Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: 25 July 2022

J.M. Jeganathan
V.S.K.R. Avula
Juniper Networks

21 January 2022

BGP Peer Auto-Configuration
draft-minto-idr-bgp-autodiscovery-01

Abstract

This document describes a layer 3 protocol (Service advertisement) to help bgp to advertise service availability and local configurations . This enables bgp speakers to discover bgp peers transport endpoints and peer's configuration within link. With Service advertisement, receivers could successfully bring up bgp protocol session without mundane configurations.

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 25 July 2022.

Copyright Notice

Table of Contents

1. Introduction

This document describes a layer 3 protocol (Service advertisement)to help bgp to advertise service availability and local configurations . This enables bgp speakers to discover bgp peer's transport endpoints and peer's configuration within link. With Service advertisement, receivers could successfully bring up bgp protocol session without

mundane configurations.

1.1. Requirements Language

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].

2. Protocol overview

This is a simple protocol to periodically send and receive UDP multicast PDU that contains bgp transport information in the form of messages and TLVs. Receiver could use this information to bootstrap the single hop bgp and/or loopback address bgp between directly connected bgp speakers. The advertised information gets expired if it is not refreshed before the lifetime ends.

This protocol does not provide any reliability of delivery and relies on UDP multicast and periodic send. The current version of this protocol assumes the link MTU is good enough to encode BGP transport information or underlying IP implementation is able to fragment and reassemble for link local multicast PDU. But this protocol is flexible enough to implement a future version of fragment TLV attachment. This is to bypass smaller link MTU for a system or environment preventing IP fragment.

Service Advertisement (SA) PDU has multiple types of messages. This document defines 2 types of messages. The primary/base messages are required for SA to operate and secondary type messages for BGP service advertisement.

3. PUD layers

The PDU contains a header followed by variable number of messages. Each message contains variable number of TLVs.

SA uses type-length-value format.

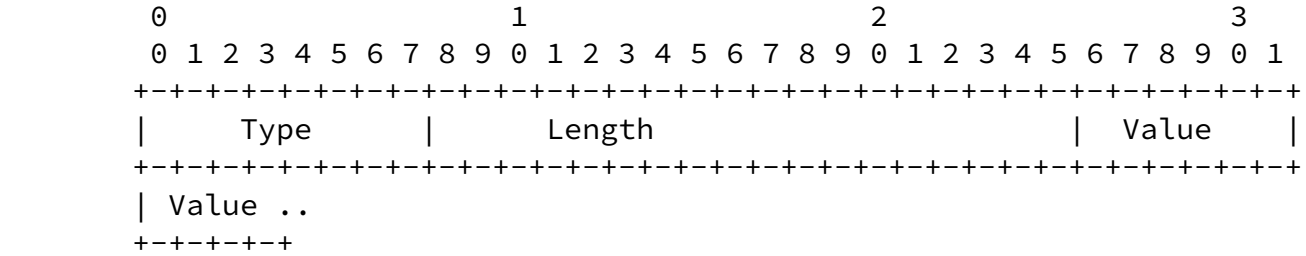


Figure 1

Type: 1-octet value to interpret the value with in message. Same type value could be reused in different message.

Length: Specifies length in octets of the value field.

Value: Octet string that encodes information to be interpreted as specified by the Type field.

SA uses message to group set of TLVs

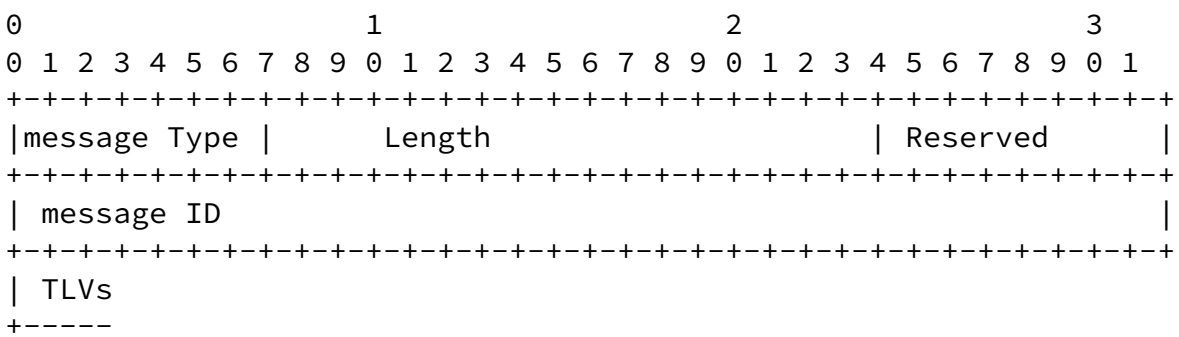


Figure 2

message Type: This 1-octet value identifies type of message.

Message Length: Specifies the length in octets of the Message ID and TLVs.

Message ID :32-bit value used to identify this message. Used for logging purpose.

SA PDU

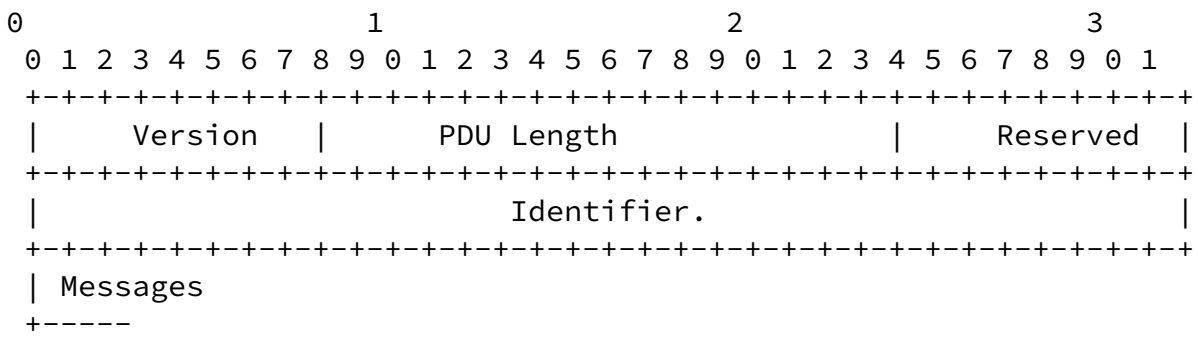


Figure 3

Version: This 1-octet unsigned integer indicates the protocol version. This version of the specification specifies Service Advertisement version 0.

PDU Length: This 2-octet unsigned integer specifies the length

Identifier: 4 octet field that uniquely identifies PDU sender. BGP id could be used for this purpose. This helps to uniquely identify sender across the parallel links between same nodes.

4. Messages

The document defines following messages.

1. SA Base message

2. BGP service advertisement message

4.1. SA Base message

The SA Base message is mandatory message and mainly used for the protocol operation.

The document defines following TLVs for SA Base message.

1. Remaining lifetime TLV

2. Config sequence TLV

3. Authentication TLV

4. Refresh request TLV

4.1.1. Remaining lifetime TLV

Remaining lifetime describes how long receiver should keep the state without seeing a PDU from the sender. The lifetime gets updated when receiver accepts the PDU.

Type : 17

Length: 2 octets

Value: Remaining lifetime in seconds

4.1.2. Config sequence TLV

Specifies a 4-octet configuration sequence number. Receiver could make use the number to detect config change. This will be useful to restart the bgp session with new parameters.

Type : 18

Length: 4 octets

Value: unsigned sequence number

4.1.3. Authentication TLV

Specifies authentication.

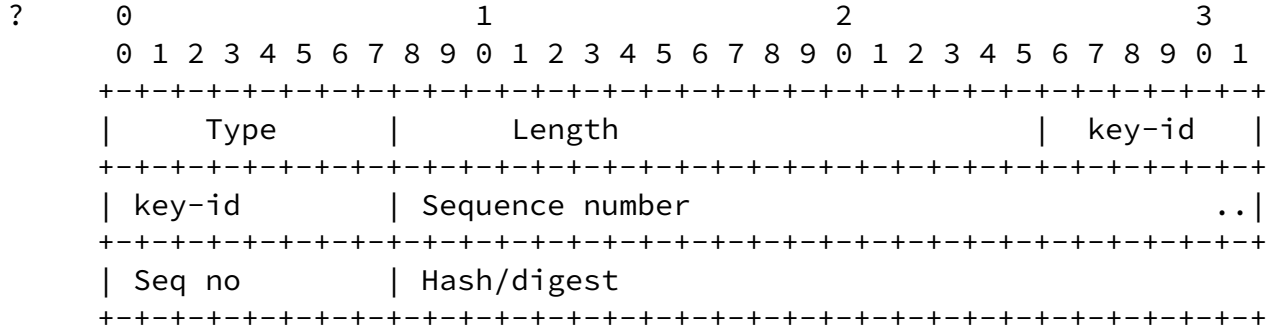


Figure 4

Type : 19

Length: variable

key-id : keychain id

Sequence-number - 4 byte sequence number of this SA base message.

Digest - Hash computed for this message using key-id mapped algorithm

4.1.4. Refresh request TLV

Optional TLV to trigger receivers to immediately send SA PDU. Presence of the TLV indicates sender request refresh. This will be used during the restart to learn about services quickly from connected devices to speed up service discovery.

Type : 20

Length: 0 octets

4.2. BGP service advertisement message

BGP Service Advertisement message provides transport information to bring up the bgp session. This document defines transport information TLVs and session information TLVs for BGP Service Advertisement messages.

Message type of BGP Service Advertisement message: 2.

Following are the Transport information TLVs

1. Local Address
2. Security TTL
3. Security Authentication

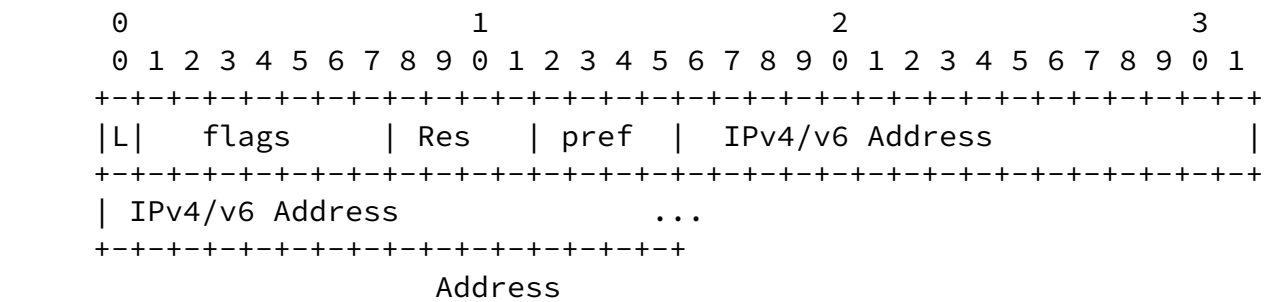
4. Link Address

5. Transport Preference.

6. TCP MSS

4.2.1. Local address

Specifies a local address used for bgp transport connection. Address encoding uses a below format. 2 octets describe the address and followed by address value.



L bit - Address of loopback interface.
pref - Preference value of 4 bits. Value from 1 to 15.
0 indicates dont care.
1 higly preferred and 15 means least preferred.
Address - For IPv4 4 octets and IPv6 16 octets

Figure 5

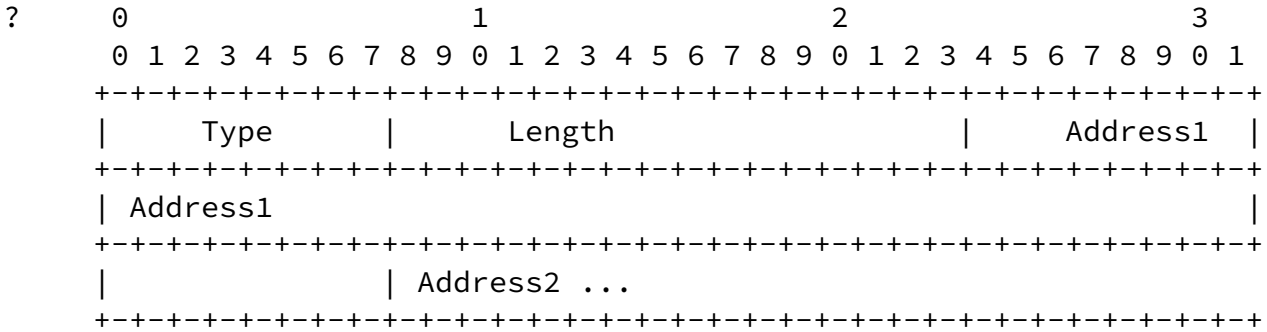


Figure 6

Type : 17

Length: variable. multiples of 6 octets

Value: IPv4 addresses with address encoding.

4.2.2. Local IPv6 address

Specifies a IPv6 local address used for bgp transport connection.

Type : 18

Length: multiples of 18 octets

Value: IPv6 addresses with address encoding used for transport connection.

4.2.3. Security TTL

TTL be accepted for bgp messages

Type : 19

Length: 0

Value: Presence of this TLV indicates that receiver accepts only packets with 255 TTL.

4.2.4. Security Authentication

Type : 20

Length: 1

Value:This supports only two values 0 and 1.

0 indicates TCP md5. 1 indicates TCP-AO Absence of this TLV indicates, no authentication used for connection.

4.2.5. TCP MSS

TCP MSS used for the connection

Type : 21

Length: 4

Value:Value in bytes

Indicates the preference of TCP MSS for the transport connection.

4.2.6. Link Address

This could be used for receiver to get nexthop information for local address TLV when sender's running IPv4 PDU and prefer IPv6 transport and vice-a-versa. This could also be used to provide reachability to loopback addresses with link address.

Type : 22

Length: 4 for IPv4 address and 16 for IPv6 address 6 for mac address.

Value:interface IPv4 or IPv6 or mac address .

5. Protocol operation

A sender should periodically send PDU to refresh the advertised information before its lifetime expires. An implementation may send PDU well before the lifetime expires based on specific events. These events could be a local config change or discovering a new advertiser. Also, implementation could switch to fast refresh when content of the pdu changes and move back to regular refresh interval. The fast refresh will help in quicker discovery and may help update content in case of auto order delivery. As stated above, this is purely an implementation technique than the protocol mandate.

To discover multi-data(IPv4/IPv6) protocol environment(mixed transport mode in a single link) sender shall send both data-protocol pdu based on local configuration. When sender choose to send both data protocol PDU it should make sure that semantic content of the messages should be same. An implementation may choose to use preferred data protocol PDU as primary send PDU and only send other data protocol PDU during the interesting events. This optimization is only possible when all the known advertisers participates in both data-protocol.

A sender should send PDU to refresh before previously advertised lifetime expires. If bgp is configured with only one transport address family(IPv4/v6) then sender shall only send corresponding data protocol PDU. If both addresses are configured, then it shall use both data protocol PDUs. PDUs are sent with source address as link primary address and destination is link local all- routers with TTL 255. If authentication is enabled then add authentication TLV using the authentication procedure described in authentication section. Populate other TLVs based on local preference and send the

PDU on configured link. Semantic content (transport and session information) of the PDU should be same irrespective of data protocol.

Receiver resets the state when it accepts new PDU irrespective of the data protocol. Receiver shall add a route for the address in local address TLV with nexthop as source address of the PDU if PDU(PUD) data protocol and local address is same address family. Otherwise if link address is available, it could be used as nexthop for the address in local address TLV. Receivers consolidate state from various TLVs and pass it on to BGP for the session opening. An implementation could only notify if the state change from previous reported state to bgp or the configuration sequence number changes from the receiver. How bgp uses this information is beyond the scope of the document.

5.1. Transmit procedure

PDUs are sent with source address as link's primary address and destination is link local all-routers with TTL 255. PDU is sent to SA UDP port(179 if assigned). After the header, SA Base message should be first message. If authentication is enabled then add authentication TLV using the authentication procedure described in authentication section. This authentication TLV should be first TLV of PDU. Add lifetime and config sequence TLVs defined in this document. Both these TLVs are mandatory TLVs. After the SA base message, add bgp service advertisement message with appropriate TLVs.

5.2. Receiver procedure

When a SA PDU received, following sanity procedure must be followed.

If TTL is not 255 then discard the PDU.

If the version is not compatible (Only compatible version is 0) then discard the PDU.

If the PDU length is greater than IP header length, then discard it.

If the first message is not SA Base, then discard the pdu.

If authentication is enabled and first TLV in the SA base message, then discard the PDU.

If authentication is enabled, then follow the authentication procedure.

If authentication is failed, then discard the PDU.

With above steps, sanity of the PDU header is verified. Receiver should start decoding the TLV information. Once all the TLV sanity checked receiver shall keep the decoded information. If the receiver decides to keep the information, then it should start a timer with specified lifetime or refresh lifetime with newer one.

The identifier in PDU header uniquely identifies the advertisement. An implementation could either implement neighbor semantic or state semantic from the advertised information along with identifier. This document does not recommend one or other.

Received and decoded information shall be passed on to bgp if the content does not match with last received or the local config has changed. This is a desired optimization, so that SA does not unnecessarily trigger failed bgp session open attempts. How bgp uses this information is beyond the scope of the document.

5.3. Transport endpoint reachability

Advertised local address reachability can either be gathered from the source address or a link address TLV. Source address of the PDU may not give reachability for all deployment(Sender using the IPv6 data protocol but prefer v4 transport). In those cases, link address TLV will provide reachability.

5.4. Protocol Authentication operation

A sender that wants to authenticate Service messages should include Authentication TLV as part of SA base message.

Sender needs to include all the fields of Authentication TLV as shown in section 4.1.3. It needs to assign a unique KEY-ID to each

authentication combination configured on the device. Key-length needs to be set to configured key's length in bytes. Sequence number is a 32-bit unsigned integer that may increment by one each time a new message is sent. Any change in TLVs for a previously advertised local address needs to be sent with an incremented TLV. Digest value can be of variable length depending upon type of authentication being used. This value is calculated over all the contents of service message.

Receiver on receiving this TLV has a sequential processing of individual fields of TLV. Sequence number is read from TLV and is compared against any existing state from this sender. If sequence number is lesser than previously received, this packet is dropped except when bgp session goes down.If last received sequence number was m and current received sequence number is n, n needs be in range of [m+1, m + 2^(32 - 1)]. This exception is needed to handle a

restarted sender who is unable to retrieve earlier sequence number due to restart. This is required when SA uses bigger lifetime. After getting KEY-id, it checks for a matching KEY-ID on it. If it does not exist, packet is dropped. Next Key-length of locally configured key is compared against key-length received in this TLV, if they do not match packet is dropped. Similarly, a comparison is done for authentication types of locally configured key and received TLV. If they do not match, packet is dropped. After above checks, hash is computed for all the contents of service with locally configured key and compared against received hash value. If they are same, authentication information matches with local configuration and messages can be further processed with protocol operations depending on type of this message.

6. Acknowledgements

Jeffrey Hass provided many useful technical and editorial comments and suggestions for improvement.

7. IANA Considerations

This document requests IANA to allocate a new UDP port (179 is the preferred number) and 2 message type code for service advertisments.

Value TLV Name Reference -----

---------- ---------- Service Name: Service advertisements Transport Protocol: UDP Assignee: IESG iesg@ietf.org Description: Service advertisments for auto configuration. Reference: This document -- draft-minto-idr-bgp-autodiscovery.txt Port Number: 179 -- To be assigned by IANA.

Figure 7

7.1. Message of SA

This document requests IANA to create a new registry following messages "Messages of SA " with the following registration procedure:

? Registry Name: Messages of SA protocol

Value	Message name	Reference
0	Reserved	This document
1	Base message	This document
2	BGP Service Advertisement	This document

Figure 8

7.2. TLVs of SA base Message

This document requests IANA to create a new registry following messages "TLVs of SA base Message" with the following registration procedure:

Registry Name: TLVs of SA base Message.

Value	TLV Name	Reference
0-16	Reserved	This document
17	Remaining lifetime TLV	This document
18	Config sequence TLV	This document
19	Authentication	This document
20	Refresh request TLV	This document
224-255	Experimental	

Figure 9

7.3. TLVs of BGP service advertisement message

This document requests IANA to create a new registry following messages "TLVs of BGP Service Advertisement" with the following registration procedure:

? Registry Name: TLVs of BGP Services.

Value	TLV Name	Reference
0-16	Reserved	This document
17	Local Address	This document
18	Local IPv6 Address	This document
19	Security TTL	This document
20	Security Authentication	This document
21	TCP MSS	This document
22	Link Address	This document
224-255	Experimental	This document

Figure 10

8. Security Considerations

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

BGP sessions transport end points discovered over this protocol can be protected against various attacks by using authentication for packets as described in Section 5.4.

Usage of seqeunce number and authentication reduces likelihood of replay attacks. As the protocol is not connection-oriented, it makes it feasible to change authentication parameters for protocol messages. This further reduces the likelihood of replay-attacks.

9. References

9.1. Normative References

[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>.

9.2. Informative References

[bgp-autoconf-considerations]
Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 5226, DOI 10.17487/RFC5226, May 2008, <https://datatracker.ietf.org/doc/draft-ietf-idr-bgp-autoconf-considerations/>.

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, DOI 10.17487/RFC2629, June 1999, <https://www.rfc-editor.org/info/rfc2629>.

[RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, DOI 10.17487/RFC3552, July 2003, <https://www.rfc-editor.org/info/rfc3552>.

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

[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>.

Appendix A. Additional Stuff

This becomes an Appendix.

Authors' Addresses

Jeyananth Minto Jeganathan
Juniper Networks
Juniper Networks, 1133 Innovation Way
Sunnyvale, CA 94089

United States of America

Email: minto@juniper.net

Venkata Shiva Krishna Reddy Avula
Juniper Networks
Juniper Networks, 1133 Innovation Way
Sunnyvale, CA 94089
United States of America

Email: venkatashiva@juniper.net