

PIM Working Group
Internet Draft
Intended status: Standards Track
Expires: May 22, 2022

Yisong Liu
China Mobile
M. McBride
T. Eckert
Futurewei
Z. Zhang
ZTE
Nov 22, 2021

PIM Assert Message Packing
draft-ietf-pim-assert-packing-05

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 22, 2022.

Copyright Notice

Copyright (c) 2021 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

Internet-Draft

PIM Assert Packing

November 2021

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

In PIM-SM shared LAN networks, there is typically more than one upstream router. When duplicate data packets appear on the LAN from different routers, assert packets are sent from these routers to elect a single forwarder. The PIM assert packets are sent periodically to keep the assert state. The PIM assert packet carries information about a single multicast source and group, along with the metric-preference and metric of the route towards the source or RP. This document defines a standard to send and receive information for multiple multicast sources and groups in a single PIM assert message. This can be particularly helpful when there is traffic for a large number of multicast groups.

Table of Contents

1.	Introduction.....	3
1.1.	Requirements Language.....	3
1.2.	Terminology.....	3
2.	Use Cases.....	3
2.1.	Enterprise network.....	4
2.2.	Video surveillance.....	4
2.3.	Financial Services.....	4
2.4.	IPTV broadcast Video.....	4
2.5.	MVPN MDT.....	4
2.6.	Summary.....	5
3.	Solution.....	5
3.1.	PIM Assert Packing Hello Option.....	5
3.2.	PIM Assert Packing Simple Type.....	5
3.3.	PIM Assert Packing Aggregation Type.....	6
3.4.	PIM Assert Timer.....	6
3.5.	PIM Assert Format Selection.....	7
4.	Packet Format.....	7
4.1.	PIM Assert Packing Hello Option.....	7
4.2.	PIM Assert Simple Packing Format.....	8
4.3.	PIM (S,G) Assert Aggregation Packing Format.....	9
4.4.	PIM (*,G) Assert Aggregation Packing Format.....	11

5.	IANA Considerations.....	14
6.	Security Considerations.....	14
7.	References.....	15
7.1.	Normative References.....	15
7.2.	Informative References.....	15

8.	Acknowledgments.....	15
	Authors' Addresses.....	16

[1.](#) Introduction

In PIM-SM shared LAN networks, there is typically more than one upstream router. When duplicate data packets appear on the LAN, from different upstream routers, assert packets are sent from these routers to elect a single forwarder according to [\[RFC7761\]](#). The PIM assert packets are sent periodically to keep the assert state. The PIM assert packet carries information about a single multicast source and group, along with the corresponding metric-preference and metric of the route towards the source or RP.

This document defines a standard to send and receive information for multiple multicast sources and groups in a single PIM assert message. It can efficiently pack multiple PIM assert messages into a single message and reduce the processing pressure of the PIM routers. This can be particularly helpful when there is traffic for a large number of multicast groups.

[1.1.](#) Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [\[RFC2119\]](#) [\[RFC8174\]](#) when, and only when, they appear in all capitals, as shown here.

[1.2.](#) Terminology

RPF: Reverse Path Forwarding

RP: Rendezvous Point

SPT: Shortest Path Tree

RPT: RP Tree

DR: Designated Router

BDR: Backup Designated Router

[2.](#) Use Cases

PIM Assert will happen in many services where multicast is used and not limited to the examples described below.

Liu, et al.

Expires May, 2022

[Page 3]

Internet-Draft

PIM Assert Packing

November 2021

[2.1.](#) Enterprise network

When an Enterprise network is connected through a layer-2 network, the intra-enterprise runs layer-3 PIM multicast. The different sites of the enterprise are equivalent to the PIM connection through the shared LAN network. Depending upon the locations and amount of groups there could be many asserts on the first hop routers.

[2.2.](#) Video surveillance

Video surveillance deployments have migrated from analog based systems to IP-based systems oftentimes using multicast. In the shared LAN network deployments, when there are many cameras streaming to many groups there may be issues with many asserts on first hop routers.

[2.3.](#) Financial Services

Financial services extensively rely on IP Multicast to deliver stock market data and its derivatives, and current multicast solution PIM is usually deployed. As the number of multicast flows grows, there are many stock data with many groups may result in many PIM asserts on a shared LAN network from publisher to the subscribers.

[2.4.](#) IPTV broadcast Video

PIM DR and BDR deployments are often used in host-side network for IPTV broadcast video services. Host-side access network failure scenario may be benefitted by assert packing when many groups are being used. According to [\[RFC7761\]](#) the DR will be elected to forward multicast traffic in the shared access network. When the DR recovers from a failure, the original DR starts to send traffic, and the current DR is still forwarding traffic. In the situation multicast

traffic duplication maybe happen in the shared access network and can trigger the assert progress.

[2.5.](#) MVPN MDT

As described in [[RFC6037](#)], MDT (Multicast Distribution Tree) is used as tunnels for MVPN. The configuration of multicast-enabled VRF (VPN routing and forwarding) or interface that is in a VRF changing may cause many assert packets to be sent in a same time.

Liu, et al.

Expires May, 2022

[Page 4]

Internet-Draft

PIM Assert Packing

November 2021

[2.6.](#) Summary

In the above scenarios, the existence of PIM assert state depends mainly on the network topology. As long as there is a layer 2 network between PIM neighbors, there may be multiple upstream routers, which can cause duplicate multicast traffic to be forwarded and assert process to occur.

Moreover as the multicast services become widely deployed, the number of multicast entries increases, and a large number of assert messages may be sent in a very short period when multicast data packets trigger PIM assert processing in the shared LAN networks. The PIM routers need to process a large number of PIM assert small packets in a very short time. As a result, the device load is very large. The assert packet may not be processed in time or even is discarded, thus extending the time of traffic duplication in the network.

Additionally, future backhaul, or fronthaul, networks may want to connect L3 across an L2 underlay supporting Time Sensitive Networks (TSN). The infrastructure may run DetNet over TSN. These transit L2 LANs would have multiple upstreams and downstreams. This document is taking a proactive approach to prevention of possible future assert issues in these types of environments.

[3.](#) Solution

The change to the PIM assert includes two elements: the PIM assert

packing hello option and the PIM assert packing method.

There is no change required to the PIM assert state machine. Basically a PIM router can now be the assert winner or loser for multiple packed (S, G)'s in a single assert message instead of one (S, G) assert at a time.

[3.1.](#) PIM Assert Packing Hello Option

The newly defined Hello Option is used by a router to negotiate the assert message packing capability. Assert packing can only be used when all PIM routers, in the same shared LAN network, support this capability. This document defines two packing methods. One method is a simple merge of the original messages and the other is to extract the common message fields for aggregation.

[3.2.](#) PIM Assert Packing Simple Type

In this type of packing, as described in [\[RFC7761\]](#), the assert message body is used as a record. The newly defined assert message

Liu, et al.

Expires May, 2022

[Page 5]

Internet-Draft

PIM Assert Packing

November 2021

can carry multiple assert records and identify the number of records.

This packing method is simply extended from the original assert packet, but, because the multicast service deployment often uses a small number of sources and RPs, there may be a large number of assert records with the same metric preference or route metric field, which would waste the payload of the transmitted message.

[3.3.](#) PIM Assert Packing Aggregation Type

When the source or RP addresses, in the actual deployment of the multicast service, are very few, this type of packing will combine the records related to the source address or RP address in the assert message.

* A (S, G) assert only can contain one SPT (S, G) entry, so it can be aggregated according to the same source address, and then all SPT (S, G) entries corresponding to the same source address are merged into one assert record.

* A (*, G) assert may contain a (*, G) entry or a RPT (S, G) entry, and both entry types actually depend on the route to the RP. So it

can be aggregated further according to the same RP address, and then all (*, G) and RPT (S, G) entries corresponding to the same RP address are merged into one assert record.

This method can optimize the payload of the transmitted message by merging the same field content, but will add the complexity of the packet encapsulation and parsing.

[3.4.](#) PIM Assert Timer

As described in [section 4.6 in RFC7761](#), the Assert Timer function of each (S,G) and (*,G) is not changed. After the Assert Message Packing function defined in this document is enabled, when the first AT of a (S,G) or (*,G) is expired, in order to carry much more assert information in this message, some of other (S,G) or (*,G) may be included in the same message, even though their ATs have not been expired. After the assert message packing, the ATs of all the (S,G) or (*,G), that are packed in the same message, are all reset. That is after the packed assert messages is sent, the ATs of the packed (S,G) or (*,G) will be set with the same value.

[3.5.](#) PIM Assert Format Selection

An implementation MUST NOT send assert messages using a packing type, unless all routers on the LAN have indicated support for the type. If both packing types are supported, then it is left to the implementation whether to use assert packing and which packing type to use. It is RECOMMENDED to use the supported method that is most efficient. The Aggregation Packing Format is likely to be the most efficient if the assert message is to include multiple records having the same source address or RP address.

The regular [RFC7761](#) assert format is still allowed to be used. For example the assert only needs to be sent for a single (S, G).

[4.](#) Packet Format

This section describes the format of new PIM messages introduced by this document.

[4.1.](#) PIM Assert Packing Hello Option

```

      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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      OptionType = TBD      |      OptionLength = 1      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Packing_Type |
+---+---+---+---+---+---+

```

- OptionType: TBD
- OptionLength: 1
- Packing_Type: The specific packing mode is determined by the value of this field:

When the first bit from the right is set to 1: indicates simple packing type as described in [section 2.2](#)

When the second bit from the right is set to 1: indicates aggregating packing type as described in [section 2.3](#)

3-8 bits: reserved for future

The node may support multiple packing types. The node MUST set the bits indicating which types they support. They MUST set both bits if they support both type 1 and type 2. The format used MUST be a format supported by all routers on the LAN, see [section 3.5](#) for details.

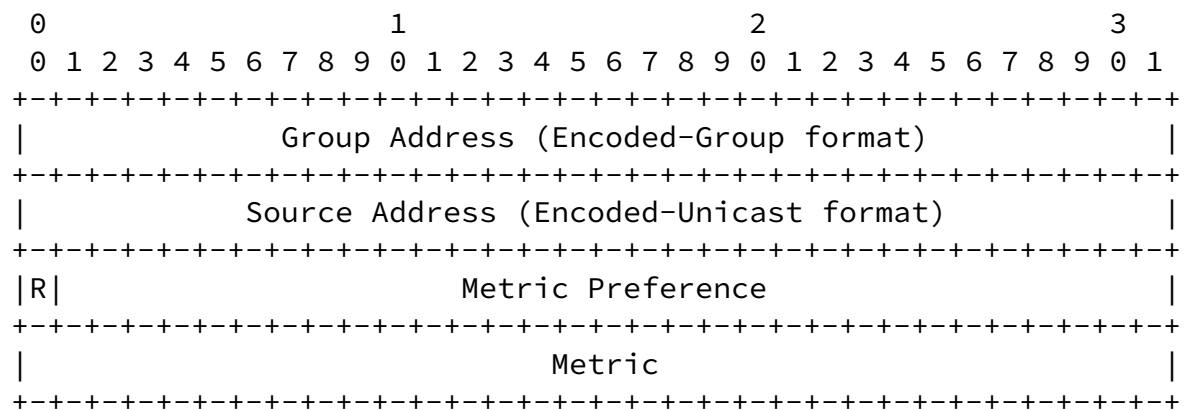
Once the packing format type is selected, this type of coding is used for Assert message packing.

If not all nodes support a same packing format, then Assert message format defined in [[RFC7761](#)] is used.

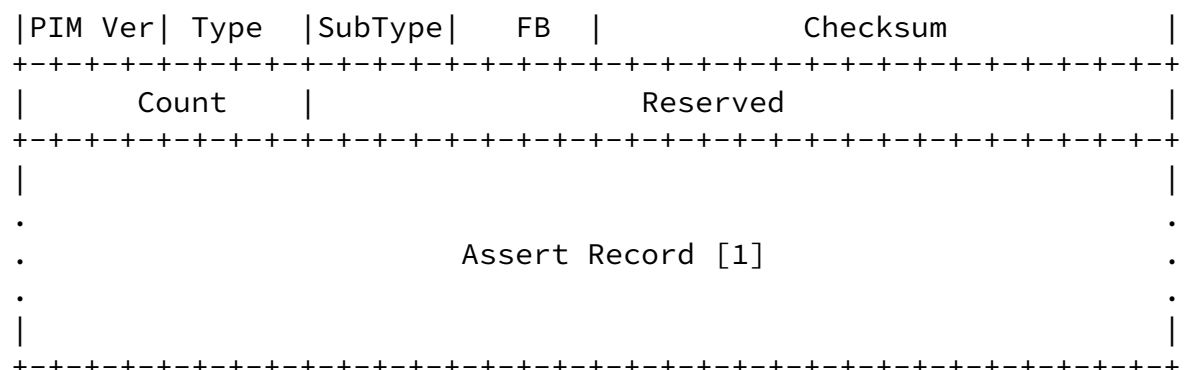
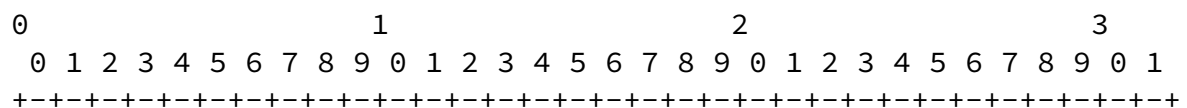
[4.2.](#) PIM Assert Simple Packing Format

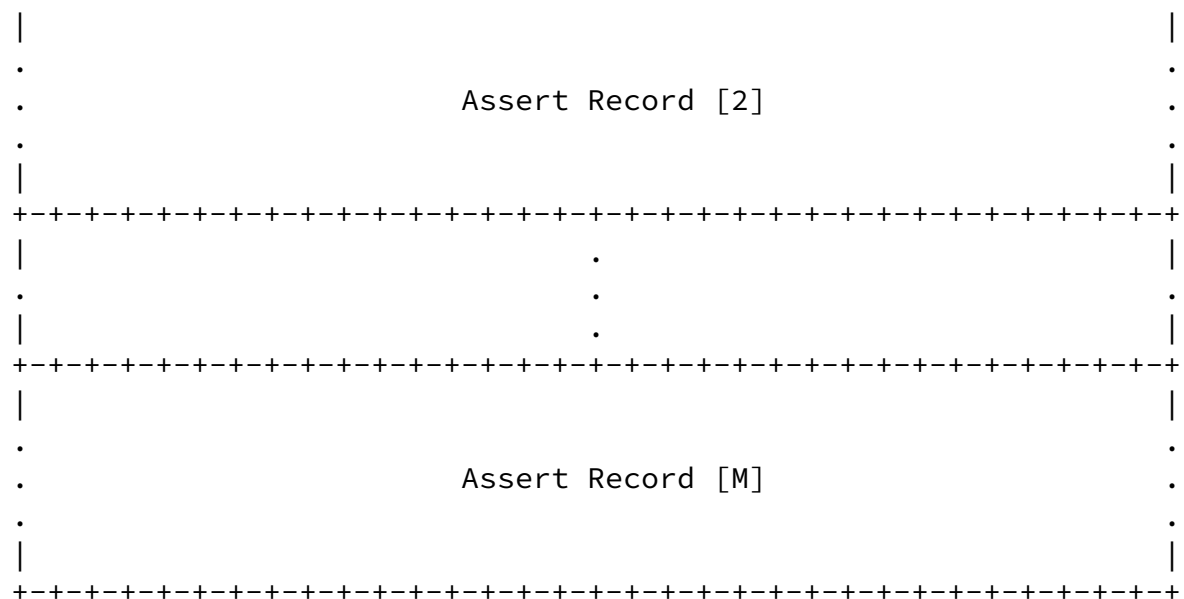
Count

The format of each record is the same as the PIM assert message body of [section 4.9.6 in \[RFC7761\]](#).



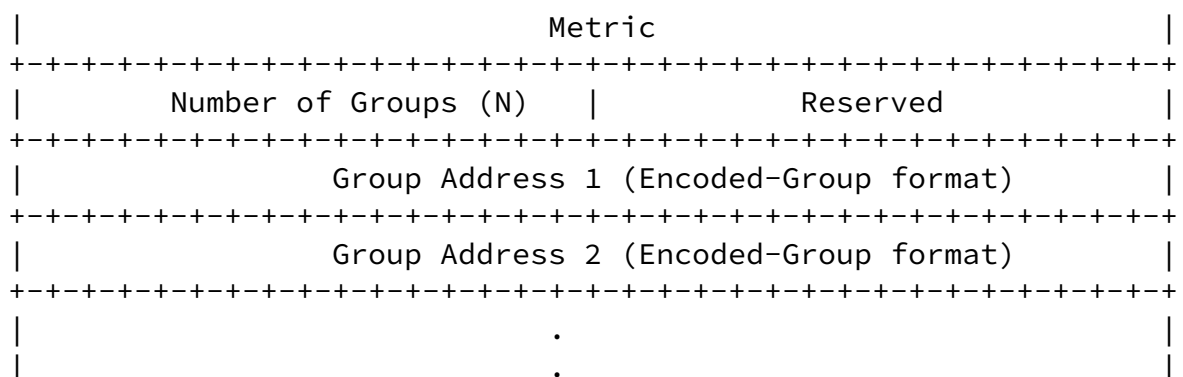
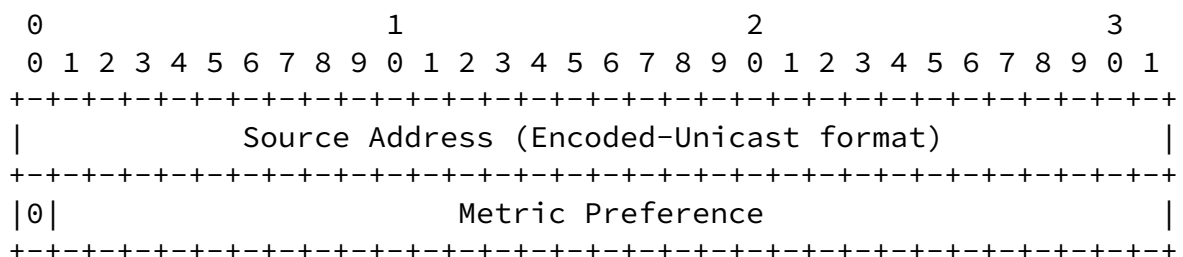
This method also extends PIM assert messages to carry multiple records. But the records are divided for (S, G) and (*, G).

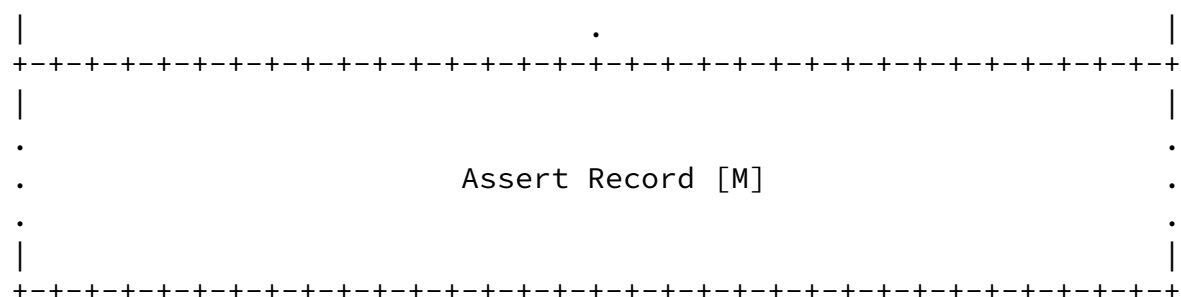




Most fields of the specific assert message format is the same as [section 4.2](#), except for the subType fields and records. When aggregated (S, G) records is carried in the message, the subType field is set to TBD3.

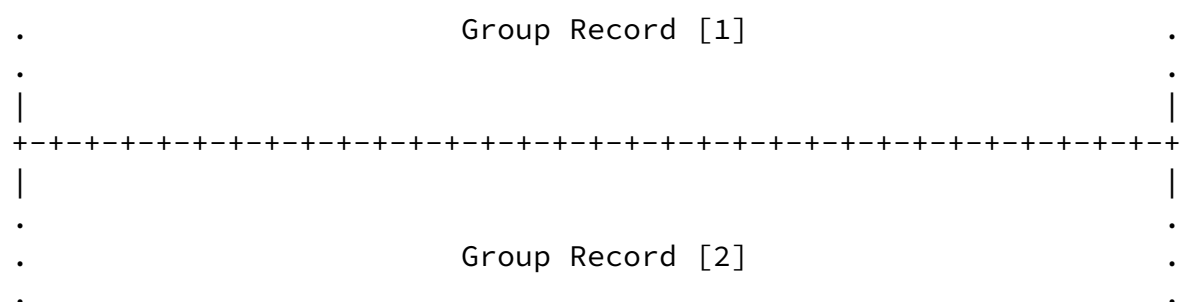
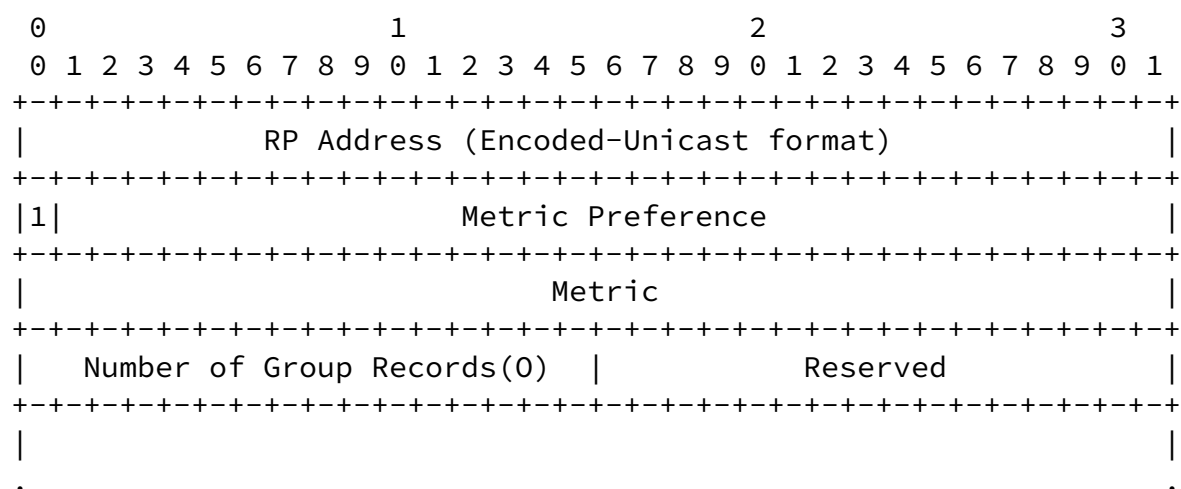
The (S, G) assert records are organized by the same source address, and the specific message format is:

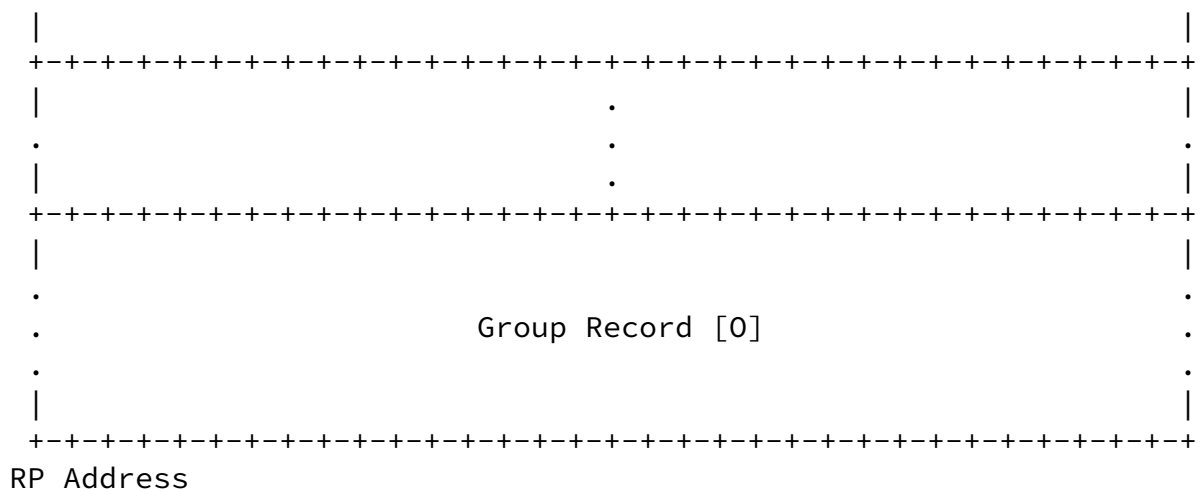




Most fields of the specific assert message format is the same as [section 4.2](#), except for the subType fields and records. When aggregated (*, G) records is carried in the message, the subType field is set to TBD4.

The (*, G) assert records are organized in the same RP address and are divided into two levels of TLVs. The first level is the group record of the same RP address, and the second level is the source record of the same multicast group address, including (*, G) and RPT (S, G), and the specific message format is:





The address of RP corresponding to all of the contained group records. The format for this address is given in the encoded unicast address in [\[RFC7761\] Section 4.9.1](#)

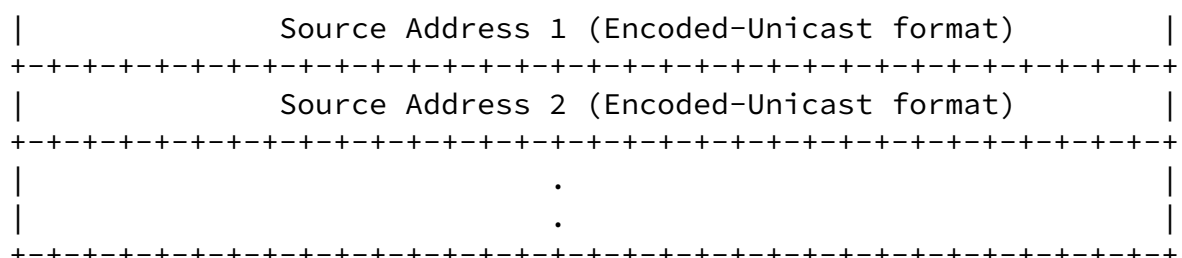
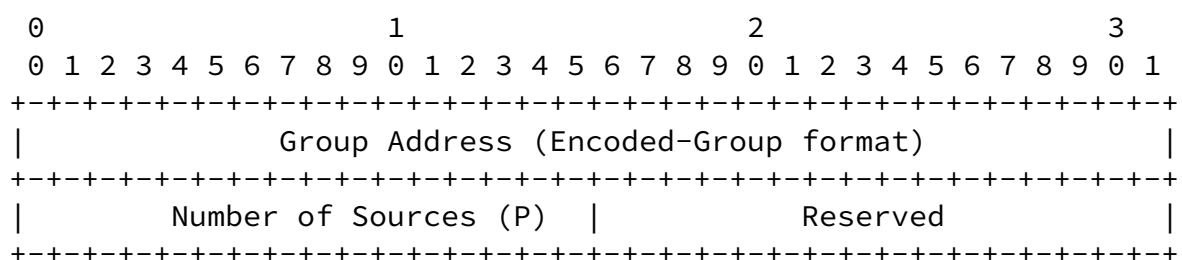
Metric Preference, Metric and Reserved

Same as [\[RFC7761\] Section 4.9.6](#)

Number of Group Records

The number of packed group records. A record consists of a group address and a source address list.

The format of each group record is:



	Source Address P (Encoded-Unicast format)	
+	+	+

Group Address and Reserved

Same as [\[RFC7761\] Section 4.9.6](#)

Number of Sources

The number of source addresses corresponding to the group address field in the group record.

Source Address

Same as [\[RFC7761\] Section 4.9.6](#), but there are multiple source addresses in the group record.

5. IANA Considerations

This document requests IANA to assign a registry for PIM assert packing Hello Option in the PIM-Hello Options and new PIM assert packet type and subtype. The assignment is requested permanent for IANA when this document is published as an RFC. The string TBD should be replaced by the assigned values accordingly.

6. Security Considerations

As described in [section 6.1 in RFC7761](#), the forged assert message may cause the legitimate designated forwarder to stop forwarding traffic to the LAN. And the packed function defined in this document, may make this situation worse, because it will affect multiple flows with a single packed assert. That is in case one forged packed assert message is accept, all the (S,G) or (*,G) carried in this message may be stopped forwarding from one or more legitimate designated forwarders. The general security function, such as authentication function defined in [RFC7761](#), or the necessary PIM filtering method, will do good help to avoid the forged assert message.

Liu, et al.

Expires May, 2022

[Page 14]

Internet-Draft

PIM Assert Packing

November 2021

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification(Revised)", [RFC 7761](#), March 2016
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), May 2017
- [RFC8736] Venaas, S., Retana, A., "PIM Message Type Space Extension and Reserved Bits", [RFC 8736](#), February 2020.

[7.2](#). Informative References

- [RFC6037] Rosen, E., Cai, Y., Wijnands, I., "Cisco Systems' Solution for Multicast in BGP/MPLS IP VPNs", [RFC6037](#), October 2010

[8](#). Acknowledgments

The authors would like to thank Stig Venaas for the valuable contributions of this document.

Authors' Addresses

Yisong Liu
China Mobile
China

Email: liuyisong@chinamobile.com

Mike McBride
Futurewei
USA

Email: michael.mcbride@futurewei.com

Toerless Eckert
Futurewei
USA

Email: tte+ietf@cs.fau.de

Zheng(Sandy) Zhang
ZTE Corporation
China

Email: zhang.zheng@zte.com.cn