backbone ,due to the fact that PIM-NG specifications uses its special unicast introduction

method to make C-MAPPERS in different domains peer alongside its RPF check method unique to PIM-NG ,and doesn't use MSDP.

- 6- If a multicast source inside a PIM-NG-DOMAIN needs to be globally accessible it MUST be first advertised to a peer C-MAPPER inside an existing PIM-NG-CORE-DOMAIN.
- 7- If an enterprise needs to divide its multicast domain into 2 or more multicast domains, without the need of applying to get more than one locally unique PIM-NG-DOMAIN number from the Regional Internet Registry, it can do so by dividing its multicast domain into 2 or more domains as sub-domains of the main domain which is assigned by the Regional Internet Registry. The related concepts will be discussed later.
- 8- A PIM-NG-DOMAIN can be connected to a PIM-NG-CORE-DOMAIN through another PIM-NG-DOMAIN or even an existing PIM-SM domain and does not need to be directly connected to send join/prune messages or receive traffic for a multicast destination group. But in order to advertise its multicast sources globally it MUST be connected to either the neighbor PIM-NG-DOMAIN which is connected to a core domain or directly to the PIM-NG-CORE-DOMAIN and by connected PIM-NG specifications dictates that at least ONE C-MAPPER inside the PIM-NG-DOMAIN MUST be a peer with either a C-MAPPER inside the neighbor domain or inside the core domain.
- 9- Normal domains can have one or more Tree-Roots(TR).if TR exists in a domain the join/prune messages MUST first be forwarded towards the existing TR, and then towards the source.
- 10- If the network is consisted of PIM-NG-CORE-DOMAINS (CD) and a client is in need of sending join/prune towards a source which can be reached through an existing PIM-NG-CORE-DOMAIN, and TR or TRs are considered in each PIM-NG-DOMAIN, each client MUST first send its join request for a source towards the TR residing inside its domain and then after the join/prune message reached the local TR it MUST be forwarded towards the TR inside the PIM-NG-CORE-DOMAIN . So all clients residing inside PIM-NG-DOMAIN MUST have knowledge about the existence of the CORE-DOMAIN TR or TRs.
- 11- When a PER/PPER receives join/prune message on an external interface which is connected to either another domain or public networks, destined for a source inside its domain or a TR inside an existing PIM-NG-CORE-DOMAIN or a source in another domain it MUST forward it first toward the closest TR inside its own domain by setting the R-BIT inside the source address field of

the received join/prune and putting the address of the closest TR in the appropriate field and then the TR will forward the traffic towards the source or existing TR inside the PIM-NG-CORE-DOMAIN if it is dictated inside the join/prune message. This process will eliminate unnecessary join message traffic towards a source in case many CLIENTS will need the same specific traffic. Or a new CLIENT may ask to receive the same traffic.

- 12- A C-MAPPER MUST add its domain number to the domain-set of any received multicast sources received from PEER-C-MAPPERS in other domains before advertising them to other PEER-C-MAPPERS in other domains.
- 13- A C-MAPPER inside a PIM-NG-DOMAIN MUST add its domain number to the domain-set of any received multicast sources from other domains when advertising to a PEER-C-MAPPER inside the same domain. These additional prepended domain values will be use by C-MAPPERS inside the domain for RPF check.
- 14- A C-MAPPER inside a PIM-NG-DOMAIN MUST add its domain number to the domain-set of any received multicast sources from peer C-MAPPERS inside the domain to a PEER-C-MAPPER inside the same domain. These additional prepended domain values will be use by C-MAPPERS inside the domain for RPF check.
- 15- If a C-MAPPER loses its connectivity with a PEER-C-MAPPER inside other domains whether a PIM-NG-DOMAIN or a PIM-NG-CORE-DOMAIN it MUST NOT immediately remove the multicast sources received from that PEER-C-MAPPER .instead in such a case the C-MAPPER MUST set a 5 minute timer by-default which is called the source-suspension timer and inform other PEER-C-MAPPERS about the incident by sending AMMT with the sources it has been receiving from that PEER being entered as suspended. This way the C-RPs inside other domain will see the suspend flag and will know that they MUST NOT answer to client requests sent for those sources until further notice ,which will be either activating the sources again or removing them.
- 16- If a C-MAPPER loses its connectivity with a PEER-C-MAPPER inside the same domain, it MUST only set the suspend flag for the sources that the lost PEER was receiving from outside domains. This is due to the fact that when inside a domain a C-MAPPER is dead C-RPs will use other existing C-MAPPERS immediately. So there will be no need to set the suspend flag for the sources that are being generated inside the domain.

- 17- C-MAPPERS MUST remove any private DOMAIN or sub-domain number before advertising any multicast source to peer C-MAPPERS inside the domain or globally to a peer C-MAPPER in another domain.
- 18- C-MAPPER or PPER with peers in other domains MUST remove the additional prepended

The bellow specifications and rules apply to PIM-NG-CORE-DOMAIN (CD):

- 1- PIM-NG-CORE-DOMAINS number MUST be globally unique to be able to advertise any multicast source meant to be used globally by any user connected to the World Wide Web.
- 2- PIM-NG-CORE-DOMAIN number assignment MUST be done by an international organization such as IANA.
- 3- A PIM-NG-CORE-DOMAIN can be simply consisted of one or more C-MAPPERS and PERs in the core of an enterprise network to advertise the multicast sources received from peer C-MAPPERS in connected PIM-NG-DOMAINS of the enterprise. Or can be the internet backbone of a country connecting the countries network infrastructure to other countries internet backbone.
- 4- A PIM-NG-CORE-DOMAIN can be the backbone of an Internet Service Provider (ISP) from its customer's point of view. in this case if a customer has 2 different networks connected through an ISP, if it needs to advertise its multicast sources globally it will need to do it through the ISP's PIM-NG-CORE-DOMAIN if it doesn't wish or doesn't need to receive a unique core domain number, by peering at least one C-MAPPER in one of its domains with a C-MAPPER inside ISP backbone.
- 5- A PIM-NG-CORE-DOMAIN can be either a public network or a private network which is connected to the public network by using the NETWORK ADDRESS TRANSLATION. And by private PIM-NG specification means that all the components of the PIM-NG domain such as C-MAPPERS and even sources can use private unicast addresses.
- 6- If an enterprise does not wish to advertise its multicast sources globally but needs to use one or more PIM-NG-CORE-DOMAIN'(s) it can freely use a domain value from the private range. But in case they need to advertise any multicast source globally to be used by users connected to other PIM-NG-CORE-DOMAINS, it MUST apply to get a unique domain number.

- 7- If an enterprise is using private PIM-NG-CORE-DOMAIN numbers for its internal use and needs to advertise its multicast sources globally without getting a unique PIM-NG-CORE-DOMAIN number, it can do so by applying to get a locally unique PIM-NG-DOMAIN number to advertise its multicast sources through the Regional Internet Registry's PIM-NG-CORE-DOMAIN. In this case the private core domain number must be set to be a sub-domain of the assigned PIM-NG-DOMAIN number, which the related concepts will be discussed later.
- 8- CORE-DOMAIN C-MAPPERS are responsible for connecting different CORE-DOMAINS to each other.
- 9- A C-MAPPER MUST add its domain number to the domain-set of any received multicast sources received from PEER-C-MAPPERS in other domains before advertising them to other PEER-C-MAPPERS in other domains.
- 10- A C-MAPPER inside a PIM-NG-CORE-DOMAIN MUST add its domain number to the domain-set of any received multicast sources from other domains when advertising to a PEER-C-MAPPER inside the same domain. These additional prepended domain values will be use by C-MAPPERS inside the domain for RPF check.
- 11- A C-MAPPER inside a PIM-NG-CORE-DOMAIN MUST add its domain number to the domain-set of any received multicast sources from peer C-MAPPERS inside the domain to a PEER-C-MAPPER inside the same domain. These additional prepended domain values will be use by C-MAPPERS inside the domain for RPF check.
- 12- C-MAPPERS MUST remove any private DOMAIN or sub-domain number before advertising any multicast source globally or to be more specific to a peer C-MAPPER in another CORE-DOMAIN.
- 13- A C-MAPPER MUST remove additional domains in domain-set of a received multicast source from a PIM-NG-DOMAIN, except the first domain number in the domain-set which shows the domain in which a source is generated, when advertising a multicast source to a neighbor PIM-NG-CORE-DOMAIN.
- 14- A C-MAPPER MUST NOT remove additional domains from the domain-set of a received multicast source when advertising to PEER-C-MAPPERS inside a PIM-NG-DOMAIN.
- 15- If a C-MAPPER loses its connectivity with a PEER-C-MAPPER either inside the same domain or other domains whether a PIM-NG-DOMAIN or a PIM-NG-CORE-DOMAIN it MUST NOT immediately remove

the multicast sources received from that PEER-C-MAPPER .instead in such a case the C-MAPPER MUST set a 5 minute timer by-default which is called the source-suspension timer and inform other PEER-C-MAPPERS about the incident by setting the SUSPEND flag inside the A-MULTICAST-MAPPING-TABLE for the sources it has been receiving from that PEER and send the table with the suspend flag to other peers. This way the C-RPs inside other domain will see the suspend flag and will know that they MUST NOT answer to client requests sent for those sources until further notice ,which will be either activating the sources again or removing them.

16- If a C-MAPPER loses its connectivity with a PEER-C-MAPPER inside the same domain, it MUST only set the suspend flag for the sources that the lost PEER was receiving from outside domains. This is due to the fact that when inside a domain a C-MAPPER is dead C-RPs will use other existing C-MAPPERS immediately. So there will be no need to set the suspend flag for the sources that are being generated inside the domain.

17- CORE-DOMAIN C-MAPPER'(s) MUST introduce any existing CORE-DOMAIN TR'(s) to normal domains for further use.

18- CORE-DOMAINS MUST have one or more TR'(s).

4.6.1.2. PIM-EDGE-ROUTER (PER/PPER)

As described in the previous section each domain is distinguished from the other domains by a unique domain number.

Each domain MUST be completely isolated from the other domains, and in case for example all domains are located under one roof inside one network infrastructure and using one united dynamic routing protocol it is strongly advised to use PERs to reduce the propagation of unnecessary multicast introduction message of C-MAPPERS or C-RP'(s) in search of their peers.

An example of such network can be an enterprise network trying to divide its current network infrastructure in to 2 or more PIM-NG multicast domains or Sub-Domains. If these connected multicast domains are not isolated somehow we will see unnecessary multicast traffic inside each domain, which in the PIM-NG case will be the multicast traffics related to C-MAPPER introduction messages to ALL-PIM-CLIENTS(239.0.1.190) or C-RP introduction messages(239.0.1.189) sent out to find PEER-C-RPs which are intended to circulate inside one domain and since PIM-NG-ROUTERS inside other domains wont react

to such messages from other domains it will be a waste of network resources.

Another example could be the World Wide Web itself, in which each multicast domain is connected to other multicast domains by either using one or more edge routers connected to other multicast domains by using MP-BGP or a simple tunnel between each 2 multicast domain to give multicast traffic transfer ability. And such designs are due to the fact that in reality not all routers between each multicast domain support the needs of multicast traffic transfer.

Because of the above facts PIM-NG introduces the concept of PIM-EDGE-ROUTER with 2 different classifications:

- 1- PIM-EDGE-ROUTER(PER): a PER is simply a PIM-NG-AWARE router that acts as the boundary between either 2 PIM-NG domains, Sub-Domain or a domain and outside networks and will control the propagation of unnecessary or unwanted PIM-NG introduction messages.
- 2- PRIVATE-PIM-EDGE-ROUTER (PPER): a PPER is a PIM-NG-AWARE router at the edge of the network which is responsible for NETWORK ADDRESS TRANSLATION (NAT) operations. PPER is responsible of becoming peer with other PPERs in other domains in order to exchange AMMT'(s) between different private PIM-NG domains. It is seen as a normal C-MAPPER by other domains, and is seen as PPER by internal C-MAPPERS in a way that an existing C-MAPPER inside the domain will only exchange the contents of AMMT with a PPER and won't introduce it to the domain.

Below specifications and rules apply to PIM-EDGE-ROUTERS (PER):

- o A PIM-NG-AWARE router connecting 2 or more PIM-NG multicast domains to each other is considered a PIM-EDGE-ROUTER(PER)
- o A PIM-EDGE-ROUTER connecting a public PIM-NG domain to other PIM-NG multicast domains is called a PER. And by public PIM-NG specifications means that either every PIM-NG-AWARE router inside the domain is using public class IPs or at least C-MAPPERS or ALL-SOURCES inside the domain use public IP addresses. In the case of public PIM-NG multicast domain ,if the domain is connected to other domains by a protocol capable of transferring multicast traffic ,such as MP-BGP it is advise to put the PER at the edge of the network. If PERs in separate domains are connected by using a tunnel, they can be placed anywhere.

- PERs can use an interface loopback as their unicast address when introducing to C-MAPPERS or any other interface connected to the DOMAIN.
- Since all interfaces of a PIM-NG-AWARE ROUTER are considered internal interfaces by default, a PIM-NG PER MUST have one or more external interfaces connected to the outside networks or domain. so a PER can have one or more interface inside domain[X] and one or more inside domain[Y].
- A PER MUST NOT forward any multicast introduction messages received on an internal interface to its external interface. Unless it is dictated to the PER to forward a specific multicast introduction which MUST BE done based on the domain number indicated in the introduction message.
- A PER can act as the boundary between to C-MAPPER Mesh-Group Areas inside a domain to prevent the propagation of ACTIVE C-MAPPER introductions from one area to the other area. In this case the PER MUST BE configured to have one or more interface in each area. These areas are ONLY meaningful to the PER, so that the PER will not forward multicast introductions from one area to the other one.
- A PER which is acting as the boundary between Sub-Domains, MUST know the main domain number in which it is resided and also the Sub-Domain numbers to which it is connected. This is duo to the fact that in some designs the need for multicasting in PIM-NG will need the PER to forward the ACTIVE C-MAPPER introduction messages sent to 239.0.1.190 in the main domain by checking the domain number of the received introduction messages. The concepts related to Sub-Domain are explained in section 4.6.5. .
- A PER acts as the boundary between 2 domains, Sub-Domains and areas, so it MUST NOT forward multicast introduction messages sent to 239.0.1.188, 239.0.1.189, 239.0.1.190 received from one domain, Sub-Domain and area to the other. ONLY in case of Sub-Domain and Area a PER acting as the boundary between the 2 can forward such traffic and ONLY if it is dictated to do so.
- A PER can act as a BORDER-PIM-NG-ROUTER (BPR) when connecting a PIM-NG domain to a PIM-SM domain. In such designs PER MUST have one or more hands or interfaces in PIM-NG domain and one or more hands in PIM-SM domain.
- A PER MUST introduce itself to the closest C-MAPPER only in case, it has the role of a BPR too, so that the C-MAPPER starts

sending AMMT to the PER. PER will use the contents of AMMT in the process of forwarding join/prune messages received from the connected PIM-SM network to the PIM-NG network.

- o If in a network design such as a network containing PIM-NG SUB-DOMAINS ,it is needed that the PER forwards the multicast traffic destined for 239.0.1.188 and 239.0.1.190 in the main domain ,it MUST become aware of the situation to be able to forward that specific traffic.
- o In case TR exists in the domain, each PER MUST set the R-BIT of the Source Unicast Address which shows that the join/prune message MUST be forwarded towards the TR first and put the address of the closest TR in the appropriate field before forwarding it in the domain.

The bellow specifications and rules apply to PRIVATE-PIM-EDGE-ROUTERS (PPER):

- o A PIM-EDGE-ROUTER connecting a private PIM-NG multicast domain to other domains ,whether private or public ,and typically is responsible for doing the NETWORK ADDRESS TRANSLATION will introduce itself as a PRIVATE-PIM-EDGE-ROUTER(PPER) to the C-MAPPERS inside the domain. It will be done by the EDGE-ROUTER at the time of introducing itself to the C-MAPPER, and by setting the PRIVATE (P)-BIT in the introduction message. And by private PIM-NG specifications means that all components of the domain such as C-MAPPERS and even sources use private IP addresses, or at least some of the sources inside the domain are using private addresses which need to be accessed from the outside.
- o In a network design in which all C-MAPPERS inside the domain are using private unicast addresses, PPER is responsible of becoming PEER with other PPERs or C-MAPPERS in other domains in order to provide the ability of exchanging the information regarding the existing multicast sources in different domains, through exchanging the AMMT between PPERs or C-MAPPERS in other domains and C-MAPPERS inside their own domain. So in such a design a PPER is seen as a PEER-C-MAPPER by PPERs or C-MAPPERS in other domains and MUST also be configured as a C-MAPPER.
- o A PPER is seen as only a PPER by C-MAPPERS inside the domain and the information related to it such as the unicast address, MUST NOT be advertised inside the domain in C-MAPPER introduction messages sent to 239.0.1.190. Such information is only usable by existing C-MAPPER'(s). And C-MAPPER uses the

information regarding PPER to send the contents of AMMT to the PPER'(s).

- In designs in which at least one C-MAPPER inside the domain is using publicly routable IP address, and is configured to directly become peer with C-MAPPER'(s) or PPER'(s) in other domains on behalf of the domain a PPER may not need to play the role of C-MAPPER too, and will only play the role of PPER and receive the A-MULTICAST MAPPING TABLE for further use.
- In a network design in which all C-MAPPERS inside the domain are using private unicast addresses ,PPER is responsible of becoming MSDP-PEER with RP'(s) inside any neighboring PIM-SM domain , in case each domain needs to become aware of the sources that are being generated in the other domain.
- A PPER can act as a BORDER-PIM-NG-ROUTER (BPR) when connecting a private PIM-NG domain to a PIM-SM domain. In such designs PER MUST have one or more hands or interfaces in PIM-NG domain and one or more hands in PIM-SM domain.
- Each PPER MUST send unicast introduction messages to the unicast address of the closest C-MAPPER, Introducing itself so that the C-MAPPERS can start sending their A-MULTICAST MAPPING TABLES.
- To find the closest C-MAPPER, a PPER MUST listen to the introduction messages sent by the C-MAPPER in the domain or in case of existing C-MAPPER Mesh-Groups the introduction messages sent by the Active C-MAPPER.
- In case of a multicast domain with more than one C-MAPPER Mesh-Group isolated and separated in to 2 or more Mesh-Group areas by setting up boundaries using PERs, it is STRONGLY advised to statically introduce at least one C-MAPPER from other Mesh-Groups to the PPER. This is a MUST due to the fact that the PPER will definitely hear the introduction message of the active C-MAPPER in area it is resided in, but it needs to also be aware of at least one C-MAPPER from the other Mesh-Groups.
- Each PPER MUST inform the C-MAPPERS inside the domain about the received multicast sources from other domains by sending the AMMT to the closest C-MAPPER to which it has introduced itself to in the first place.
- When a PPER receives information regarding multicast sources inside its own domain, since it is at the edge of a private

network it MUST first save the information related to the unicast address of each source with private IP address in a table specific to PPER called INTERNAL-MULTICAST-SOURCE TABLE (IMST) and then it MUST change the unicast address of the sources with its public unicast address before sending the AMMT to PEER PPERs or MSDP-PEERs. The process of changing source unicast address of multicast groups MUST be done based on an existing ACL or route-map, due to the fact that some multicast groups may use public unicast addresses, and we don't want ending up changing those addresses too. So if a source unicast address needs to be changed it MUST be chosen by an ACL.

- o If a PPER receives a join/prune message on an external interface with its own unicast address as the source of the multicast group, it MUST check the contents of IMST and find the real unicast address of the source that is originating traffic for multicast group (G) and replace it with the address inside the join/prune message and forward it to the domain.
- o If a domain is connected to outside networks through more than one PPERs , and each PPER is acting as a C-MAPPER to become peer with PPER'(s) or C-MAPPER'(S) in other domains on behalf of C-MAPPERS inside its domain , a mechanism MUST BE used so that all the PPERs use one united publicly routable IP address as the originator address of the sources that are originated inside the domain, and also in case there are multicast sources using private IP addresses that must be advertised to other domains, as the source unicast address of such multicast sources. This is due to the fact that we don't want to end up advertising our multicast sources with different originator address or source unicast addresses to other domains. So we advise that on each PPER a united publically routable address be configured as the originator address if the PPERs also have the role of C-MAPPER.
- o In case TR exists in the domain, each PPER MUST set the R-BIT of the Source Unicast Address which shows that the join/prune message MUST be forwarded towards the TR first and put the address of the closest TR in the appropriate field before forwarding it in the domain.
- o EACH PPER MUST change the originator unicast address of the sources created inside its domain with its own public address. This address or the originator unicast address will be used in case a PIM-SM domain is needed to become connected to a network of PIM-NG domains in the process of RPF check.

- o PPER MUST maintain connectivity with C-MAPPER through its unicast introduction messages and in case it loses its connectivity with the C-MAPPER and if SC-MAPPER exists, it MUST immediately query the SC-MAPPER. This is due to the fact that a PPER is acting as the C-MAPPER and any existing SC-MAPPER from the peer C-MAPPER's point of view. PPER MUST send periodic introductions every 30 seconds.

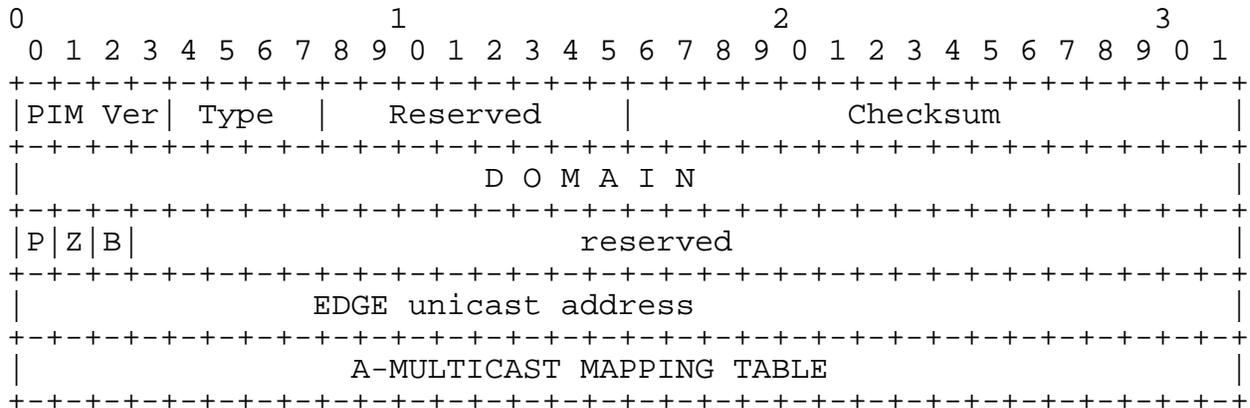


Figure 39 PER introduction message

- o TYPE :EDGE
- o Z-BIT: whenever a PPER needs to inform a change about its connected neighbor DOMAINS to C-MAPPER it will set this BIT and then send the information.
- o P-BIT: when set indicates to the receiving C-MAPPER that it is a PPER.
- o B-BIT: Border-PIM-NG-ROUTER BIT is set when a PER has the role of BPR.

| source address | Destination Group | source host |
|----------------|-------------------|-------------|
| | | |
| | | |

Figure 40 Internal Multicast Source Table

4.6.1.3. Tree Root (TR)

Up to this point of explaining PIM-NG specifications, although the existence of a component named TR was mentioned, clients were to make direct contact with a source and send their join message directly to the desired source. The processes have been explained this way to give a good understanding of the underlying processes.

They still can, but to support the needs of multicasting, PIM-NG introduces the concept of TREE-ROOT (TR). This concept has been introduced, due to the fact that, there might be many clients in each domain in need of listening to a particular multicast traffic and if they were all to send their join messages directly to that source or the TR residing in the core domain, the final result will be the waste of bandwidth and the waste of source and CORE-DOMAIN-TR resources.

The below specifications apply to TR:

- o Each normal domain can have one or more TR'(s) which it is strongly advised to be used.
- o Each core-domain MUST have one or more TR'(s).
- o ALL-PIM-NG-TR'(s) MUST introduce themselves to the closest C-MAPPER by sending unicast introduction messages, so that the C-MAPPER learns the TR's unicast address and introduce it to the domain.
- o If any core-domain is considered, the C-MAPPERS inside the core-domain MUST introduce any existing TR inside the core-domain to all normal domains connected to it so that the entire PIM-NG multicast network will become aware of the existence of core-domain TR'(s).
- o If a TR is receiving multicast traffic (G), and other TR'(s) exist in the domain, it MUST make the other TR'(s) aware of the traffic it is receiving by sending an introduction message containing the JOINED-GROUP-TABLE, to the unicast address of the other TR'(s) it receives from C-MAPPER inside PDTT. This is done due to fact that other clients may become interested in receiving such traffic later, so if they send their join to another TR which is closer to them and not aware of the multicast traffic, it can be double act of sending joins out of a domain and a waste of resources.

- When a TR receives a join/prune message with the R-BIT of Source Unicast address being set, which means that the message MUST reach the desired TR first, it MUST check the address of the TR and if the address is its own address it MUST clear the R-BIT, which means that the message has reached the desired TR in the domain and then forward it towards the desired source.
- If in a PIM-NG multicast domain more than one TR exists, and a TR receives a join/prune message for (S, G), before clearing the R-BIT and forwarding the message to the next hop, it MUST first check its JOINED-GROUPS TABLE to see if other TR'(s) has joined the SPT for that (S, G). and if an entry is found in the JOINED-GROUP table received from other TR'(s), then if and ONLY IF the cost of reaching the TR which already joined the SPT for (S, G) based on the calculations done by the underlying unicast routing protocol is better than the cost to reach the source, it MUST leave the R-BIT and MUST NOT clear it and MUST put the address of the TR which has already joined the SPT for (S,G) in the appropriate field and forward the join/prune message to the next hop in the best path towards that TR. Otherwise in case the cost of reaching the source is better than the cost of reaching the next TR which is already joined the SPT for (s, G), it MUST clear the R-BIT and forward the join/prune message to the next hop in the best path towards the source.
- If an RP has the role of TR, receiving a join/prune message of (*, G, RPT) with the RP address as the TR address means that the RP MUST NOT send back the (S) for (G) to the client and MUST find that source unicast address of (G) in either MMT or AMMT and put the address in the appropriate field and forward the packet towards the source.
- If a TR is also a C-RP, then TR introduction messages MUST NOT BE sent and only a C-RP introduction message with the TR-BIT set WILL be enough.
- Each TR MUST maintain connectivity with the C-MAPPER by sending periodic introductions every 30 seconds.
- If more than 1 TR exists within a domain in order to reduce the size of PDTT in C-MAPPER introduction messages sent to 239.0.1.190 it is advised to use the ANYCAST concept in the Domain. to use the ANYCAST concept each TR MUST be configured with both the ANYCAST address which is used by all clients and introduced by C-MAPPER in its instruction messages and an additional unicast address which will be used in TR to C-MAPPER and TR to TR communications. If the ANYCAST concept is used the

ANYCAST address MUST be defined on the C-MAPPER or Active C-MAPPER to be used in PDTT by C-MAPPER when sending introductions to 239.0.1.190.

- o If the ANYCAST concept is used the C-MAPPER MUST send the unicast address of all existing TR'(s) that it finds to existing TR'(s) within its unicast introduction messages sent to TR'(s) for further use by TR'(s).

At the end it must be noted that the TR do not need to be a separate component, and each existing C-MAPPER or C-RP in the domain is the best choice to also play the role of an TR. But PIM-NG specifications STRONGLY advise the use of C-RP if it is a must.

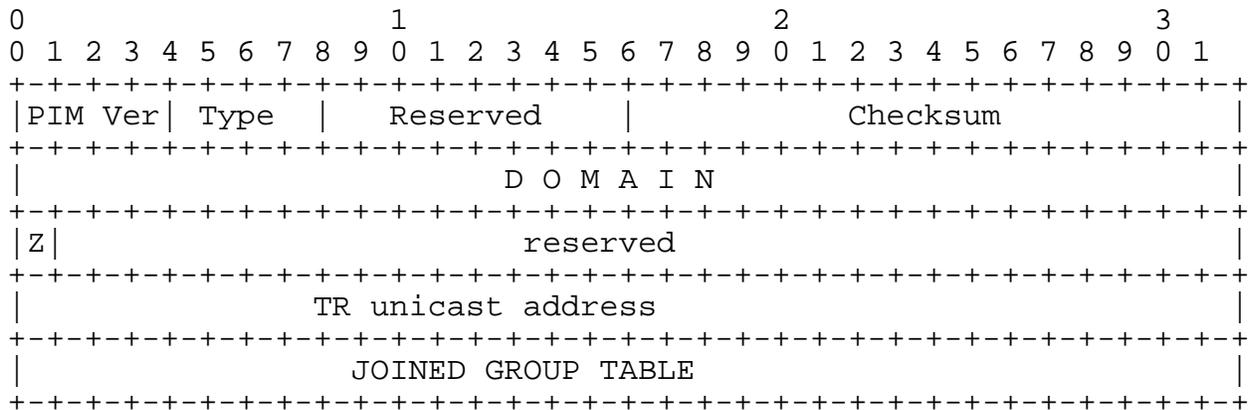


Figure 41 TR introduction message

- o Type : TR
- o Z-BIT: if set indicates to the receiver that there has been a change in the joined-group-table.

```

+-----+
|      JOINED-GROUP1      |
+-----+
| MULTICAST GROUP (G)    |
+-----+
| SOURCE UNICAST ADDRESS |
+-----+
|      .                  |
+-----+
|      JOINED-GROUP N    |
+-----+
| MULTICAST GROUP (G)    |
+-----+
| SOURCE UNICAST ADDRESS |
+-----+
    
```

- Source Unicast address Field holds the unicast address of either a TR or Client who is joined a tree for (G). it is set to all 0s if the TR itself is joined to the Tree for (G)

Figure 42 Joined-Groups table used by both TR'(s) and Clients

4.6.1.4. Domain-set and RPF check

PIM-NG Domain-Set is similar to BGP AS-Path sequence in a way that each C-MAPPER MUST add its PIM-NG domain number to the domain-set of any multicast source before advertising it to a peer C-MAPPER .so at the end Domain-Set of a multicast source is a string of domains a multicast source has passed to reach its current domain or better to say to reach the C-MAPPER inside the current domain. From this perspective Domain-Set can be called Domain-Path sequence.

Domain-set is what a PIM-NG C-MAPPER uses to RPF check a multicast source it receives from its peer in other domains. Unlike the RPF check method used in PIM-SM and MSDP which has some limitations like:

- o The network of multicast domains cannot have a transitory AS, due to the fact that for a Source Active message to pass the RPF check, the first AS in best path to the originator RP MUST BE equal to the MSDP peer. So if a customer's Autonomous System is connected to an ISP's Autonomous system, the RP inside the customer network must be MSDP peer with the RP inside the ISP network.
- o Filtering SA messages can be a hard task, because of the RPF check rule in a way that, you cannot have a transitory AS

between to autonomous systems. such a case can be seen when administrators does not wish to update all of the SAs or some of the SAs created inside their PIM-SM domain to the neighbor Autonomous System , but they need to update them to a MSDP-PEER in an Autonomous System which is (i.e.)2 AS away.

o . . .

The RPF check method used in PIM-NG provides the ability of having transitory Autonomous Systems. Examples of such designs are:

- o Different PIM-NG domains are connected by an ISP's backbone and the ISP is only needed to provide UNICAST-REACHABILITY between different PIM-NG domains or different Autonomous Systems.
- o an enterprise does not wish to update all of its multicast sources to the neighbor Autonomous System or PIM-NG multicast domain , but it needs all of its multicast sources to be advertised to an AS which can be reached through this AS. consider the bellow illustration in which (i.e.) AS1 needs to advertise its multicast sources to AS3 and AS4 but doesn't want AS2 to become aware of those multicast sources or at least part of the multicast sources :

AS1 (D1) ----AS2 (D2) ----AS3 (D3) ----AS4 (D4)

MSDP specifications suggest that such work cannot be done and there will be a RPF check failure. But in PIM-NG such a work can be done.

So if a source (S) is generating multicast traffic (G) in AS1 which is assigned a PIM-NG domain number of 1(D1) in the above illustration, the C-MAPPER inside D1 will advertise the (S,G) with the domain-set (1) to the peer C-MAPPER in D2 and likewise the C-MAPPER in D2 will advertise the (S,G) to its peer in D3 with domain-set of (1,2) and at the end the C-MAPPER inside D4 will receive the (S,G) with domain-set (1,2,3).

The bellow rules apply to the DOMAIN-SET:

- o Each C-MAPPER or PPER MUST add its domain number to the domain-set of a multicast source before advertising it to its peers in other domains.
- o If inside a domain , more than one C-MAPPER is considered, each C-MAPPER MUST add its domain number to the domain-set of a multicast source before advertising it to its peers inside the

As shown in 0 a multicast source (s) in AS1 with PIM-NG domain number 1(D1) is generating multicast traffic destined to (G) and the C-MAPPER in the domain is advertising the (S, G) to its peers in D2 and D3. C-MAPPERS in D2 and D3 receive the update. C-MAPPER in D3 advertises the (S, G) to its peer in D2 with the domain-set (D1, D3). C-MAPPER in D2 receives 2 updates for (S,G) with different domain-sets , one is (D1) and the other is (D1,D3) . So at this point the update received from D1 passes the RPF check and the update received from R3 fails the RPF check.

- o If there is a tie in the domain-set of a multicast source received from peers in 2 different domains , the update received from the sender in the best path towards the originator according to the unicast routing information MUST pass the RPF check. In the above example if D3 is connected to D4 by making a C-MAPPER from each domain peer with the other one, then C-MAPPER in D4 will receive the update about (S, G) with equal domain sets of (D1, D2) and (D1, D3), which in this case the update received from the C-MAPPER in D2 which is in the best path towards the originator C-MAPPER in D1 passes the RPF check.
- o If inside a domain more than one C-MAPPER exists and the MESH-GROUP concept is not used, the above rules MUST be followed.

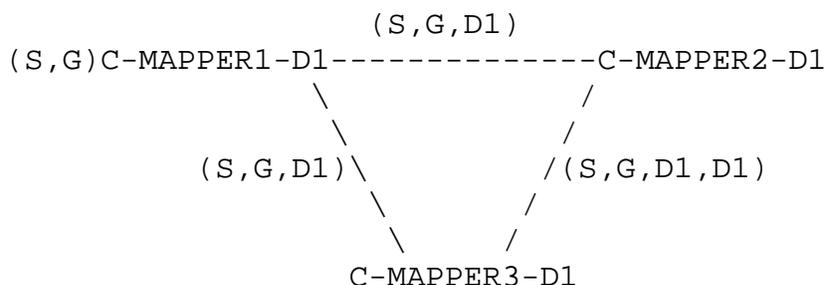


Figure 44 Intra-Domain connectivity and RPF check

As illustrated in 0, in which C-MAPPER1 is advertising an update about a newly originated source (S,G) with domain-set (D1) to its peers C-MAPPER2 and C-MAPPER3. C-MAPPERS 2 and 3 receive the update and start advertising the update to each other , so at this point and after the advertisement C-MAPPER2(i.e.) receives

another update for (S,G) from C-MAPPER3 with domain-set (D1,D1) and since the domain sequence of the second received update from C-MAPPER3 is longer than the one received from C-MAPPER1 , it fails the RPF check.

- o The above rules are applicable in different situations regarding RPF check, and bring so much flexibility to work with multicast sources and the updates regarding those sources.
- o If a C-MAPPER inside a CORE-DOMAIN receives an update about a multicast source from a peer C-MAPPER in a CORE-DOMAIN and ALSO from a peer C-MAPPER in a PIM-NG-DOMAIN , the update received from the peer in CORE-DOMAIN MUST pass the RPF check without considering the DOMAIN-SET.

4.6.2. Inter-domain connectivity concepts

PIM-NG Inter-domain connectivity is mainly focused on the processes involved in exchanging the information regarding multicast sources being originated in different multicast domains, as for now like PIM-SM, PIM-NG uses the underlying unicast routing protocol information in order to send join/prune messages from a client in need of receiving a multicast traffic to the final destination which can be a server generating the multicast traffic.

Although, PIM-NG specification processes and features are designed in a way that makes it possible to use a multicast routing protocol unique to PIM-NG in order to send join/prune messages from a receiver client to the originator of a multicast traffic, but since it is out of the scope of this document ,we are going to only explain the processes related to connecting 2 multicast domains and exchange the 'A-MULTICAST MAPPING TABLE'(s) between them ,which shall be enough to find a source and through using the underlying unicast routing protocols information reach the source.

The related concepts are going to be explained in 2 different sections, which are Public Domain connectivity and Private Domain Connectivity.

4.6.2.1. Public Domains

A public domain is considered by PIM-NG specifications to be either a domain in which C-MAPPERS and sources use public IP addresses or a domain divided in to 2 domains or sub-domains in which the unicast address of C-MAPPERS and existing multicast sources are known to the

population, and there is no need to use a PPER and domains are separated by PER'(s) which act as the domain boundary.

Concepts are explained in 2 different sections called inter-domain concepts and intra-domain concepts. Inter-domain section is focused of connectivity of 2 different PIM-NG domains in terms of exchanging information regarding originated multicast sources in each domain and the process in which a CLIENT sends join/prune message to a source. And intra-domain section which is the following section is focused mainly on the communication between any existing TR with C-MAPPER and finally clients, which is actually the only feature of PIM-NG specifications that hasn't been described yet.

4.6.2.1.1. Intra-domain concepts

As said before, TR is responsible of sending join messages on behalf of CLIENTs inside the domain to a source inside the domain or other domains.

If CORE-DOMAIN implementations are considered and CORE-DOMAINS do exist, and the domain-set of a multicast source indicates that the source of the multicast traffic can be reached by passing a PIM-NG-CORE-DOMAIN then the client creating the join/prune message MUST set the C-BIT of source unicast address and put the address of the CORE-DOMAIN-TR in the appropriate field and then send the message to the next hop.

But first a PIM-NG-AWARE-ROUTER inside each domain must be chosen to act as the TR. and as mentioned before; each existing C-MAPPER can take the role of a TR to, which eliminates the need for a separate router to be used.

Commands such as the commands bellow are initiated on the chosen router:

```
<#IP PIM-NG TR>
```

```
<#IP PIM-NG SOURCE INTERFACE LOOPBACK X>
```

```
<#IP PIM-NG DOMAIN [X]>
```

```
<#IP PIM-NG INTERFACE X,Y,Z>
```

The above command lines will tell the router that it is the TR of PIM-NG domain X ,and it should use its interface loopback defined in the command as the unicast address it is going to introduce itself with it to the C-MAPPER.

Although it may seem useful to have backup TR in the domain, we advise the use of a new TR instead of backup TR for bringing both redundancy and high availability at the same time to the domain. So no back-up TR is considered .the only case which there might be a back up TR implementation is when a C-RP has the role of TR and a SC-RP is considered.

In case C-RP'(s) is going to have the role of TR too, and the concept of ANYCAST-RP is in use in the domain, the following rules MUST be obeyed:

- o Since ANYCAST-RP concept causes the clients to contact the closest RP unicast address according to their unicast routing tables, if C-RP is going to act as a TR, then ALL-C-RPs in a multicast domain MUST be configured to act as a TR too.
- o The active C-MAPPER MUST become aware of the situation which can be done by a command line initiated on ALL-C-MAPPERS. So the active C-MAPPER introduces only one TR unicast address which will be the ANYCAST address used for C-RPs.
- o Each TR in this particular design MUST use 2 different addresses. One for the ANYCAST address and one for introducing itself to the closest C-MAPPER. The addresses MUST be the addresses used for C-RP processes which had been explained before.
- o C-MAPPERS then MUST introduce ALL the other TRs to each TR they find, inside the PDTT sent with the unicast-encapsulated introduction messages they send to each TR. This MUST be done so that TRs can send their JOIN-GROUP TABLE to each other.

After the TR is configured it will do the following processes:

- 1- It waits to hear the introduction message of the C-MAPPER sent to 239.0.1.190 to learn its unicast address.
- 2- As soon as TR receives the C-MAPPER introduction it will update its internal PDTT with the address of C-MAPPER and other information included in the table.
- 3- TR starts to send unicast introduction messages to the unicast address of C-MAPPER, which includes its unicast address.
- 4- TR will send these introductions as keep-alive messages every 30sec to the C-MAPPER.

- 5- If it's joined to any multicast group (G), and by joining we mean receiving traffic for that group, it MUST put an entry for that group in its JOINED-GROUP-TABLE for further use.
- 6- If in the introduction received from the C-MAPPER it receives the address of other TRs, it MUST send its JOINED-GROUP-TABLE to other TRs by sending unicast TR introduction messages to the address of those TRs. This MUST be done so that if a new client in another side of the domain sends a join for the same group to another TR, the TR will only send a join for that group to the TR which is already receiving the traffic for the group (G).
- 7- The introduction messages between different existing TRs are only sent at the time, there is a change in the JOINED-GROUP-TABLE by setting the Z-BIT. This is due to the fact that TR'(s) communicate with C-MAPPER and this communication and periodical introductions acts as a keep-alive so there will be no need for the TRs to send periodical introductions as Keep-Alive to each other.

After the C-MAPPER receives an introduction from the TR it will:

- 1- Update its PIM domain topology table, and send an introduction message to ALL-PIM-NG-CLIENTS for further use.
- 2- C-MAPPER maintains connectivity with the TR by sending unicast acknowledgement every 30 second back to the TR in response to TR introduction.
- 3- If ANYCAST concept is needed to be implemented in a domain with more than one TR , then the C-MAPPER MUST know the address used as the ANYCAST address and only send that address to ALL-PIM-NG-CLIENTS.

4.6.2.1.2. Inter-domain concepts

In a Public PIM-NG Multicast domain a C-MAPPER can directly become peer with a C-MAPPER in another domain. And ALL Multicast Sources use public IP addresses. The processes regarding to peering the C-MAPPERS are much like the processes related to peering C-MAPPERS inside the same domain. The only difference between the concepts is:

- o A C-MAPPER MUST NOT send a multicast C-MAPPER introduction message to find its peer in another domain. Instead a C-MAPPER'(s) MUST send a unicast-encapsulated C-MAPPER introduction message to the unicast address of its peer in another domain.

0 which illustrates a multicast network with 2 public PIM-NG multicast domains will be used for further explanations.

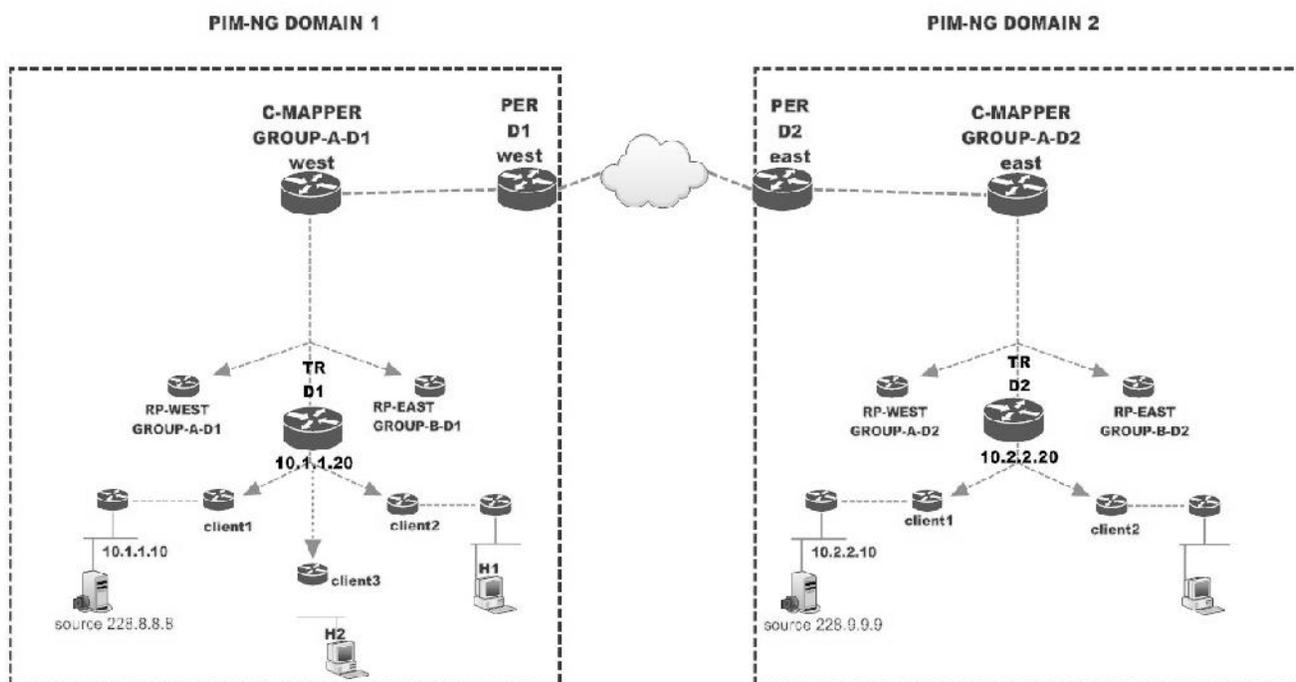


Figure 45 network with 2 multicast domains

For the sake of simplicity in explaining the processes lets accept the fact that , there is unicast reachability between the 2 domains by using either a dynamic routing protocol capable of carrying multicast traffic like MBGP or a tunnel between the 2 domains.

The processes to connect the 2 multicast domains are as follows:

- o In Each domain a C-MAPPER which uses a publically routable unicast address MUST be chosen, so makes it possible to directly make the C-MAPPERS peers.
- o C-MAPPER-WEST (WEST from now one) becomes peer with C-MAPPER-EAST (EAST from now one) using the unicast address of EAST to send the introduction message. the same is done on east:

<#IP PIM-NG PEER MAPPER DOMAIN 2 ADDRESS EAST>

And on east:

<#IP PIM-NG PEER MAPPER DOMAIN 1 ADDRESS WEST>

- o WEST starts sending unicast-encapsulated introduction messages to EAST every 30 seconds until it receives the unicast-encapsulated introduction message of EAST.
- o After receiving the first introduction message, both C-MAPPER'(s) MUST send their introduction messages periodically every 60 seconds to maintain connectivity with each other as keep-alive messages.
- o If any SC-MAPPER is considered in a domain, its address MUST be introduced to the peer in the introduction message.
- o If WEST needs to inform its peer which is EAST about a change in the contents of AMMT, it MUST set the ZTCN-BIT in the introduction message and send the AMMT.
- o WEST MUST add its domain number to the domain-set of any newly originated multicast sources before sending an update about it to EAST. The same MUST be done by EAST.
- o Each C-MAPPER will then send the received update from the peer to the C-RPs residing in the domain

Now bellow the processes related to communication between clients and sources are explained:

- o A host (H1) in domain1 (D1) wants to receive the multicast traffic destined for 228.9.9.9(i.e.) which is being generated in domain2 (D2).
- o H1 shows its desire to receive the traffic through IGMP messages it sends to the upstream router.
- o Client 2 becomes aware that a host inside the network behind it needs to receive the traffic destined for 228.9.9.9.
- o Client2 checks its PDTT to find the closest C-RP, and chooses C-RP-EAST(i.e.).since the RP and existing TR are not the same , H1 MUST send a unicast-encapsulated request for source message to C-RP.
- o Client2 sends a unicast-encapsulated PIM-NG request message to the C-RP. And receives a unicast acknowledge from the C-RP containing the unicast address of the source generating the multicast traffic in the format of (10.2.2.10, 228.9.9.9).

- o Client2 creates a join/prune message for (10.2.2.10,228.9.9.9) and since a TR exists in the domain and the domain-set of the source received from the C-RP doesn't contain any CORE-DOMAIN in the path , H1 sets the R-BIT of source unicast address and puts the address of the TR(10.1.1.20) in the appropriate field which shows that the join/prune message MUST first be sent towards the TR and sends the join/prune message towards the next hop or PIM-NG neighbor and joins the (S,G,RPT) rooted at TR which is considered to be the best path to reach the source.
- o TR receives the join/prune message and clears the R-BIT and its unicast address from the TR unicast address field, and forwards the message towards the next hop in the best path towards the source.
- o PER-EAST in domain2 (D2) receives a join/prune message on its external interface for multicast group 228.9.9.9 and must forward it to the next hop in the best path towards the source which is 10.2.2.10.
- o Since a TR exists in the domain PER-EAST MUST set the R-BIT again and put the address of TR(10.2.2.20) in the appropriate field and forward the join/prune message towards the source and actually join the (S,G,RPT) rooted at TR.
- o The join/prune message reaches the TR in D2, and TR again clears both the R-BIT and the TR unicast address field and forwards the message to the next hope.
- o Finally join/prune message for multicast destination 228.9.9.9 arrives at 10.2.2.10 and source becomes aware that a client somewhere is in need of receiving the traffic and it MUST join the Shortest path tree which is actually the current (S,G,RPT) rooted at the TRs in each domain. So at the end 10.2.2.10 starts forwarding the traffic.

Now a new host in D1 which is going to be H2 needs to receive the same traffic:

- o H2 shows its interest through IGMP to the upstream router which is client3.
- o Client3 goes through the same process as client2 to find the source unicast address of (G) and receives it from closest C-RP.
- o Client3 sends the join/prune message towards the TR.

- o TR receives a join/prune message for (S,G) which it already joined the shortest path tree for it , so without forwarding the join/prune message any further ,TR starts to forward the traffic for (10.2.2.10,228.9.9.9) down the new SPT towards client3.

4.6.2.2. Private domains

A private domain is considered by PIM-NG specifications to be a multicast domain in which all multicast sources or at least some of the multicast sources use private IP addresses.

In such a domain:

- o If the C-MAPPERS inside the domain use private IP addresses, then one or more PPER'(s) MUST act as the C-MAPPER to become peer with C-MAPPERS in other domains on behalf of the C-MAPPERS inside the domain. And also a mechanism MUST be used so that PPER'(s) use a united public IP address as the originator of the multicast sources originated inside the domain and also as the source unicast address of the sources using private IP addresses.
- o If C-MAPPER'(s) inside the domain use public IP addresses, then there is no need for PPER'(s) to have the role of C-MAPPER too. Although, there might be some multicast designs in which it is needed for the PPER'(s) to act as C-MAPPER too. In such domains with a C-MAPPER using public IP address, a mechanism MUST be used so that the PPER'(s) become aware of the unicast address that is used by C-MAPPER'(s) as the source unicast address of the multicast sources using private IP addresses.

In 0 an example network, including 2 private PIM-NG multicast domains is illustrated. In this example that is going to be used to explain the processes involved, Domain1 (D1 for simplicity) is a PIM-NG domain in which ALL multicast sources and the C-MAPPER use private IP addresses and Domain2 (D2 for simplicity) is a PIM-NG domain in which the C-MAPPER uses public IP address and is responsible to directly become peer with a C-MAPPER in D1 and a group of multicast sources use Public IP Addresses and a group use Private IP Addresses.

In D1 PPER-WEST with the unicast address 192.168.1.10 is going to act as the C-MAPPER and will become peer with C-MAPPER-EAST (EAST for simplicity) with the unicast address 192.168.2.10. And in D2 EAST will become peer with PPER-WEST.

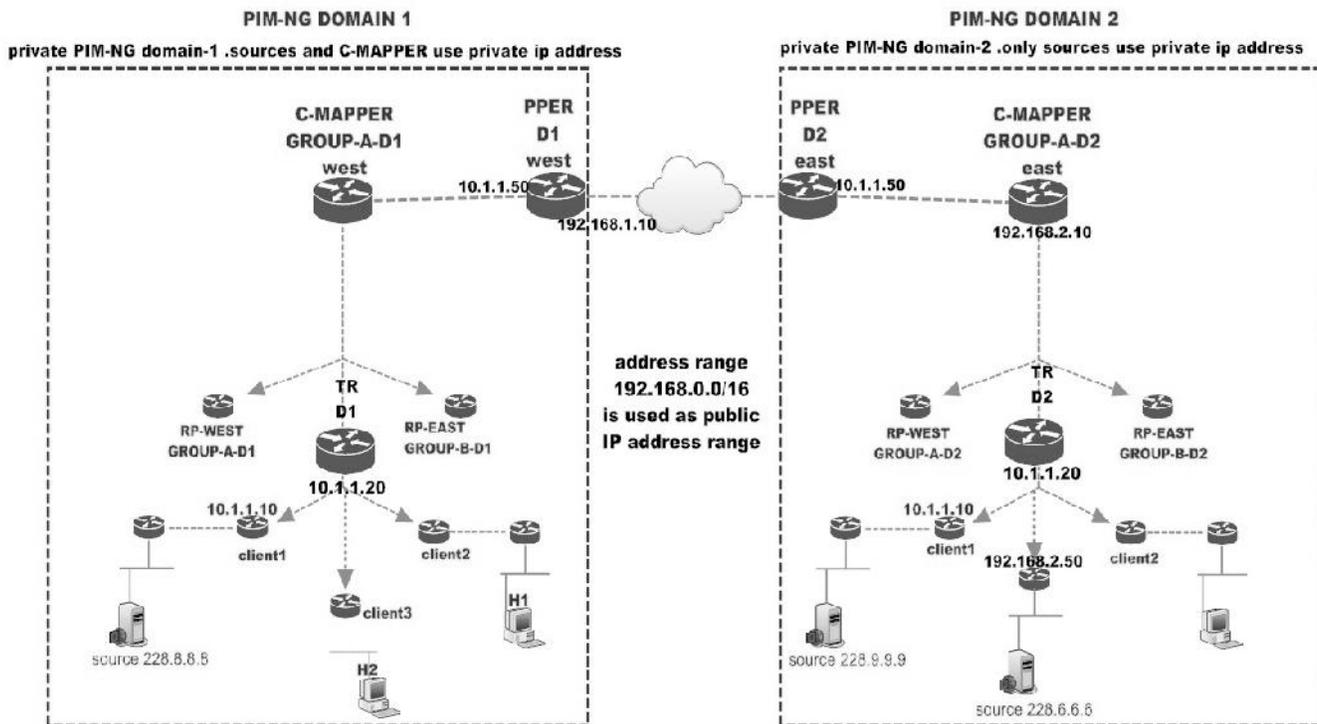


Figure 46 Multicast network with private PIM-NG domains

Different processes regarding the advertisement of multicast sources being generated in each domain to the other domain, and how the join/prune message is going to reach its final destination, are going to be explained in the following sections.

4.6.2.2.1. Intra-Domain processes

Since most of the processes and concepts regarding the connection between C-MAPPERS, clients, C-RP and TR in a PIM-NG domain have been explained up to this point ,processes referred to as Intra-domain processes will be related to the unexplained parts which are processes a C-MAPPER and the PPER'(s) are involved in within the Domain which in our example are D1 and D2 .Before starting to explain the processes a C-MAPPER and a PER or PPER are involved in ,we are going to take a look at an example of the steps related to make the PER ready and after that an example of how to make a PPER ready :

Bellow the processes related to preparing a PER is explained:

- o A PIM-NG-AWARE router is chosen to act as the PER.
- o a set of commands are initiated on the router:

```
<#IP PIM-NG EDGE>
```

```
<#IP PIM-NG DOMAIN [X]>
```

```
<#IP PIM-NG EDGE SOURCE INTERFACE [TYPE] [NUMBER]>
```

```
<# IP PIM-NG INTERFACE [type] [number] INTERNAL>
```

```
<#IP PIM-NG INTERFACE [TYPE] [NUMBER] EXTERNAL>
```

Or in case the PER is connected directly to 2 domains

```
<#IP PIM-NG DOMAIN [X] INTERFACE [TYPE] [NUMBER]>
```

```
<#IP PIM-NG DOMAIN[Y] INTERFACE [TYPE] [NUMBER]>
```

Or incase a PER is acting as a BPR:

```
<#IP PIM-NG DOMAIN [X] INTERFACE [TYPE] [NUMBER]>
```

```
<#IP PIM-SM INTERFACE [TYPE] [NUMBER]>
```

- o The above set of commands tells the router that it is a PER inside domain-X .and it MUST bring the interfaces mentioned in the command set in to the game of PIM-NG.
- o Interfaces marked as internal are going to be the ones connected to the internal domain and the one marked with external are going to be the ones connected to the outside networks.
- o In case PER is placed right between 2 domains or sub-domains it MUST have one or more interface inside one domain and one or more interfaces inside the other domain.
- o If a PER is placed between a PIM-NG domain and PIM-SM domain, It MUST have one or more interfaces inside PIM-NG domain and one or more interfaces in the PIM-SM domain. Such design may be seen in enterprise networks migrating from PIM-SM t PIM-NG over a period of time or in network designs where a PIM-SM domain is going to be connected to a network of PIM-NG domains or 2 or more PIM-NG domains are suppose to be connected through either an ISP's backbone which is still using PIM-SM or a transitory AS

using PIM-SM and the PIM-SM domain'(s)are to be used to forward the join/prune messages. Such a PER is called a BPR, and the PER MUST introduce itself to the closest C-MAPPER with the B-BIT in the introduction message being set. BPR concepts will be discussed in the appropriate section.

Bellow the processes related to preparing a PPER are explained:

- o A PIM-NG-AWARE router at the edge of the domain is chosen to play the role of PPER. PPER-WEST and PPER-EAST in 0.
- o The router MUST be configured and become aware of its role inside the domain and in case it is going to act as the C-MAPPER to become peer with other PPERs or C-MAPPERS in other domains. As before we are going to go through the steps by initiating some commands on the router to make the process of explanation easier:

```
<#IP PIM-NG EDGE-PRIVATE>
```

```
<# PIM-NG DOMAIN [X]>
```

```
<# PIM-NG INTERFACE [TYPE] [NUMBER] INTERNAL>
```

```
<# PIM-NG INTERFACE [TYPE] [NUMBER] EXTERNAL>
```

```
<# PIM-NG EDGE SOURCE INTERFACE [TYPE] [NUMBER]>
```

The above commands tells the router that it is PPER in a PIM-NG domain[X](D1/D2) .also the above commands dictates to the router that it has some internal interfaces which are connected to the inside network and some external interfaces which are connected to the outside world .

One other thing that the router must become aware of is the address it must use when introducing itself as the PPER to the C-MAPPERS inside the domain which in our example is going to be 10.1.1.50 for PPER-WEST and 10.2.2.50 for PPER-EAST. This address MUST be an address known to the domain in which the PPER resides. And it is advised to use a loopback interface as the source.

- o If the PPER is supposed to act as a C-MAPPER too (PPER-WEST in 0), it MUST become aware of the new role and also MUST know the address of its peers. one other thing that MUST be configured is the address that must be used as the C-MAPPER address which MUST be a public IP address :

<#IP PIM-NG MAPPER>

<# PIM-NG MAPPER SOURCE {INTERFACE [TYPE] [NUMBER]}|ADDR>

<# PIM-NG PEER MAPPER DOMAIN[Y] PEER-ADDRESS [A.B.C.D]>

Or if the peer PPER is a connected PIM-NG-AWARE router

<# PIM-NG PEER MAPPER DOMAIN[Y] INTERFACE [TYPE] [NUMBER]>

- o The address that is going to be used in the process of finding and communicating with a peer PPER or C-MAPPER, will be used as the Originator Unicast Address and source unicast address of the multicast sources that are originated inside the domain the PPER resides and use a private IP address. In Figure 44 the address used for PPER-WEST is 192.168.1.10 in D1 and the address used for the C-MAPPER-EAST in D2 is 192.168.2.10.
- o If the domain is connected to other domains by more than one PPER a mechanism MUST be used so that if PPERs has the role of C-MAPPER they all use one united public IP address as the Originator and source unicast address of the sources that are originated inside the domain and use private addresses. This is due to the fact that we don't want to end up advertising our sources with different source unicast addresses or originator addresses ,so PIM-NG specification suggests to configure on all PPERs an IP address from the public IP address range that is assigned to the AS or network. This can be done by a command like:

<#PIM-NG ORIGINATOR-ADDRESS [A.B.C.D]>

- o if a C-MAPPER inside the domain with public IP address is used to become peer with C-MAPPERS or PPERs in other domains and the PPERs doesn't have the role of C-MAPPER they MUST become aware of the IP address that the C-MAPPER is using as the originator and source unicast address of the multicast sources that are generated inside the domain and use private IP addresses. This MUST be done due to the fact that a PPER may receive a join/prune message for a multicast source inside the domain which its unicast address is the address of the C-MAPPER and will be routed towards the C-MAPPER and that MUST NOT be done. So we suggest dictating to PPERs that if they receive a join/prune message on an external interface with the address of the C-MAPPER as the source unicast address, it IS NOT meant to be forwarded towards the C-MAPPER and the address MUST be

changed with the real address of the multicast source. this can be done by a command like the one mentioned above or:

<#IP PIM-NG SOURCE-ADDRESS [A.B.C.D]>

In 0 this address is going to be 192.168.2.10 which is the address of C-MAPPER-EAST and MUST be set on PPER-EAST.

Processes related to PER will be discussed later when a PIM-NG domain needs to be connected to a PIM-SM domain. So in this section only PPER related processes are going to be explained.

- o As soon as PPER is configured and becomes aware of its role in the domain, it MUST communicate with the closest C-MAPPER according to the information received in the C-MAPPER introduction message and incase there are more than one C-MAPPER in domain based on the information it finds inside the PDTT, by sending an EDGE introduction message. PPER MUST set the P-BIT in EDGE introduction message which indicates to the C-MAPPER that it is sent from a PPER. PPER MUST send its unicast address inside the message so that C-MAPPER'(s) can use it to communicate with the PPER. So both PPERs in figure 44 start introducing themselves to the C-MAPPERS inside the domain.
- o Each PPER MUST send the introduction messages every 30 seconds to the C-MAPPER which acts as the keep-alive message.
- o Both C-MAPPERS in D1 and D2 MUST start sending A-MMT'(s) to the PPERs of their domain. The received information will be used by PPERs differently, due to the fact that PPER-WEST is also a C-MAPPER and PPER-EAST is only PPER.
- o Both PPERs MUST save the information regarding multicast sources that are originated inside their domain in IMST for further use.
- o PPER-WEST which has the role of a C-MAPPER too MUST send any received information regarding the multicast sources being generated in D1 which it receives from C-MAPPER-WEST to C-MAPPER-EAST by setting the Z-BIT in the introduction message and sending the updates inside the A-MMT. If only some of multicast sources MUST be advertised, it can be done through filtering process.
- o Since ALL multicast sources inside D1 are using private IP addresses, those multicast sources that are needed to be advertised to D2 MUST be chosen through an ACL, so that at the

time of updating to C-MAPPER-EAST, PPER-WEST changes the originator address and source unicast address of those sources with its own public unicast address which is 192.168.1.10 in 0.

This can be done by choosing the multicast source addresses needed to be advertised in an ACL and initiating a command line (i.e.) on PPER-WEST:

```
<#IP PIM-NG ACL X ORIGINATOR-ADDRESS [SELF| (A.B.C.D)]>
```

Which dictates to the PPER that it MUST change the originator address and source unicast address of multicast sources chosen by the ACL X(i.e.) with its own public IP address or incase there are more than one PPER with the address indicated by the command.

- o Since C-MAPPER-WEST is not peer with any C-MAPPER in other domains it MUST maintain connectivity with PPER'(s) of its domain in order to receive information regarding multicast sources in other domains and advertise multicast sources in its domain to other domains.
- o In D2, C-MAPPER-EAST is directly involved in the process of becoming peer with a C-MAPPER in another domain which is PPER-EAST, so it MUST only send the information regarding multicast sources being originated inside its domain to PPER-EAST, UNLESS it becomes aware that PPER-EAST is also BPR and connected to a PIM-SM domain, which in this case it MUST send the full A-MULTICAST MAPPING TABLE which may include multicast sources being originated in other domains.
- o Since C-MAPPER-EAST is going to advertise multicast sources originated inside the domain (D2) to its peer in D1, it MUST change the source unicast address of multicast sources originated inside D2 with private IP addresses, with its own public IP address which is 192.168.2.10 in Figure 44. It MUST be done based on an ACL or route-map like what had been done on PPER-WEST , since there are multicast sources in D2 with public source unicast address that their address must not be changed.
- o Since PPER-EAST is only acting as PPER it MUST become aware about the address the C-MAPPER is using as source unicast address when advertising multicast sources with private IP addresses to other domains, so that when it receives a join/prune message on its external interface with that source unicast address, it can act properly. This can be done by initiating a command line(i.e.) on PPER-EAST like:

<#IP PIM-NG SOURCE-ADDRESS [A.B.C.D]>

Instead of A.B.C.D, the public IP address being used by C-MAPPER in such process MUST BE used, which in our example is 192.168.2.10.

4.6.2.2.2. Inter-domain concepts

Inter-domain connectivity concepts are mostly focused on the process of exchanging information regarding multicast sources between the 2 domains and the process through which a CLIENT will send join/prune message for a multicast group inside another domain which is using a private IP address inside the destination domain.

First the processes regarding the exchange of A-MULTICAST MAPPING TABLES between the 2 domains of 0 will be explained and then to explain the process of sending the join/prune message and receiving multicast traffic, an example in which hosts in D1 are in need to receive multicast traffic which is generated in D2 will be explained.

1- Exchanging AMMT between the 2 domains :

- o Since PPER-WEST has the role of C-MAPPER, it will become peer with C-MAPPER-EAST on behalf of the C-MAPPER-WEST. And C-MAPPER-EAST will become peer with PPER-WEST which is seen by C-MAPPER-EAST as C-MAPPER-WEST or the C-MAPPER residing in D1.
- o Both PPER-WEST and C-MAPPER-EAST MUST use the public unicast address of the other one to send unicast-encapsulated introduction messages to both introduce themselves to each other and exchange the contents of their A-MULTICAST MAPPING TABLES which holds the information regarding multicast sources being originated in either of the domains. So PPER-WEST MUST use the unicast address 192.168.2.10 (0) which is the address of C-MAPPER-EAST, as its peer address. And C-MAPPER-EAST MUST use unicast address 192.168.1.10 (0) as its peer address.
- o As explained before both PPER-WEST and C-MAPPER-EAST MUST do some modifications on the information regarding multicast sources generated inside their domains with private IP addresses before advertising them to the other domain. So considering 0 , C-MAPPER-EAST MUST change the source unicast address of the source generating traffic for 228.9.9.9 which

is a private IP address with its own public address which is 192.168.2.10 , and send it to PPER-WEST.

- o The rest of the process of becoming peer C-MAPPERS is as before, in terms of sending unicast-encapsulated introduction message to the peer and maintaining connectivity with the peer by sending introduction messages every 60 seconds as Keep-Alive messages.
- o If any changes occur inside (i.e.) D1 regarding the multicast sources, PPER-WEST MUST first set the ZTCN-BIT and inform the change by sending its AMMT to C-MAPPER-EAST.
- o Also both PPER-WEST and C-MAPPER-EAST MUST add their domain number to the domain-set of any multicast source they are going to advertise to the other one.

2- Sending join/prune message for a multicast group inside a private PIM-NG domain:

- o Considering 0 as our example network, H1 and H2 in D1 need to receive multicast traffic destined for 228.9.9.9 and 228.6.6.6 which are being originated in D2.
- o So H1 starts the process by sending an IGMP message to the upstream router. And the request will reach CLIENT2.
- o So CLIENT2 starts the process of finding the unicast address of the source generating traffic destined for 228.9.9.9 by sending a unicast-encapsulated request to the closest C-RP.
- o As the source generating the traffic destined for 228.9.9.9 is using Private IP address within D2, C-MAPPER-EAST had changed the unicast address of the source with its own public IP address at the time of advertising to PPER-WEST. So C-RP will answer to the CLIENT2's request by sending back the unicast address of the source in the format of (S, G, domain-set) or (192.168.2.10, 228.9.9.9, 2).
- o Client 2 receives the C-RP acknowledge and sends a (S, G, RPT) join/prune message rooted at the TR to the next hop in the best path towards the TR.
- o TR receives the join/prune message. Clears both the R-BIT and its address from TR unicast address field and forwards the message to the next hop in the best path toward the source of multicast group 228.9.9.9 in D2.

- o PPER-EAST receives a join/prune message on its external interface with the SOURCE UNICAST ADDRESS of 192.168.2.10, and since PPER-EAST is told that this is the source address that is used for sources generating multicast traffic inside its domain (D2) which are using private IP addresses and not the real address of the source, it knows that some modifications MUST be done to the message before forwarding it to the next hop in the best path towards the existing TR in D2.
- o PPER-EAST MUST check the contents of IMST to find the real unicast address of the source generating traffic destined to 228.9.9.9, which is 10.2.2.10.
- o After finding the real unicast address of the source, PPER-EAST modifies the join/prune message by changing the source unicast address inside the message which is 192.168.2.10 with the real address which is 10.2.2.10.
- o Since a TR exists inside the domain PPER-EAST MUST set the R-BIT of source unicast address and put the address of TR in the appropriate field and forwards the message to the next hop in the best path towards the TR.
- o TR receives the join/prune message and joins the (S, G, RPT) tree and after clearing both R-BIT and TR unicast address, forwards the message to the next hop in the best path towards 10.2.2.10.
- o 10.2.2.10 receives the join/prune message and joins the (S, G, RPT) tree and starts forwarding the traffic destined for 228.9.9.9 down the tree towards the receiver which is client2 in D1.
- o Now H2 behind client3 in D1 shows interest in receiving multicast traffic destined for 228.6.6.6.
- o Client3 goes through the same process to receive the unicast address of the source generating traffic for 228.6.6.6. And after receiving the unicast address which is 192.168.2.50 in the format of (S, G, Domain-set) or (192.168.2.50, 228.6.6.6, 2), goes through the same process as client2 to send the join/prune message.
- o The join/prune message reaches PPER-EAST. Since the source unicast address is not changed, PPER-EAST MUST forward the

packet to the next hop in the best path towards the existing TR, without any modifications.

- o TR receives the join/prune message and after clearing both R-BIT and TR unicast address field sends the message to the next hop in the best path towards 192.168.2.50.
- o 192.168.2.50 receives the join/prune message, and joins the (S, G, RPT) tree and starts forwarding the multicast traffic destined for 228.6.6.6 down the (S, G, RPT) tree which is considered to be the SPT towards the receiver in D1 which is client3.

The above processes are involved in connecting 2 PIM-NG domains in order to both advertise multicast sources from one private domain to another and sending join/prune messages from one private domain to another domain. The related processes were explained through an example to both simplify the explanation process and understanding the need for PPER existence in PIM-NG specifications.

In the next sections the concepts regarding PIM-NG-CORE-DOMAIN through an example will be explained, and after that 2 new concepts called PIM-NG SUB-DOMAIN and PIM-NG STUB-DOMAIN will be explained and defined.

4.6.3. Core-Domain implementation

Up to this point of PIM-NG specifications process explanation, almost all of the related concepts have been explained. Bellow we are going to list them to review what has been explained:

- o Interaction between a source and C-RP and how a source registers.
- o Interaction between a Client and a C-RP to find a source for a multicast group (G).
- o Interaction between a Client and a source.
- o The processes by which the PIM-NG population find the C-RP'(s) in a multicast domain. And the interaction between C-RP and C-MAPPER.
- o Interaction between TR and C-MAPPER.

- o C-MAPPER and C-RP interaction to exchange the information regarding registered multicast sources in a domain.
- o Interaction between C-MAPPERS inside the same domain and different PIM-NG domains, regarding the exchange of information related to existing multicast source in each domain.
- o Processes involved in connecting multiple PIM-NG multicast domains.
- o . . .

One thing that has remained unexplained is related to a multicast network design in which PIM-NG-CORE-Domain with TR is considered.

As explained earlier:

- o ONLY when a PIM-NG-DOMAIN is connected to the outside network through a PIM-NG-CORE-DOMAIN
- o And ONLY when it is receiving the information regarding existing multicast sources in other domains through the Core-Domain, and Core-Domain is consisted of TR'(s).

a client MUST send, its join/prune message for a multicast source that can be reached through the Core-Domain cloud, first towards the TR inside to Core-Domain to join the (S,G,RPT) rooted at both any existing TR inside client's domain and then the TR inside the Core-Domain.

The above mechanism will eliminate future join/prune messages for the same multicast source which might be sent from Clients in other domains.

So with the above being said, a series of processes will be involved which are as follows:

- o A C-MAPPER'(s) inside a Core-Domain MUST introduce at least one TR to its peer C-MAPPER in connected PIM-NG-DOMAINS.
- o If a CORE-DOMAIN is consisted of more than one TR,PIM-NG specifications suggests the use of ANYCAST concept, so that the C-MAPPER'(s) inside the CORE-DOMAIN will only send one unified TR unicast address to peers in connected PIM-NG-DOMAINS, which reduces the amount of data being sent in C-MAPPER introduction message.

- o PIM-NG specifications suggests that the unicast address of the desired TR'(s) to be introduced to peers in PIM-NG-DOMAINS be dictated to C-MAPPER'(s). This process is actually a MUST as there might be TRs inside the CORE-DOMAIN cloud that are using private addresses or must not be introduced to connected PIM-NG-DOMAINS.
- o C-MAPPER'(s) inside the CORE-DOMAIN MUST introduce existing TR'(s) to peers in PIM-NG-DOMAINS, by sending the unicast address of the TR'(s) inside the CORE TOPOLOGY TABLE (CTT), which only contains the information regarding existing TR'(s) in the CORE-DOMAIN.
- o If a PIM-NG-DOMAIN is connected to a PIM-NG-CORE-DOMAIN through another PIM-NG-DOMAIN, the CTT MUST reach that domain too. For instance, if D1 is connected to D2 and D2 is connected to CORE-DOMAIN D10001, then C-MAPPER inside D2 MUST pass the received CTT to peer C-MAPPER in D1. This process is a MUST if D1 is receiving information regarding multicast sources in other domains through D2 and needs to reach outside sources.
- o A client in a PIM-NG-DOMAIN in need to receive a multicast traffic , that the associated DOMAIN-SET shows it can be reached through a CORE-DOMAIN , MUST use the information regarding the existing TR'(s) in CORE-DOMAIN at the time of creating join/prune message . This MUST be done by setting the C-BIT of source unicast address and putting the address of the CORE-DOMAIN-TR in the appropriate field, so that it will reach the TR inside the CORE-DOMAIN. The logic behind this process will be explained through an example multicast network.
- o After the join/prune message reaches the desired CORE-DOMAIN-TR, the TR'(s) MUST clear both the C-BIT and information inside the CORE DOMAIN Tree Root ADDRESS field and then forward the message to the next hop in the best path towards the source.

Bellow the above processes and the logic behind the above behavior of PIM-NG specifications at the presence of CORE-DOMAIN-TR will be explained through an example.

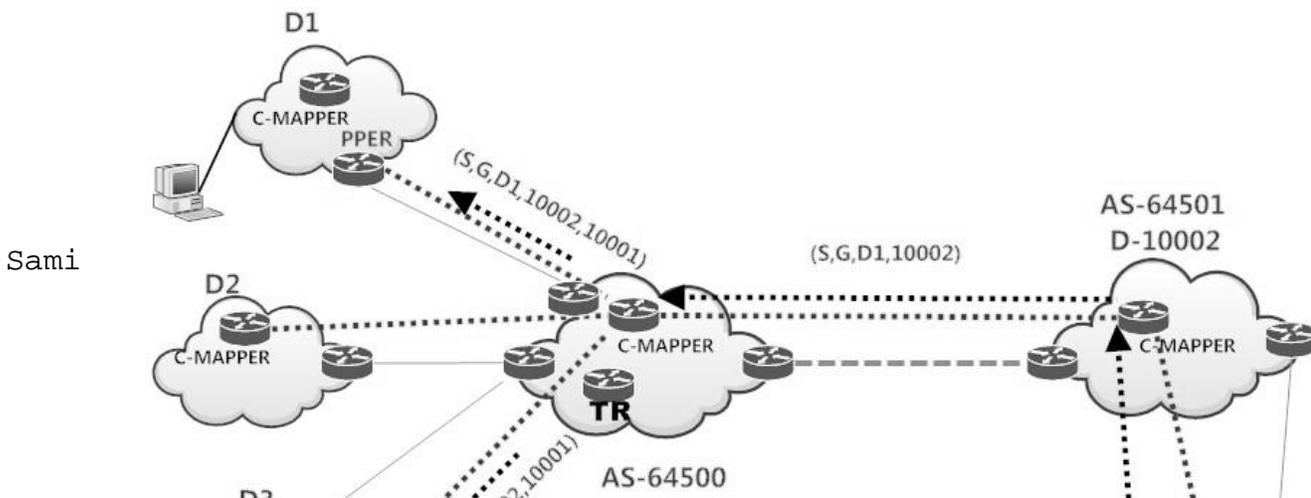


Figure 47 core domain implementation

In the multicast network design illustrated in 0 as an example we have 2 CORE-DOMAINS with domain numbers 10001 and 10002. D10001 is connected to D1, D2 and D3 and also connected to D10002, while D10002 is connected to D1.

A multicast source starts generating multicast traffic destined for (G) in D1 behind 10002 and the information regarding the source is passed from C-MAPPER in D1 to C-MAPPER in D10002 as (S, G) with domain-set (D1). (S, G) is advertised to D10001 by C-MAPPER in D10002 as (S, G) with domain-set (D1, 10002). And finally the information regarding the multicast source reaches D1, D2 and D3 behind D10001 as (S, G) with domain-set (D1, 10002, 10001).

As it is illustrated in 0, D3 behind D10001 has direct unicast connectivity to D1 behind D10002 through a tunnel which is capable of carrying multicast traffic.

As an example, a client/host in D3 shows interest in receiving multicast traffic destined for (G) and after receiving the unicast address of the source generating that traffic, it is going to create a join/prune message and send it towards the next hop in the best path towards the source. It is assumed in this example that the best path is the tunnel interface between D3 and D1. So if the join/prune

message is going to be directed through the tunnel, what happens if for instance another client in (i.e.) D1 behind D10001 comes up later and asks to receive the same traffic?. Let's assume that the join/prune from client in D3 is directed through the tunnel and both the source in D1 and client in D3 join the SPT for (S, G). Now a new client in D1 behind D10001 comes up and sends a join/prune message towards the source in D1 behind D10002, and finally both the source in D1 and client in D1 behind D10001 join SPT for (S, G). At this point we see that the source has joined 2 different SPTs to send the traffic to almost the same destination up to D10001, which is considered by PIM-NG specifications a waste of resources.

Instead of the above processes , since the domain-set associated with (S,G) indicates that the update regarding (S,G) passed through CORE-DOMAIN10001, the client in D3 MUST send its join/prune message in a way ,so that it will reach the CORE-DOMAIN ,by setting the C-BIT of the source unicast address filed in join/prune message and putting the address of the CORE-DOMAIN-TR it had received from the C-MAPPER in D3 in the appropriate field and then send the message to the next hop in the best path first towards any existing TR inside D3 and then after the message reaches the D3-TR and the R-BIT is cleared towards the CORE-DOMAIN-TR.

This way, as the message reaches the next hop and the next hope sees that the R-BIT is set in case of existing TR in D3, and also sees that the C-BIT is set which indicates that the message MUST be first forwarded towards the CORE-DOMAIN-TR.

the message reaches the CORE-DOMAIN-TR and the TR clears the C-BIT and forwards the message to the next hop in the best path towards the source .and finally the client in D3, the TR in D3, the TR in CORE-DOMAIN and the source join the (S, G, RPT) rooted at existing TRs in each domain which is considered the Shortest Path Tree by PIM-NG specifications, and source starts sending traffic to (G) down the SPT.

Now if the client in D1 behind D10001 comes up and sends the join/prune message to receive the same traffic, the join/prune message will only need to go up to the CORE-DOMAIN-TR, and since the CORE-DOMAIN-TR has already joined the SPT for (S, G) it will start forwarding traffic destined for (G) down the SPT towards the client in D1.

Although the above process may seem unnecessary, it MUST be done to reduce the amount of join/prune messages that might be sent towards a source from different domains behind a CORE-DOMAIN and ONLY MUST BE

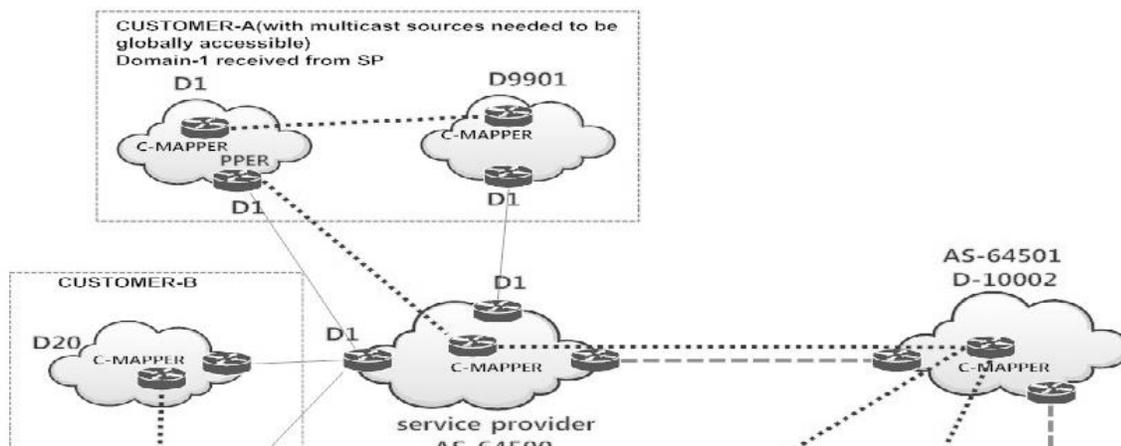


done whenever the domain-set associated with a multicast source indicates that it can be reached through a CORE-DOMAIN.

Figure 48 join/prune message being sent towards the source

4.6.4. Multiple multicast domains scenario

Up to this point of explaining PIM-NG specifications almost all of the concepts regarding registering a new multicast source, connecting 2 multicast domains, advertising the information regarding a new multicast source to another domain and many other concepts unique to PIM-NG have been explained. Now in this section we are going to check most of the concepts regarding a multicast network consisted of multiple PIM-NG domains through an example. In 0 a multicast network consisted of multiple PIM-NG-CORE-DOMAINS and PIM-NG-DOMAINS is illustrated.



Sami

143]

Figure 49 network with multiple multicast domains

In the illustrated network, you will see 5 PIM-NG-CORE-DOMAINS, 2 of which connected to some PIM-NG-DOMAINS. we have a PIM-NG-CORE-DOMAIN which is actually a SERVICE PROVIDER Autonomous System, a transitive multicast domain which is unique to PIM-NG as the nature of MSDP and its RPF check method doesn't allow PIM-SM to have a transitive multicast domain or Autonomous System, an AS which can be considered an enterprise network that divided its multicast domain in to multiple PIM-NG-DOMAINS which are all connected to outside world through the CORE-DOMAIN, and 2 other CORE-DOMAINS that are needed to make the explanation of RPF check concept and transitory AS/multicast domain easier.

In 0, the Service Provider network is shown as AS-64500 and DOMAIN-10001(D10001 from now on) which is a CORE-DOMAIN number, and is connected to its customer A with PIM-NG-DOMAIN number 1 which is assigned by the Service Provider to make the customer able of receiving multicast updates from outside and also send multicast

updates to the outside world. Customer A has divided its network to 2 domains using a private domain number for the second domain. Also the Service Provider network has another customer connected to it, which is named Customer B, which is using the SP's network to only connect its multicast domains and doesn't wish to advertise any multicast source to the outside world or receive any, and to be more specific it has a closed and private multicast domain. This D10001 is connected to another CORE-DOMAIN which is D10002 with AS number 64501.

D10002 is then connected to AS-64502 with DOMAIN number D-10003 which is assumed to be a transitive AS or multicast domain. D-10003 is only providing unicast reach ability so that the join/prune messages or multicast traffic can reach from one domain to other domains. Also please be noted that this transitive domain can be a PIM-SM domain which is only providing the unicast reach ability between PIM-NG domains. It is connected to AS-64503 with Domain number D-10004 and AS-64504 with Domain number D-10005. D-10004 is only being used to review the RPF check method used in PIM-NG and D-10005 is considered to be an organization's Autonomous System which has divided its multicast domain to multiple domains to have a better control over its multicast network.

With the above being explained , now it's time to overview the involved concepts through 0, which shows the network in 0 but with the multicast sources being advertised from one domain to other domains.

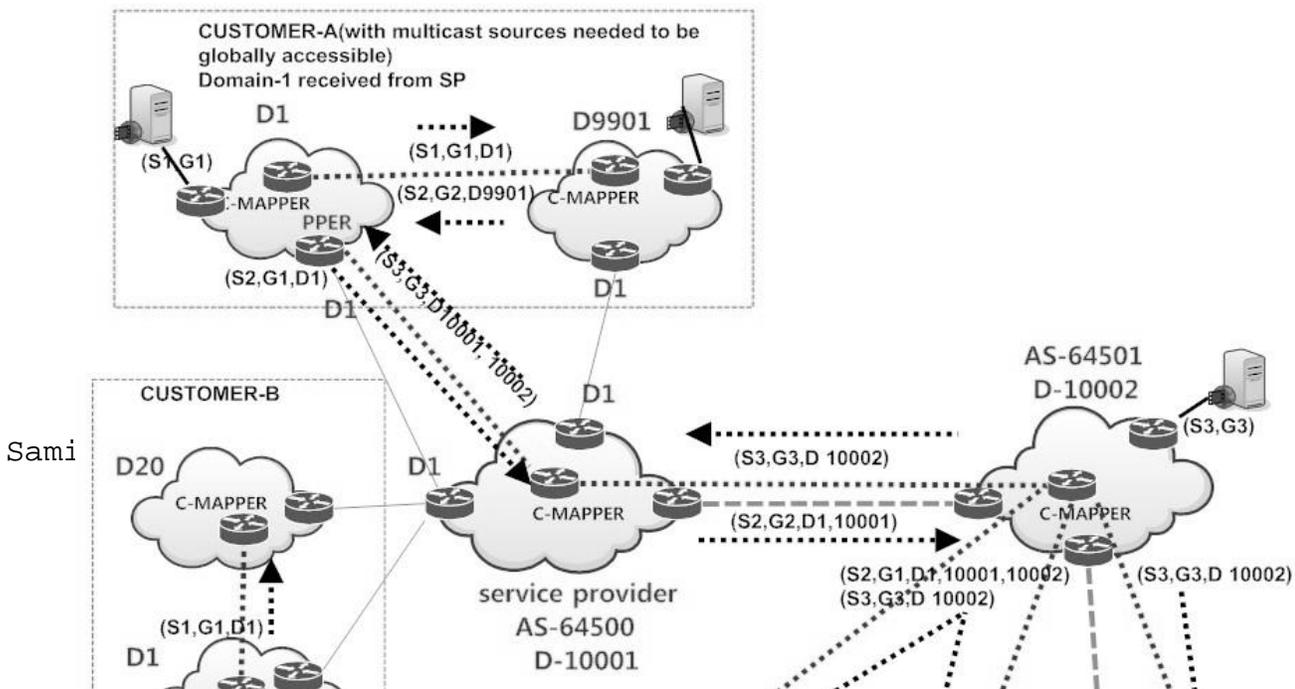


Figure 50 multicast network and multicast advertisements

As illustrated in 0:

- o Since customer-B's multicast domain behind D10001 is assumed to be a private and closed multicast domain, the multicast sources originated inside its D1 will only be advertised to D20 which is also customer-B's multicast domain, and will not be advertised to outside world or better to say to the C-MAPPER inside D10001. And customer-B is only using multicast network of D10001 as a mean to exchange multicast traffic between its multicast domains.

- o As it is illustrated in 0, customer-A's multicast domain is divided in to 2 multicast domains both connected to D10001 and since customer-A has some sources that needed to be globally accessible, it has received PIM-NG-DOMAIN number 1(D1) from D10001 (Service Provider). Customer-A uses a private domain number (D9901) for its second domain, and as it is illustrated the C-MAPPERS inside D1 and D9901 are directly connected as a mean to exchange information regarding multicast sources internally. And D1 is using a PPER to connect to a C-MAPPER in D10001 in order to exchange information regarding multicast sources globally/externally.
- o A multicast source inside D9901 is originating multicast traffic destined for (G), and the update about it is advertised to the peer C-MAPPER in D1 as (S2, G2, D9901). And since the domain number 9901 is a private domain number, the PPER in D1 will remove D9901 from the domain-set of (S2,G2) before advertising it to the C-MAPPER in D10001, and sends the update as (S2,G2,D1).
- o C-MAPPER in D10001 receives the update and advertises it to its peer in D10002 as (S2, G2, D1, 10001). The C-MAPPER in D10002 is also peer with C-MAPPERS in D10003, D10004 and D10005. As illustrated in 0, a multicast source (S3) comes up in D10002 and starts generating traffic destined for (G3). So the C-MAPPER in D10002 starts advertising information about this new source too its peers as (S3, G3, D10002).
- o D10003 is assumed to be a transitive domain only providing multicast reachability between D10002, 10004 and 10005. And as it is illustrated in 0, C-MAPPER in D10002 is only advertising information about multicast sources that are generated inside its domain and not those generated in D10001. This being said, C-MAPPER in D10002 advertises the multicast source created in its domain as (S3, G3, 10002) to C-MAPPER in D10003.
- o D10004 and D10005 need to receive the multicast updates completely, so the C-MAPPER in D10002 peers with C-MAPPER'(s) or PPER'(s) in D10004 and D10005 and starts sending updates about multicast sources it knows as (S2,G2,D1,D10001,D10002) and (S3,G3,D10002) to them.
- o Up to this point, C-MAPPER in D10004 and PPER in D10005 as illustrated in 0 have received the updates. Now the C-MAPPER in D10004 starts to advertise the received updates to its peer in D10005. So it advertises them as

(S2,G2,D1,D10001,D10002,D10004) and (S3,G3,D10002,D10004) to D10005.

- o PPER in D10005 receives the updates from its peer in D10004 and since it has received an update from its peer in D10002, about the same multicast sources, it starts the RPF check mechanism.
- o PPER in D10005, first, compares the DOMAIN-SET of the 2 received updates and the result of the comparison will be that, the updates received from D10004 will fail the RPF check. This is due to the fact that the domain set of the received updates from D10004 are longer than the ones received from D10002.
- o At the end the PPER in D10005 sends the updates to the closest C-MAPPER inside its domain and the C-MAPPER in D10005 advertises those updates to its peers in D1 and D2.

As explained above, PIM-NG specifications allow the existence of transitive multicast domain or Autonomous Systems, which due to the MSDP RPF check nature hasn't been totally implementable.

And also because of the unique RPF check method used by PIM-NG, the control over filtering the updates about multicast sources is easier and more enhanced. Which is also mostly because of the RPF check method used by PIM-NG which allows the existence of transitory domains or an intermediary multicast domain which only needs to receive partial updates, like D10003 in Figure 48.

In the following sections, the remaining concepts of PIM-NG regarding multiple multicast domain connectivity such as SUB-DOMIAN and PIM-SM-COMPATIBILITY, and controlling the updates received from a PIM-NG-DOMAIN to improve the security of the multicast network by STUB-DOMAIN concept will be explained.

4.6.5. PIM-NG Sub-Domain

PIM-NG Sub-Domain concept is introduced to cover 2 areas:

- o To control the propagation of multicast traffic inside one multicast domain, which from this perspective the Sub-Domain concept is more like dividing a multicast domain in to different areas.
- o To give the administrators a robust control over a multicast domain, by dividing it in to different sub-domains. And make modifications in one sub-domain in a way that no multicast

source updates are accepted from that part of the domain through using the concept of Stub-Domain, which will be discussed later, or do FILTERING on the multicast sources that are usable in a Sub-Domain.

Bellow specifications and rules apply to PIM-NG Sub-Domain:

- o A PIM-NG Domain can be divided to up to 254 Sub-domains. This number is derived from the maximum number of available C-MAPPERS inside a PIM-NG domain. As said before C-MAPPERS can have up to 255 different groups and the fact that each C-MAPPER is already a member of the main PIM-NG domain, gives us the equation $255-1=254$.
- o The sub-domain numbers MUST be different from the main domain number, so that, the C-MAPPER introduction messages meant for the main domain won't be read by the components of each Sub-Domain.
- o If a PIM-NG domain is divided in to Sub-Domains, all the C-RPs, TR'(s) and clients resided in a Sub-Domain MUST only be configured with Sub-Domain number to use it as their domain number.
- o The only components of a PIM-NG domain which MUST be aware of both the main PIM-NG domain number and the Sub-Domain number they reside in are C-MAPPERS, PER'(s), PPER'(s) and TR'(s).
- o Since the main purpose of dividing a PIM-NG Domain to Sub-Domains is to have a better control over the propagation of the information regarding the existing multicast sources, in a way that either a Sub-Domain doesn't receive the information regarding specific multicast sources or in a Sub-Domain no sources are allowed to do exist and send register messages to C-RPs in the Sub-Domain, any mechanism used such as filtering or defining a Sub-Domain as an STUB-Domain, MUST BE applied within Sub-Domains and at the time the information regarding multicast sources are advertised within the Sub-Domain to existing C-RP'(s).
- o PIM-NG specification suggests the use of one or more PER'(s) as the boundary between each 2 Sub-Domain to limit the propagation of multicast introductions sent by C-MAPPER'(s) and C-RP'(s) in each Sub-Domain.
- o If in a PIM-NG multicast domain, only one C-MAPPER is considered, then the multicast domain can be divided to ONLY 2

Sub-Domains and the C-MAPPER MUST act as the PER or boundary between the 2 Sub-Domains with one hand in each Sub-Domain.in this case the C-MAPPER MUST NOT forward the multicast introduction messages of one Sub-Domain to the other one.

- o If in PIM-NG SUB-DOMAINS, redundancy and high availability of the C-MAPPERS are needed to be considered, then each SUB-DOMAIN CAN use as many C-MAPPERS within the Sub-Domain to support the needs of multicasting in the Sub-Domain But ONLY 2 C-MAPPERS, MUST BE and ARE allowed to be both a member of the main domain and the Sub-Domain.and the rest of the C-MAPPERS will only be a member of the Sub-Domain. So if it is not needed, PIM-NG specifications STRONGLY suggests using 2 C-MAPPERS in each Sub-Domain.so if redundancy must be considered the number of Sub-Domains will reduce to 127.
- o In the case of redundant C-MAPPERS in each Sub-Domain, the Mesh-Group concept described in section 4.5.2 MUST BE used inside each Sub-Domain so only the active C-MAPPER will be able to send introduction messages destined to 239.0.1.190. so first C-MAPPERS MUST send their multicast introductions destined to 239.0.1.188 using the Sub-Domain number, to find their peers if the dynamic methods are in use and elect the active C-MAPPER within the Sub-Domain and after that they MUST send multicast introductions using their main domain number to find their peers in the main domain if the dynamic methods are in use.
- o As explained in section 4.5.2. , a PIM-NG domain CAN use up to 25 C-MAPPER Mesh-Groups. In this case each 10 C-MAPPER within either 10 or 5 Sub-Domains, CAN form a C-MAPPER Mesh-Group in the main domain. for instance, if a PIM-NG multicast domain with domain number D1, is divided in to 10 Sub-Domains(Sub-Domain 2-11), with each Sub-Domain including 2 C-MAPPERS for redundancy and high availability, C-MAPPERS inside Sub-Domains 2-6 can form Mesh-Group1 in D1 and C-MAPPERS in Sub-Domains 7-11 can form Mesh-Group2 in D1.
- o When a C-MAPPER is a member of the main domain and also a Sub-Domain, The concept of Active and Standby C-MAPPER in a mesh group described in section 4.5.2.1. , MUST BE applied to each Sub-Domain and depending on the needs of the PIM-NG multicast domain the Active and Standby C-MAPPER concept CAN be applied in the main domain. So for instance, if D1 is divided into 2 Sub-Domains (Sub-Domain 2-3) with 2 C-MAPPERS in each Sub-Domain being a member of both the PIM-NG domain and the Sub-Domain, C-MAPPERS in Sub-Domains2 will form Mesh-Group1 in

Sub-Domain2 and C-MAPPERS in Sub-Domain3 will form Mesh-Group1 in Sub-Domain3. And ALL 4 C-MAPPERS as C-MAPPERS of Domain1 will either form Mesh-Group1 in Domain1 to begin the process of Active C-MAPPER election in Domain1 between the 4 of them or simply will become peer C-MAPPERS in Domain1 to exchange their A-Multicast Mapping Tables and other information needed.

- o The Mesh-Groups in the main domain (i.e. D1) are formed to elect the ACTIVE C-MAPPER ONLY under these circumstances:
 1. If a PIM-NG domain which is divided to Sub-Domains, is connected to outside world or other PIM-NG domains or a PIM-SM domain by using the PPER concept, and since PPER'(s) MUST communicate with the closest C-MAPPER in the PIM-NG domain to receive the full AMMT which is being exchanged between the C-MAPPERS in the main domain, ONLY IF the needs of multicast Domain DICTATES TO USE the dynamic method for the communication of PPER'(s) with C-MAPPER'(s) in the Domain, then the C-MAPPERS MUST form Mesh-Group'(s) in the PIM-NG Domain and the active C-MAPPER MUST ONLY introduce existing C-MAPPERS in the main domain in its introduction messages sent to 239.0.1.190.
 2. If a PIM-NG domain which is divided to Sub-Domains, is connected to a PIM-SM domain by using the PER concept, and since in such design PER'(s) MUST communicate with the closest C-MAPPER in the PIM-NG domain to receive the full AMMT which is being exchanged between the C-MAPPERS in the main domain, ONLY IF the needs of multicast Domain DICTATES TO USE the dynamic method for the communication of PER'(s) with C-MAPPER'(s) in the Domain, then the C-MAPPERS MUST form Mesh-Group'(s) in the PIM-NG Domain and the active C-MAPPER MUST ONLY introduce existing C-MAPPERS in the main domain in its introduction messages sent to 239.0.1.190.
- o If Mesh-Group'(s) are formed in the main domain so that through an election the ACTIVE C-MAPPER starts introducing in the main domain, then ONLY the PER'(s) acting as the boundary between Sub-Domains MUST be dictated to pass the introduction messages sent to 239.0.1.190 which are generated ONLY BY the ACTIVE C-MAPPER in the main domain and NOT Sub-Domains.
- o If a PIM-NG Domain which is divided to Sub-Domains is connected to other Domains using PPER concept or is connected to a PIM-SM domain, To reduce the bandwidth and network resources consumption IF the needs of multicasting in the Domain does

not dictate to use the dynamic methods for the communication between PPER'(s) or PER'(s) and C-MAPPERS, PIM-NG specifications Strongly SUGGESTS to statically introduce C-MAPPER'(s) to each PPER or PER through command initiation on PPER'(s) and PER'(s).

- o If the C-MAPPERS are statically introduced to PER'(s) and PPER'(s), then a mechanism MUST be used so that the C-MAPPERS which are member of a Mesh-Group in the main domain, don't elect the Active C-MAPPER. This way the C-MAPPERS which are members of the Mesh-Group in the main domain will only exchange their AMMT'(s) and if needed the information regarding the existing TR'(s). and Since no Active C-MAPPER is elected the PER'(s) acting as boundary between the Sub-Domains MUST NOT be dictated to forward the C-MAPPER introductions sent to 239.0.1.190 which are generated by the ACTIVE in the main domain. So PIM-NG specification SUGGESTS doing above through command initiation as an option.
- o Also in case that the C-MAPPER'(s) are introduced to PER'(s) or PPER'(s) statically, PIM-NG specifications SUGGESTS that the C-MAPPERS in the main domain simply become peer and don't form Mesh-Groups, which eliminates the need to use the above mentioned mechanism.
- o With the above concepts, each C-MAPPER within each Mesh-Group in the main domain, will inform ALL the other members of the Mesh-Group as soon as any changes occurs within a Sub-Domain.
- o If in a PIM-NG Domain divided to Sub-Domains, the C-MAPPERS are to form Mesh-Group'(s) in the main domain through dynamic method explained in section 4.5.2. , the PER'(s) acting as the boundary between Sub-Domains MUST BE dictated to forward the C-MAPPER introduction messages destined to 239.0.1.188 and generated ONLY by C-MAPPERS in the main domain. so if Domain1 is divided to Sub-Domains 2 and 3, then the PER'(s) between Sub-Domain2 and Sub-Domain3 MUST BE dictated to forward the C-MAPPER introduction messages destined to 239.0.1.188 with Domain-Value set to1 (Domain1).
- o To reduce the bandwidth and network resources consumption in the above case, in which C-MAPPERS are to form Mesh-Group'(s) in the main domain, PIM-NG Specifications Suggests to peer the C-MAPPERS using the unicast Address of each C-MAPPER and not the GROUP number of C-MAPPERS which will eventually make the C-MAPPERS to send multicast introductions to 239.0.1.188 to find their peers. This being said, the dynamic method is still

the preferred method IF the network infrastructure and resources can support ALL the needs of PIM-NG multicast Domain.

- o If in the main domain, the C-MAPPER Mesh-Group concept is implemented with more than one Mesh-Group, then it is advised to ONLY peer 2 C-MAPPERS from each Mesh-Group with 2 C-MAPPERS from the other Mesh-Group'(s). The peering of C-MAPPERS inside different Mesh-Groups formed in the Main domain MUST BE done using the static method and there MUST be a boundary between the 2 Mesh-Groups. For instance, if D1 is divided in to 10 Sub-Domains (Sub-Domains 2-11) and C-MAPPERS in Sub-Domains2-6 are forming Mesh-Group1 and C-MAPPERS in Sub-Domains7-11 are forming Mesh-Group2, then 2 C-MAPPERS from Mesh-Group1 will become peer with 2 C-MAPPERS in Mesh-Group2 using the static method to bring redundancy to the entire domain. and it is STRONGLY advised by PIM-NG specifications to use a PER as the boundary between Sub-Domains 2-6 and Sun-Domains 7-11 so that no multicast introduction messages will be forwarded from the side including Sub-Domains 2-6 to the side including Sub-Domains 7-11. The above method of connecting different Mesh-Groups will become necessary in designs with more than 2 Sub-Domains and many C-MAPPERS in use to reduce the amount of introduction message traffic in the domain.
- o Each Sub-Domain MUST BE treated as a PIM-NG Domain, in which all the components of the domain including Clients, C-RP'(s), TR'(s) and C-MAPPER'(s) use the same domain or better to say Sub-Domain number with the C-MAPPER'(s) and TR'(s) as the only components being the member of the main domain and Sub-Domain. So clients, C-RPs and TR'(S) inside a Sub-Domain MUST only listen to C-MAPPER introduction messages with the same domain number.
- o Since each Sub-Domain MUST BE treated as a separate Domain, each Sub-Domain can have up to 255 C-RP and TRs if needed, and the all the PIM-NG specifications and rules that apply to exiting C-RP'(s) and TR'(s) in a single multicast domain, are applicable here.
- o C-MAPPERS in a Sub-Domain MUST become peer with other C-MAPPERS in other Sub-Domains using their main domain number, in order to advertise the information regarding multicast sources inside their Sub-Domain to other Sub-domains. So if (i.e.) if we have C-MAPPER-A and C-MAPPER-B inside D1, and C-MAPPER-A is a member of sub-Domain-2(SD2) and C-MAPPER-B is a member of

SD-3 , then C-MAPPER-A MUST become peer with C-MAPPER-B which is inside D1 and not D3.

- o As explained in section 4.5.2. ,when there are more than one C-MAPPER in a PIM-NG domain , each STC-MAPPER MUST send the information regarding newly found C-RPs or TR'(s) to the active C-MAPPER in the domain ,so that the active C-MAPPER will be able to introduce the C-RPs or TRs to the PIM-NG population by sending introduction messages to 239.0.1.190. but PIM-NG specifications dictates that in multicast design including Sub-Domain implementation ,C-MAPPERS MUST NOT send the information regarding the existing C-RPs in their Sub-Domain to other C-MAPPERS from different Sub-Domain'(s). And this is due to the fact that each Sub-Domain MUST BE treated as a separate Domain. so such information MUST only be exchanged within the Sub-Domain.
- o If TR'(s) are considered in a PIM-NG domain which is divided into Sub-Domains, since the communication between the TR and a client is limited to the join/prune messages, then C-MAPPERS in different Sub-Domains MUST send the information regarding TR'(s) they find inside their Sub-Domain to their peer C-MAPPERS which are resided in other Sub-Domains so that :
 1. Clients in other Sub-Domains become aware of the existence of TR and use its unicast address to send the join/prune message towards the closest TR.
 2. TR'(s) will find each other to communicate regarding the exchange of JOINED-GROUP TABLE.
- o In case of existing TR'(s) in a PIM-NG domain divided to Sub-Domains, PIM-NG specifications STRONGLY suggests :
 1. To Use a separate router as the TR if applicable
 2. To Use the ANYCAST concept to introduce the existing TR'(s) to the clients.
 3. NOT TO use C-RP'(s) as the TR if possible, due to the fact that in this case because of the method used by PIM-NG to send a join/prune message when C-RP and TR are both the same, there might be some issues if the C-RPs in one Sub-Domain go offline and after this the clients in this Sub-Domain may chose to send their join/prune messages toward an TR/C-RP resided in another Sub-Domain which doesn't have the

full A-MULTICAST MAPPING TABLE, because of multicast source filtering in that Sub-Domain.

4. That if the design needs dictates to use C-RP'(s) as TR'(s) in each Sub-Domain, the TR ANYCAST address used in each Sub-Domain be different from other Sub-Domains, and C-MAPPERS be dictated through command initiation as an option NOT TO send the information regarding the TR'(s) in their Sub-Domain to other C-MAPPERS which are resided in other Sub-Domains. This way each single Sub-Domain will totally act as a separate Domain.
- o If more than one TR exists in a PIM-NG Domain which is divided to Sub-Domains, and the information regarding the TRs is being exchanged among C-MAPPER so that the clients become aware of the existence of TR'(s) and also TR'(s) can communicate with each other using the dynamic method, then TR'(s) MUST use their main domain number when sending introductions containing JOINED-GROUPS TABLE to each other and not the SUB-Domain number, as the only thing that 2 TR CAN share and HAVE in common is the main domain number.
 - o TR'(s) MUST introduce themselves to the ACTIVE C-MAPPER in the Sub-Domain and not an ACTIVE C-MAPPER in the main domain.
 - o Each C-MAPPER MUST send the information regarding new multicast sources being originated inside its Sub-Domain to their peer C-MAPPER'(s) by exchanging the contents of A-MULTICAST MAPPING TABLE.
 - o C-MAPPER'(s) that are both a member of a Sub-Domain and the main domain, MUST remove the Sub-Domain number from the domain-set of multicast sources inside their Sub-Domain, before advertising them to their peer C-MAPPER(s) in other Sub-Domains or outside their multicast domain. After the deletion of Sub-Domain number, ALL the rules and specifications related to Domain-Set and RPF check will be applied. This is due to the fact that from other C-MAPPERS point of view, whether in the same PIM-NG domain (i.e. D1) or another PIM-NG domain (i.e. D2) the updates MUST be seen as generated in D1.
 - o If any FILTERING in regards to multicast sources needed to be done, so that a Sub-Domain does not receive the information regarding an specific multicast source or a group of multicast sources, it MUST ONLY BE done within each Sub-Domain in a way that the C-MAPPER'(s) within a Sub-Domain which ARE a member of the main domain do the filtering on the contents of A-

MULTICAST MAPPING TABLE at the time of advertising to C-MAPPERS that are only a member of the Sub-Domain or C-RP'(s) within the Sub-Domain.

- o C-MAPPERS in a PIM-NG Domain which is divided to Sub-Domains MUST exchange the full A-MULTICAST MAPPING TABLE between themselves without any filtering, and as said above any filtering MUST ONLY BE done within each Sub-Domain at the time of advertising to C-RPs within a Sub-Domain. This MUST be done so that any PER'(s) or PPER'(s) connecting the PIM-NG Domain to other domain'(s) or PIM-SM domain'(s) receive the full A-MULTICAST MAPPING TABLE.
- o If a PER is placed between 2 PIM-NG Sub-Domains to isolate them, then the PER MUST know the main domain number it is resided in. this is a MUST so that where ever the needs of multicasting dictates the PER will forward the C-MAPPER introduction messages which are meant to be heard by C-MAPPERS and PER or PPERs in the main domain.
- o If Sub-Domains are isolated from each other by using PER'(s) and the dynamic C-MAPPER peering methods are used among different Sub-Domains, then each PER MUST be dictated to let the C-MAPPER introduction messages destined to 239.0.1.188 and are meant to finding other C-MAPPER in the main domain ,to be forwarded from one Sub-domain to the other. As explained above, up to 2 C-MAPPERS from each Sub-Domain with total number of 10 C-MAPPERS are allowed to form C-MAPPER Mesh-Group in the main domain to propagate the information regarding multicast sources in each Sub-Domain, so it is vital that the C-MAPPER introduction messages meant for the main Domain are forwarded by the PER.
- o If the C-MAPPERS in existing Sub-Domain have formed more than one Mesh-Group in the main domain , PIM-NG specifications STRONGLY suggests that the PER'(s) between the 2 Mesh-Group areas DO NOT pass the C-MAPPER introduction messages sent to 239.0.1.188 from one Mesh-Group area to the other. This is needed to be taken in to consideration when the number of C-MAPPERS and mesh groups in a multicast domain increase, to limit the propagation of such traffic. So as explained above it is STRONGLY suggested to choose at least 2 C-MAPPERS from each Mesh-Group and make them peer with C-MAPPERS in other Mesh-Groups using the static methods. consider the bellow design in which Domain1(D1) is divided in to 4 Sub-Domains(SD2-5) :

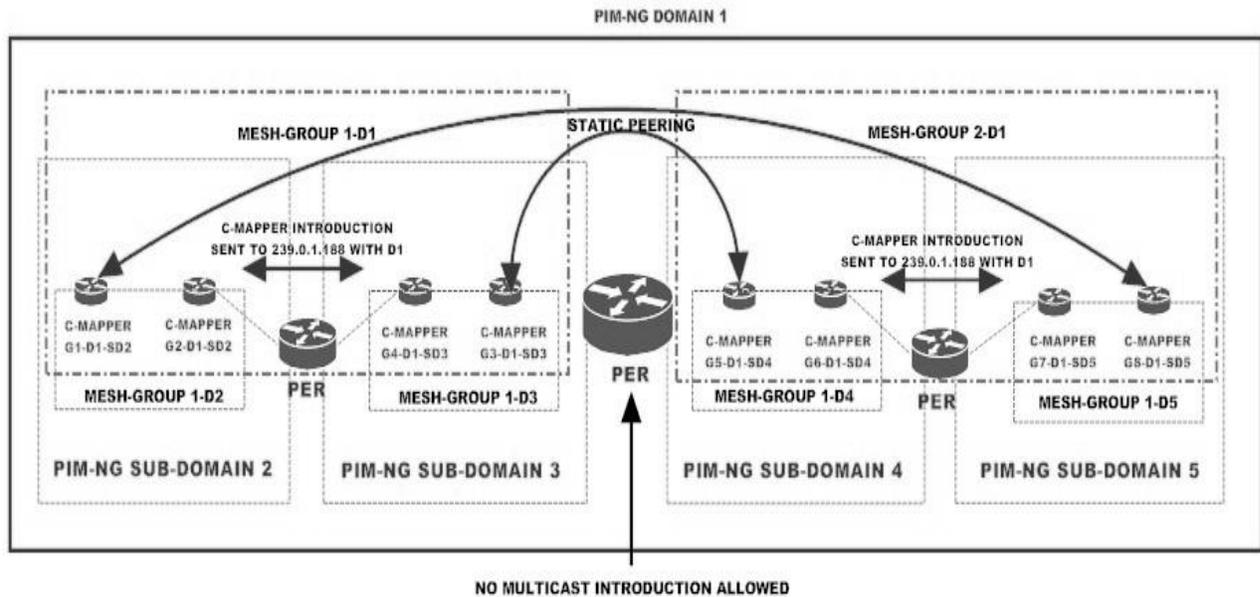
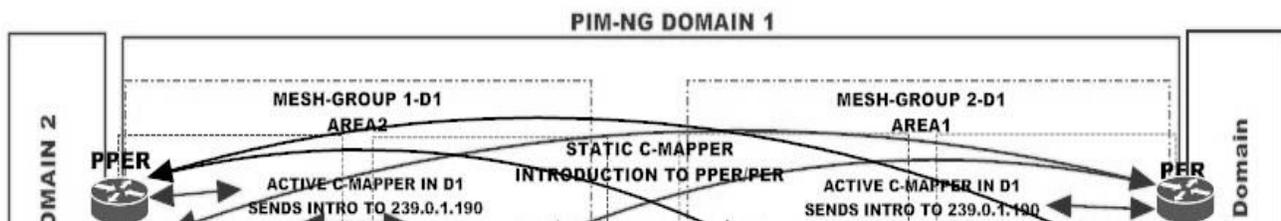


Figure 51 Mesh-Groups and C-MAPPER peering

- o If a PIM-NG domain divided to Sub-Domains is connected to a PIM-SM domain and the 2 domains are separated by a PER which is acting as the BORDER-PIM-ROUTER(BPR) too, then the PER MUST become aware of both main PIM-NG domain number and the PIM-NG Sub-Domain number it resides in, and introduce itself to the closest C-MAPPER in the main domain if the dynamic method is in use or to any C-MAPPER that is statically introduced to the PER using the main domain number and not the Sub-Domain number.
- o If a PIM-NG domain is divided to Sub-Domains, and the PIM-NG domain is connected to other PIM-NG domains using PPER'(s), then the PPER MUST know both the main domain number and the Sub-Domain number it is resided in, and MUST automatically communicate with the closest C-MAPPER in the main Domain and not the Sub-Domain, if the dynamic method is in use and the ACTIVE C-MAPPER within the main domain is chosen to introduce all existing C-MAPPERS in the PIM-NG domain by sending introductions destined to 239.0.1.190, or communicate with any C-MAPPER that is introduced to it in the main domain.
- o When a C-MAPPER in a PIM-NG Domain which is divided to Sub-Domains with multicast source filtering being applied to Sub-Domains, receives a PER or PPER introduction, it MUST exchange the full A-MULTICAST MAPPING TABLE with the PER or PPER so that the PER or PPER has a complete knowledge about existing

multicast sources. This is why such PER or PPER MUST introduce itself to the C-MAPPERS within the main domain and not the sub-domain. as explained earlier each Sub-Domain can have as many C-MAPPERS as needed but only 2 of them are allowed to be a member of both main domain and Sub-Domain, and if such PER or PPER introduces itself to the closest C-MAPPER within a Sub-Domain that multicast source filtering is applied in it, there is a possibility that the PER or PPER introduces itself to a C-MAPPER which is only a member of the Sub-Domain and doesn't have the full A-MULTICAST MAPPING TABLE.

- o If a PIM-NG domain is divided to Sub-Domains and the existing C-MAPPERS within each Sub-Domain are to form C-MAPPER Mesh-Group with other C-MAPPERS in other Sub-Domains in the main domain using the dynamic method used by PIM-NG specifications , the PERs acting as the boundary between Sub-Domains MUST forward the C-MAPPER introduction messages sent to 239.0.1.188 and 239.0.1.190(active C-MAPPER introduction). This is advised to be done through command initiation as an option.
- o If a PIM-NG domain which is divided to Sub-Domains is connected to other PIM-NG domains using PPER concept, and the C-MAPPERS in PIM-NG domain are forming Mesh-Groups then the PERs acting as the boundary between the Sub-Domains MUST pass the introduction message of ACTIVE C-MAPPER in the main domain sent to 239.0.1.190, so that the PPER connecting the PIM-NG domain to other PIM-NG domains will automatically learn the address of the C-MAPPERS and communicate with them.
- o If a PIM-NG domain which is divided to Sub-Domains is connected to a PIM-SM domain, and the C-MAPPERS in PIM-NG domain are forming Mesh-Groups then the PERs acting as the boundary between the Sub-Domains MUST pass the introduction message of ACTIVE C-MAPPER in the main domain sent to 239.0.1.190, so that the PER connecting the PIM-NG domain to PIM-SM domain will automatically learn the address of the C-MAPPERS and communicate with them.
- o If in a PIM-NG Domain which is divided to Sub-Domains and connected to outside Domains using PPER'(s) or PIM-SM domains using PER'(s) which act as BPR, more than one C-MAPPER Mesh-Group is considered , Since it is STRONGLY suggested by PIM-NG specifications to use a PER between the 2 Mesh-Group area, so that the ACTIVE C-MAPPER introductions of one area IS NOT forwarded to other areas, PIM-NG specifications SUGGESTS to statically introduce at least 2 C-MAPPERS from other Mesh-Groups that such PER or PPER is not resided in the associated area to the PER or PPER so that if the connection between the PER or



PPER with the C-MAPPERS inside the closer Mesh-Group area is lost, the PER or PPERs will be able to communicate with the C-MAPPERS in other areas.

Figure 52 PPER and PER communication with C-MAPPER

As explained above throughout different rules and specifications that apply to PIM-NG Sub-Domain'(s), the Sub-Domain concept is introduced to give administrators and designers the ability to divide the multicast domain to different areas and provide a robust control over the propagation or advertisement of the multicast sources within each area or better to say Sub-Domain or as a security feature treat a Sub-Domain in a way that no multicast source registration is allowed in an specific area or Sub-Domain by applying the STUB-DOMAIN concept to that Sub-Domain . In 0 a PIM-NG domain divided to Sub-Domains is illustrated.

Sam

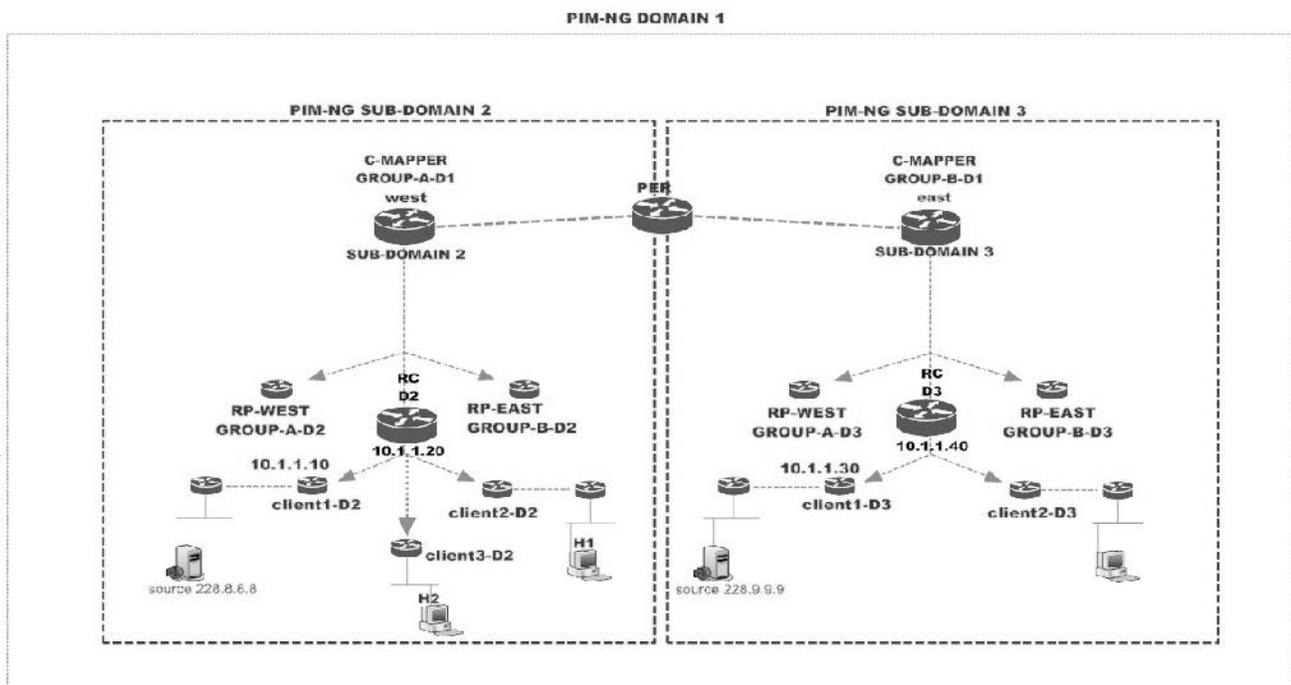


Figure 53 Sample PIM-NG domain Divided to
Sub-Domains

4.6.6. PIM-NG Stub-Domain

PIM-NG introduces the STUB-DOMAIN concept as a security feature when dealing with multiple PIM-NG Domains or PIM-NG Sub-Domains. Through implementing the STUB-DOMAIN concept, a PIM-NG Domain or Sub-Domain is treated as a multicast domain in which only receivers for multicast traffic exist.

The bellow specifications and rules apply to a PIM-NG STUB-DOMAIN:

- o A STUB-DOMAIN is a domain in which, no multicast source exists.
- o A STUB-DOMAIN is a domain in which, ONLY multicast traffic receivers MUST exist.
- o If a Domain is considered to be STUB-DOMAIN, the C-RP'(s) within the Domain MUST BE dictated NOT TO accept any source register messages. To do so PIM-NG specifications Suggests that, at the time of configuring the domain number in which a C-RP exists, the C-RP be dictated that the domain it is resided in is a STUB-DOMAIN.
- o If a Domain is considered to be a STUB-DOMAIN, the C-MAPPERS inside the Domain MUST BE dictated NOT TO accept any multicast source advertisement from existing C-RP'(s), or better to say the C-MAPPER'(s) inside a STUB-DOMAIN MUST NOT receive any AMMT from the C-RP'(s) inside the domain, and if any A-MULTICAST MAPPING TABLE is received from C-RP'(s) it MUST NOT be accepted.
- o If a PIM-NG Domain is divided to PIM-NG Sub-Domains, and one or more Sub-Domain'(s) MUST be treated as a STUB-DOMAIN, then ALL C-RP'(s) inside such Sub-Domain'(s) MUST become aware about the situation, so that they WILL NOT accept any source register messages.
- o If a PIM-NG Domain is divided to PIM-NG Sub-Domains, and one or more Sub-Domain'(s) MUST be treated as a STUB-DOMAIN, then C-MAPPER'(s) inside such Sub-Domain'(s) MUST NOT accept any AMMT from the C-RP'(s) inside the Sub-Domain.
- o If 2 separate PIM-NG multicast domains (i.e. D1 and D2) are connected and one of them (i.e. D1) is considered by the other Domain (i.e. D2) a STUB-DOMAIN, then the C-MAPPER'(s) or PPER'(s) in the Domain which is not a Stub-Domain (i.e. D2) MUST NOT accept any multicast source advertisements from the peer C-MAPPERS in the Domain that is considered a STUB-DOMAIN (i.e. D1). PIM-NG specifications suggests that this be done at the time of introducing the peer C-MAPPERS.for instance by initiating a command like this :

<#IP PIM-NG PEER MAPPER DOMAIN[value] MAPPER-ADDR[X.Y.Z.W] STUB>

- o No AMMT MUST BE received or accepted from a peer C-MAPPER which is considered to be in a STUB-DOMAIN.
- o The C-RP'(s) inside a STUB-DOMAIN MUST only receive AMMT from C-MAPPER'(s) in the domain and MUST NOT generate any AMMT.

- o Since as dictated by PIM-NG specifications NO advertisements regarding multicast sources can be accepted from a STUB-DOMAIN, In a multicast network design with a STUB-DOMAIN in the middle of 2 normal domains PIM-NG specifications suggest 2 different approaches:
 1. If possible, C-MAPPER'(s) from the normal Domains MUST become peer with each other and also the C-MAPPER'(s) in the STUB-DOMAIN, so that the C-MAPPER'(s) in the normal Domains can exchange information regarding multicast sources originated in their domains and only advertise them to the C-MAPPER'(s) in the middle domain. in this method the STUB-DOMAIN acts as a transitory multicast domain and actually is used by the normal domains as a path to send join/prune messages and receive the desired multicast traffic.
 2. If it is not possible to make the C-MAPPERS in normal domains peer with each other, PIM-NG specifications suggest to use a mechanism as an optional feature in regards to STUB-DOMAIN, so that the C-MAPPER'(s) in normal domains become peer with the C-MAPPER'(s) in the STUB-DOMAIN and accept receiving advertisements regarding multicast sources or better to say accept the received AMMT from C-MAPPER'(s) in the STUB-DOMAIN, BUT filter any information regarding multicast sources that are generated within the STUB-DOMAIN.
- o The above approach is unique to PIM-NG because of its unique RPF check method, which allows the existence of transitory multicast domains or Autonomous Systems. The above explanations can be seen in 0.
- o If a PIM-NG domain is connected to a PIM-SM domain, and the PIM-SM domain is considered to be a STUB-DOMAIN, PIM-NG specifications follows 2 different approaches:
 1. If the PIM-SM domain is between 2 PIM-NG domains, the C-MAPPER'(s) which is becoming MSDP-PEER with RP'(s) in the PIM-SM domain MUST BE dictated not to accept any Source Active messages from their MSDP-PEER'(s). And C-MAPPERS within the PIM-NG domains MUST become peer with each other.
 2. if the PIM-NG domain is connected to a network of PIM-SM domains and the directly connected PIM-SM domain must be considered a STUB-DOMAIN, the C-MAPPER'(s) in the

PIM-NG domain MUST be dictated as an optional feature to perform filtering on the received Source Active messages received from the MSDP-PEER in the STUB-DOMAIN, so that any source with the originator address equal to the address of the MSDP-PEER is filtered.

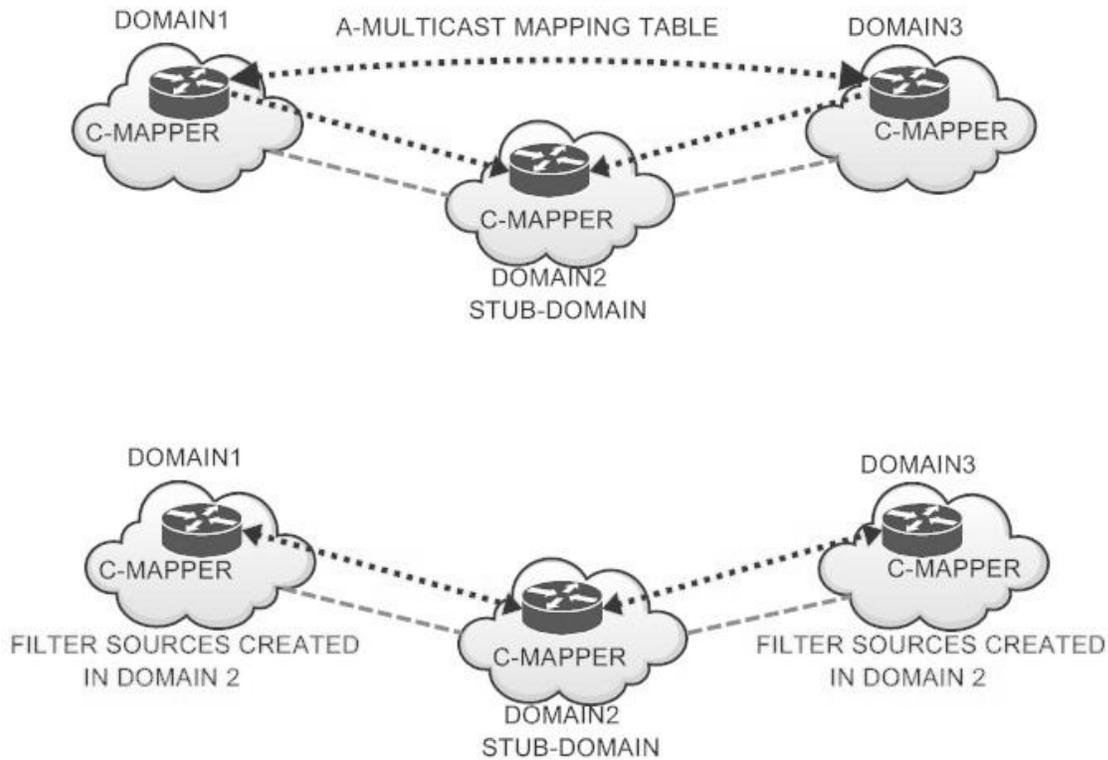


Figure 54 Stub-Domain as the transitory multicast domain

The above specifications and rules apply to a PIM-NG STUB-DOMAIN. and the reason that the Stub-Domain concept is explained as a part of the concepts related to PIM-NG and multiple multicast domains, is that this concept is only meaningful and applicable when we are dealing with multiple multicast domains, and the STUB-DOMAIN concept can be used as a security measure when dealing with multicast domains that no multicast source'(s) are expected to be advertised from them, to eliminate the possibility of existing attackers.

4.7. PIM-NG Bidirectional logic

One of PIM-NG's features that can make it by far a suitable candidate for scenarios with huge number of sources and receivers for the same multicast group and specially data centers is its Bidirectional PIM logic. Current implementation and concept cannot allow for the existence of redundant roots per bidirectional tree which is a desirable factor especially in data centers where we can face a huge number of receivers and sources for the same multicast group (G). Other than redundant roots being able to benefit from redundant trees per bidirectional group can be considered another desirable factor which currently needs to use other protocols in conjunction with PIM-SM.

PIM-NG can provide both redundant roots per bidirectional tree and redundant trees per multicast group by making use of some of the processes and specifications which has been explained up to this point alongside its unique Bidirectional logic.

In addition to the above PIM-NG introduces a new method of Bidirectional multicast group discovery unique to PIM-NG called Bidirectional Group Auto Sense mechanism which allows a multicast domain to automatically sense the existence of Bidirectional Groups and change the logic of the domain for those groups to Bidirectional.

In the following sections different concepts and specifications of PIM-NG bidirectional logic will be discussed.

4.7.1. Requirements

- For Bidirectional PIM-NG logic to be implemented in a well organized manner at least 1 TR MUST do exist in a PIM-NG multicast domain.
- All PIM-NG aware routers MUST be Bidirectional PIM aware too by default.

4.7.2. Bidirectional Multicast Group Discovery

Before getting involved in how Bidirectional trees are formed, we need to understand how Bidirectional Multicast Groups within a Domain are discovered so we will be able to deliver the desired traffic destined for (G) from sources to receivers.

PIM-NG uses 2 different methods for Bidirectional source discovery which are called Manual mode and autosense mode. No matter what type is used it is mandatory that a PIM-NG domain in which the bidirectional logic is going to be used have 1 or more TR'(s). That being said at least 1 TR MUST exist in such domains.

4.7.2.1. Manual Mode

Bidirectional Multicast source discovery manual mode specifications are as follows:

- 1- In this mode The Bidirectional Multicast Group'(s) are defined and configured manually on the existing C-MAPPER or Active C-MAPPER if more than 1 C-MAPPER is considered.
- 2- For this mode to be activated the Active C-MAPPER MUST be configured appropriately by the administrator and the C-MAPPER MUST inform all the population by setting the B-BIT in its introduction messages sent to 239.0.1.190.
- 3- After the Groups with Bidirectional logic are defined and configured on the C-MAPPER, the C-MAPPER MUST notify the entire population of PIM-NG aware routers about the existence of such Groups so that the domain's logic will change to Bidirectional for those group'(s). This is done by the C-MAPPER through sending a special table called "Bidirectional Groups Table (BGT)" in its introduction messages sent to 239.0.1.190.

```

+-----+
|Bidirectional Group| TR Group |R-flag |
+-----+
|                   |          |       |
+-----+
|                   |          |       |
+-----+

```

Figure 55 Bidirectional Groups Table (BGT)

- R-Flag: set by a C-RP or C-MAPPER to inform the C-MAPPER to remove a multicast group from its BGT table. If this flag is set it means that the group is not active anymore.

0 shows the format of BGT in which C-MAPPER announces the Multicast Groups for which the logic of the Domain MUST change to Bidirectional. Also the C-MAPPER announces a TR group which defaults to 0 and its use will be discussed further when there are lots of TR'(s) in a multicast domain and the needs of multicasting dictates to use a handful of TRs for different sets of

Bidirectional Multicast groups as the ROOT. So for the sake of simplicity it assumed that there are not so many TRs.

- 4- After the C-MAPPER sends BGT in its introduction messages all the PIM-NG aware routers will know about any multicast group that must be treated as a Bidirectional Group and sources and receivers will immediately join the SPT rooted at the closest TR for that group'(s) without registering with C-RP'(s).
- 5- In scenarios with SUB-Domains implemented it is possible to filter any desired Bidirectional Multicast group from being updated to the Active C-MAPPER within each Sub-Domain by the Active C-MAPPER in the main domain.
- 6- BGT MUST BE exchanged between neighbor clients when a new client is added at the time of synchronization.
- 7- If a Group (G) is removed the C-MAPPER MUST announces this event by sending the BGT and setting the R-Flag for that (G).
- 8- In scenarios with Core Domain implementations, in which 1 or more Domains are connected to a Core domain and senders and receivers are distributed in different domains, Bidirectional Groups MUST be configured and defined on an Active C-MAPPER within the Core Domain and then be updated to C-MAPPERS inside connected Domains.
- 9- When Bidirectional groups are defined on C-MAPPERS within a core domain, PIM-NG strongly advises to update it based on domains in which any sender or receiver of Group G exists. to make it more clear it must be said that it is suggested to use a mechanism through which it will be possible to send the BGT to desired C-MAPPER neighbors within connected domains or if needed to all the C-MAPPERS. This way senders and receivers within any desired domain can join the Bidirectional tree for group G within their domain and eventually the tree in the core domain.
- 10-If 2 PIM-NG multicast domains are going to communicate using Bidirectional PIM, then at least one C-MAPPER in any one of the domains MUST be configured with appropriate Bidirectional information and pass the information to the other C-MAPPER by exchanging BGT.
- 11-When a Bidirectional Group is removed from the BGT of a C-MAPPER the C-MAPPER MUST notify all the population of PIM-NG aware routers within its domain and in case of domains connected to a core domain all the C-MAPPERS within each domain or desired domain

about the change by sending a BGT with the R-Flag of the removed Multicast Group being set to 1.

This process is much like what is happening in current implementations except for the fact that it allows for the existence of many TR's within a domain and thus redundant roots per Bidirectional group, which will be discussed later.

4.7.2.2. Autosense mechanism

PIM-NG allows a multicast domain to automatically sense the existence of Bidirectional groups and change the logic of the Domain for such groups. This is accomplished due to the fact that the Source registry mechanism and the processes through which a client finds the source of a multicast address are performed in a complete different and innovative way than current implementations.

As explained in chapter 4.2. PIM-NG uses a new mechanism for registering the source and the way a client or end host finds its desired multicast source. Involved mechanisms like saving the information of active sources in MMT by C-RP'(s) and also the unicast nature of Source discovery messages sent from clients to C-RP'(s) provides the ability of keep track of existing multicast sources and receivers or better to say how many receivers and sources do exist for the same multicast group.

By definition a Bidirectional Multicast Group is a group for which lots of simultaneous senders and receivers do exist or better to say it is a group for which a sender is also a receiver.

PIM-NG Bidirectional specifications with regards to Autosense are as follows

- 1- Autosense is the default mode of PIM-NG Bidirectional logic.
- 2- Autosense is suitable for any type of multicast domain and MUST only be used when there is only one multicast domain. That being said it is not advised to use it in scenarios with multiple domains connected to or through a core domain. For such scenarios it is suggested to ONLY use the Manual mode.
- 3- In PIM-NG a multicast group G for which at least 1 unicast address is seen by C-RP as both the source and receiver IS considered a Bidirectional multicast group. So in a PIM-NG domain with Bidirectional Source Autosense discovery mechanism activated on

its C-RPs and C-MAPPER'(s) when a C-RP receives a Request for source type of message for multicast group G from a client whose unicast address is also registered as the source of that same multicast Group G the C-RP becomes aware that it is dealing with a Bidirectional multicast group and thus it immediately informs the closest C-MAPPER about it by sending the so called Bidirectional Groups Table in its unicast introductions to the C-MAPPER containing the multicast Group with both the TR group and R-Flag fields being set to 0.

- 4- As soon as a Multicast Group (G) is announced as a Bidirectional Multicast Group, the Logic of the Domain for that (G) MUST change to Bidirectional and Sources and receivers MUST immediately join the SPT rooted at closest TR or better to say receivers which are also considered sources will send their join to the TR as the Source.
- 5- The C-RP'(s) MUST send the BGT to C-MAPPER immediately whenever it senses a change by setting the ZTCN bit in its message.
- 6- Receiving the message containing the BGT from a C-RP, the C-MAPPER MUST immediately inform all the existing Clients by sending an introduction message containing the BGT to 239.0.1.190.
- 7- When the Active C-MAPPER receives an introduction from either a C-RP or other C-MAPPERS containing BGT it MUST immediately inform all the other C-MAPPERS it is in contact with so that all the C-MAPPERS become aware of the situation.
- 8- If more than 1 C-MAPPER exists and due to the fact that in such cases and with regards to chapter 4.5.2. The Active C-MAPPER is responsible for all the introductions sent to 239.0.1.190, any C-MAPPER that receives a BGT showing a change MUST inform the Active C-MAPPER. So in case of multiple C-MAPPERS each C-MAPPER receiving a BGT with the R-Flag being set MUST inform the Active and now if the Active doesn't receive a BGT with the associated R-Flag of Bidirectional group (G) not set which shows that there are still senders and receivers in the next 15+2 minutes it MUST inform all the population about the change and removal of that group from Bidirectional logic.
- 9- If the SUB-Domain concept is implemented the Active C-MAPPER in the main domain MUST inform all C-MAPPERS inside Sub-Domains by sending BGT to them.
- 10- In scenarios with SUB-Domains implemented it is possible to filter any desired Bidirectional Multicast group from being

updated to the Active C-MAPPER within each Sub-Domain by the Active C-MAPPER in the main domain.

- 11- BGT MUST BE exchanged between neighbor clients when a new client is added at the time of synchronization.
- 12- When the Autosense mechanism is used, it is possible to remove a multicast group from the Bidirectional logic or announce it deactivated. This can be done by defining the group (G) on the C-MAPPER or Active C-MAPPER and introduce it to population by setting the R-Flag in the BGT.
- 13- If a PIM-NG aware router senses that it is in ideal state for a Bidirectional Group for 60 minutes or better to say it hasn't received any updates from the C-MAPPER about group (G) nor it has been part of the Bidirectional Tree for that (G) it MUST remove it from its internal BGT and change its logic until further notice from the C-MAPPER.
- 14- Each TR MUST send an empty Joined Groups Table to the closest C-MAPPER every 10 minutes to the C-MAPPER. This empty table is sign that all the groups that are announced by the C-MAPPER are active. And eventually C-MAPPERS MUST inform the Active C-MAPPER about the situation.
- 15- If a TR is pruned from the Bidirectional Tree of group (G) or better to say it is not part of the Bidirectional tree of (G) for 20 minutes it announces this to C-MAPPER by sending the Joined Groups table and putting an entry for that group (G) in the table. This is the only time the logic of using Joined Groups Table is different.
- 16- If C-MAPPER or Active C-MAPPER doesn't receive a joined Groups Table from at least one TR showing that all the announced Bidirectional Groups are active for 50 minutes or 5 times to the default time TRs must send their periodical messages containing Joined Groups Table to C-MAPPER , it MUST announce that (G) not active by sending a BGT and setting the associated R-Flag of that Group so that all the population will remove that group until further notice or better to say until at least one source for that group becomes active.
- 17- If a PIM-NG aware router doesn't need to be anymore part of the Bidirectional Tree for (G) it will send a Prune to upstream routers and all the rules with regards to a prune message and shared networks in which more than one router are connected through a shared media are applicable.

4.7.3. Redundant TRs

This section is dedicated to talk about a special feature of PIM-NG Bidirectional logic which makes it possible to use multiple redundant Tree Roots per bidirectional Multicast group and eventually as it is described later to have redundant Bidirectional Trees per multicast group rooted at existing TRs.

In a PIM-NG multicast domain with at least 2 TRs, it is possible to use either all existing TRs in a redundant manner or use a handful of TRs and associate them with 1 set of bidirectional multicast groups and associate other TRs to other groups depending on the needs of multicasting.

Both of the above mentioned factors are desirable in dense environments with many senders and receivers for the same G and residing in different parts of the domain. In the following sections I am going to explain the 2 different approaches PIM-NG provides with regards to existing TRs and redundancy of Tree Roots.

4.7.3.1. ANYCAST Approach

As the name suggests, this method uses the ANYCAST concept with regards to the available TRs within a multicast domain and how they are introduced by the C-MAPPER.

By using this method it doesn't matter how many TRs exist in a PIM-NG multicast domain, because the C-MAPPER updates only one unicast address for existing TRs in its PDTT with the TR group field being set to 0.

As soon as Clients receive the information about the unicast address of the TRs and with regards to Bidirectional logic explained so far, each client immediately joins the SPT for G rooted at the closest TR by setting the source unicast address to the unicast address of the TR and also setting the R-bit of the source address which indicates that the join message should be forwarded towards the TR.

PIM-NG specifications dictates that a join message in which the source unicast address and Tree Root addresses are the same is a sign of a Bidirectional join message and thus in addition to the Multicast group address received in BGT tells intermediary clients that their logic MUST change to Bidirectional for such messages.

At the end of above procedure each client that needs to be part of the Bidirectional Tree for G joins the SPT rooted at the closest TR.

4.7.3.2. TR Grouping

There are scenarios in which it is not desirable to use all the available TRs in Bidirectional processes or it is needed to use some of the available TRs as the redundant roots of a set of Bidirectional Groups and other TRs for other sets.

PIM-NG makes such implementation possible through its well mannered design in a way that it is now possible to consider any one of the bellow approaches that suits well:

- 1- Use the ANYCAST method of TR introduction alongside the Grouping of TRs. In this approach all the TRs are introduced for normal multicast routing using the ANYCAST address and a selected group of TRs can be introduced in addition to the ANYCAST with their unicast address used for communication between TRs and C-MAPPERS alongside their TR group which is manually configured on the Active C-MAPPER. In such case clients MUST only use the TRs with TR Group number other than 0 when dealing with Bidirectional multicast groups.
- 2- Use only Grouping of TRs. In this approach only the unicast address of available TRs are updated by the C-MAPPER alongside their TR group which is set manually on the Active C-MAPPER.

Both of the above methods are usable although, the second approach will result in huge amount of data in PDTT updated by C-MAPPER. No matter which method is used clients will use the groupings ONLY when dealing with Bidirectional multicast groups.

For this method to be implemented a series of processes MUST take place which are listed below:

- 1- TRs MUST be manually grouped on the C-MAPPER or the Active C-MAPPER. This is done by simply assigning a group number with a value between 0 and 255 to each desired set of TRs, with value 0 being reserved for introducing TRs using ANYCAST method and value 255 reserved for introducing a TR which is part of all Bidirectional Groups. This group value is used later by clients at the time of choosing which TR they must send their join towards as the root of the Bidirectional tree.

- 2- A TR which is part of all TR groups which is introduced to the population by TR group value of 255, is called Bidirectional Translator TR and is used later in inter-domain Bidirectional connectivity.
- 3- Then the C-MAPPER MUST update the groupings in its PDTT when sending its introduction to 239.0.1.190 by setting the TR Group field associated with each TR to the configured value and send it to all the population of PIM-NG aware routers so that even existing TRs will become aware of this groupings.
- 4- A mapping between existing and activated Bidirectional Multicast groups and the configured TR Groups on the Active C-MAPPER MUST be done either through manual configuration or an algorithm which through a hashing mechanism assigns and maps each set of Bidirectional multicast group to a TR group. It is suggested to use the automated method through the proposed algorithm which is useful when dealing with the Autosense mechanism of Bidirectional source discovery or big multicast domains.
- 5- Finally when the C-MAPPER is going to notify the PIM-NG population about activated Bidirectional Groups by sending the BGT in its introduction messages, it MUST set the TR Group field associated to each Bidirectional Multicast group with the related and mapped TR group.
- 6- Clients will receive the mappings in the BGT and since they know which TR belongs to which Group because they have received it before in PDTT sent by C-MAPPER, they eventually know which TR should be used as the root of each Tree and will use the unicast address of the appropriate TR which is the closest one to them to join SPT.
- 7- A TR CAN BE a member of multiple groups and a C-MAPPER MUST send the groupings as individual entries in its PDTT. The only exception is for Bidirectional translator TR'(s) that are identified by TR group 255.
- 8- If a TR which is member of a group is lost and by lost PIM-NG specifications means the loss of connectivity between TR and C-MAPPER in a way it is sensed by C-MAPPER, the C-MAPPER MUST choose a suitable TR to replace the lost TR and automatically update it by sending the unicast address of the new TR and associated group in its PDTT. By suitable, PIM-NG specifications means the closest TR to the current TR based on the information found in RIB.

- 9- If the above happens and a TR from a group is lost and a new TR is chosen automatically and the chosen TR is already member of other TR group'(s), then C-MAPPER MUST update all the Groups the TR is a member of in its PDTT as separate entries again. This is a MUST as if a TR is a member of group A and it is also chosen to become a member of group B, if the C-MAPPER send an update with only an entry for group B, clients MUST and will take it as if the TR can be used ONLY for Bidirectional Groups with associated TR group B.
- 10-If a lost TR part of a TR group becomes alive again, the C-MAPPER MUST send an update including PDTT with an entry for the activated TR and removing the TR which was automatically chosen as its replacement.

4.7.4. Bidirectional Tree formation

Tree formation can be considered the most important part of PIM-NG Bidirectional logic as it provides the ability to benefit from redundant Roots per Bidirectional Tree and eventually redundant Bidirectional trees per multicast group.

The basic rule of Bidirectional PIM with regards to tree formation and joining the SPT rooted at the TR applies to PIM-NG too, except for the fact that in PIM-NG actually no shared path tree exists and join messages carry the unicast address of a TR as the source of multicast group in (S, G) format. So in the following sections I will explain the process through which a Bidirectional Tree is formed between TRs and the how clients join the Tree rooted at the closest TR.

4.7.4.1. Tree formation between TRs

As described before in a PIM-NG domain each TR knows the unicast address of other TRs because they receive it from the C-MAPPER. This address is not the ANYCAST address but the address used by TRs to introduce themselves to the closest C-MAPPER and communicate with the C-MAPPER and other TRs to exchange their joined groups table.

Because of the above feature of PIM-NG with regards to TRs, it is very easy to make the TRs join a Tree using the unicast address of other TRs as the source address of a multicast group (G).

The above approach works well in normal PIM-NG multicast communications because it might be needed to forward a received join message towards a TR which already joined the SPT for (S, G) and is closer than the source itself and can be placed anywhere in the domain, but when it comes to Bidirectional communications and tree formation between the TRs PIM-NG's approach is a little different.

PIM-NG specifications and processes with regards to Tree formation between TRs are as follows:

- 1- If a TR receives a join message for Bidirectional Group (G) it MUST first check the contents of its internal Joined Groups Table which indicates all the multicast groups other TRs residing in the domain already joined the Tree for.
- 2- If an entry matching Bidirectional group (G) is found in the Joined Groups table, TR MUST forward the received join message towards the TR which is already joined the Bidirectional Tree for (G) by changing the Source unicast address and TR unicast address of the join message with the unicast address of the TR which is already joined the Bidirectional Tree for (G) without touching the R-Bit.
- 3- If multiple entries matching Bidirectional group (G) are found, TR MUST forward the join message towards the closest TR.
- 4- If a TR receives a join message for a Bidirectional Group (G) and ONLY if there are no entries matching Bidirectional Group (G) in joined Groups table, it MUST inform other TRs by updating its Joined Groups table and send it in its introduction message to other TRs.
- 5- If the ANYCAST TR approach is used the Joined Groups table MUST be sent to all the existing TRs.
- 6- If the TR grouping approach is used, the Joined Groups Table MUST be sent to all the TRs that are member of the same group and specifically to Bidirectional Translator TR'(s) which are a member of all existing Groups and are used in Inter-Domain connectivity.
- 7- The above method will form a distributed Bidirectional Tree for group (G) with multiple TRs as the Roots of the tree and more importantly this tree is formed whenever it is demanded and won't waste much bandwidth.
- 8- Please refer to section 4.7.5. For information about Bidirectional Tree formation between different PIM-NG multicast domains.

Through the above approach and process at the end a Bidirectional Tree with multiple TRs as its redundant roots forms nicely between existing TRs whether the ANYCAST approach or the TR grouping approach is used. All the TRs are working redundantly which provides higher Bidirectional communications speed due to the fact that each client can simply get connected to its closest TR. This approach is mostly beneficial in dense environments such as data centers where we are dealing with high number of senders and receivers who are in different parts of the domain.

Also needless to say this approach provides the possibility of inter-domain Bidirectional connectivity wherever applicable with bellow specifications:

- 1- A C-MAPPER in the remote multicast domain or core domain MUST become aware of the activated Bidirectional Groups through receiving a BGT from its neighbor or manual configuration. To make it clear the 2 domains MUST be neighbors or better to say a neighbor ship between C-MAPPERS in desired domains MUST be established.

4.7.4.2. Completing the Tree formation

Final stage of a full Bidirectional Tree formation will be clients to join the Tree for (G).

Each Client with need to send and receive the traffic of Bidirectional Multicast group (G) MUST join the SPT rooted at the closest TR.

If the ANYCAST approach is used then each client MUST use the ANYCAST address as the Root address and also unicast address of the source for (G) when sending the join message. And in case of TR grouping the unicast address of the closest TR associated to (G) with respect to the information it has received from C-MAPPER in BGT and PDTT.

At the end a distributed redundant Bidirectional tree with regards to the Roots and Tree formation between clients and Roots will form which covers an entire domain and is a feature desirable in dense environments.

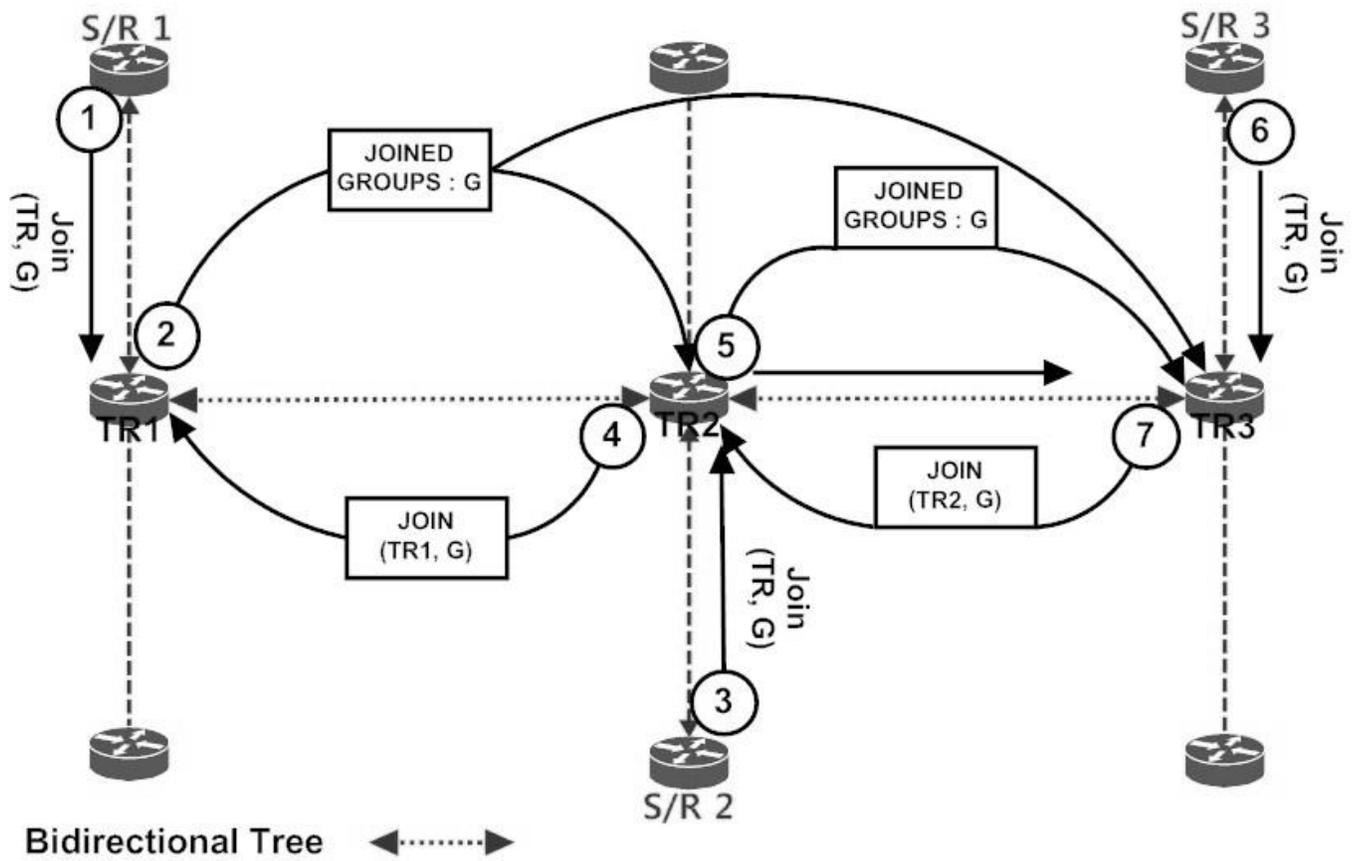


Figure 56 Bidirectional Tree formation with ANYCAST TR Address

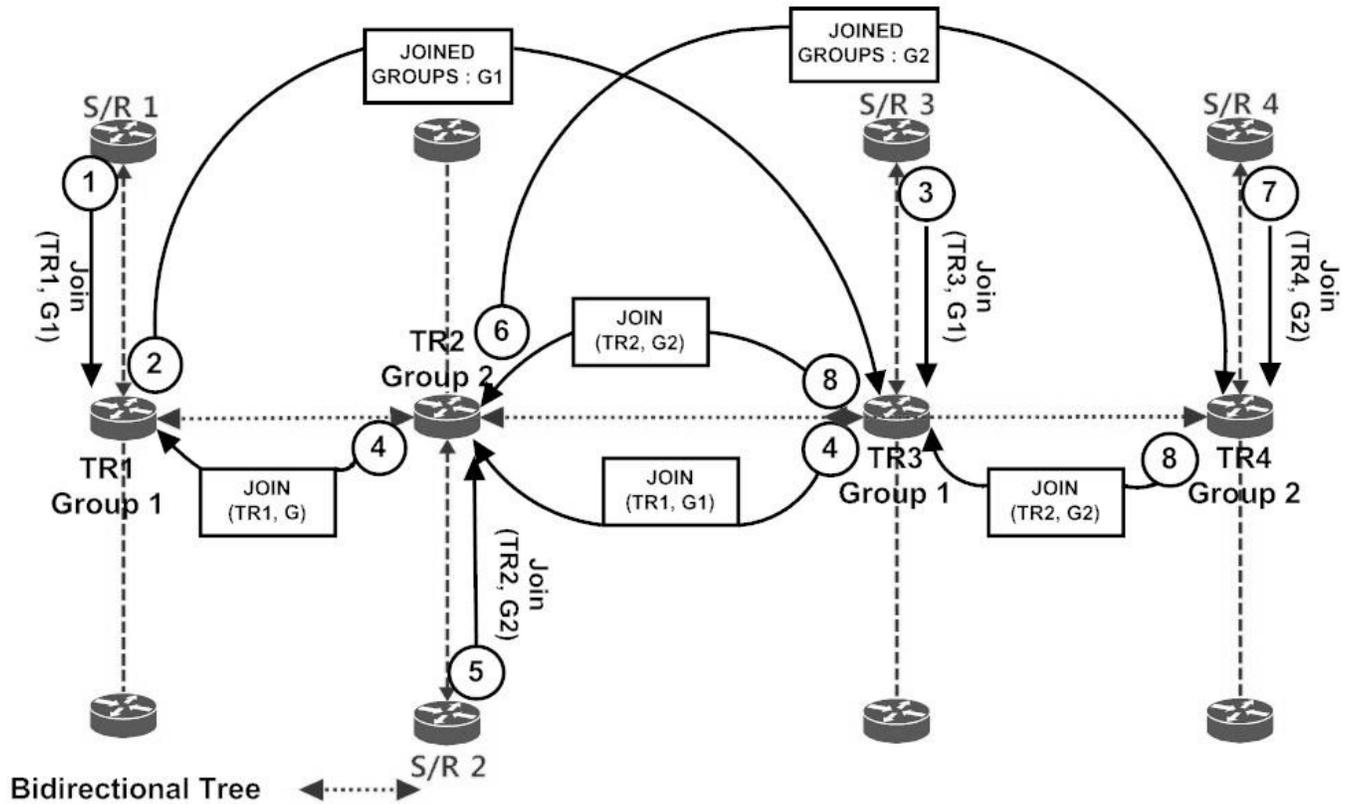


Figure 57 Bidirectional Tree formation with TR grouping

4.7.5. Inter-Domain Bidirectional connectivity rules

- 1- As described previously a C-MAPPER in a core domain MUST introduce existing TR'(s) within its domain to their neighboring C-MAPPERS in connected domain'(s). This procedure is a MUST in Bidirectional logic to help TR'(s) in domains communicate with TR'(s) in core domains to exchange their joined groups table.
- 2- C-MAPPER'(s) within PIM-NG domains MUST introduce at least one TR to their neighbors alongside the associated Domain number so that

the TR is distinguishable. This is done by exchanging the unicast address of selected TR'(s) in PDTT. PIM-NG specifications dictates that the ONLY time a C-MAPPER uses its PDTT to send any information to its neighbors in other domains is when Bidirectional logic is being used and only the information of TR'(s) MUST be exchanged between C-MAPPER'(s) in normal domains or between a C-MAPPER in a normal domain and a C-MAPPER in the core domain.

- 3- Unicast address of any TR that is received in PDTT from a neighbor C-MAPPER in another domain MUST only be sent to existing TRs and not all the population alongside the Domain number value associated to the TR to make the TR distinguishable. This is different from what happens with TRs in a core domain because as it had been explained the unicast address of TR'(s) within a core domain received in Core Domain topology Table MUST be sent to all the population by the C-MAPPER or Active C-MAPPER.
- 4- When dealing with inter-domain Bidirectional connectivity scenarios, PIM-NG STRONGLY advises to use ANYCAST method explained in chapter 4.7.3.1. In all the involved domains to ease the processes.
- 5- As explained before it is also possible to update selected TRs to other domains through manual configuration. This being said, if in a domain TR grouping method is and must be used, then one or more TRs must be chosen to act as Bidirectional Translator TR with TR group number of 255 and the unicast address of the Bidirectional translators MUST be updated in either PDTT or Core Domain Topology Table. This is due to the fact that only a TR which is part of all TR groups or better to say can be aware of all active or to be active Bidirectional Groups can be involved in inter-domain connectivity.
- 6- The joined groups table exchanged between TR's in different multicast domains MUST only contain the information regarding Bidirectional Groups.
- 7- If the TR grouping approach is used, ONLY the Bidirectional Translator TR'(S) MUST send Joined Groups Table to TR'(s) in other domains when dealing with Bidirectional Groups. This is due to the fact a Translator TR is the only TR within a PIM-NG domain to know about all the active multicast groups and will receive the Joined Groups Table from all existing TRs member of any available TR group.

- 8- When a TR receives a join for (G) from a client and there is no entry for (G) in its Joined Groups Table matching (G) it MUST inform other TRs.
- 9- If the ANYCAST approach is used then TRs in other domains will receive the Joined Groups table too and if they receive a join from a client they will forward the join towards TR'(s) in other domains which are already joined the Bidirectional Tree for (G).
- 10- If TR grouping is used and since only the Translator TR is allowed to communicate with TRs in other Domains, the joined Groups table MUST be sent to the Translator TR and Translator TR MUST inform TRs in other Domains no matter if it is already joined the Bidirectional tree for (G) or not.
- 11- When a PER or PPER receives a join for a Bidirectional Group form outside its domain and the TR grouping is used, it MUST forward it towards the closest Translator TR.
- 12- When a PER or PPER receives a join for a Bidirectional Group form outside its domain and the ANYCAST approach is used, it MUST forward it towards the closest TR.
- 13- When a Translator TR receives a join for Bidirectional Group (G) it MUST forward the join towards the closest TR already joined the tree for (G) based on the TR grouping information associated with the TRs and its internal Joined Groups Table mappings. This is why it is called a Translator as it allows to connect a domain in which ANYCAST is used to a domain in which TR grouping is used or the other way around or to connect domains with different TR groupings in use.
- 14- When a Translator TR receives a join message for (G) and ONLY if there are no entries matching (G) in its Joined Groups Table it MUST inform other TRs based on TR grouping and mappings to Bidirectional Groups.
- 15- Since by definition a TR only accepts control plain packets from TRs inside its domain, a mechanism MUST be taken into consideration with regards to ONLY Bidirectional logic so that TR's accept Joined Groups Table from desired TRs with regards to the domain in which they reside. This can be simply a manual configuration through which a TR becomes aware of to which TR in which Multicast domain it can send Joined Groups table and from which TRs it can accept such table.

- 16- It is STRONGLY advised not to use the Autosense mechanism explained in(4.7.2.2.) when dealing with Inter-domain connectivity.
- 17- If in involved domains the Autosense mechanism (4.7.2.2.) is needed to be implemented, then C-MAPPERS in involved domains MUST exchange their BGTs to inform each other about any changes.
- 18- If Autosense mechanism is implemented and Bidirectional group (G) is needed to become deactivated totally, it MUST only be done within the domain the C-MAPPER or Active C-MAPPER is resided and it MUST NOT be announced to C-MAPPERS in other domains..

At the end through all the above specifications, definitions and concepts PIM-NG provides the ability to easily benefit from having a distributed Bidirectional Tree for group (G) with redundant Tree ROOTs and Redundant Trees. This behavior is one of the many features it provides and is specifically beneficial for data centers or whatever multicast domain with huge number of distributed multicast sources and receivers for the same group (G).

Also as explained it also provides the possibility to bring the Inter-Domain connectivity concept with regards to Bidirectional Tree which makes it a good choice for such scenarios.

4.8. PIM-SM compatibility

Up to this point of explaining different concepts of PIM-NG as a new multicast protocol, almost ALL aspects of PIM-NG specifications have been covered.

Now it is time to make it compatible with its predecessor PIM-SM, so that PIM-NG multicast domains can be connected to PIM-SM domains or a network of PIM-SM domains. Compatibility of PIM-NG with PIM-SM is related to the bellow fields:

- o Exchanging the information regarding the multicast sources originated in either the PIM-NG Domain'(s) or PIM-SM Domain'(s), so that receivers can find the source for their desired multicast traffic and send join/prune messages towards the desired multicast source, whether it is inside a PIM-NG Domain or PIM-SM Domain.
- o The transformation of PIM-NG join/prune messages to PIM-SM messages so that routers that are only PIM-SM aware will be able to forward the join/prune messages to the final destination.
- o The transformation of PIM-SM join/prune messages to PIM-NG join/prune messages so that the join/prune message can be forwarded according to PIM-NG specifications within the network of PIM-NG Domains.

In the following sections, different concepts, specifications and rules in regards to connecting a PIM-NG multicast domain to a PIM-SM multicast domain will be discussed. First the concepts regarding the exchange of information regarding originated multicast sources in the form of Source Active (SA) messages will be discussed and after that the concepts related to sending join/prune messages will be discussed.

4.8.1. PIM-SM compatibility and SA messages

As described by RFC 3610[9], the information regarding originated multicast sources MUST be exchanged between RPs that are MSDP peer. And such information is sent from one RP to another RP, in a Source Active (SA) message, which contains:

- o Originator address or the unicast address of the originating RP
- o Source address or the unicast address of the source generating the traffic

- o Group address the source sends data to

And In PIM-NG as described in section 4.5. , the information regarding newly originated multicast sources is carried inside AMMT and between peer C-MAPPERS using unicast-encapsulated C-MAPPER introductions, which contains:

- o Originator address or the unicast address of the originating RP
- o Source address or the unicast address of the source generating the traffic
- o Group address the source sends data to
- o Domain-Set

As it has been described, the format of AMMT and the information related to newly originated multicast sources is similar to the contents of SA messages used by PIM-SM MSDP peers, which makes it easier to connect a PIM-NG Domain or a network of PIM-NG multicast domains to a PIM-SM Domain or a network of PIM-SM multicast Domains. And by connecting in this section PIM-NG specifications refers to the exchange of information regarding multicast sources.

The rules, concepts and specifications regarding compatibility of PIM-NG and PIM-SM are as follows:

- o Depending on whether we are dealing with a public or private PIM-NG domain, a C-MAPPER or PPER in the PIM-NG Domain MUST BE chosen to become MSDP-PEER with an RP in the PIM-SM Domain.
- o All the rules that apply to MSDP [9] MUST be applied when a C-MAPPER or PPER becomes MSDP-PEER with an RP.
- o Whenever a C-MAPPER which is also MSDP-PEER with an RP, receives an update regarding newly originated multicast sources inside the AMMT from a PEER C-MAPPER and needs to advertise the received update to the MSDP-PEER'(s):
 1. It MUST remove the DOMAIN-SET of any multicast sources that are to be advertised to the MSDP-PEER'(s)
 2. It MUST create a Source Active message containing information regarding ALL multicast sources that are to be advertised to the MSDP-PEER'(s).
 3. The SA message contains ONLY:

- o Originator address
- o Source unicast address
- o Group destination

Which is exactly what an SA carries.

- o Then the C-MAPPER MUST send the SA message to its MSDP-PEER'(s) the way explained by MSDP specifications [9] to bring compatibility to PIM-SM.
- o Because of the RPF method used by MSDP, the C-MAPPER or PPER which is becoming MSDP-PEER with an RP MUST reside in the first Autonomous System in the best path towards the AS in which the originator C-MAPPER exists.
- o As explained by PIM-NG specifications, each multicast source has a Domain-Set associated with it, which shows that in which domain a source is being originated and also is used for PIM-NG RPF check. So when a C-MAPPER receives an SA message from a MSDP-PEER (RP), and needs to advertise the information to a PEER C-MAPPER, it MUST add the following to the Domain-Set:
 1. Its own domain number
 2. A value equal to letter "S" which shows that this source is received from a network of PIM-SM Domain'(s). So if C-MAPPER in Domain1 receives an update for (S, G) from its MSDP-PEER, and needs to update it to its PEER C-MAPPER'(s), it MUST modify the Domain-Set as explained and advertise the associated Domain-set as (D S, 1).
- o If a C-MAPPER receives an update for (S,G) from a MSDP-PEER, and an update for the same (S,G) from a peer C-MAPPER which the associated Domain-Set indicates that (S,G) is originated inside a PIM-NG Domain and not a PIM-SM domain, then the update received from the peer C-MAPPER MUST pass the RPF check.
- o If a C-MAPPER receives an update for (S,G) from a MSDP-PEER, and an update for the same (S,G) from a peer C-MAPPER which the associated Domain-Set indicates that (S,G) is originated inside a PIM-SM Domain and not a PIM-NG domain, then the update received from the MSDP-PEER MUST pass the RPF check.
- o If a C-MAPPER receives an update from a peer C-MAPPER regarding sources that MUST be considered as SUSPENDED, the C-MAPPER MUST

NOT send any SA message to MSDP-PEER'(s) until the suspension time is over and the multicast sources are either deleted or active again.

- o If a C-MAPPER loses its connectivity with its MSDP-PEER, it MUST start the suspension timer and send an update about the suspended multicast sources to peer C-MAPPER, and if after the suspension duration which defaults to 5 minutes, the connection with the MSDP-PEER is not established again, it MUST delete the received multicast sources from that MSPD-PEER and inform peer C-MAPPER'(s).or if it has been receiving updates regarding those sources from a peer C-MAPPER or MSDP-PEER it MUST put the received updates from those peers in it's a-MULTICAST MAPPING TABLE and inform its PEER'(s).
- o .

4.8.2. PIM-SM compatibility and join/prune messages

A PIM-NG multicast Domain is connected to a PIM-SM multicast Domain, using a PER or PPER to isolate the 2 multicast Domains and prevent the propagation of multicast introduction messages of PIM-NG Domain in to PIM-SM Domain and also to prevent the propagation of PIM-SM multicast traffic related to BSR[9] and RP'(s) within the PIM-SM domain.

As mentioned in different parts of PIM-NG specifications, a PER or PPER which acts as boundary between a PIM-NG Multicast Domain and a PIM-SM Multicast Domain is called a BORDER-PIM-ROUTER(BPR). And the responsibility of exchanging join/prune messages between the 2 Domains is on the BPR. But due to the fact that PIM-NG uses its own version of join/prune message which is different from that of PIM-SM in parts related to the Tree Root UNICAST ADDRESS, a BPR MUST modify the join/prune messages received from a PIM-SM Domain and likewise it MUST do the same when forwarding join/prune messages from a PIM-NG Domain to a PIM-SM Domain.

Bellow specifications, concepts and rules apply when a join/prune message is forwarded from a PIM-NG domain to a PIM-SM domain and likewise from a PIM-SM Domain to a PIM-NG Domain:

- o If a PER is chosen to act as a BPR, then as soon as the PER is configured and becomes aware that it is connected to a PIM-SM Multicast Domain, it MUST introduce itself to the closest C-MAPPER within the PIM-NG Multicast Domain. The BPR introduction is done by the PER and through sending a unicast-encapsulated introduction message (4.6.1.2.) to the closest C-MAPPER or any

C-MAPPER that is introduced to the PER. The type of the introduction message is set to BPR, and the B-BIT in the introduction message MUST BE set which indicates to the receiving C-MAPPER that it's Domain is connected to a PIM-SM Domain and it MUST start sending the full AMMT to the PER.

- o If a PPER is chosen to act as BPR the introduction process as the BPR is not needed due to the fact that the PPER MUST introduce itself to the closest C-MAPPER or any C-MAPPER that is introduced to it as soon as it is configured to be a PPER.
- o A BPR has 2 types of interfaces:
 1. Internal: an internal interface is connected to the PIM-NG Domain.
 2. External: an external interface is connected to the PIM-SM Domain.
- o A BPR MUST convert any PIM-NG join/prune messages for (S,G) it receives from within the PIM-NG Domain or better to say on an internal interface to a PIM-SM join/prune message, before forwarding it on an external interface which is connected to a PIM-SM Domain. The conversion process is not time consuming, due to the fact that the PIM-NG join/prune messages are designed to be similar to PIM-SM join/prune messages as much as possible.
- o A BPR MUST convert the join/prune messages it receives on an external interface connected to a PIM-SM Domain, ONLY under these conditions:
 1. At least one TR MUST do exist in the PIM-NG Domain.
 2. If no TR exists inside the PIM-NG Domain, then For each (S,G) in the join/prune message, the BPR MUST first check it's AMMT which it receives from the C-MAPPER'(s) inside the PIM-NG Domain. And ONLY IF:
 - o It finds an entry inside the AMMT for the (S,G).
 - o The Domain-Set associated with the (S, G) indicates that the (S, G) is reachable via a connected PIM-NG-CORE-DOMAIN, or better to say the update regarding the (S,G) is passed through a PIM-NG-CORE-DOMAIN.

The BPR MUST convert the PIM-SM join/prune message to PIM-NG join/prune message, and fill out the required fields related to the TR UNICAST ADDRESS and CORE TR UNICAST ADDRESS.

- o If none of the above conditions are met, PIM-NG specifications STRONGLY ADVISES that the BPR DO NOT convert the PIM-SM join/prune message and only forward it to the next hop in the best path towards the source.
- o The above being said, ALL PIM-NG-AWARE routers MUST BE PIM-SM Compatible in parts mostly related to forwarding join/prune messages.
- o if a BPR receives a join/prune message for (G) on an external interface connected to PIM-SM with the S-BIT of SOURCE UNICAST ADDRESS being set, which means that the join/prune is a PIM version 1 type of message it MUST NOT convert the PIM-SM join/prune to PIM-NG join/prune.

4.9. Loop prevention

In order to have a loop free multicast domain PIM-NG suggest the use of Reverse Path Forwarding check(RPF)to prevent any loops from occurring.

Such loops can be under one of the bellow categories:

- o C-MAPPER multicast introduction or notification messages destined to 239.0.1.190 and sent to ALL-PIM-NG-AWARE routers within a multicast domain.
- o C-MAPPER multicast introductions destined to 239.0.1.188 which is used by existing C-MAPPER'(s) to find either a backup or peer C-MAPPER.
- o C-RP multicast introductions destined to 239.0.1.189 which is used by existing C-RP'(s) to find either a backup or peer C-RP.
- o Loops occurring due to a join/prune message which is sent hop by hop towards an existing source by a client which needs to join the SPT for (S, G) in scenarios which clients are connected to each other through switches and within a LAN.

To eliminate and prevent such loops PIM-NG uses the RPF check method and concept used by PIM-SM which is the best current practice. And

since no shared tree (*, G, rpt) rooted at RP forms in PIM-NG there won't be any need for such RPF check although since PIM-NG-AWARE routers are by default PIM-SM compatible it MUST be considered.

The bellow rules MUST be applied to prevent any loops by the unwanted propagation of such multicast introductions:

- o PIM-NG boundary routers such as PER, PPER, BPR or EDGE-CLIENT'(S) MUST NOT forward such introduction messages to other domain'(s) in case of PER, PPER and BPR and down the MULTI-ACCESS network in case of EDGE-CLIENT'(S).
- o C-MAPPER'(S) which receive an introduction message destined to either 239.0.1.190 or 239.0.1.188 with the source address being equal to C-MAPPER'(s) address MUST discard the packet and NOT forward any further.
- o C-RP'(S) which receive an introduction message destined to either 239.0.1.189 with the source address being equal to C-MAPPER'(s) address MUST discard the packet and NOT forward any further.

4.10. DR election and PIM Assert Message

PIM-NG specifications suggest using the processes and concepts introduced and defined by PIM-SM [7] with regards to Assert messages and DR election as the best current practice when dealing with scenarios and topologies involving multi-access LAN'(s).

The only place that the DR election is different from that of PIM-SM is in implementations and topologies in which EDGE-CLIENT'(s) are considered at the edge of a multi-access network which allows the administrators to dictate the DR for the Multi-Access network by manipulating the priority of Edge-Clients. So the Edge-Client with higher priority becomes the DR for the entire Multi-Access network.

5. Security Considerations

This section is going to cover some of the security concerns related to PIM-NG specifications covered in this document, and possible solutions for those security issues. As this document is an earlier version of PIM-NG specifications, only related security issues are going to be covered.

5.1. Attacks based on forged messages

The extent of possible damages depends on the type of messages that are forged. PIM-NG processes use different kinds of messages like link-local messages, multicast messages and unicast messages. And each type of message will be discussed separately.

5.1.1. Unicast forged messages

As Register, join and leave messages alongside C-RP introduction messages sent to C-MAPPER are forwarded by intermediate routers to their destination using normal IP forwarding, without authentication there is a high possibility for an attacker anywhere inside the network, to forge these messages .the effects of such forgery can be as follows:

- 1- By forging a register message an attacker can inject forged traffic into the RP and to the entire PIM-NG domain.
- 2- By forging join message ,an attacker may become able to act as the man in the middle and receive a traffic that is not meant for that receiver .or traffic can be delivered to parts of the network that no legitimate receivers exists ,which can cause waste of bandwidth.
- 3- By forging leave message, an attacker can prevent a legitimate receiver from receiving the traffic it needs.
- 4- By forging a C-RP introduction message sent to the C-MAPPER, an attacker can become a real threat to the entire domain, by injecting false information to the domain.
- 5- By forging a TR introduction message sent to C-MAPPER, an attacker can become a real threat to the domain, by becoming a man in the middle and receive a copy of all the multicast traffic that is passing through the TR or by dropping the received join/prunes which can cause the connectivity problems.

5.1.2. Forged link local messages

As Forged Hello messages are sent to link-local ALL-PIM-ROUTERS and are not forwarded by the compliant router, they can cause problems such as:

- 1- If the source of forged message is inside a Multi-access LAN or to be more specific a local client, it can give an attacker the

possibility of playing the role of an EDGE-CLIENT, and prevent the legitimate receivers from receiving the desired traffic or reaching the desired sources.

- 2- By forging a Hello message an unauthorized client may become able to play the role of designated router (DR) for a LAN and become responsible for forwarding traffic on behalf of local members or hosts. This can prevent hosts from receiving the desired traffic.
- 3- By forging a Hello message, an unauthorized router in a PIM-NG domain can become part of the domain and cause damages such as preventing its neighbors from receiving C-MAPPER introductions or injecting false information inside the PIM domain topology table.

5.1.3. Forged multicast messages

C-MAPPER introduction messages sent to ALL-PIM-NG clients , MAPPER introduction messages sent to ALL-PIM-NG-MAPPERS in order to finding the PEER-C_MAPPERS or SC-MAPPERS and C-RP introduction messages sent to ALL-PIM-NG-RPs are PIM-NG multicast messages required for the processes of PIM-NG. But an attacker might become able to forge such messages and cause damages. The damages that can be done to a PIM-NG domain are as follows:

- 1- By forging a C-MAPPER introduction message sent to ALL-PIM-NG-CLIENTS (239.0.1.190), an attacker can inject false information in to the domain, by either injecting false data into the PIM domain topology table.
- 2- By forging a C-MAPPER introduction message sent to ALL-PIM-NG-CLIENTS, an attacker can take the role of C-MAPPER and introduce itself to C-RPs and finally , can cause the clients not to be able to find the existing C-RPs , and prevent them from receiving the desired traffic.
- 3- By forging MAPPER introduction message sent to ALL-PIM-MAPPERS ,an attacker can become able to take the role of ACTIVE C-MAPPER in the process of ACTIVE-C-MAPPER election, and also become peer with other C-MAPPER'(s) and cause damage to the domain by injecting false data into the A-MULTICAST MAPPING table.
- 4- By forging RP introduction message sent to ALL-PIM-RPs, an attacker can be able to take the role of C-RP in the process of C-RP election or become the ACTIVE C-RP in a C-RP Mesh-Group and also become peer with other existing C-RP'(s) in search of its peer and cause problems by injecting false data in to either MULTICAST

MAPPING table or A-MULTICAST MAPPING TABLE. And preventing legitimate receivers from receiving the desired traffic.

5.2. Non-cryptographic authentication mechanisms

A PIM-NG router should provide an option to limit the set of neighbors from which it accepts join/prune, Assert, and hello messages, by either static configuration of IP addresses or an IPSEC security association CAN be used. And a PIM-NG router should not accept protocol messages from a router from which it has not yet received a valid hello message. Also a PIM-NG router SHOULD NOT accept any hello message from a router that is not within the same PIM-NG Domain unless it is a PIM-EDGE-ROUTER acting as the boundary between different PIM-NG Domains.

A Designated Router MUST NOT register-encapsulate a packet and send it to the C-RP if the source address of the packet is not a legal address for the subnet on which the packet was received. Similarly, a PIM EDGE-CLIENT MUST NOT accept a register message from its downstream PIM-NG Clients, if the source address of the register message is not a legal address for the subnet on which the register message is received.

A Designated Router MUST NOT accept a C-RP acknowledge packet whose IP source address is not a valid C-RP address for the local Domain. Similarly, a PIM EDGE-CLIENT MUST NOT accept a C-RP acknowledge packet whose IP source address is not a valid C-RP address for the local Domain.

A mechanism MUST be considered as an option so that a C-RP restricts the range of source addresses from which it accepts Register-Encapsulated packets. Also it is STRONGLY advised to consider a mechanism through which a C-RP restricts the range of source addresses from which it accepts C-MAPPER or C-RP introduction messages, so that a possible attacker cannot send such messages or such messages from unknown ranges are not accepted. Also as explained throughout PIM-NG specifications a C-RP MUST NOT accept source register, C-RP, C-MAPPER messages with different Domain number from the one dictated to the C-RP.

A mechanism MUST be considered as an option so that a C-MAPPER restricts the range of source addresses from which it accepts unicast-encapsulated C-MAPPER, C-RP, PER, PPER, and TR messages within the same Domain, due to the fact that if dynamic methods explained by PIM-NG specifications are used if this ranges are not

restricted an attacker may try to introduce itself to the C-MAPPER as a legitimate component of a PIM-NG Domain. As explained throughout the PIM-NG specifications, a C-MAPPER MUST ONLY accept C-RP, TR, PER and PPER introduction messages that carry the same Domain number as the Domain number dictated to the C-MAPPER.

All options that restrict the range of addresses from which packets are accepted MUST default to allowing all packets.

5.3. Authentication

Like PIM-SM [7], PIM-NG specifications recommends to use IPSEC [4] transport mode using the Authentication Header (AH) to prevent the above attacks against PIM.

The proposed methods of protecting:

- o Link-Local Multicast Messages
- o Unicast register messages

By PIM-NG, are the methods recommended by PIM-SM Specifications [7] as the best current method of protecting the above PIM messages.

Since PIM-NG uses processes different from that of PIM-SM in parts related to:

- o The unicast C-RP acknowledge to the DR instead of sending register stop message.
- o Unicast introduction messages, sent from C-RP to C-MAPPER.
- o Unicast introduction messages, sent between the C-RPs.
- o Unicast introduction messages, sent between the C-MAPPERS.
- o Unicast introduction messages, sent from BPR, PPER, and any existing TR to C-MAPPER.
- o Unicast introduction messages, sent between existing TRs.

a mechanism MUST BE taken in to consideration to protect such messages.

With the above being said, PIM-NG recommends to use, the Register Message protection method and Register-Stop protection mechanism

recommended by PIM-SM[7] which is considered the best current method of protecting unicast Messages.

5.3.1. Protecting Multicast Introduction Message

One important security threat to a PIM Domain that is not covered is related to the multicast messages sent from a C-MAPPER to all PIM-NG routers which is similar to the process related to BSR[9]. since when in a PIM Domain the dynamic methods of finding RP (PIM-SM), C-RP (PIM-NG) are in use, the multicast messages sent from a C-MAPPER or BSR play the main role in the process of introducing existing C-RP or RP to all PIM-NG or PIM-SM routers, a mechanism MUST be taken in to consideration so that a PIM-NG CLIENT or simply put a PIM router doesn't accept an unauthorized C-MAPPER introduction message.

in order to protect such a message, PIM-NG recommends that an SA and SPI be defined on existing legitimate C-MAPPERS and on all PIM-NG routers by the network administrator to authenticate such multicast messages. So if an unauthorized C-MAPPER multicast introduction message is received by the first PIM-NG-AWARE router, it will be rejected (dropped without process) and won't be forwarded any further.

5.4. Denial-Of-Service attacks

There are number of denial-of-service attacks against PIM-NG that can be caused by generating false PIM-NG protocol messages or false traffic. Using the authentication methods can prevent some, but not all, of these attacks. Some of the most possible attacks are:

- o Sending packets to many different group addresses quickly can be considered a denial-of-attack, which can cause many register packets, loading the DR, the C-RP, the C-MAPPERS when more than one C-MAPPER exists and finally the routers between these components.
- o Many forged join messages can cause many multicast trees to be set up and consume network resources.
- o With regards to PIM-NG, many forged join messages can cause many Request For Source messages that will be sent from a CLIENT to the C-RP, and can be considered a denial of service attack.

To reduce the possibility of the creation of unwanted register messages, if applicable, PIM-NG specifications STRONGLY suggest using the STUB-DOMAIN concept (4.5.6) in Domains that no multicast source is supposed to exist.

6. IANA Considerations

6.1. PIM-NG multicast destination group addresses

PIM-NG processes require to use 3 multicast addresses from the internet network control block (RFC 5771[2]).for the simplicity of explanation process in this documents these 3 addresses are chosen from the scoped multicast ranges. The addresses are needed for the bellow processes:

- o Destination group address used for C-MAPPER introduction process. The multicast introduction message is sent to ALL-PIM-NG routers.
- o Destination group address used for C-MAPPER introduction process so that C-MAPPERS can find each other dynamically. The multicast introduction message is sent to ALL-PIM-NG C-MAPPERS.
- o Destination group address used for C-RP introduction process so that C-RPs can find each other dynamically. The multicast introduction message is sent to ALL-PIM-NG C-RPs.

These addresses are needed to be assigned by IANA after this document is approved.

6.2. PIM-NG packets and values of type field

New type values in PIM-NG packet header and new packet formats designed specifically for PIM-NG to support the needs of the different processes of PIM-NG, and will need to be reviewed by the experts and assignments need to be made.

6.3. PIM-NG Domain numbers

The PIM-NG Domain number field is considered as a 32 bit field to support the future needs of multicasting, in regards to PIM-NG Multicast Protocol.

Domain Number (4.6.1.1) has a vital role in PIM-NG processes and functionality, which provides the possibility of using Sub-Domain concept as a security feature. Also it MUST BE noted that through using the Domain numbers, PIM-NG is able to use a unique RPF method and a simple multicast domain isolation and separation method, which provides many features in comparison to the previous versions of PIM protocol.

CORE-DOMAIN assignments are needed to be done by IANA as the controlling entity, so that, no conflict can happen as explained throughout this document.

7. Conclusions

PIM-NG is a multicast protocol that , although may seem to have lots of features compared to previous protocols like PIM-SM, but through its many features and processes, provides a robust and sound control over the propagation of the information regarding the existing multicast sources. Its processes are enhanced, so that a host in search of the source of a multicast traffic can communicate with the desired source as fast as possible, by eliminating the need to necessarily join the RPT rooted at an RP within a multicast domain.

Also duo to its unique method of RPF check provides the ability of implementing transitory multicast domains which was not implementable before. And because of using the Domain concept provides many features in parts related to security and controlling over the propagation of multicast information inside a PIM-NG multicast Domain.

8. References

8.1. Normative References

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [2] Cotton, M., L. Vegoda and D. Meyer, "IANA Guidelines for IPV4 Multicast Address Assignments", RFC 5771, March 2010.
- [3] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, version 3", RFC 3376, October 2002.
- [4] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.
- [5] Narten, T. and H. Alvetrand, "Guildlines for Writing an IANA considerations section in RFCs", BCP 26, RFC 5226, May 2008.

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

8.2. Informative References

- [7] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006.
- [8] Fenner, B. and D. Meyer, "Multicast Source Discovery Protocol (MSDP)", RFC 3618, October 2003.
- [9] Bhaskar, N., Gall, A., Lingard, J., and S. Venaas, "Bootstrap Router (BSR) Mechanism for PIM Sparse Mode", Work in Progress, May 2006.
- [10] Hardjono, T. and B. Weis, "The Multicast Group Security Architecture", RFC 3740, March 2004.
- [11] Handley, M., Kouvelas, I., Speakman, T., and L. Vicisano, "Bidirectional Protocol Independent Multicast" (BIDIR-PIM), RFC 5015, October 2007.
- [12] Kim, D., Meyer, D., Kilmer, H., and D. Farinacci, "Anycast Rendezvous Point (RP) mechanism using Protocol Independent Multicast (PIM) and Multicast Source Discovery Protocol (MSDP)", RFC 3446, January 2003.
- [13] Savola, P., Lehtonen, R., and D. Meyer, "Protocol Independent Multicast - Sparse Mode (PIM-SM) Multicast Routing Security Issues and Enhancements", RFC 4609, August 2006.

9. Acknowledgments

The author would like to thank Mr.Reza Izadi for his review and kind comments, corrections and support throughout the process of designing and writing different concepts of PIM-NG Multicast Protocol. Also the author would like to thank Mr.Stig Venas for his review and comments which helped to make this draft complete.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Saeed Sami
P.O.BOX 1466955316
Unit.4
No.47
Saadatabad 35th.St
Khovardin.Blvd
Sanat.SQ
Tehran. Iran

Phone: +989123844205
Email: sami@ssami.biz
sami.biz.email@gmail.com