Workgroup: Internet Engineering Task Force
Internet-Draft:
draft-minto-idr-bgp-autodiscovery
Published: 12 July 2021
Intended Status: Standards Track
Expires: 13 January 2022
Authors: J.M. Jeganathan   V.S. Avula
         Juniper Networks   Juniper Networks

BGP Peer Auto-Configuration

Abstract

This document describes a layer 3 protocol (Service advertisement) to help bgp to advertise service availability and local configurations . This enable 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

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 enable 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.

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, 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 1hop bgp and/or loopback address bgp between directly connected bgp speakers. The advertised information gets expired if not refreshed before the lifetime.

This protocol does not provide any reliability of delivery and relay 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 able to do fragment and reassembly for link local multicast PDU. But this protocol flexible enough implements a future version of fragment TLV attachment to bypass the smaller link MTU for the system or environment prevent 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 operates 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.

```
  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                       |  Value    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Value ..
 +-+-+-+-+
```

Figure 1

Type: This 1-octet value to define how to interpret the value with in message. The 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

```
 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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |message Type |      Length                     | Reserved      |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | message ID                                                    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | 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

```
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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |     Version   |      PDU Length               |     Reserved  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                          Identifier.                          |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Messages
 +-----
```

Figure 3

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

Message Length: Length of the PDU.

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 help 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

4.1.1. Remaining lifetime TLV

Remaining lifetime describes how long receiver keep the state without seeing a PDU from the sender. The lifetime gets updated whenever receiver accept the PDU.

Type : 1

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 open message again.

Type : 2

Length: 4 octets

Value: unsigned sequence number

4.2. BGP service advertisement message

BGP service advertisement message enable BGP to use the information to successfully bring up bgp session. The document defines transportport information TLVs and session information TLVs for BGP service advertisement message.

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.

Type : 1

Length: 4 for IPv4 transport and 16 for IPv6 transport

Value: local IPv4 or IPv6 address used for transport connection.

4.2.2. Security TTL

TTL be accepted for bgp messages

Type : 2

Length: 0

Value: The presence of the TLV indicate receiver only accept TTL with 255.

4.2.3. Security Authentication

TTL be accepted for bgp messages

Type : 3

Length: 1

Value:This support only two values 0 and 1.

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

4.2.4. TCP MSS

TCP MSS used for the connection

Type : 4

Length: 4

Value:Value in bytes

Indicates the preference of TCP MSS for the transport connection.

4.2.5. link address

This will be useful to receiver for nexthop for local address TLV when sender running IPv4 PDU and prefer IPv6 transport and vice versa.

Type : 5

Length: 4 for IPv4 transport and 16 for IPv6 transport.

Value:local IPv4 or IPv6 address used for transport connection.

4.2.6. Transport address family preference

When both IPv4 and IPv6 transport are available this, this tlv indicates sender preference

Type : 6

Length: 1 for IPv4 transport and 16 for IPv6 transport.

Value:0 IPv4 and 1 IPv6, 2 for both

5. Protocol operation

A sender should periodically send PDU to refresh the advertised information before remaining life become zero.

A sender should send the PDU to refresh the before the advertised remaining lifetime expire. If bgp is only 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 PDU. PDUs are send with source address as link primary address and destination is link local all-routers with TTL 255. if the authentication is enabled then add authentication TLV using the authentication procedure described in [TBD]. Populated other TLVs based on local preference and send the PDU in configured link. The sematic content (transport and session information) of the PDU should be same irrespective the data protocol.

Receiver reset refresh of the state whenever it accepts the 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 PUD data protocol and local address is same address family. Otherwise if link address is available and link address could be used as nexthop for the address in local address TLV. Receivers consolidate state from various TLV and pass on 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.

6. Acknowledgements

Jeff 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 advertisments
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 4

7.1. Message of SA

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

Registry Name: Messages of SA protocol

Value	Message name	Reference
0	reserved	This document
0	SA Base message	This document
1	BGP service advertisement message	This document

Figure 5

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	Reserved	This document
1	Remaining lifetime TLV	This document
2	Config sequence TLV	This document
3	Authentication	This document
224-255	Experimental	

Figure 6

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 service advertisement Message.

Value	TLV Name	Reference
0	Reserved	This document
1	Local Address	This document
2	Security TTL	This document
3	Security Authentication	This document
4	link address	This document
5	Transport preference	This document
6	TCP MSS	This document
224-255	Experimental	

Figure 7

8. Security Considerations

All drafts are required to have a security considerations section. See RFC 3552 [RFC3552] for a guide.

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

[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
Juniper Networks
Juniper Networks, 1133 Innovation Way
Sunnyvale, CA 94089
United States of America

Email: venkatashiva@juniper.net