Abstract

This document proposes a solution for BGP deployments in some specific environments to automatically establish BGP sessions without need for manual peer configuration.

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 June 10, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
1. Definitions of Terms Used in This Memo

NLRI - Network Layer Reachability Information.

RIB - Routing Information Base.

AS - Autonomous System number.

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. Introduction

The popularity of use of BGP in number of data centers or campus
deployments where BGP is being used as/instead of IGP brings
operational challenges associated with setup of BGP peering relations

This proposal aims on automating the BGP session bring-up without
aprior knowledge of the peer’s IP addresses by use of existing BGP
protocol.

3. Proposed Solution

When BGP attempts to establish the EBGP sessions with unknown peers
router will start sending BGP Session Explorer (BSE) packets which
are to be regular BGP session establishment packets containing plain
BGP OPEN Message except instead of regular peer address a multicast
address will be used 224.0.0.2 "All Routers on this Subnet" as UDP destination address. Destination port of this UDP packets is to be assigned by IANA. Source address will be selected by the operator either local interface address (when establishing plain p2p relation) or loopback address (when establishing peering between loopback addresses).

Authors leave this as open discussion point to use instead of well known multicast address of 224.0.0.2 a new IANA assigned multicast address dedicated for the purpose of BGP Auto Session Setup.

Unidirectional reception of BGP Session Explorers will allow for peer to respond with standard unicast TCP BGP OPEN Message using destination address indicated previously as source address of BSE packets. Source address in the BGP OPEN Message attempt will be peer’s local operator’s selected source address.

The above procedure will allow for automated BGP session bring-up without aprior knowledge of the peer’s BGP peering address for any AFI/SAFI.

BGP Sessions Explorers are unidirectional and only are to be sent out on those interfaces where there is no direct EBGP session established on or which would be otherwise used as recursive members of L3 group of parallel links with already recursive routes installed to correspond EBGP session between loopback addresses.

4. Loopbacks reachability bootstrapping

Upon reception of BGP Session Explorer packet on a BGP designated port BGP will parse the BGP OPEN Message in an informational mode to record peer’s interest in EBGP session establishment. However at this point there is an assumption that loopback addresses are unreachable on both sides. When session is p2p session over connected interface the reachability to session endpoints is by default in place and no further work is needed.

In the case of loopbacks after successful parsing of BGP Session Explorer packets BGP is to install in RIB BGP reachability towards the source address of the BSE source address with the outgoing interface BSE packet arrived on.

Such reachability is of temporary nature till BGP session is established between peers and peers exchange in their corresponding BGP UPDATE Messages loopback reachability with at least one next hop belonging to local connected address.
It is recommended that in the event of no session being established such temporary reachability will time out after configurable timer interval (default 180 seconds).

5. ECMP routes recursion

The described session establishment process will result in either point to point EBGP sessions or EBGP sessions between loopback addresses.

In the former case the direct point to point connected subnet is used as peering address and there is no need for any additional procedures.

In the case however when peering is established between loopbacks - typically the case in the ECMP based deployments when multiple L3 interface interconnect given pair of routers the loopback address used both as peering address and next hop of advertised routes need to recursively resolve via all directly connected subnets in order to effectively perform load balancing of traffic. For this task authors recommend regular BGP UPDATE Message to be used along with new BGP ATTRIBUTE MULTIPLE_HOP containing the list of all connected local addresses configured to be used as ECMP paths towards non connected next hop. The detailed proposal for this attribute has been described in the former work: draft-bhatia-bgp-multiple-next-hops [I-D.bhatia-bgp-multiple-next-hops].

An alternative methods for learning connected addresses towards not connected next hop can also be used. The choice of methods of accomplishing this reachability propagation is purposely made out of scope of this specification allowing both operator’s choice as well as technology evolution in this space.

6. Scalability

Parsing received to port 179 TCP packets and fixed size BGP OPEN Messages from all potential EBGP peers in applicable deployment scenarios of the target space of this proposal may result in very limited and contained need for additional processing. When port 179 is not open BGP OPEN Messages will be dropped. Upon establishing BGP session no new BGP OPEN Messages will be send on a given subnet.

7. Advantages to alternative proposals

The solutions described does not require any new message on the wire other then standard BGP OPEN Message already defined in other documents.
The proposed solution does not require any extra efforts for route installation in RIB and FIB other than via standard BGP route insertion and deletion procedures.

The proposed solutions reuses all of the existing BGP mechanisms in the space of session establishment and session maintenance and does not result in any race conditions or conflicts between existing and new procedures.

The proposed solution by its design does not allow any additional functionality like interface_ids or node/link topology discovery as it is authors believe that there are much better methods to accomplish those tasks outside of BGP protocol.

8. Changes to BGP Finite State Machine

The following changes to BGP FSM are proposed:

To be completed when/if document gets traction in the WG.

9. Security Considerations

No new security issues are introduced to the BGP protocol by this specification.

All operational security procedures which are applicable to standard BGP operation apply here.

It is highly recommended that TCP authentication when establishing unicast TCP sessions is used.

10. IANA Considerations

This document request IANA to allocate UDP port number for BGP Session Explorer messages.

11. Acknowledgments

Authors would like to thank ... for their valuable input.

12. References

12.1. Normative References


12.2. Informative References

[I-D.bhatia-bgp-multiple-next-hops]

Author’s Address

Robert Raszuk (editor)
Bloomberg LP
731 Lexington Ave
New York City, NY 10022
USA

Email: robert@raszuk.net