Network Working Group   D. Walton
Internet Draft   A. Retana
Expiration Date: January 2009   E. Chen
                                                     Cisco Systems
                                                        J. Scudder
                                                  Juniper Networks

                  Advertisement of Multiple Paths in BGP

                    draft-walton-bgp-add-paths-06.txt

Status of this Memo

Abstract

   In this document we propose a BGP extension that allows the
   advertisement of multiple paths for the same address prefix without
   the new paths implicitly replacing any previous ones. The essence of
   the extension is that each path is identified by a path identifier in
   addition to the address prefix.

1. Introduction

The BGP specification [RFC4271] defines an "Update-Send Process" to
advertise the routes chosen by the Decision Process to other BGP
speakers. No provisions are made to allow the advertisement of
multiple paths for the same address prefix, or Network Layer
Reachability Information (NLRI). In fact, a route with the same NLRI
as a previously advertised route implicitly replaces the previous
advertisement.

In this document we propose a BGP extension that allows the
advertisement of multiple paths for the same address prefix without
the new paths implicitly replacing any previous ones. The essence of
the extension is that each path is identified by a path identifier in
addition to the address prefix.

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. How to Identify a Path

As defined in [RFC4271], a path refers to the information reported in
the path attribute field of an UPDATE message. As the procedures
specified in [RFC4271] allow only the advertisement of one path for a
particular address prefix, a path for an address prefix from a BGP
peer can be keyed on the address prefix.

In order for a BGP speaker to advertise multiple paths for the same
address prefix, a new identifier (termed "Path Identifier" hereafter)
needs to be introduced so that a particular path for an address
prefix can be identified by the combination of the address prefix and
the Path Identifier.

The assignment of the Path Identifier for a path by a BGP speaker is
purely a local matter. However, the Path Identifier MUST be assigned
in such a way that the BGP speaker is able to use the (prefix, path
identifier) to uniquely identify a path advertised to a neighbor. A
BGP speaker that re-advertises a route MUST generate its own Path
Identifier to be associated with the re-advertised route. A BGP
speaker that receives a route SHOULD NOT assume that the identifier
carries any particular semantics; it SHOULD be treated as an opaque
value.

4. Extended NLRI Encodings

   In order to carry the Path Identifier in an UPDATE message, the
   existing NLRI encodings are extended by prepending the Path
   Identifier field, which is of four-octets.

   For example, the NLRI encodings specified in [RFC4271, RFC4760] are
   extended as the following:

```
            +-------------------------------+
            | Path Identifier (4 octets)    |
            +-------------------------------+
            | Length (1 octet)              |
            +-------------------------------+
            | Prefix (variable)             |
            +-------------------------------+
```

   and the NLRI encoding specified in [RFC3107] is extended as the
   following:

```
            +-------------------------------+
            | Path Identifier (4 octets)    |
            +-------------------------------+
            | Length (1 octet)              |
            +-------------------------------+
            | Label (3 octets)              |
            +-------------------------------+
            | ...                           |
            +-------------------------------+
            | Prefix (variable)             |
            +-------------------------------+
```

   The usage of the extended NLRI encodings is specified in the
   Operation section.

5. ADD-PATH Capability

The ADD-PATH Capability is a new BGP capability [RFC3392]. The
Capability Code for this capability is specified in the IANA
Considerations section of this document. The Capability Length field
of this capability is variable. The Capability Value field consists
of one or more of the following tuples:

```
        +-----------------------------------------------+
        | Address Family Identifier (2 octets)          |
        +-----------------------------------------------+
        | Subsequent Address Family Identifier (1 octet) |
        +-----------------------------------------------+
        | Send/Receive (1 octet)                        |
        +-----------------------------------------------+
```

The meaning and use of the fields are as follows:

    Address Family Identifier (AFI):

        This field is the same as the one used in [RFC4760].

    Subsequent Address Family Identifier (SAFI):

        This field is the same as the one used in [RFC4760].

    Send/Receive:

        This field indicates whether the sender is (a) willing to
        receive multiple paths from its peer (value 1), (b) would
        like to send multiple paths to its peer (value 2), or (c)
        both (value 3) for the <AFI, SAFI>.

6. Operation

The Path Identifier specified in the previous section can be used to
advertise multiple paths for the same address prefix without
subsequent advertisements replacing the previous ones. Apart from
the fact that this is now possible, the route advertisement rules of
[RFC4271] are not changed. In particular, a new advertisement for a
given address prefix and a given path identifier replaces a previous
advertisement for the given address prefix and the given path
identifier.

A BGP speaker that is willing to receive multiple paths from its

peer, or would like to send multiple paths to its peer, SHOULD
advertise the ADD-PATH Capability to the peer using BGP Capabilities
advertisement [RFC3392].

A BGP speaker MUST follow the existing procedures in generating an
UPDATE message for a particular <AFI, SAFI> to a peer unless the BGP
speaker advertises the ADD-PATH Capability to the peer indicating its
desire to send multiple paths for the <AFI, SAFI>, and also receives
the ADD-PATH Capability from the peer indicating its willingness to
receive multiple paths for the <AFI, SAFI>, in which case the speaker
MUST generate a route update for the <AFI, SAFI> based on the
combination of the address prefix and the Path Identifier, and use
the extended NLRI encodings specified in this document. The peer
SHALL act accordingly in processing an UPDATE message related to a
particular <AFI, SAFI>.

7. **Applications**

The BGP extension specified in this document can be used by a BGP
speaker to advertise multiple paths in certain applications. The
availability of the additional paths can help reduce or eliminate
persistent route oscillations [RFC3345]. It can also help with
optimal routing and routing convergence in a network. The
applications are detailed in separate documents.

8. **Deployment Considerations**

The extension proposed in this document provides a mechanism for a
BGP speaker to advertise multiple paths over a BGP session. Care
needs to be taken in its deployment to ensure consistent routing and
forwarding in a network, the details of which will be described in
separate application documents.

9. **IANA Considerations**

IANA needs to assign a capability number for the ADD-PATH Capability
described in this document.

10. Security Considerations

This document introduces no new security concerns to BGP or other specifications referenced in this document.

11. Acknowledgments

We would like to thank David Cook and Naiming Shen for their contributions to the design and development of the extension, and for co-authoring drafts that lead to the current document.

Many people have made valuable comments and suggestions, including Dave Meyer, Srihari Sangli, Eric Rosen, Dan Tappan, Robert Raszuk, Mark Turner, Danny McPherson, Eugene Kim, Pradosh Mohapatra, Rex Fernando, and Keyur Patel.

12. Normative References

[RFC4271] Rekhter, Y., T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)," RFC 4271, January 2006.

[RFC3392] Chandra, R. and J. Scudder, "Capabilities Advertisement with BGP-4," RFC 3392, November 2002.

[RFC4760] Bates, T., Chandra, R., Rekhter, Y., and D. Katz, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[RFC3107] Rekhter, R. and E. Rosen, "Carrying Label Information in BGP-4," RFC 3107, May 2001.

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

13. Informative References

[RFC3345] McPherson, D., V. Gill, D. Walton, and A. Retana, "Border Gateway Protocol (BGP) Persistent Route Oscillation Condition", RFC 3345, August 2002.

14. Authors' Addresses

    Daniel Walton
    Cisco Systems, Inc.
    7025 Kit Creek Rd.
    Research Triangle Park, NC 27709

    Email: dwalton@cisco.com

    Alvaro Retana
    Cisco Systems, Inc.
    7025 Kit Creek Rd.
    Research Triangle Park, NC 27709

    Email: aretana@cisco.com

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

    Email: enkechen@cisco.com

    John Scudder
    Juniper Networks

    Email: jgs@juniper.net

15. Intellectual Property Considerations

    This section is taken from Section 5 of RFC 3668.

    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

16. Copyright Notice