Network Working Group   Quaizar Vohra
Internet Draft   Nuova Systems
Expiration Date: July 2007   Enke Chen
                                                               Cisco Systems

BGP Support for Four-octet AS Number Space

draft-ietf-idr-as4bytes-13.txt

Status of this Memo

Abstract

Currently the Autonomous System number is encoded as a two-octet entity in BGP. This document describes extensions to BGP to carry the Autonomous System number as a four-octet entity.

1. Introduction

Currently the Autonomous System number is encoded as a two-octet entity in BGP [BGP]. To prepare for the anticipated exhaustion of the two-octet AS numbers, this document describes extensions to BGP to carry the Autonomous System number as a four-octet entity.

More specifically, this document defines a new BGP capability, Four-octet AS Number Capability, that can be used by a BGP speaker to indicate its support for the four-octet AS numbers. Two new attributes, AS4_PATH and AS4_AGGREGATOR, are introduced that can be used to propagate four-octet based AS path information across BGP speakers that do not support the four-octet AS numbers. This document also specifies mechanisms for constructing the AS path information from the AS_PATH attribute and the AS4_PATH attribute.

The extensions proposed in this document allow a gradual transition from 2-octet AS numbers to 4-octet AS numbers.

2. Specification of Requirements

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 [RFC2119].

3. Protocol Extensions

For the purpose of this document we define a BGP speaker which does not support the new 4-octet AS number extensions as an OLD BGP speaker, and a BGP speaker which supports the new 4-octet AS number extensions as a NEW BGP speaker.

BGP carries the Autonomous System number in the "My Autonomous System" field of the OPEN message, in the AS_PATH attribute of the UPDATE message, and in the AGGREGATOR attribute of the UPDATE message. BGP also carries the Autonomous System number in the BGP Communities attribute.

A NEW BGP speaker uses BGP Capability Advertisements [RFC2842] to advertise to its neighbors (either internal or external) that it supports 4-octet AS number extensions, as specified in this document.

The Capability that is used by a BGP speaker to convey to its BGP peer the 4-octet Autonomous System number capability, also carries the 4-octet Autonomous System number of the speaker in the Capability Value field of the Capability Optional Parameter. The Capability

Length field of the Capability is set to 4.

NEW BGP speakers carry AS path information expressed in terms of 4-octet Autonomous Systems numbers by using the existing AS_PATH attribute, except that each AS number in this attribute is encoded not as a 2-octet, but as a 4-octet entity. The same applies to the AGGREGATOR attribute - NEW BGP speakers use the same attribute, except that the AS carried in this attribute is encoded as a 4-octet entity.

To preserve AS path information with 4-octet AS numbers across OLD BGP speakers, this document defines a new AS path attribute, called AS4_PATH. This is an optional transitive attribute that contains the AS path encoded with 4-octet AS numbers. The AS4_PATH attribute has the same semantics as the AS_PATH attribute, except that it is optional transitive, and it carries 4-octet AS numbers.

To prevent the possible propagation of confederation path segments outside of a confederation, the path segment types AS_CONFED_SEQUENCE and AS_CONFED_SET [RFC3065] are declared invalid for the AS4_PATH attribute.

Similarly, this document defines a new aggregator attribute called AS4_AGGREGATOR, which is optional transitive. The AS4_AGGREGATOR attribute has the same semantics as the AGGREGATOR attribute, except that it carries a 4-octet AS number.

Currently assigned 2-octet Autonomous System numbers are converted into 4-octet Autonomous System numbers by setting the two high-order octets of the 4-octet field to zero. Such a 4-octet AS number is said to be mappable to a 2-octet AS number.

To represent 4-octet AS numbers (which are not mapped from 2-octets) as 2-octet AS numbers in the AS path information encoded with 2-octet AS numbers, this document reserves a 2-octet AS number. We denote this special AS number as AS_TRANS for ease of description in the rest of this specification. This AS number is also placed in the "My Autonomous System" field of the OPEN message originated by a NEW BGP speaker if the speaker does not have a (globally unique) 2-octet AS number.

4. Operations

4.1. Interaction Between NEW BGP Speakers

A BGP speaker that supports 4-octet Autonomous System numbers SHOULD advertise this to its peers using the BGP Capability Advertisements. A BGP speaker that advertises such capability to a particular peer, and receives from that peer the advertisement of such capability MUST encode Autonomous System numbers as 4-octet entities in both the AS_PATH and the AGGREGATOR attributes in the updates it sends to the peer, and MUST assume that these attributes in the updates received from the peer encode Autonomous System numbers as 4-octet entities.

The new attributes, AS4_PATH and AS4_AGGREGATOR SHOULD NOT be carried in the UPDATE messages between NEW BGP peers. A NEW BGP speaker that receives the AS4_PATH and AS4_AGGREGATOR path attributes in an UPDATE message from a NEW BGP speaker SHOULD discard these path attributes and continue processing the UPDATE message.

4.2. Interaction Between NEW and OLD BGP Speakers

4.2.1. BGP Peering

Note that peering between a NEW BGP speaker and an OLD one is possible only if the NEW BGP speaker has a 2-octet AS number. However, this document does not assume that an Autonomous System with NEW speakers has to have a globally unique 2-octet AS number - AS_TRANS could be used instead (even if multiple Autonomous System would use it).

4.2.2. Generating Updates

When communicating with an OLD BGP speaker, a NEW speaker MUST send the AS path information in the AS_PATH attribute encoded with 2-octet AS numbers. The NEW speaker also MUST send the AS path information in the AS4_PATH attribute (encoded with 4-octet AS numbers), except for the case where the entire AS path information is composed of 2-octet AS numbers only. In this case the NEW speaker SHOULD NOT send the AS4_PATH attribute.

In the AS_PATH attribute encoded with 2-octet AS numbers, non-mappable 4-octet AS numbers are represented by the well known 2-octet AS number, AS_TRANS. This will preserve the path length property of the AS path information; and will also help in updating the AS path

information received on a NEW BGP speaker from an OLD speaker, as explained in the next section.

The NEW speaker constructs the AS4_PATH attribute from the information carried in the AS_PATH attribute. In the case where the AS_PATH attribute contains either AS_CONFED_SEQUENCE or AS_CONFED_SET path segments, the NEW speaker, when constructing the AS4_PATH attribute from the AS_PATH attribute, MUST exclude such path segments. The AS4_PATH attribute will be carried across a series of OLD BGP speakers without modification and will help preserve the truly 4-octet AS numbers in the AS path information.

Similarly, if the NEW speaker has to send the AGGREGATOR attribute, and if the aggregating Autonomous System's AS number is truly 4-octets, the speaker constructs the AS4_AGGREGATOR attributes by taking the attribute length and attribute value from the AGGREGATOR attribute and placing them into the attribute length and attribute value of the AS4_AGGREGATOR attribute, and sets the AS number field in the existing AGGREGATOR attribute to the reserved AS number, AS_TRANS. Note that if the AS number is 2-octets only, then the AS4_AGGREGATOR attribute SHOULD NOT be sent.

4.2.3. Processing Received Updates

When a NEW BGP speaker receives an update from an OLD one, it should be prepared to receive the AS4_PATH attribute along with the existing AS_PATH attribute. If the AS4_PATH attribute is also received, both the attributes will be used to construct the exact AS path information, and therefore the information carried by both the attributes will be considered for AS path loop detection.

Note that a route may have traversed a series of autonomous systems with 2-octet AS numbers and OLD BGP speakers only. In that case, if the route carries the AS4_PATH attribute, this attribute must have remained unmodified since the route left the last NEW BGP speaker. The trailing AS path information (representing autonomous systems with 2-octet AS numbers and OLD BGP speakers only) is contained only in the current AS_PATH attribute (encoded in the leading part of the AS_PATH attribute).

Under certain conditions it may not be possible to reconstruct the entire AS path information from the AS_PATH and the AS4_PATH attributes of a route. This occurs when two or more routes that carry the AS4_PATH attribute are aggregated by an OLD BGP speaker, and the AS4_PATH attribute of at least one of these routes carries at least one 4-octet AS number (as oppose to a 2-octet AS number that is encoded in 4 octets). Depending on the implementation, either the

AS4_PATH attribute would be lost during route aggregation, or both the AS_PATH attribute and the AS4_PATH attribute would contain valid, partial information that can not be combined seamlessly, resulting in incomplete AS path information in these cases.

A NEW BGP speaker should also be prepared to receive the AS4_AGGREGATOR attribute along with the AGGREGATOR attribute from an OLD BGP speaker. When both the attributes are received, if the AS number in the AGGREGATOR attribute is not AS_TRANS, then the AS4_AGGREGATOR attribute and the AS4_PATH attribute SHALL be ignored, and the AGGREGATOR attribute SHALL be taken as the information about the aggregating node, and the AS_PATH attribute SHALL be taken as the AS path information; otherwise the AGGREGATOR attribute SHALL be ignored, and the AS4_AGGREGATOR attribute SHALL be taken as the information about the aggregating node, and the AS path information would need to be constructed, as in all other cases.

In order to construct the AS path information, it would be necessary to first calculate the number of AS numbers in the AS_PATH attribute and in the AS4_PATH attribute using the method specified in Sect. 9.1.2.2 [BGP] and [RFC3065] for route selection.

If the number of AS numbers in the AS_PATH attribute is less than the number of AS numbers in the AS4_PATH attribute, then the AS4_PATH attribute SHALL be ignored, and the AS_PATH attribute SHALL be taken as the AS path information.

If the number of AS numbers in the AS_PATH attribute is larger than or equal to the number of AS numbers in the AS4_PATH attribute, then the AS path information SHALL be constructed by taking as many AS numbers and path segments as necessary from the leading part of the AS_PATH attribute, and then prepending them to the AS4_PATH attribute so that the AS path information has an identical number of AS numbers as the AS_PATH attribute. Note that a valid AS_CONFED_SEQUENCE or AS_CONFED_SET path segment SHALL be prepended if it is either the leading path segment or adjacent to a path segment that is prepended.

5. Handling BGP Communities

As specified in [RFC1997], when the high-order two-octets of the community attribute is neither 0x0000 nor 0xffff, these two octets encode the Autonomous System number. Quite clearly this would not work for BGP speakers that use 4-octets Autonomous System numbers. Such BGP speakers should use the Four-octet AS Specific Extended Communities [AS-EXT-COMM] instead.

6. Transition

The scheme described in this document allows a gradual transition from 2-octet AS numbers to 4-octet AS numbers. One can upgrade one Autonomous System or one BGP speaker at a time.

To simplify transition this document assumes that an Autonomous System could start using a 4-octet AS number only after all the BGP speakers within that Autonomous System have been upgraded to support 4-octet AS numbers.

An OLD BGP speaker MUST NOT use AS_TRANS as its Autonomous System number.

A non-mappable 4-octet AS number can not be used as a "Member AS Number" of a BGP Confederation until all the BGP speakers within the Confederation have transitioned to support 4-octet AS numbers.

In an environment where an Autonomous System that has OLD BGP speakers peers with two or more Autonomous Systems that have NEW BGP speakers and use AS_TRANS (rather than having a globally unique AS number), use of Multi-Exit Discriminators by the Autonomous System with the OLD speakers may result in a situation where Multi-Exit Discriminator will influence route selection among the routes that were received from different neighboring Autonomous Systems.

Under certain conditions it may not be possible to reconstruct the entire AS path information from the AS_PATH and the AS4_PATH attributes of a route. This occurs when two or more routes that carry the AS4_PATH attribute are aggregated by an OLD BGP speaker, and the AS4_PATH attribute of at least one of these routes carries at least one 4-octet AS number (as oppose to a 2-octet AS number that is encoded in 4 octets). When such aggregation results in creating a route that is less specific than any of the component routes, (route whose NLRI covers NLRI of all the component routes), loss of the AS path information does not create a risk of a routing loop. In all other cases loss of the AS path information does create a risk of a routing loop.

7. IANA Considerations

This document expands the pool for AS numbers from 0 - 65535 to 0 - 4294967295. The AS numbers are managed by the IANA "Autonomous System Numbers" registry. Other than expanding the AS number pool, this document does not propose any modifications to the existing policies and procedures pertaining to the AS number allocation.

This document uses a BGP Capability code to indicate that a BGP speaker supports the 4-octet AS numbers. The Capability Code 65 has been assigned by IANA per RFC 2842.

In addition, this document introduces two new BGP optional transitive attributes, and their type codes have been assigned by the IANA. The first one is the AS4_PATH attribute, value 17, which preserves the AS path information with 4-octet AS numbers across old BGP speakers. The second one is the AS4_AGGREGATOR attribute, value 18, which is similar in use to the current AGGREGATOR attribute but it carries a 4-octet AS number.

Finally, this document introduces a reserved 2-octet AS number - AS_TRANS. The AS number 23456 has been assigned by the IANA for AS_TRANS.

8. Security Considerations

This extension to BGP does not change the underlying security issues inherent in the existing BGP, except for the following:

The inconsistency between the AS_PATH attribute and the AS4_PATH attribute can create loss of the AS path information, and potential routing loops in certain cases as discussed in the document. This could be exploited by an attacker.

9. Acknowledgments

The authors would like to thank Yakov Rekhter, Chaitanya Kodeboyina, and Jeffrey Haas for the numerous discussions which went into the making of this document.

10. Normative References

[BGP] Rekhter, Y., Li, T., and Hares, S., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[RFC1997] Chandra, R., Traina, P. and Li, T., "BGP Communities Attribute", RFC 1997, August 1996.

[RFC2842] Chandra, R., and Scudder, J., "Capabilities Advertisement with BGP-4", RFC 2842, May 2000.

[RFC3065] Traina, P., McPherson, D., Scudder, J., "Autonomous System Confederations for BGP", RFC 3065, February 2001.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

11. Non-normative References

[AS-EXT-COM] Rekhter, Y., Ramachandra, S., and Tappan, D. "Four-octet AS Specific BGP Extended Community", draft-rekhter-as4octet-ext-community-00.txt, October 2005.

12. Authors' Information

Quaizar Vohra
Nuova Systems
2600 San Tomas Expressway
Santa Clara, CA 95051

Email: qv@nuovasystems.com

Enke Chen
Cisco Systems, Inc.
170 W. Tasman Dr.
San Jose, CA 95134

Email: enkechen@cisco.com

13. Intellectual Property Considerations

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

14. Full Copyright Notice