

Network Working Group
Internet-Draft
Expires: May 3, 2009

D. Ward
R. White
D. Farinacci
A. Banerjee
Cisco Systems
R. Perlman
Sun Microsystems
November 3, 2008

Carrying Attached Addresses in IS-IS
draft-ward-l2isis-04

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of 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 3, 2009.

Abstract

This draft specifies the IS-IS extensions necessary to support multi-link IPv4 and IPv6 networks, as well as to provide true link state routing to any protocols running directly over layer 2. While supporting this concept involves several pieces, this document only describes extensions to IS-IS. We leave it to the systems using these IS-IS extensions to explain how the information carried in IS-IS is used.

1. Overview

There are a number of systems (for example, [[RBRIDGES](#)]) which have proposed using layer 2 addresses carried in a link state routing protocol, specifically IS-IS [[IS-IS](#)] [[RFC1195](#)], to provide true layer 2 routing in specific environments. This draft proposes a set of TLVs and sub-TLVs to be added to [[IS-IS](#)] level 1 PDUs, and three new PDU types, to support these proposed systems.

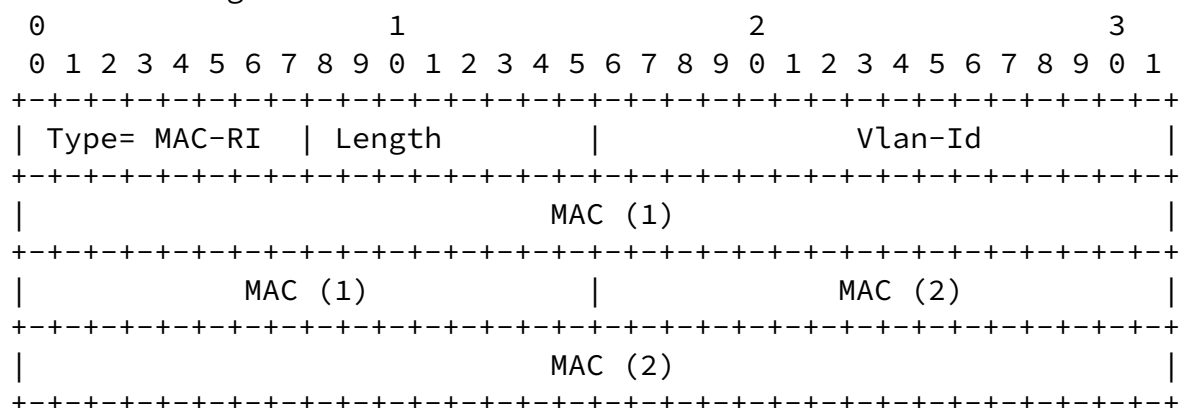
This draft does not propose new forwarding mechanisms using this additional information carried within IS-IS. There is a short section included on two possible ways to build a shortest path first tree including this information, to illustrate how this information might be used.

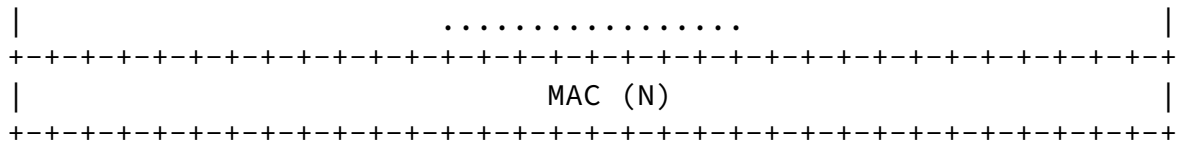
2. Proposed Enhancements to IS-IS

This draft proposes additional TLVs, to carry unicast and multicast attached address information. It also proposes additional sub-tlvs to carry information regarding building trees for Layer 2 networks. This draft proposes three new IS-IS PDUs, the Multicast Group (MGROUP) PDU, for carrying a list of attached or joined multicast groups. The Multicast Group Complete Sequence Number (MGROUP-CSNP) PDU and the Multicast Group Partial Sequence Number (MGROUP-PSNP) PDU packets are also defined to be used with the new MGROUP-PDU to perform database exchange on the MGROUP PDU packets.

2.1. The MAC-Reachability TLV

The MAC-Reachability (MAC-RI) sub-TLV is IS-IS TLV type 141 and has the following format:



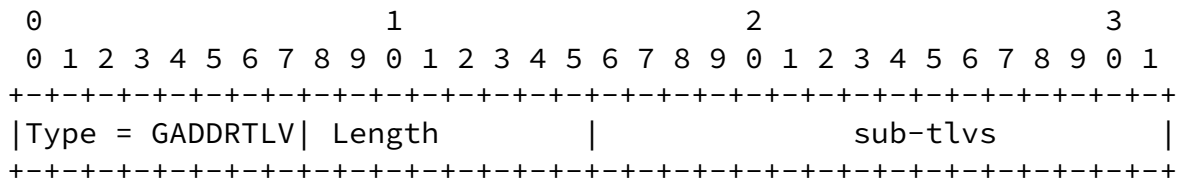


- o Type: TLV Type, set to 141 (MAC-RI).
- o Length: Total number of octets contained in the TLV.
- o Vlan-id: This carries a 16 bit VLAN identifier that is valid for all subsequent MAC addresses in this TLV.
- o MAC(i): This is the 48-bit MAC address reachable from the IS that is announcing this TLV.

The MAC-RI TLV is carried in a standard Level 1 link state PDU. It MUST contain only unicast addresses.

2.2. The Group Address TLV

The Group Address (GADDR) TLV is IS-IS TLV type 142 [TBD] and has the following format:



- o Type: TLV Type, set to GADDR-TLV 142 [TBD].
- o Length: Total number of octets contained in the TLV, including the length of the sub-tlvs carried in this TLV.
- o sub-tlvs: The following sub-TLVs are defined.

The GADDR TLV is carried within Multicast Group Level 1 link state PDU.

2.2.1. The Group MAC Address sub-TLV

The Group MAC Address (GMAC-ADDR) sub-TLV is IS-IS TLV type 1 and has the following format:



```

| Length | (1 byte)
+-----+
| Vlan-Id | (2 byte)
+-----+
| Num Group Recs | (1 byte)
+-----+
| GROUP RECORDS (1) |
+-----+
| ..... |
+-----+
| GROUP RECORDS (N) |
+-----+

```

where each group record is of the form:

```

+-----+
| RESERVED | (1 byte)
+-----+
| Num of Sources | (1 byte)
+-----+
| Group Address | (6 bytes)
+-----+
| Source 1 Address | (6 bytes)
+-----+
| Source 2 Address | (6 bytes)
+-----+
| Source M Address | (6 bytes)
+-----+

```

- o Type: TLV Type, set to 1 (GMAC-ADDR) of length 1 byte.
- o Length: Total number of octets contained in the TLV.
- o Vlan-id: This carries a 16 bit VLAN identifier that is valid for all subsequent MAC addresses in this TLV.
- o Number of Group Records: This is of length 1 byte and lists the number of group records in this TLV.
- o Group Record: Each group record has a reserved space and followed by the number of sources, each of length 1 byte. It then has a 48-bit multicast Group Address followed by 48-bit source MAC addresses. An address being a group multicast address or unicast source address can be checked using the multicast bit in the address.

The GMAC-ADDR sub-TLV is carried within the GADDR TLV and MUST be carried in a standard Level 1 link state MGROUP PDU.

[2.2.2.](#) The Group IP Address sub-TLV

The Group IP Address (GIP-ADDR) sub-TLV is IS-IS TLV type 2 and has the following format:

```

+---+---+---+---+
| Type=GIP-ADDR |
+---+---+---+---+
|   Length       |                               (1 byte)
+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Vlan-Id           | (2 byte)
+---+---+---+---+---+---+---+---+---+---+---+---+
| Num Group Recs |                               (1 byte)
+---+---+---+---+---+---+---+---+---+---+---+---+
|                               GROUP RECORDS (1) |
+---+---+---+---+---+---+---+---+---+---+---+---+
|                               .....             |
+---+---+---+---+---+---+---+---+---+---+---+---+
|                               GROUP RECORDS (N) |
+---+---+---+---+---+---+---+---+---+---+---+---+

```

where each group record is of the form:

```

+---+---+---+---+
|   RESERVED     | (1 byte)
+---+---+---+---+
| Num of Sources | (1 byte)

```

```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Group Address                (4 bytes) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Source 1 Address             (4 bytes) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Source 2 Address             (4 bytes) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Source M Address             (4 bytes) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- o Type: TLV Type, set to 2 (GIP-ADDR).
- o Length: Total number of octets contained in the TLV.
- o Vlan-id: This carries a 16 bit VLAN identifier that is valid for all subsequent IPv4 source or group addresses in this TLV.
- o Number of Group Records: This is of length 1 byte and lists the number of group records in this TLV.
- o Group Record: Each group record has a reserved space and followed by the number of sources, each of length 1 byte. It is followed by a 32-bit IPv4 Group Address followed by 32-bit source IPv4 addresses.

The GIP-ADDR TLV is carried within the GADDR TLV and MUST be carried in a standard Level 1 link state MGROUP PDU.

[2.2.3.](#) The Group IPv6 Address sub-TLV

The Group IPv6 Address (GIPV6-ADDR) TLV is IS-IS sub-TLV type 3 and has the following format:

```

+---+---+---+---+---+---+
|Type=GIPv6-ADDR|
+---+---+---+---+---+---+
|   Length       |                               (1 byte)
+---+---+---+---+---+---+
|           Vlan-Id           | (2 byte)
+---+---+---+---+---+---+
|Num Group Recs |                               (1 byte)
+---+---+---+---+---+---+
|                               GROUP RECORDS (1) |

```

```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |
|                                     .....
|                                     |
|                               GROUP RECORDS (N) |
|                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where each group record is of the form:

```

+---+---+---+---+---+
|   RESERVED   |   (1 byte)
+---+---+---+---+---+
| Num of Sources |   (1 byte)
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Group Address           (16 bytes) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Source 1 Address        (16 bytes) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Source 2 Address        (16 bytes) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Source M Address        (16 bytes) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- o Type: TLV Type, set to 3 (GIPV6-ADDR).
- o Length: Total number of octets contained in the TLV.
- o Vlan-id: This carries a 16 bit VLAN identifier that is valid for all subsequent IPv6 source or group addresses in this TLV.
- o Number of Group Records: This of length 1 byte and lists the number of group records in this TLV.
- o Group Record: Each group record has a reserved space and followed by the number of sources, each of length 1 byte. It is followed by a 128-bit multicast IPv6 Group Address followed by 128-bit source IPv6 addresses.

The GIPV6-ADDR sub-TLV is carried within the GADDR TLV and MUST be carried in a standard Level 1 link state MGROUP PDU.

2.3. Sub-TLVs for the Capability TLV

The Capability TLV is an optional TLV [[RFC 4971](#)] that may be generated by the originating IS. We introduce these additional sub-TLVs that are carried within it. These sub-tlvs announce the

capabilities of the router for the entire IS-IS routing domain.

2.3.1. The Device ID sub-TLV

The Device ID (DEVID) sub-TLV carries information about the identity of the advertising device, along with information about device priority. The Device-Id sub-TLV MUST be carried within the CAPABILITY TLV in a level-1 non-pseudo-node LSP generated by the originating 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								
Type										Length										Reserved										Priority									
Reserved										Device Id																													

- o Type: TLV Type, set to 5 (DEVID).
- o Length: Total number of octets contained in the TLV.
- o Reserved: Set to 0.
- o Priority: Set to application dependent values.
- o Device ID: Left padded device ID or alias.

2.3.2. The Root Priority sub-TLV

The Root Priority sub-TLV MUST be carried within the CAPABILITY TLV in a level-1 non-pseudo-node LSP generated by the originating IS. Each device announces a broadcast root-priority and the number of trees it expects all other nodes to compute if it does become the broadcast root. Once a node receives a new LSP, it runs an election algorithm, independently of the other nodes in the network, to determine the broadcast root. The node that announced the lowest broadcast priority becomes the root of the broadcast tree. If two devices advertise the same broadcast priority, the device with the lower system ID becomes the root of the broadcast tree. The elected broadcast-root decides on the multicast-roots to be used in the network domain and their roots. This announcement takes place in the roots identifier sub-TLV.


```

|Type = ROOT-PRI|
+---+---+---+---+---+---+
|   Length           | (1 byte)
+---+---+---+---+---+---+
|   Broadcast Root Pri   | (2 byte)
+---+---+---+---+---+---+
|Num of multi-destination trees | (2 byte)
+---+---+---+---+---+---+

```

- o Type: TLV Type, set to 6 (ROOT-PRI).
- o Length: Total number of octets contained in the TLV.
- o Br Root Pri: This gives the value of the priority with which this node wants to be the broadcast root node in the Layer-2 domain.
- o Num of multi-destination trees: This gives the number of distribution trees for multi-destination frames that will be in use in the Layer-2 domain, excluding the broadcast tree rooted at itself, if this device becomes the broadcast root in the domain.

2.3.3. The Root Identifier Sub-tlv

The root identifier sub-tlv is populated by the root of the broadcast tree. If this is also announced by other nodes in the network, it implies that the specific node that is advertising it will only restrict traffic to the common set of the trees in its announcement and the ones announced by the broadcast root. It is carried within the CAPABILITY TLV in a level-1 non-pseudo-node LSP and is given as:

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|Type= ROOT-IDs | Length           | Broadcast Root System Id... |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               ... Broadcast Root System Id               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               Multicast Root System Id ...               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| ... Multicast Root System Id | ...
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

- o Type: TLV Type, set to 7 (ROOT-IDs).
- o Length: Total number of octets contained in the TLV.
- o Broadcast Root System Id: The Broadcast Root System ID at which a broadcast tree is rooted. It is of length 6 bytes.
- o Multicast Root System Id: The Multicast Root System ID at which a multicast tree is rooted. It is of length 6 bytes.

A locally significant hash is used by edge devices to determine which multicast root (or set of multicast roots) is used to send traffic for a specific multicast group. If there is a discrepancy between

the number of multi-destination trees the broadcast-root has announced, and the number of roots the root-identifier carries, nodes should compute trees on the additional roots.

3. The Multicast Group PDU

The systems that this draft is concerned with want to carry not only layer-2 unicast information in the link state protocols, but also multicast information. This draft has defined a new Multicast Group (MGROUP) PDU that can be used to advertise a set of attached, or joined, multicast groups. Accordingly, it has also introduced a couple more PDUs as described in the next sections for the flooding and update process to work seamlessly.

In the Layer-2 environment, it is expected the join/leave frequency of the multicast members will be much higher than unicast topology changes. It is efficient to separate the updates for the group membership change information from the remainder of the information by placing this information in a separate PDU. This enables reachability information, that would trigger an SPF, to be not impacted at all. Furthermore, during SPF runs, TLVs being on different PDUs which do not affect SPF need not be inspected during processing.

The choice of a different PDU also opens the LSP-space to another 256 fragments to carry a large number of groups. This additional space can be used judiciously to carry only multicast information.

The Multicast Group (MGROUP) PDU can be used to advertise a set of attached, or joined, multicast groups. The MGROUP PDU is formatted identical to a Level 1 Link State PDU, as described in Section 9.3 of [\[IS-IS\]](#). One field, PDU Type, is changed to 19 [TBD], to signify this PDU is carrying multicast group information, rather than unicast reachability information.

The Multicast Group PDU carries TLVs indicating multicast membership information. There are three sub-TLVs of the GADDR TLV defined in this document, that MAY be present in this PDU, namely, GMAC-ADDR, GIP-ADDR, and GIPV6-ADDR TLVs.

One or more TLVs MAY be carried in a single MGROUP PDU. Future multicast address TLVs MAY be defined using other type codes, and be carried in an MGROUP PDU.

Internet-Draft Carrying Attached Addresses in IS-IS October 2008

[3.1.](#) The Multicast Group Partial Sequence Number PDU

The Multicast Group Partial Sequence Number (MGROUP-PSNP) PDU is used to reliably flood the MGROUP PDU following the base protocol specifications.

[3.2.](#) The Multicast Group Complete Sequence Number PDU

The Multicast Group Complete Sequence Number PDU (MGROUP-CSNP) PDU is used to reliably flood the MGROUP PDU following the base protocol specifications.

[3.3.](#) Enhancements to the flooding process

This draft proposes that the information contained in the MGROUP-PDU is in a parallel database and its update mechanisms mimic that of the regular database. Nodes running IS-IS in an L2 domain MUST support these additional MGROUP PDUs defined in this draft. In general, the flooding of the MGROUP-PDU in tandem with the MGROUP-PSNP and MGROUP-CSNP PDUs uses the same update procedures as defined for the regular LSP, PSNP, and CSNP PDUs.

For example, on P2P links CSNP is exchanged on the formation of an adjacency. In a similar fashion a MGROUP-CSNP MUST also be exchanged between the neighbors at the same time. This gets the initial MGROUP-database synchronization going. After this similar actions of the base protocol specifications for the regular database synchronization will be maintained to keep the MGROUP-database synchronized.

Similarly, on LAN links the DIS is responsible for sending periodic CSNP transmissions. The DIS in the L2 IS-IS network domain will also be responsible for sending periodic MGROUP-CSNP transmissions. The update and flooding process will work in parallel for the two databases and there is no further synchronization between them.

In general, the database synchronization is performed in parallel with no interactions between the messages. However, the initial triggers that start a CSNP exchange are correlated, in the sense it

also triggers a MGROUP-CSNP exchange. For example, during graceful restart [[RFC 3847](#)], a parallel MGROUP-CSNP and MGROUP-PSNP exchange and update process will be run for the MGROUP-PDUs and restart process completes after both databases have been received.

[4.](#) Considerations for Using L2 Information in IS-IS

While this document does not specify the way in which addresses

carried in these TLVs is used in IS-IS, two general areas of concern are considered in this section: building the SPF tree when using this information, and the election of designated intermediate systems (DIS) in an environment using this information.

[4.1.](#) Building SPF Trees with Layer 2 Information

Each IS which is part of a single broadcast domain from a layer 2 perspective will build multiple SPF trees (SPT) for every IS that is announced by the IS deemed to be the broadcast root.

We assume some mechanism for forwarding traffic to these attached addresses added to the SPT is provided for in the mechanism proposing the use of these extension TLVs.

[4.2.](#) Designated Intermediate Routers

A single DIS SHOULD be elected as described in [[IS-IS](#)] for each layer 2 broadcast domain (VLAN) for which information is being carried in IS-IS. This reduces the amount of work required to flood and maintain synchronized databases over the underlying media on which IS-IS is running and providing layer 2 forwarding information for.

[5.](#) Acknowledgements

The authors would like to thank Les Ginsberg for his useful comments.

[6.](#) Security Considerations

This document adds no additional security risks to IS-IS, nor does it

provide any additional security for IS-IS.

7. IANA Considerations

This document creates three new PDU types, namely the MCAST PDU, MCAST-CSNP PDU, and the MCAST-PSNP PDU. IANA SHOULD assign a new PDU type to the level-1 PDUs described above and reflect it in the PDU registry.

MCAST-PDU	Level-1 PDU Type: 19
MCAST-CSNP-PDU	Level-1 PDU Type: 22
MCAST-PSNP-PDU	Level-1 PDU Type: 29

This document requires the definition a set of new ISIS TLVs, the MAC-Reachability TLV (type 141), and the Group Address TLV (type 142)

that needs to be reflected in the ISIS TLV code-point registry.

This document creates a number of new sub-TLV in the numbering space for the Group Address TLV and the Capability TLV. The TLV and sub-TLVs are given below:

					IIH	LSP	SNP	MCAST LSP	MCAST SNP
MAC-RI TLV	(141)				-	X	-	-	-
GADDR-TLV	(142)				-	-	-	X	-
GADDR-TLV.GMAC-ADDR		sub-tlv	1		-	-	-	X	-
GADDR-TLV.GMAC-IP		sub-tlv	2		-	-	-	X	-
GADDR-TLV.GMAC-IPV6		sub-tlv	3		-	-	-	X	-
CAPABILITY.Device ID		sub-tlv	5		-	X	-	X	-
CAPABILITY.Root Priority		sub-tlv	6		-	X	-	-	-
CAPABILITY.Roots		sub-tlv	7		-	X	-	-	-

IANA SHOULD manage the remaining space using the consensus method.

8. References

8.1. Normative References

- [IS-IS] ISO/IEC 10589, "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)", 2005.
- [RFC 1195] Callon, R., "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments", 1990.
- [RFC 3847] Shand, M. and L. Ginsberg, "Restart Signaling for Intermediate System to Intermediate System (IS-IS)", 2004.
- [RFC 4971] Vasseur, JP. and N. Shen, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", 2007.

[8.2.](#) Informative References

- [RBRIDGES] Perlman, R. and J. Touch, "Transparent Interconnection of Lots of Links (TRILL): Problem and Applicability

Ward, et al. Expires May 3, 2009 [Page 12]

Internet-Draft Carrying Attached Addresses in IS-IS October 2008

Statement", 2008.

Authors' Addresses

David Ward
Cisco Systems

Email: wardd@cisco.com

Russ White
Cisco Systems

Email: riw@cisco.com

Dino Farinacci

Cisco Systems

Email: dino@cisco.com

Ayan Banerjee

Cisco Systems

Email: ayabaner@cisco.com

Radia Perlman

Sun Microsystems

Email: Radia.Perlman@Sun.com

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND

THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.