PIM Assert Message Packing
draft-liu-pim-assert-packing-02

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/1id-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 August 6, 2020.

Copyright Notice

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 multiple
   multicast source and group information in a single PIM assert packet
   in a shared network. This can be particularly helpful when there is
   traffic for a large number of multicast groups.

Table of Contents

[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 multiple
multicast source and group information in a single PIM assert packet
in a shared LAN network. It can efficiently pack multiple PIM assert
packets 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.

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. Summary

In the above scenarios, the existence of PIM assert process 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 process 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 packet instead of one
(S, G) assert at a time. An assert winner is now responsible for
forwarding traffic from multiple (S, G)'s out of a particular
interface based upon the multiple (S, G)'s packed in a single
assert.

3.1. PIM Assert Packing Hello Option

The newly defined Hello Option is used by a router to negotiate the
assert packet packing capability. It 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, the original assert message body is used as
a record. The newly defined assert message 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.

4. Packet Format

This section describes the format of new PIM messages introduced by
this document. The messages follow the same transmission order as
the messages defined in [RFC7761].

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 |
+-+-+-+-+-+-+-+-+
```
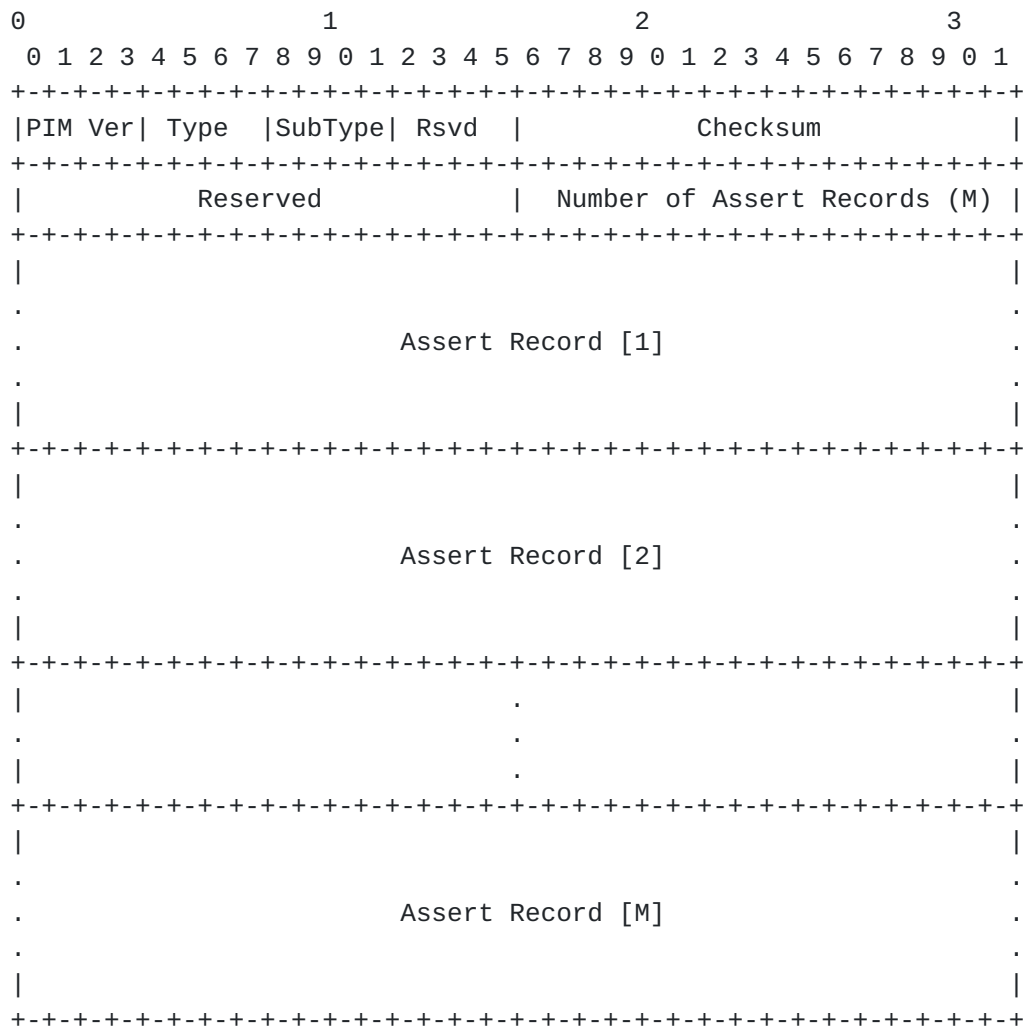
- OptionType: TBD

- OptionLength: 1

- Packing_Type: The specific packing mode is determined by the value

    of this field:

1: indicates simple packing type as described in section 2.2

2: indicates aggregating packing type as described in section 2.3

3-255: reserved for future

4.2. PIM Assert Simple Packing Format

```
 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  |SubType| Rsvd  |            Checksum           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Reserved          |  Number of Assert Records (M) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
.                                                               .
.                       Assert Record [1]                       .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
.                                                               .
.                       Assert Record [2]                       .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              .                                |
.                              .                                .
|                              .                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
.                                                               .
.                       Assert Record [M]                       .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

PIM Version, Reserved, Checksum

Same as [RFC7761] Section 4.9.6

Type

The new Assert Type and SubType values TBD

Number of Assert Records

The number of packed assert records. A record consists of a single assert message body.

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

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Group Address (Encoded-Group format)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Source Address (Encoded-Unicast format)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|                    Metric Preference                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Metric                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

4.3. PIM Assert Aggregation Packing Format

This method also extends PIM assert packets to carry multiple records. The specific assert packet format is the same as section 4.2, but the records are divided into two types.

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

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Source Address (Encoded-Unicast format)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|                    Metric Preference                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Metric                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Number of Groups (N)    |            Reserved          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Group Address 1 (Encoded-Group format)          |
```

```
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |             Group Address 2 (Encoded-Group format)           |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                            .                                 |
  |                            .                                 |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |             Group Address N (Encoded-Group format)           |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Source Address, Metric Preference, Metric and Reserved

    Same as [RFC7761] Section 4.9.6, but the source address MUST NOT
be set to zero.

Number of Groups

    The number of group addresses corresponding to the source address
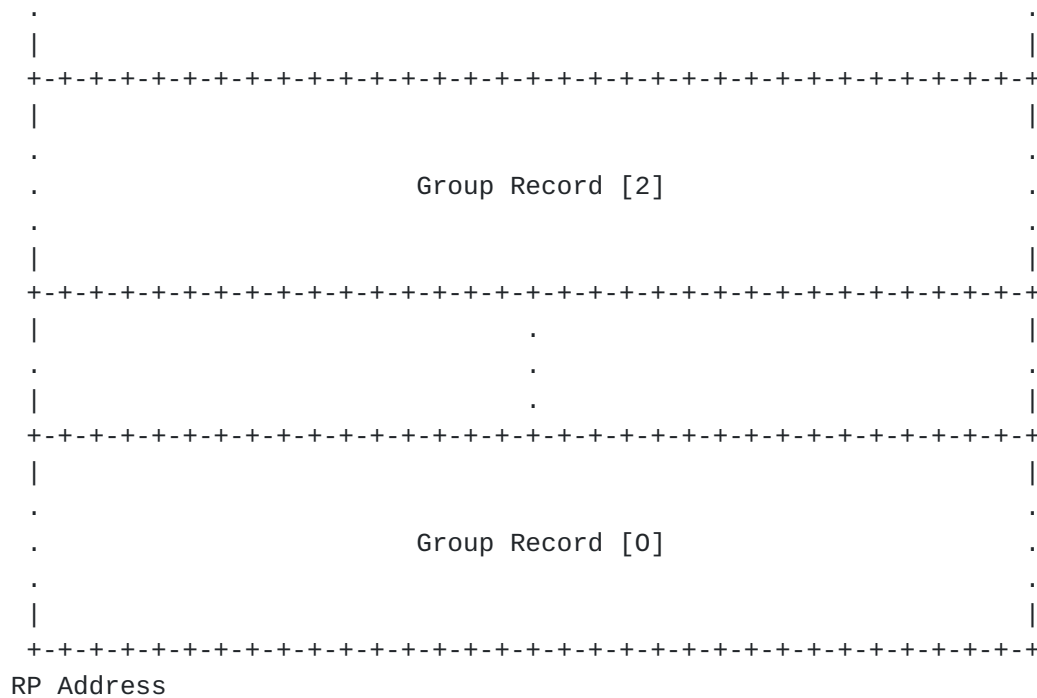field in the (S, G) assert record.

Group Address

    Same as [RFC7761] Section 4.9.6, but there are multiple group
addresses in the (S, G) assert record

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:

```
  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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |             RP Address (Encoded-Unicast format)              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |1|                    Metric Preference                       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                          Metric                              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |   Number of Group Records(O)  |           Reserved           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                              |
 .                                                              .
 .                       Group Record [1]                       .
```

```
             .                                              .
             |                                              |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                                              |
       .                                              .
       .                 Group Record [2]             .
       .                                              .
       |                                              |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                        .                     |
       .                        .                     .
       |                        .                     |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                                              |
       .                                              .
       .                 Group Record [O]             .
       .                                              .
       |                                              |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

RP Address

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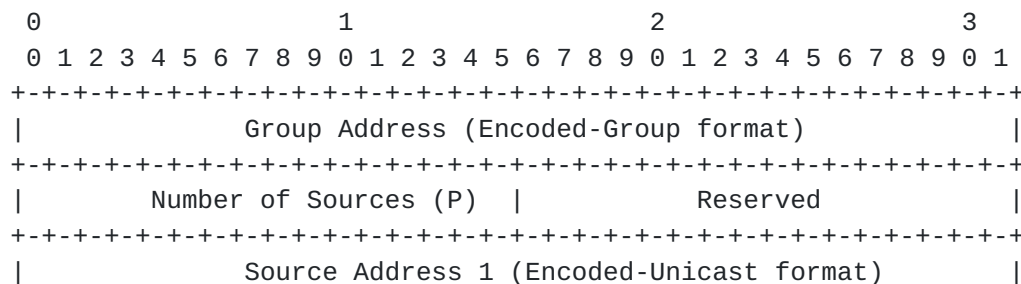unicast address in [RFC7761] Section 4.9.1

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:

```
  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
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |             Group Address (Encoded-Group format)             |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |       Number of Sources (P)   |            Reserved          |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |             Source Address 1 (Encoded-Unicast format)        |
```

```
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                 Source Address 2 (Encoded-Unicast format)     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                               .                               |
   |                               .                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                 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

   For general PIM-SM protocol Security Considerations, see [RFC7761].

   TBD

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
             IndependentMulticast - 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

7.2. Informative References

   TBD

8. Acknowledgments

   The authors would like to thank the following for their valuable
   contributions of this document:

   TBD

Authors' Addresses

    Yisong Liu
    China Mobile

    Email: liuyisong@chinamobile.com

    Mike McBride
    Futurewei

    Email: michael.mcbride@futurewei.com

    Toerless Eckert
    Futurewei

    Email: tte+ietf@cs.fau.de