Workgroup: Inter-Domain Routing
Internet-Draft: draft-uttaro-idr-bgp-oad-00
Published: 10 March 2023
Intended Status: Standards Track
Expires: 11 September 2023
Authors: J. Uttaro AT&T
A. Lingala AT&T
K. Patel Arrcus, Inc.
D. Rao Cisco Systems
B. Wen Comcast
A. Retana Futurewei Technologies, Inc.
S. Sangli Juniper Networks
P. Mohapatra Sproute Networks

One Administrative Domain using BGP

Abstract

This document defines a new External BGP (EBGP) peering type known as EBGP-OAD. EBGP-OAD peering is used between two EBGP peers that belong to One Administrative Domain (OAD).

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 11 September 2023.

Copyright Notice

Table of Contents

1. Introduction

At each EBGP boundary, BGP path attributes are modified as per standard BGP rules [RFC4271]. This includes prepending the AS_PATH attribute with the autonomous-system number of the BGP speaker and stripping any IBGP-only attributes.

Some networks span more than one autonomous system and require more flexibility in the propagation of path attributes. These networks are said to belong to One Administrative Domain (OAD). It is desirable to carry IBGP-only attributes across EBGP peering when the peers belong to OAD. This document defines a new EBGP peering type known as EBGP-OAD. EBGP-OAD peering is used between two EBGP peers that belong to OAD. This document also defines rules for route announcement and processing for EBGP-OAD peers.

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. Discussion

Networks have traditionally been demarcated by an autonomous system/ BGP border which correlates to an administrative boundary. This

paradigm no longer serves the needs of network designers or customers due to the decoupling of IGP from BGP, BGP-free core in the underlay (e.g. using BGP labeled unicast [RFC8277]), the use of BGP to facilitate multiple service overlays (e.g., L2VPN, L3VPN, etc.) spanning multiple regions and AS domains, and the instantiation of customer sites on multiple content service providers (CSPs).

For example, sites in a BGP/MPLS VPN [RFC4364] may be distributed across different AS domains. In some cases, the administrator of the VPN may prefer that some attributes are propagated to all their sites to influence the BGP decision process. An example could be LOCAL_PREF which is ignored if received on an EBGP session [RFC4271].

3. Operation

[RFC4271] defines two types of BGP peerings used during a BGP protocol session. As part of the extensions defined in this document, the EBGP peering is divided into two types:

1. EBGP as defined in [RFC4271].

2. EBGP-OAD as defined below.

The EBGP-OAD session is a BGP connection between two external peers in different Autonomous Systems that belong to OAD. In general, the EBGP-OAD speakers follow the EBGP route advertisement, route processing, path attribute announcement and processing rules as defined in [RFC4271]. However, EBGP-OAD speakers are also allowed to announce and receive any IBGP-only or non-transitive attributes that were restricted to remain within an Autonomous System [RFC4271].

Unless explicitly specified, all path attributes MAY be advertised over an EBGP-OAD session. The reception of any path attribute over an EBGP-OAD session MUST NOT result in an error, unless it is malformed. Received path attributes SHOULD NOT be ignored by the receiver, unless directed to by local policy.

Unless explicitly specified, the current processes for the advertisement of path attributes remains unchanged when advertised through an EBGP-OAD peering. The process for EBGP advertisement MUST take priority over the process for IBGP advertisement. For example, the AS_PATH attribute is modified as specified in Section 5.1.2 of [RFC4271], bullet b ("BGP speaker advertises the route to an external peer").

An EBGP-OAD speaker MUST support four-octet AS numbers and avertize the "support for four-octet AS number capability" [RFC6793] .

The following sections describe modifications to route advertisements and path attribute announcements that are specific to the EBGP-OAD peering.

3.1. Next Hop Handling

It is reasonable for EBGP-OAD peers to share a common Interior Gateway Protocol (IGP). In such a case, NEXT_HOP attribute and the Next Hop in the MP_REACH_NLRI attribute [RFC4760] MAY be left unchanged.

3.2. MULTI_EXIT_DISC (MED) Handling

The determination of the neighboring AS for the purpose of BGP Route Selection [RFC4271] MAY also consider the ASN of the EBGP-OAD peer. If so, all the peers in the receiving ASN MUST be configured to use the same criteria.

3.3. Route Reflection

BGP Route Reflection [RFC4456] is an alternative to full-mesh IBGP. The ORIGINATOR_ID and CLUSTER_LIST attributes MUST NOT be advertised over an EBGP-OAD session. If received, the procedures in [RFC7606] apply.

4. Deployment and Operational Considerations

For the EBGP-OAD session to operate as expected, both BGP speakers MUST be configured with the same session type. If only one BGP speaker is configured that way, and the other uses an EBGP session, the result is that some path attributes may be ignored and others will be discarded, but the BGP session will remain operational.

The default BGP peering type for a session that is across autonomous systems SHOULD be EBGP. BGP implementation SHOULD provide a configuration-time option to enable the EBGP-OAD session type. If the session type is changed once the BGP connection has been established, the BGP speaker MUST readvertise its entire Adj-RIB-Out to its peer. Requesting a route refresh [RFC7313] is RECOMMENDED.

The requirement that Import and Export Policies exist [RFC8212] SHOULD be disabled if both peers are configured with the EBGP-OAD session type.

If multiple peerings exist between two autonomous systems that belong to OAD, all SHOULD be configured consistently. Improper configuration may result in inconsistent or unexpected forwarding. The inconsistent use of EBGP-OAD sessions is out of scope of this document.

BGP Confederations [RFC5065] provide similar behavior, on a session by session basis, as what is specified in this document. The use of confederations with an EBGP-OAD peering is out of scope of this document.

The consideration of the ASN of the EBGP-OAD peer to determine the neighboring AS for MED comparison Section 3.2 may result in the creation persistent route oscillations, similar to the Type II Churn described in [RFC3345]. [RFC7964] provides solutions and recommendations to address this issue.

5. IANA Considerations

This memo includes no request to IANA.

6. Security Considerations

This extension to BGP does not change the underlying security issues inherent in the existing BGP protocol, such as those described in [RFC4271] and [RFC4272].

This document defines a new BGP session type which combines the path attribute propagation rules for EBGP and IBGP peering. Any existing security considerations related to existing path attributes apply to the new EBGP-OAD session type.

By combining the path attribute propagation rules, IBGP information may now be propagated to another autonomous system. However, it is expected that the new session type will only be enabled when peering with a router that also belongs to OAD. If misconfigured, the impact is minimal due to the fact that both [RFC4271] and [RFC7606] define mechanisms to deal with unexpected path attributes.

7. References

7.1. 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>.

[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>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP

(IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[RFC5065] Traina, P., McPherson, D., and J. Scudder, "Autonomous System Confederations for BGP", RFC 5065, DOI 10.17487/RFC5065, August 2007, <https://www.rfc-editor.org/info/rfc5065>.

[RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space", RFC 6793, DOI 10.17487/RFC6793, December 2012, <https://www.rfc-editor.org/info/rfc6793>.

[RFC7313] Patel, K., Chen, E., and B. Venkatachalapathy, "Enhanced Route Refresh Capability for BGP-4", RFC 7313, DOI 10.17487/RFC7313, July 2014, <https://www.rfc-editor.org/info/rfc7313>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

[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>.

[RFC8212] Mauch, J., Snijders, J., and G. Hankins, "Default External BGP (EBGP) Route Propagation Behavior without Policies", RFC 8212, DOI 10.17487/RFC8212, July 2017, <https://www.rfc-editor.org/info/rfc8212>.

7.2. Informative References

[RFC3345] McPherson, D., Gill, V., Walton, D., and A. Retana, "Border Gateway Protocol (BGP) Persistent Route Oscillation Condition", RFC 3345, DOI 10.17487/RFC3345, August 2002, <https://www.rfc-editor.org/info/rfc3345>.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <https://www.rfc-editor.org/info/rfc4272>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[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>.

[RFC7964] Walton, D., Retana, A., Chen, E., and J. Scudder, "Solutions for BGP Persistent Route Oscillation", RFC 7964, DOI 10.17487/RFC7964, September 2016, <https://www.rfc-editor.org/info/rfc7964>.

[RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017, <https://www.rfc-editor.org/info/rfc8277>.

Acknowledgements

TBD

Authors' Addresses

Jim Uttaro
AT&T

Email: ju1738@att.com

Avinash Lingala
AT&T

Email: ar977m@att.com

Keyur Patel
Arrcus, Inc.

Email: keyur@arrcus.com

Dhananjaya Rao
Cisco Systems

Email: dhrao@cisco.com

Bin Wen
Comcast

Email: bin_wen@comcast.com

Alvaro Retana
Futurewei Technologies, Inc.

Email: alvaro.retana@futurewei.com

Srihari Sangli
Juniper Networks

Email: ssangli@juniper.net

Pradosh Mohapatra
Sproute Networks

Email: pradosh@sproute.com