

Network Working Group
Internet Draft
Intended status: Standard Tack

Tissa Senevirathne
Nataraj Bacthu
Raghava Sivaramu
CISCO Systems
Sam Aldrin
Donald Eastlake
HuaWei Technologies

March 4, 2012

Expires: September 2012

IP multicast data plane failure detection
draft-tissa-pim-mcastoam-00.txt

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#). This document may not be modified, and derivative works of it may not be created, and it may not be published except as an Internet-Draft.

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#). This document may not be modified, and derivative works of it may not be created, except to publish it as an RFC and to translate it into languages other than English.

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>

Internet-Draft

[draft-tissa-pim-mcastoam-00](#)

March 2012

This Internet-Draft will expire on September 4, 2012.

Copyright Notice

Copyright (c) 2012 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.

Abstract

ICMP Based multicast messaging infrastructure is presented in this document. Messages presented in this document is intended to be utilized for troubleshooting multicast data plane connectivity issues as well as identify multicast routers performing different roles, such as Rendezvous Point (RP), Designated Router etc.

Table of Contents

1.	Introduction.....	3
2.	Conventions used in this document.....	4
3.	Extensions.....	4
3.1.	Receiver Address scope.....	4
3.1.1.	Theory of Operation.....	4
3.1.2.	C-type-Addr-scope.....	5
3.2.	Originator IP Address.....	5
3.3.	Multicast Router Role.....	6
3.3.1.	Role Request Use Case Example.....	9
3.4.	Incoming and Outgoing Interfaces.....	10
3.5.	Reverse Path Forwarding.....	11
3.5.1.	RPF Response of Data Plane information.....	12
3.5.2.	RPF Response of Control Plane information.....	14
4.	Security Considerations.....	15
5.	IANA Considerations.....	15

6. References.....	16
6.1. Normative References.....	16
6.2. Informative References.....	16
7. Acknowledgments.....	16

[1. Introduction](#)

Echo Request and Response messages are widely used in troubleshooting connectivity faults in unicast IP network. In a typical IP unicast network, there is a strict 1-1 relationship between an IP address and a device. Hence, Echo request messages can uniquely identify the liveliness of the end-device.

On the other hand single multicast address, also known as multicast group address, represents many devices which are interested in receiving multicast traffic destined to the specific group address. Echo request addressed to a specific multicast group address potentially generate large number of responses that may overwhelm the requester to the point that information may not be useful. It would be very attractive, if there exists a method that would allow a user to specify the end station address or addresses from which a response is desired. Ability to indicate such a response scope, allows user to narrow down the responses only to the desired scope (ie end stations).

In an IP unicast network, there are two classes of devices; end stations and routers. In a multicast network, based on the multicast routing protocol, devices belong to multiple different classes and they perform completely different set of functions/roles. Routers in a PIM-SM multicast routed network can belong to the following classes; Designated Router, Rendezvous Point Router. Functions of each of these classes are explained in the [RFC-PIMSM]. There are no convenient methods a user can utilize to identify different multicast roles routers are performing. Either the user has to hop between routers looking for information or use a different tools such as SNMP to identify the routers that are performing a specific function. Ability to utilize an integrated tool such as Ping to discover routers performing specific roles not only convenient but can be very efficient, especially when troubleshooting a complex problem.

Currently there are few tools such as "mtrace" that are widely

utilized in troubleshooting multicast paths. Most of these tools only validate the control plane. As such, they may not be able to detect failures associated with data plane.

In this document we propose a framework that extends ICMP messages to carry different multicast troubleshooting related extensions. The extensions specified in this document utilize approaches presented in [[RFC4884](#)] and [PING-EXT].

"ssmping" and "asmping" are commonly utilized to troubleshoot multicast connectivity. These tools require presences of a server at the multicast source. In typical deployments, especially in hosting services, network operator may be different from the administrator of the servers. In such scenarios use of tools such as above provide operational and administrative challenges. The tools presented in this document can be exercised from first hop routers and allow flexibility to masquerade the source address to the address of the multicast source. Included in the payload is the originator IP address where reply needed to be sent. Originator IP address may be different than the source IP address.

[2.](#) Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC-2119](#) [[RFC2119](#)].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying [RFC-2119](#) significance.

[3.](#) Extensions

We propose to define a new object class TBD-MUL-OBJ within the ICMP Extension Object defined in [[RFC4884](#)]. TBD-MUL-OBJ object encapsulate extensions related to multicast OAM. Different c-types are expected to be defined with the Object type TBD-MUL-OBJ to specify specific information related to multicast OAM.

[3.1.](#) Receiver Address scope

Ping (ICMP Echo request) directed to a multicast destination "G" may

trigger responses from every receiver in the network that is receiving multicast traffic for "G". It is convenient to have a method to specify subnetwork or end station from which the requester is expecting a response.

3.1.1. Theory of Operation

Requester includes one or more c-type-Addr-scope (Figure 1) within the ICMP Echo request message. Each c-type-Addr-scope represents an address from which a response is required.

Each device that has an active interface with the specified destination multicast group address "G", compares the embedded c-type-Addr-scope against its local interface IP addresses. If they

match, a response SHOULD be generated. If they do not match request MUST BE silently discarded. Absence of the c-type-Addr-scope or inclusion of value zero in the IP Address field of the c-type-Addr-scope indicates that a response is required from all devices with an active interface with address "G".

3.1.2. C-type-Addr-scope

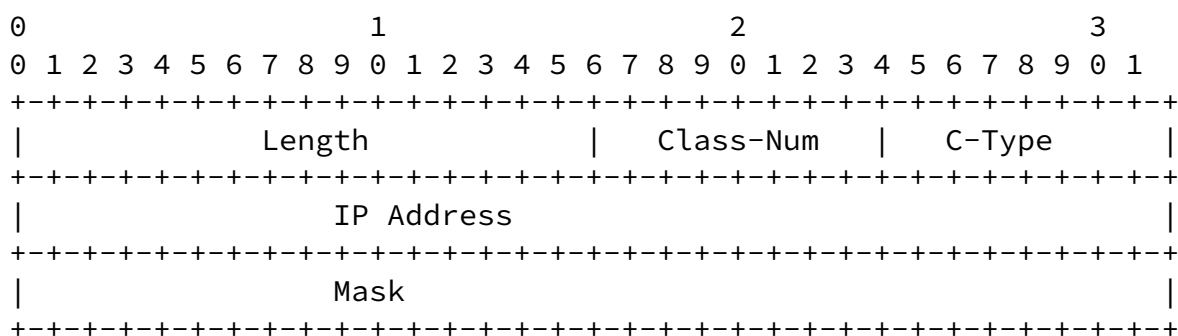


Figure 1 c-type-Addr-scope

- o Length : 8 or 32 octets. 8 octets indicate embedded is an IPv4 address. 32 octets indicate IPv6 address
- o IP Address : 4 or 16 octets

- o Mask : 4 or 16 octets, represents the associated address mask.

3.2. Originator IP Address

Source address of the IP header plays an important role in multicast packet forwarding. It is highly desirable to have ability generate multicast oam packets from the first hop Designated Router, or any intermediate routers, with the source IP address field set to the source IP address of the multicast server. Originator IP address c-type defined in this section allows the originator to include its IP address in the payload of the multicast OAM message. Responders are required to generate the response addressed to the embedded Originator IP address not to the source IP address of the multicast OAM packet. When originator IP address c-type is absent in the payload, requester may utilize the source address of the multicast OAM packet.

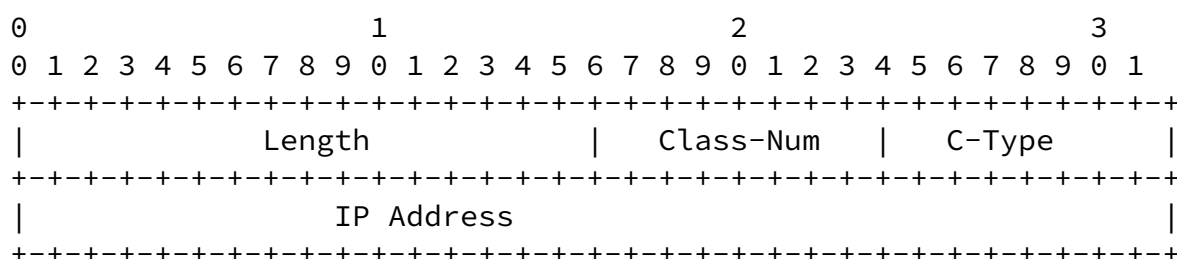


Figure 2 c-type-Originator-IP-Address

- o Length : 8 or 32 octets. 8 octets indicate embedded is an IPv4 address. 32 octets indicate IPv6 address
- o IP Address : 4 or 16 octets, IP address of the originator where a response is required to be sent.

3.3. Multicast Router Role

As it was stated earlier identification of the role played by routers within the multicast network is very important when

troubleshooting a problem.

We propose to define two categories of c-types. C-type-role-request, embedded in the request message indicates to responders to include c-type-role-response in the response. C-type-role-response MUST be included in the response message only when c-type-role-request is present in the request message. Multiple c-types are defined for the roel-response. Each role-response c-type may carry additional information that may be specific to the requested role. In the event a given router performing multiple roles, multiples of such role responses may be included in a given response.

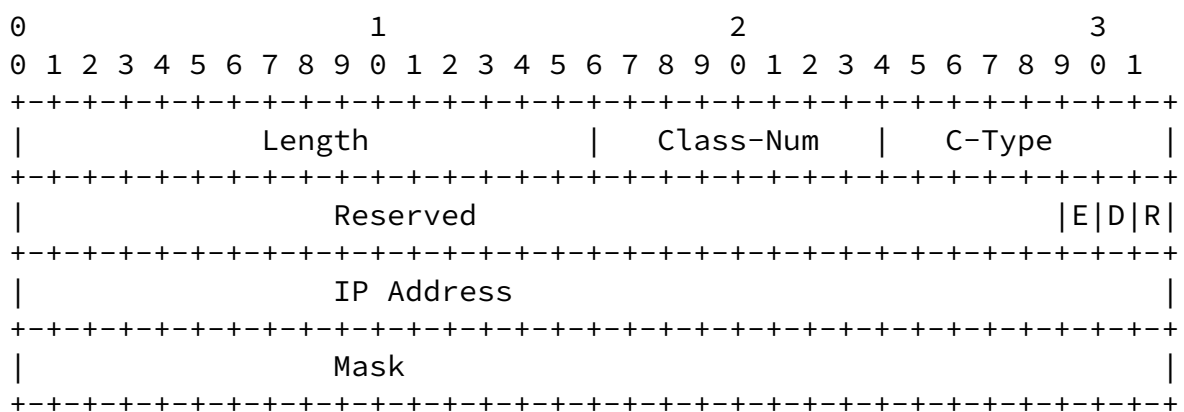


Figure 3 c-type-role-request

- o Length : 12 or 36 octets depending on IP address is IPv4 or IPv6.

- o Reserved : Transmitted zero and ignored on receipt
- o R : (1 bit), indicates response is requested from Rendezvous Point routers. IP Address specifies the Group address(es) that the RP is servicing.
- o D : (1 bit), indicates response is requested from Designated Routers. IP Address specifies the subnet that the DR is servicing.
- o E : (1 bit) indicates response is requested from an end station. IP Address specifies the subnet or the end-station address.
- o Only one of the E/D/R bits MUST be set in a single c-type. A request MAY contain multiple c-type-role-requests with different values of IP Address/Mask and/or of E/D/R flags.

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Length                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               RP IP Address                        |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Reserved                             |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Num of Groups                       |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Multicast Group Address 1           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Address Mask 1                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

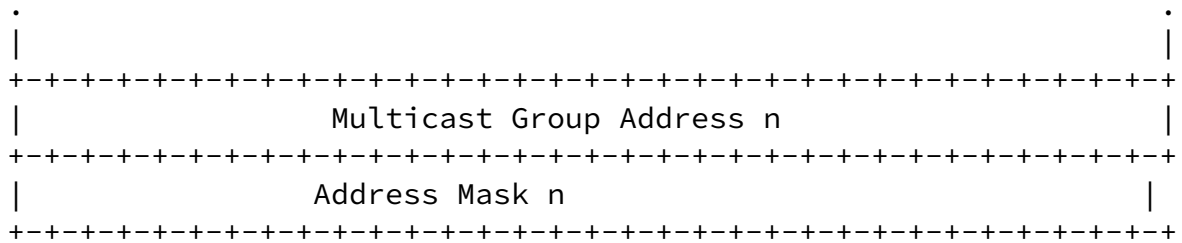
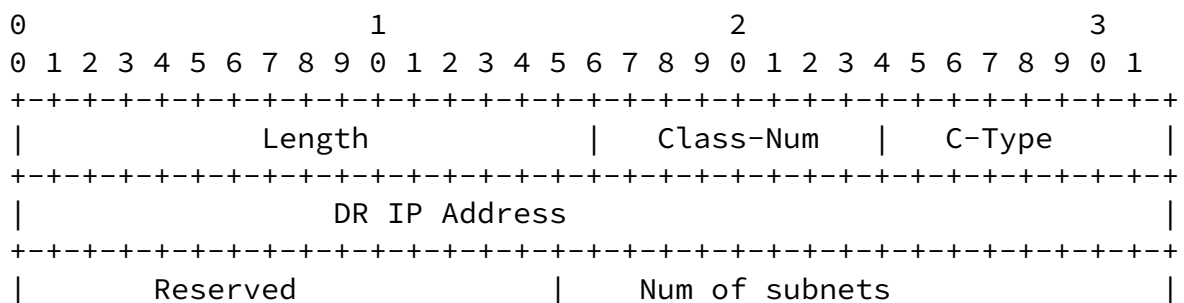



Figure 4 c-type Role Response of Rendezvous Points

- o Length : $4+4+(4+4)n$ or $16+4+(16+16)n$ octets depending on IP address is IPv4 or IPv6. Where n is the number of Multicast Group addresses.
- o RP IP Address: (4 or 16 octets). Unicast IP address of the Rendezvous Point
- o Reserved : Transmitted zero and ignored on receipt
- o Num of Groups: Number of Multicast Groups embedded in the c-type.
- o Multicast Group Address: (4 or 16 octets).
- o IP address mask associated with the Group address: (4 or 16 octets)



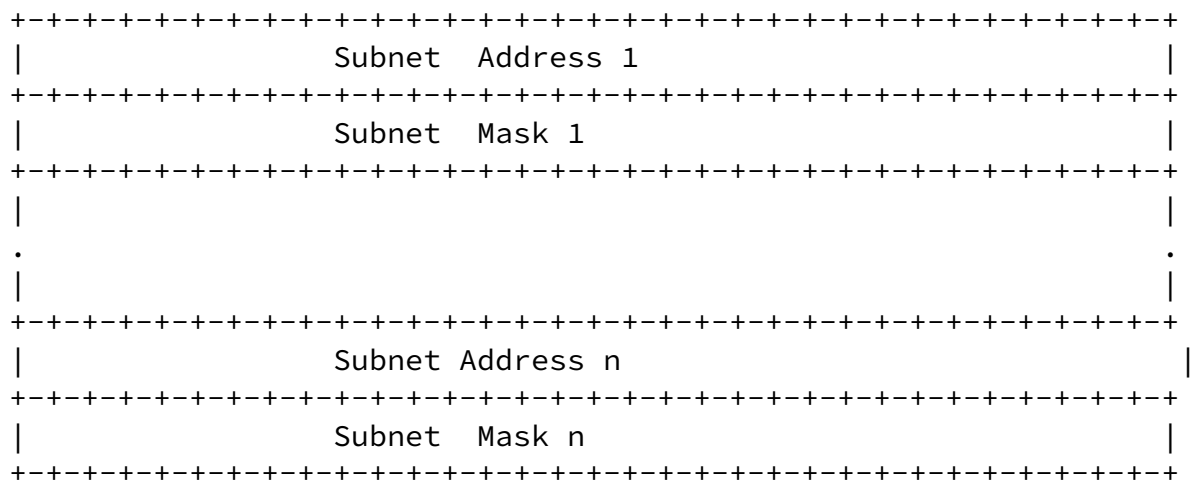


Figure 5 c-type Role Response of Designated Router

- o Length : 4+4+4n or 16++4+16n octets depending on IP address is IPv4 or IPv6. Where n is the number of Multicast Group addresses.
- o DR IP Address : (4 or 16 octets) Management IP Address of the Designated Router
- o Reserved : Transmitted zero and ignored on receipt
- o Num of Subnets: Number of subnets embedded in the c-type
- o Subnet Address: (4 or 16 octets) Subnet Address of which this Router is a DR
- o Subnet Mask: (4 or 16 octets) Subnet mask

3.3.1. Role Request Use Case Example

Let's assume an operator desires to identify, for a multicast address "G", all Rendezvous Point (RP) routers in the network and Designated router servicing subnet 10.10.10.0/24.

ICMP echo request message with destination addressed to multicast address G with router alert flag is generated. Additionally, following role-request c-types are included in the ICMP echo request

message:

1. c-type-role-request, with R flag set and IP address field set to G and subnet mask set to 255.255.255.255. This indicates response is required from RP that are servicing group G.
2. c-type-role-request with D flag set and IP address field set to 10.10.10.0/24. This indicates a response is required from routers that are functioning as DR for subnet 10.10.10.0/24.

Each intermediate router along the path receives a copy of the ICMP echo request message due to router alert flag in the header.

Intermediate routers process the ICMP echo request message extensions and follow c-type-role-request message processing.

1. RPs servicing group "G" generate ICMP echo response addressed to the requester and include ICMP extension Role Response of Rendezvous Points. In the Role Response of Rendezvous Points message, IP address is set to the IP address of the RP, Multicast Group address and applicable subnet mask are set accordingly.
2. DR that is servicing subnet 10.10.10.0/24 generate ICMP echo response addressed to the requester and include ICMP extension Role response of Designated Router. In the ICMP extension Role response of Designated Router c-type, IP address is set to the IP address of the DR and applicable Subnet Address and Mask are also included.

[3.4.](#) Incoming and Outgoing Interfaces

This c-type encodes the IP address of the upstream interface from which the ICMP Echo request was received and downstream interfaces to which the request would be forwarded.

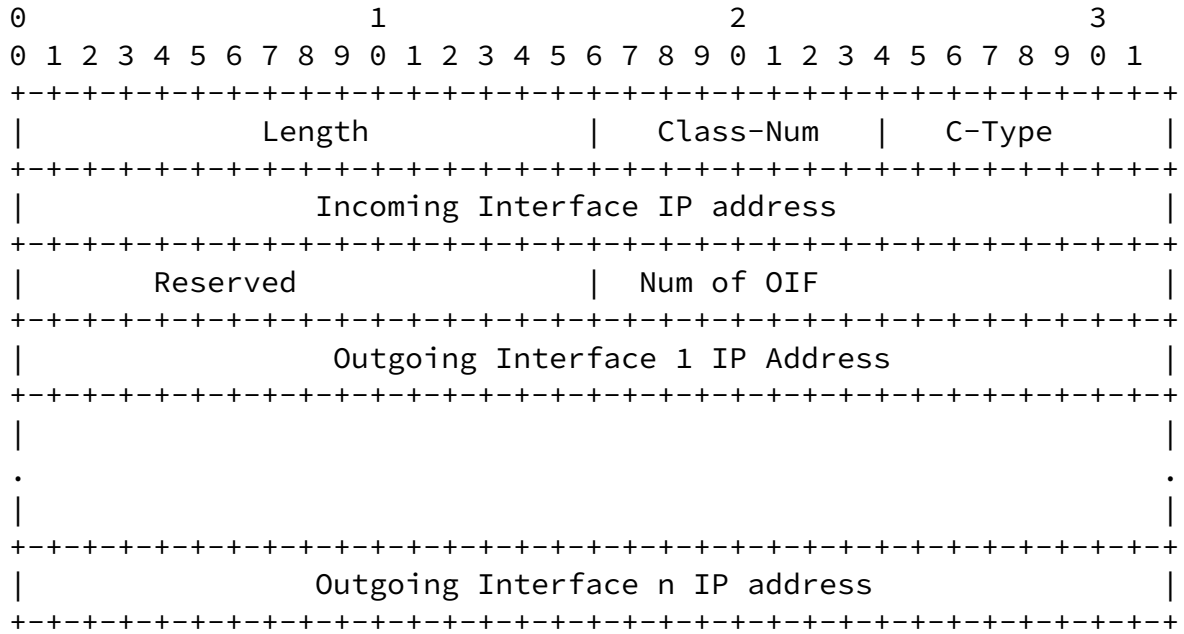


Figure 6 c-type Incoming and Outgoing Interfaces

- o Length : 4+4+4n or 16++4+16n octets depending on IP address is IPv4 or IPv6. Where n is the number of Multicast Group addresses.
- o Incoming Interface IP address: (4 or 16 octets) IP address of incoming Interface
- o Reserved : (2 octets) Transmitted zero and ignored on receipt
- o Num of OIF: (2 octets) Number of outgoing Interface.
- o Down Stream neighbor: (4 or 16 octets) IP address of the Outgoing Interfaces

[3.5](#). Reverse Path Forwarding

"mtrace" is the only tool that is currently available to identify RPF issues. "mtrace" is a control plane tool and may not always give proper fault coverage, especially when control and data plane are out of alignment.

c-type RPF-REQ presented here may be included to request RPF information. C-type RPF-RES carries the requested information.

As previously mentioned, ICMP ECHO request messages addressed to a specific group address G may not be responded by the intermediate routers as the intermediate routers may not have an active group G on the router. We propose, when response from intermediate routers are required, to generate ICMP Echo request messages with Router Alert flag set.

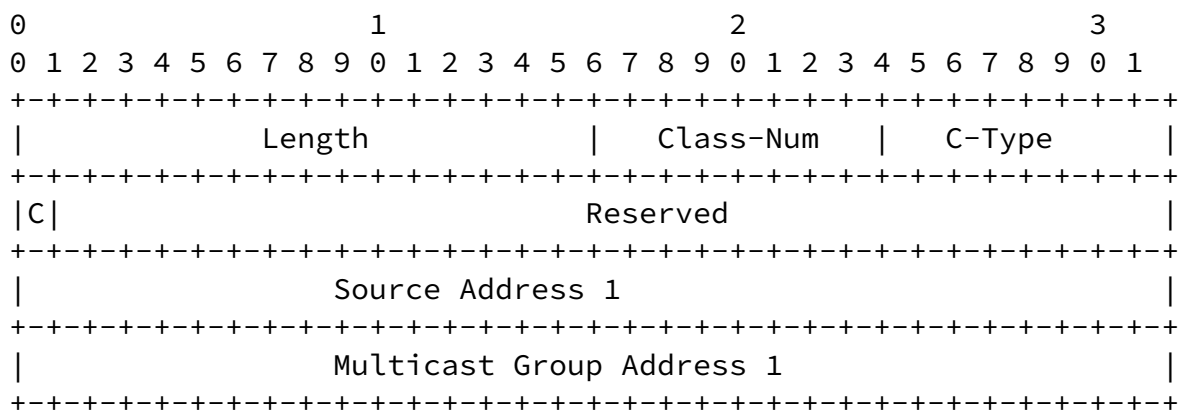


Figure 7 c-type Role RPF request

- o Length : 4+ 2*4 or 4+2*16 octets depending on IP address is IPv4 or IPv6. Where n is the number of Multicast Group addresses.
- o Reserved : (15 bits) Transmitted zero and ignored on receipt
- o C (1 bit) indicates that control plane validation requested.
- o Source Address: (4 or 16 octets) Source IP address "S" in (S,G). Zero in the source address field represents * notation in (*,G). In general this indicates request for information related to the shared tree.
- o Multicast Group Address: (4 or 16 octets) multicast group Address G

[3.5.1](#). RPF Response of Data Plane information

The c-type defined in this section provide RPF information contained in the data plane of the responding device for the specified (S,G).

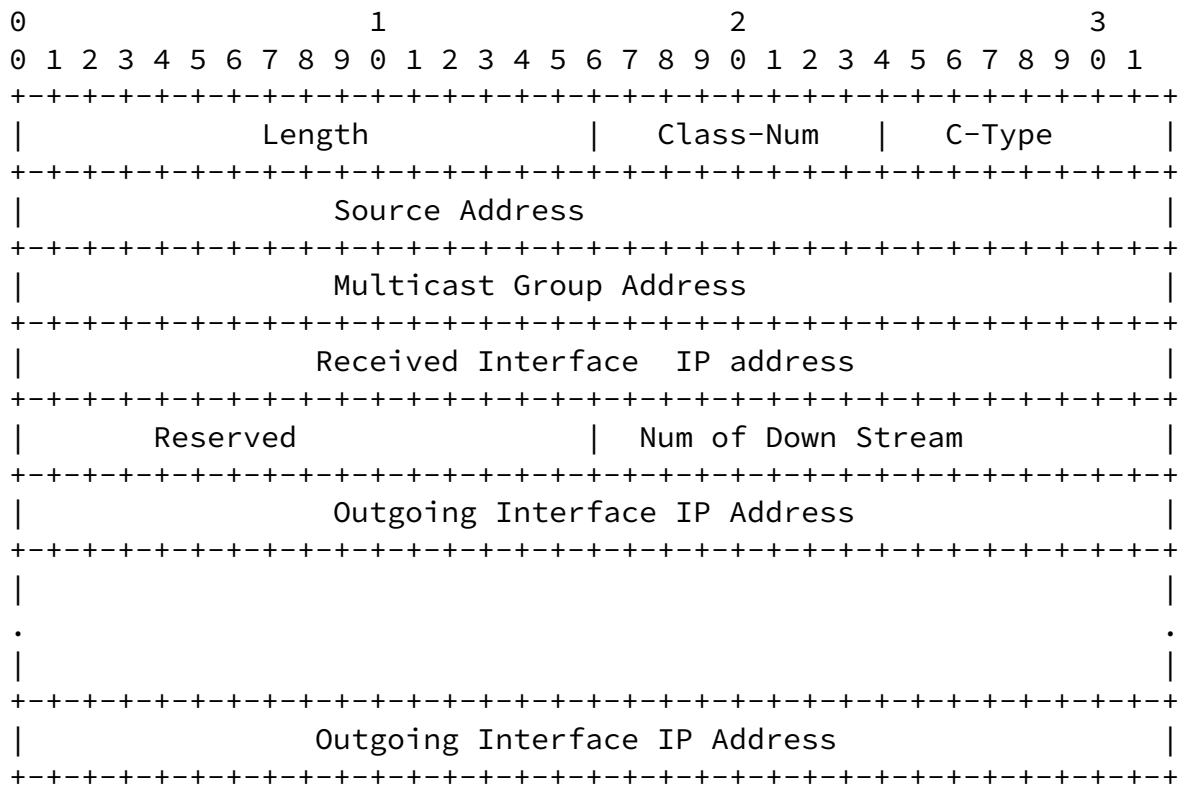


Figure 8 c-type RPF response for Data plane information

- o Length : 4+4+4+4+4n or 16+16+16+4+16n octets depending on IP address is IPv4 or IPv6. Where n is the number of Multicast Group addresses.
- o Source Address : 4 or 16 octets based on IPv4 or IPv6. Represents one of the source addresses specified in the RPF request c-type. Zero value indicate * notation in (*,G).
- o Multicast Group Address: 4 or 16 octets based on IPv4 or IPv6. Represents one of the group specified in the RPF request c-type. Source Address and Multicast Group Address pair MUST match exactly as specified in the RPF request c-type.
- o Received Interface IP address: 4 or 16 octets, indicates the IP

address of the received interface.

- o Reserved : (2 octets) Transmitted zero and ignored on receipt
- o Num of Down Stream: (2 octets) Number of Downstream neighbors.
- o Outgoing Interface IP Address: (4 or 16 octets) IP address of the outgoing Interfaces.

[3.5.2.](#) RPF Response of Control Plane information

The c-type defined in this section provide RPF information contained in the Control plane of the responding device for the specified (S,G).

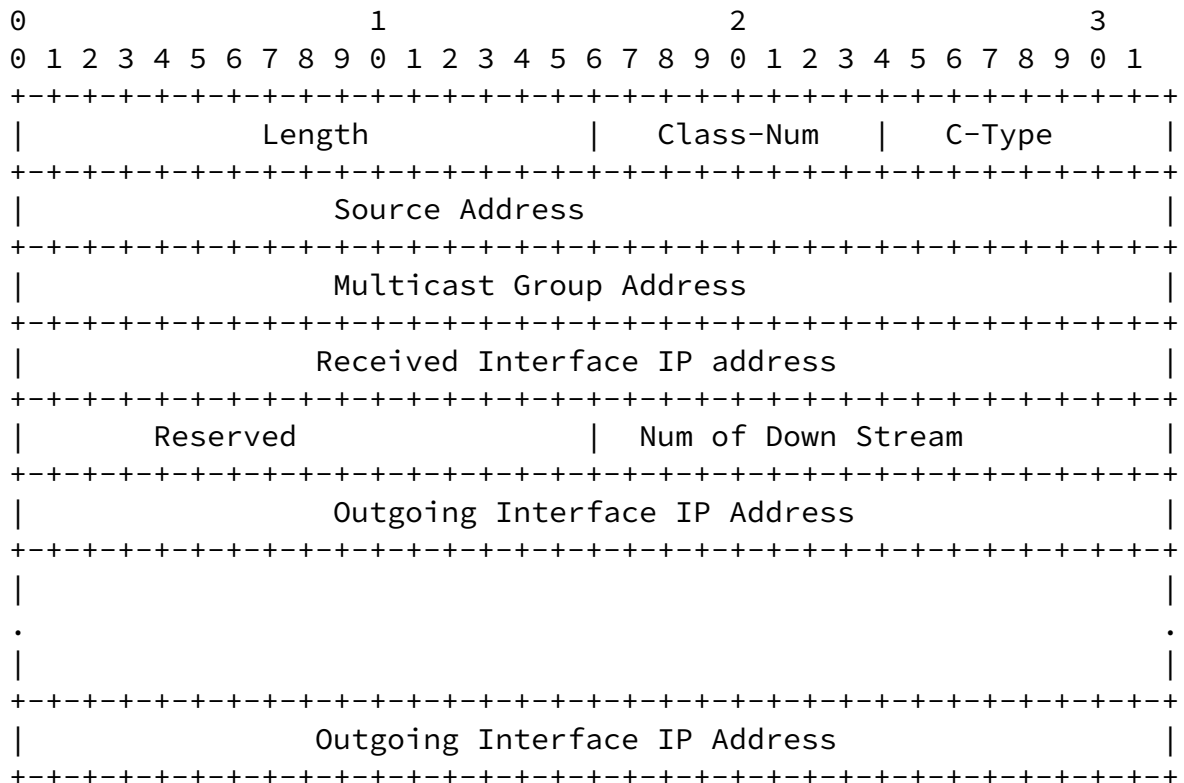


Figure 9 c-type RPF response for Control plane information

- o Length : 4+4+4+4+4n or 16+16+16+4+16n octets depending on IP address is IPv4 or IPv6. Where n is the number of Multicast Group addresses.
- o Source Address : 4 or 16 octets based on IPv4 or IPv6. Represents one of the source addresses specified in the RPF request c-type. Zero value indicate * notation in (*,G).
- o Multicast Group Address: 4 or 16 octets based on IPv4 or IPv6. Represents one of the group specified in the RPF request c-type. Source Address and Multicast Group Address pair MUST match exactly as specified in the RPF request c-type.

- o Received Interface IP address: 4 or 16 octets, indicates the IP address of the received interface.
- o Reserved : (2 octets) Transmitted zero and ignored on receipt
- o Num of Down Stream: (2 octets) Number of Downstream neighbors.
- o Outgoing Interface IP Address: (4 or 16 octets) IP address of the outgoing Interfaces.

[4.](#) Security Considerations

TBD

[5.](#) IANA Considerations

IANA is requested to provide a new object class to represent ICMP Object extension for multicast OAM.

IANA is requested to maintain a registry within the multicast ICMP extension object to include c-types defined under the multicast OAM object type.

c-types defined in this document are:

c-type name	number	Reference
Address Scope	1	3.1.2.
Originator IP Address	2	3.2.
Role Request	3	3.3.
Role Response Rendezvous Point	4	3.3.
Role Response Designated Router	5	3.3.
Incoming and Outgoing Interfaces	6	3.4.
RPF Information Request	7	3.5.
RPF Response Dataplane Infor	8	3.5.1.
RPF Response Control Plane Info	9	3.5.2.

Figure 10 C-types defined in this document

[6.](#) References

[6.1.](#) Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC4884] Bonica, R., et.al, "Extended ICMP to support multipart Messages", [RFC 4884](#), April 2007.
- [PINGEXT] Shen, N., et.al, "Traceroute and Ping Message Extensions", [draft-shen-traceroute-ping-ext-04](#), Work in Progress, February, 2012.

[6.2.](#) Informative References

- [RFC6450] Venaas, S. "Multicast Ping Protocol", [RFC 6450](#), December 2011.

[7.](#) Acknowledgments

<Add any acknowledgements>

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

Authors' Addresses

Tissa Senevirathne
CISCO Systems
375 East Tasman Drive
San Jose, CA 95134, USA

Phone: +1-408-853-2291
Email: tsenevir@cisco.com

Nataraj Bacthu
CISCO Systems
425 East Tasman Drive
San Jose, CA 95134, USA

Email: nbatchu@cisco.com

Senevirathne

Expires September 4, 2012

[Page 16]

Internet-Draft

[draft-tissa-pim-mcastoam-00](#)

March 2012

Raghava Sivaramu
CISCO Systems
425 East Tasman Drive
San Jose, CA 95134, USA

Email: raghavas@cisco.com

Sam Aldrin
HuaWei Technologies
2330 Central Expressway
Santa Clara, CA 95051, USA

Email: aldrin.ietf@gmail.com

Donald Eastlake
HuaWei Technologies

155 Beaver Street
Milford, MA 01757, USA

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com