Workgroup: Network Working Group
Internet-Draft:
draft-acee-idr-lldp-peer-discovery-14
Published: 6 December 2022
Intended Status: Standards Track
Expires: 9 June 2023
Authors: A. Lindem
Cisco Systems
K. Patel
Arrcus, Inc
S. Zandi
LinkedIn
J. Haas
Juniper Networks, Inc
X. Xu
Capitalonline

BGP Logical Link Discovery Protocol (LLDP) Peer Discovery

Abstract

Link Layer Discovery Protocol (LLDP) or IEEE Std 802.1AB is implemented in networking equipment from many vendors. It is natural for IETF protocols to avail this protocol for simple discovery tasks. This document describes how BGP would use LLDP to discover directly connected and 2-hop peers when peering is based on loopback addresses.

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). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on 9 June 2023.

Copyright Notice

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

Table of Contents

1. Introduction
  1.1. Requirements Notation
    1.1.1. Requirements Language
2. LLDP Extensions
  2.1. LLDP IETF Organizationally Specific TLV Format
  2.2. BGP Config OS-TLV Format
    2.2.1. BGP Config OS-TLV - Peering Address Sub-TLV
    2.2.2. BGP Config OS-TLV - BGP Local AS Sub-TLV
    2.2.3. BGP Config OS-TLV - BGP Identifier Sub-TLV
    2.2.4. BGP Config OS-TLV - Session Group-ID Sub-TLV
    2.2.5. BGP Config OS-TLV - BGP Session Capabilities Sub-TLV
    2.2.6. BGP Config OS-TLV - Key Chain Sub-TLV
    2.2.7. BGP Config OS-TLV - Local Address Sub-TLV
    2.2.8. BGP Config OS-TLV - BGP State Version Sub-TLV
3. BGP LLDP Peer Discovery Operations
  3.1. Advertising BGP Speaker
  3.2. Receiving BGP Speaker
  3.3. Updating or Deleting Auto-Discovery Parameters
4. LLDP Authentication/Encryption
5. Security Considerations
6. IANA Considerations
  6.1. IANA Assigned LLDP Subtype
  6.2. BGP Config LLDP OS-TLV Sub-TLVs
7. Contributors
8. References
  8.1. Normative References
  8.2. Informative References
Appendix A. Acknowledgments
Authors' Addresses

1. Introduction

Link Layer Discovery Protocol (LLDP) [LLDP] or IEEE Std 802.1AB is implemented in networking equipment from many vendors. It is natural for IETF protocols to avail this protocol for simple discovery tasks. This document describes how BGP [RFC4271] would use LLDP to discover directly connected and 2-hop peers when peering is based on loopback addresses.

1.1. Requirements Notation

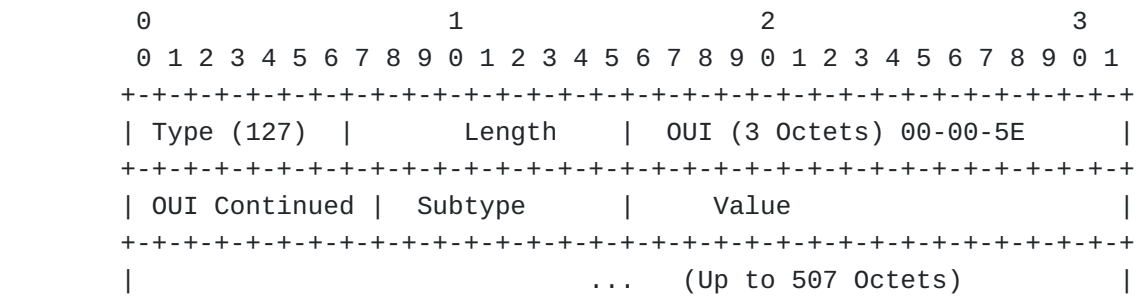
1.1.1. Requirements Language

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.

2. LLDP Extensions

2.1. LLDP IETF Organizationally Specific TLV Format

The format of the LLDP IETF Organizationally Specific TLV (OS-TLV) is defined in [LLDP]. It is shown below for completeness.



Type IETF Organizationally Specific TLV type value, 127.

Length The length of the remainder of the TLV.

OUI IETF Organizationally unique identifier for the organization's OUI. For IANA, this is value is 00-00-5E as specified in [IEEE-802-IANA].

Subtype IETF specific subtype

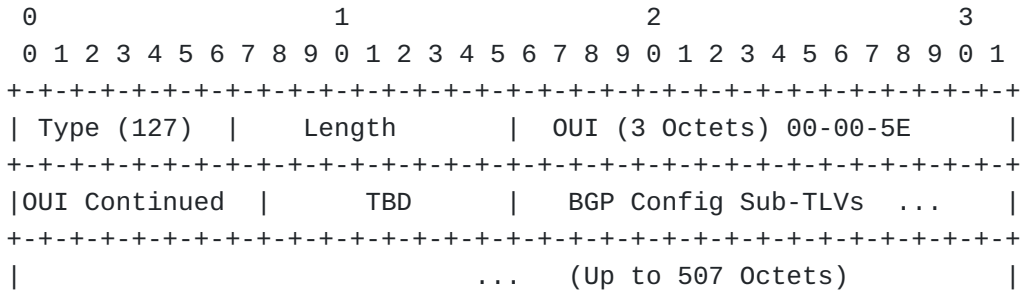
Value Value for organizationally specific TLV. The Length of the value is 4 octets less than the TLV length.

Figure 1: LLDP IETF Organizationally Specific TLV

The OUI for IANA was allocated in section 1.4 of [RFC7042]. This document requests creation of a registry for IETF specific sub-types for LLDP IETF Organizationally Specific TLVs.

2.2. BGP Config OS-TLV Format

The BGP Config IETF Organizationally Specific TLV (OS-TLV) will be used to advertise BGP configuration information. The configuration information will be composed of Sub-TLVs. Since the length is limited to 507 octets, multiple BGP Config OS-TLVs could be included in a single LLDP advertisement.



Length — The length of the BGP TLV.

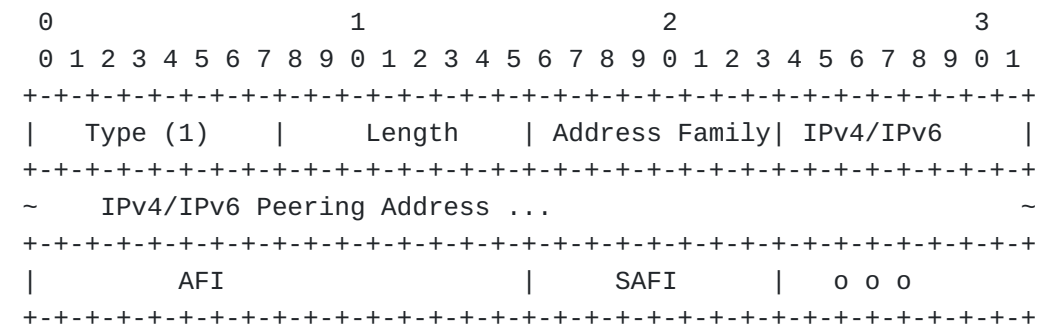
Subtype — IETF specific subtype for BGP Config OS-TLV. The value shall be TBD.

Value — BGP Config Sub-TLVs each with a 1 byte Type and Length. The Length will include solely the value portion of the TLV and not the Type and Length fields themselves.

2.2.1. BGP Config OS-TLV - Peering Address Sub-TLV

The BGP OS-TLV Peering Address Sub-TLV will be used to advertise the local IP addresses used for BGP sessions and the associated address families specified by AFI/SAFI tuples. The AFI/SAFI tuple, 0/0, indicates to use the associated peering address for all locally configured address families without an explicit peering address specification. As always, the address families supported for a given BGP session will be determined during capabilities negotiation [RFC4760]. It is RECOMMENDED that the wildcard AFI/SAFI be used in deployments with fairly homogenous address family usage.

The format of the BGP Peering Address Sub-TLV is shown below.



Type — The Sub-TLV Type value shall be 1.

Length — The Sub-TLV length in octets will be 4 for IPv4 or 16 for IPv6 plus 3 times the number of AFI/SAFI tuples.

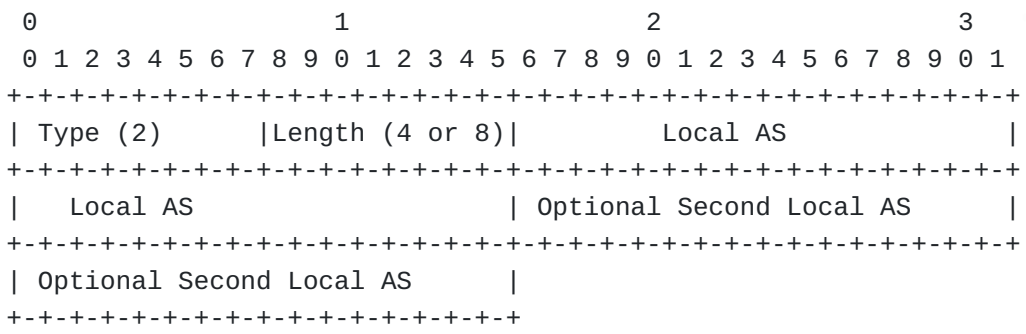
Address Family — IANA Address family (1 for IPv4 or 2 for IPv6)

Peering Address — An IPv4 address (4 octets) or an IPv6 address (16 octets)

AFI/SAFI Pairs — One or more AFI/SAFI tuples for BGP session using this peering address. The AFI/SAFI tuple, 0/0, is a wildcard indicating to attempt negotiation for all AFI/SAFIs.

2.2.2. BGP Config OS-TLV - BGP Local AS Sub-TLV

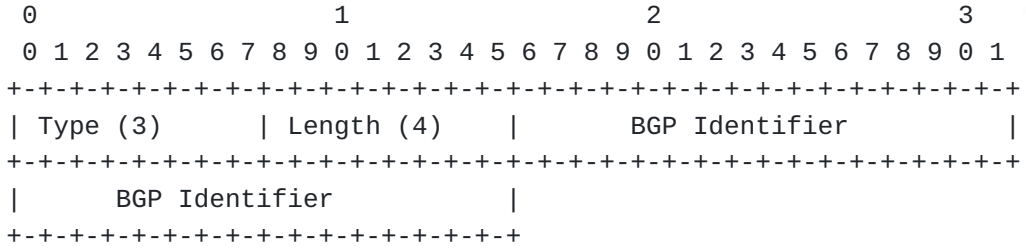
The BGP Config OS-TLV Local AS Sub-TLV will be used to advertise the 4-octet local Autonomous System (AS) number(s). For AS transitions, a second local AS number may be specified. The format of the BGP Local AS Sub-TLV is shown below.



Type	The Sub-TLV Type value shall be 2.
Length	The Sub-TLV Length will be 4 or 8 octets.
Local AS	Local Autonomous System (AS)
Second Local AS	Local Autonomous System (AS)

2.2.3. BGP Config OS-TLV - BGP Identifier Sub-TLV

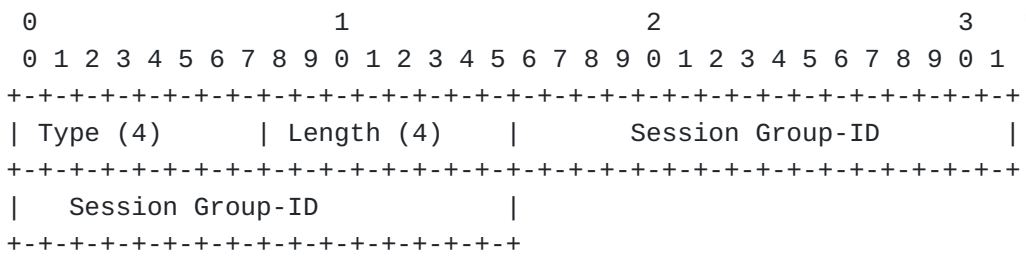
The BGP Config OS-TLV BGP Identifier Sub-TLV will be used to advertise the 4-octet local BGP Identifier. The BGP Identifier is used for debugging purposes and possibly to reduce the likelihood of BGP connection collisions. The format of the BGP Identifier Sub-TLV is shown below.



Type	The Sub-TLV Type value shall be 3.
Length	The Sub-TLV Length will be 4 octets.
BGP Identifier	Local BGP Identifier (aka, BGP Router ID)

2.2.4. BGP Config OS-TLV - Session Group-ID Sub-TLV

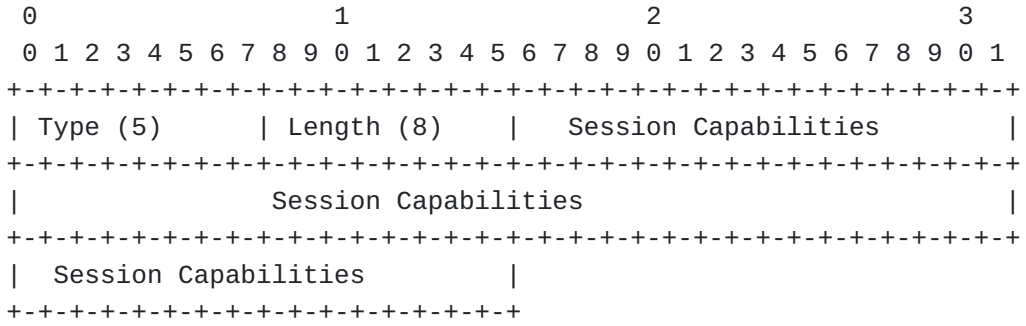
The BGP Config OS-TLV Session Group-ID Sub-TLV is an opaque 4-octet value that is used to represent a category of BGP session that is supported on the interface. The format of the Session Group-ID Sub-TLV is shown below.



Type	The Sub-TLV Type value shall be 4.
Length	The Sub-TLV Length will be 4 octets.
Session Group-ID	The session group-id used to indicate a class or category of BGP session supported on the interface.

2.2.5. BGP Config OS-TLV - BGP Session Capabilities Sub-TLV

The BGP Config OS-TLV Session Capabilities Sub-TLV will be used to advertise an 8-octet Session Capabilities field. The session capabilities are represented as bit flags identifying the supported BGP session capabilities. The format of the BGP Session Capabilities Sub-TLV is shown below.



Type	The Sub-TLV Type value shall be 5.
Length	The Sub-TLV Length will be 8 octets.
Session Capabilities	Bit fields identify BGP session capabilities

The BGP Session Capabilities is an 8-octet bit field. The most significant bit is the first bit (Bit 1) of the Session Capabilities. The following bits are defined:

Bit 1: This bit indicates support for TCP MD5 authentication [TCP-MD5].

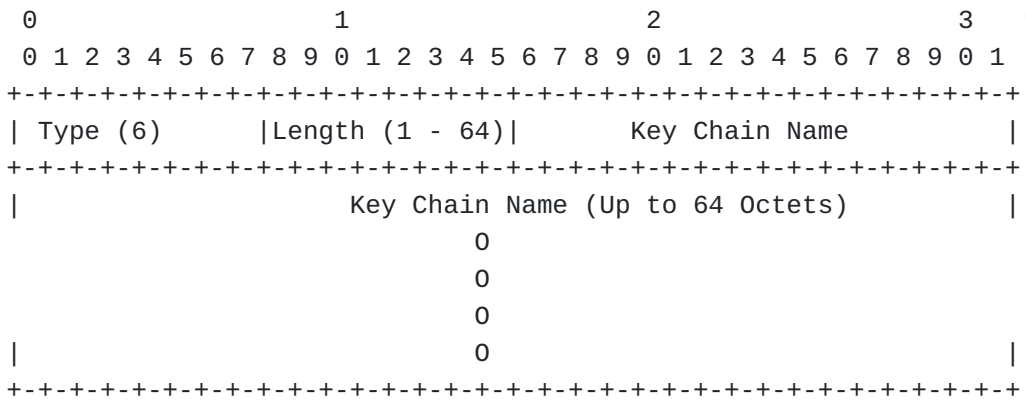
Bit 2: This bit indicates support for TCP-AO authentication [TCP-AO].

Bit 3: This bit indicates support for Generalized TTL Security Mechanism (GTSM) [GTSM] with a configured TTL range of 254-255.

TCP MD5 authentication is described in [RFC2385]. The TCP Authentication Option (TCP-AO) is described in [RFC5925]. The Generalized TTL Security Mechanism (GTSM) is described in [RFC5082]. If both TCP MD5 authentication and TCP-AO authentication are specified and TCP-AO is supported, it will take precedence.

2.2.6. BGP Config OS-TLV - Key Chain Sub-TLV

The BGP Config OS-TLV Key Chain Sub-TLV is a string specifying the name for the key chain used for session authentication. Key chains [RFC8177] are a commonly used for protocol authentication and encryption key specification. Given the limited length of all BGP configuration information, the key chain name will be limited to 64 characters and will not include a trailing string delimiter. The format of the Session Group-ID Sub-TLV is shown below.



Type — The Sub-TLV Type value shall be 6.

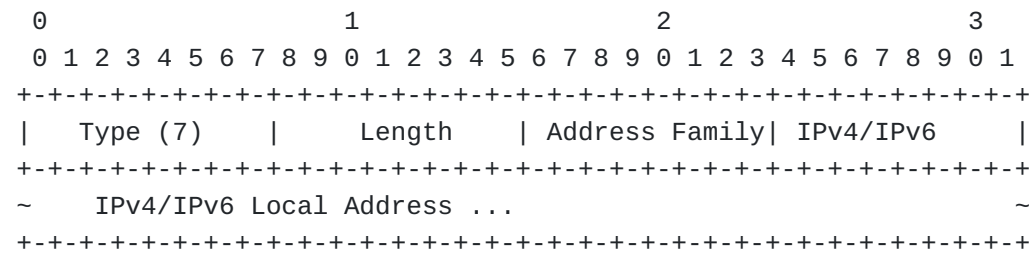
Length — The Sub-TLV Length will be 1 - 64 octets.

Key Chain Name — The name of a key chain to be used for MD5 or TCP-AO authentication.

2.2.7. BGP Config OS-TLV - Local Address Sub-TLV

The BGP OS-TLV Local Address Sub-TLV will be used to advertise a local IP addresses used for BGP next-hops. Advertising a local interface address is useful when the address family is different from the advertised BGP peering address.

The format of the BGP Local Interface Address Sub-TLV is shown below.



Type — The Sub-TLV Type value shall be 7.

Length — The Sub-TLV length in octets will be 4 for IPv4 or 16 for IPv6 plus 3 times the number of AFI/SAFI tuples.

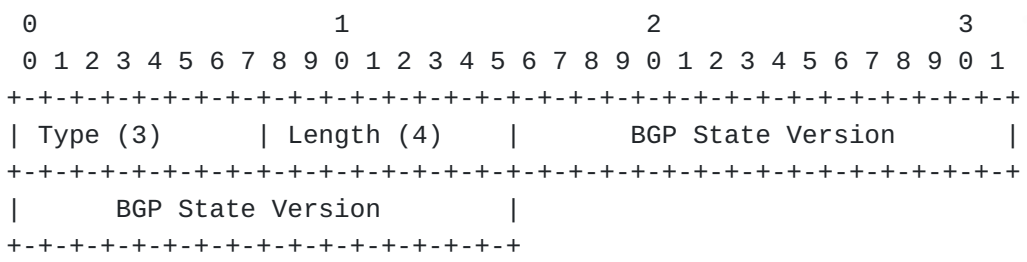
Address Family — IANA Address family (1 for IPv4 or 2 for IPv6)

Local Address — An IPv4 address (4 octets) or an IPv6 address (16 octets)

2.2.8. BGP Config OS-TLV - BGP State Version Sub-TLV

The BGP OS-TLV Version Sub-TLV will be used to advertise a monotonically increasing version. This version will indicate if any local BGP state that may impact BGP session establishment has changed. Changes can range from anything as obvious a change in local peering address to more indirect changes such as the modification of the key-chain being advertised.

The format of the BGP State Version Sub-TLV is shown below.



Type	The Sub-TLV Type value shall be 8.
Length	The Sub-TLV Length will be 4 octets.
BGP State Version	BGP State Version - Monotonically increasing version number indicating if any local state that may effect BGP session establishment has changed.

3. BGP LLDP Peer Discovery Operations

The simple use case is to just use the peer address advertised in the LLDP Packet Data Unit (PDU) to establish a 1-hop BGP peer session. This can be used in data centers using BGP as described in [RFC7938]. The use case where a loopback address or other local address is advertised as the peering address is also supported. However, reachabiliy to a peering address other than the interface address is beyond the scope of this document.

3.1. Advertising BGP Speaker

A BGP speaker MAY advertise its BGP peering address in an LLDP PDU for a link using the BGP Local Address Sub-TLV of the BGP-OS TLV. This can be an IPv4 or IPv6 local address associated with the LLDP link for 1-hop peering. For 2-hop peering, it could be a loopback address or any other address that is local to the node but not the LLDP link. As noted above, reachability to the loopback address is beyond the scope of this document.

A BGP speaker MAY advertise its local AS number using the BGP Local AS Sub-TLV of the BGP-OS TLV. During AS transitions, a second local AS number may be included in the Local AS Sub-TLV. The local BGP identifier may also be advertised using the BGP Identifier Sub-TLV of the BGP-OS TLV. While not specifically required for session establishment, the values may be used for validation, trouble-shooting, and connection collision avoidance. A BGP speaker may also announce a Session Group-ID indicating the class or category of session(s) supported and/or mapping to a set of session parameters. Additionally, a BGP speaker MAY also announce relevant capabilities using BGP Session Capabilities Sub-TLV of the BGP-OS TLV.

If TCP MD5 authentication [RFC2385] or TCP Authentication Option (TCP-AO) [RFC5925] is to be used on the session, the Key Chain Sub-TLV of the BGP-OS TLV MAY be used to specify the key chain name.

3.2. Receiving BGP Speaker

A BGP speaker configured for LLDP peer discovery WILL attempt to establish BGP sessions using the address in the BGP Local Address Sub-TLV of BGP-OS TLV format. If the peering address is directly accessible over the link on which the LLDP PDU is received, the BGP speaker will attempt to establish a 1-hop BGP session with the peer.

If the received BGP Peering Address is not directly accessible over the link, the peer must be reachable for the session to be established and the mechanisms for establishing reachability are beyond the scope of this specification. If the BGP speaker receives the same BGP peering address in LLDP PDUs received on multiple links, it will not establish multiple sessions. Rather, a single 2-hop session will be established.

When the deployment of address families is fairly homogenous across the deployment, the wildcard AFI/SAFI can be utilized to simplify LLDP advertisement. When there is variance in the address families supported, usage of the wildcard could result in session establishment delay due to capabilities negotiation [RFC5492].

A BGP speaker MAY receive a remote neighbor's local AS number(s) in an LLDP PDU in the BGP Local AS Sub-TLV of the BGP-OS TLV. A BGP speaker MAY use the received local AS number(s) to perform validation checking of the AS received in the OPEN message. A BGP speaker MAY receive a remote neighbor's BGP Identifier in the BGP Identifier Sub-TLV of the BGP-OS TLV. This can be used to avoid connection collisions by delaying session establishment if the remote BGP Identifier is greater than the receiving speaker's BGP Identifier.

A BGP speaker MAY receive a Session Group-ID Sub-TLV in the LLDP BGP-OS TLV. This Session Group-ID may be used for validation and/or mapping the session to a particular set of session parameters. For example, the Session Group-ID could be mapped to a spine, leaf, or Top-of-Rack (ToR) session in a data center deployment and can be used to detect cabling problems when an unexpected Session Group-ID is received.

Additionally, A BGP speaker MAY receive a remote neighbor's capabilities in LLDP in the BGP Session Capabilities Sub-TLV of the BGP-OS TLV. A BGP speaker MAY use the received capabilities to ensure appropriate local neighbor configuration in order to facilitate session establishment.

If TCP MD5 authentication [RFC2385]. or TCP Authentication Option (TCP-AO) [RFC5925] is to be used on the session as determined either via the Session Capabilities Sub-TLV, Session Group-ID, or local policy, the key chain name in the Key Chain Sub-TLV of the BGP-OS TLV MAY be used to identify the correct key chain [RFC8177].

The BGP State Version associated with the LLDP peer SHOULD be retained to determine whether anything impacting BGP session establishment has changed. When session establishment fails, this can be used to avoid back-off on attempting to establish a BGP session when nothing has changed on the peer or locally.

3.3. Updating or Deleting Auto-Discovery Parameters

A BGP speaker MAY change or delete any BGP LLDP auto-discovery parameter by simply updating or removing the corresponding Sub-TLV previously advertised in the BGP-OS TLV. Additionally, the BGP State Version Sub-TLV should be advertised with the version incremented from the previous version. The BGP speaker(s) receiving the advertisement will update or delete the changed or deleted auto-discovery parameters. However, there will be no change to existing BGP sessions with the advertising BGP Speaker. Changes to existing BGP sessions are the purview of the BGP protocol and are beyond the scope of this document.

Since LLDP information is cumulative, reception of an LLDP PDU without the BGP-OS TLV indicates that BGP LLDP auto-discovery has been disabled for the BGP speaker and all parameters learnt during BGP LLDP auto-discovery SHOULD be deleted. As above, changes to existing BGP sessions are beyond the scope of this document.

The LLDP Multi-Frame extention [LLDP-MULTIFRAME] removes the limit on a LLDP PDU being fitting in a single layer 2 frame. This will increase the number of TLVs which can be contained in LLDP PDU and the applicablity of LLDP as a BGP discovery protocol. The specificaiton of the LLDP API for BGP and other applications is beyond the scope of this document. However, it is RECOMMENDED that the LLDP BGP TLVs only be delivered to BGP when a complete LLDP PDU is received.

4. LLDP Authentication/Encryption

The IEEE 802.1AE [MACsec] standard can be used for encryption and/or authentication to provide privacy and integrity. MACsec utilizes the Galois/Counter Mode Advanced Encryption Standard (AES-GCM) for authenticated encryption and Galois Message Authentication Code (GMAC) if only authentication, but not encryption is required.

The MACsec Key Agreement (MKA) is included as part of the IEEE 802.1X-20200 Port-Based Network Access Control Standard [MKA]. The

purpose of MKA is to provide a method for discovering MACsec peers and negotiating the security keys needed to secure the link.

5. Security Considerations

This security considerations for BGP [RFC4271] apply equally to this extension.

Additionally, BGP peering address discovery should only be done on trusted links (e.g., in a data center network) since LLDP packets are not authenticated or encrypted [LLDP].

LLDP Authentication and/or encryption can provided as described in section Section 4.

6. IANA Considerations

6.1. IANA Assigned LLDP Subtype

IANA is requested to assign a code point in the IANA Link Layer Discovery Protocol (LLDP) TLV Sub-Types Registry for BGP configuration. The value is TBD.

6.2. BGP Config LLDP OS-TLV Sub-TLVs

IANA is requested to create a registry for Sub-TLVs of the BGP Config LLDP OS-TLV. Assignments are requested as specified in the table below.

Range	Assignment Policy
0	Reserved (not to be assigned)
1	Peering Address
2	Local AS
3	BGP Identifier
4	Session Group-ID
5	Session Capabilities
6	Key Chain Name
7	Local Address
8	BGP State Version
9-127	Unassigned (IETF Review)
128-254	Reserved (Not to be assigned now)
255	Reserved (not to be assigned)

Figure 2: LLDP BGP Config OS-TLV Types

*Types in the range 9-127 are to be assigned subject to IETF Review. New values are assigned only through RFCs that have been shepherded through the IESG as AD-Sponsored or IETF WG Documents [RFC5226].

*Types in the range 128-254 are reserved and not to be assigned at this time. Before any assignments can be made in this range, there MUST be a Standards Track RFC that specifies IANA Considerations that covers the range being assigned.

7. Contributors

Contributors' Addresses

8. References

8.1. Normative References

[LLDP]

IEEE, "IEEE Standard for Local and metropolitan area networks-- Station and Media Access Control Connectivity Discovery Corrigendum 2: Technical and Editorial Corrections", IEEE 802.1AB-2009/Cor 2-2015, DOI 10.1109/ieeestd.2015.7056401, 9 March 2015, <https://doi.org/10.1109/ieeestd.2015.7056401>.

[LLDP-MULTIFRAME] IEEE, "IEEE Standard for Local and metropolitan area networks-- Station and Media Access Control Connectivity Discovery Amendment 2: Support for Multiframe Protocol Data Units", IEEE 802.1ABdh-2021, DOI 10.1109/IEEESTD.2022.9760302, 19 April 2022, <https://doi.org/10.1109/IEEESTD.2022.9760302>.

[MACsec] IEEE, "IEEE Standard for Local and metropolitan area networks - Media Access Control (MAC) Security", IEEE Standard 802.1AE-2018, 27 September 2018.

[MKA] IEEE, "IEEE Standard for Local and metropolitan area networks - Port Based Network Access Control", IEEE Standard 802.1X-2020, 30 January 2020.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

8.2. Informative References

[RFC2385] Heffernan, A., "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC 2385, DOI 10.17487/RFC2385, August 1998, <https://www.rfc-editor.org/info/rfc2385>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[RFC5082] Gill, V., Heasley, J., Meyer, D., Savola, P., Ed., and C. Pignataro, "The Generalized TTL Security Mechanism

(GTSM)", RFC 5082, DOI 10.17487/RFC5082, October 2007, <https://www.rfc-editor.org/info/rfc5082>.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 5226, DOI 10.17487/RFC5226, May 2008, <https://www.rfc-editor.org/info/rfc5226>.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <https://www.rfc-editor.org/info/rfc5492>.

[RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <https://www.rfc-editor.org/info/rfc5925>.

[RFC7042] Eastlake 3rd, D. and J. Abley, "IANA Considerations and IETF Protocol and Documentation Usage for IEEE 802 Parameters", BCP 141, RFC 7042, DOI 10.17487/RFC7042, October 2013, <https://www.rfc-editor.org/info/rfc7042>.

[RFC7938] Lapukhov, P., Premji, A., and J. Mitchell, Ed., "Use of BGP for Routing in Large-Scale Data Centers", RFC 7938, DOI 10.17487/RFC7938, August 2016, <https://www.rfc-editor.org/info/rfc7938>.

[RFC8177] Lindem, A., Ed., Qu, Y., Yeung, D., Chen, I., and J. Zhang, "YANG Data Model for Key Chains", RFC 8177, DOI 10.17487/RFC8177, June 2017, <https://www.rfc-editor.org/info/rfc8177>.

Appendix A. Acknowledgments

Thanks to Sujay Gupta and Paul Congdon for review and comments.

Thanks to Donald Eastlake for guidance on IANA LLDP TLV Subtype assignment. Thanks to Dan Romascanu for review of the IANA considerations.

Authors' Addresses

Acee Lindem
Cisco Systems
301 Midenhall Way
Cary, NC 27513
United States of America

Email: acee@cisco.com

Keyur Patel

Arrcus, Inc

Email: keyur@arrcus.com

Shawn Zandi
LinkedIn
222 2nd Street
San Francisco, CA 94105
United States of America

Email: szandi@linkedin.com

Jeff Haas
Juniper Networks, Inc
1133 Innovation, Inc.
Sunnyvale, CA 94089
United States of America

Email: jhaas@juniper.net

Xiaohu Xu
Capitalonline

Email: xiaohu.xu@capitalonline.net