XXX
XXX WARNING: old character encoding and/or character set
XXX

Network Working Group   D. Walton
Internet Draft   E. Chen
Intended Status: Standards Track   Cisco Systems
Expiration Date: December 18, 2012   A. Retana
                                               Hewlett-Packard Co.
                                                         J. Scudder
                                                   Juniper Networks
                                                     June 17, 2012

Advertisement of Multiple Paths in BGP

draft-ietf-idr-add-paths-07.txt

Status of this Memo

Copyright Notice

INTERNET DRAFT   draft-ietf-idr-add-paths-07.txt   June 2012

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](1). Introduction

    The BGP specification [[RFC4271](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.

[1.1](1.1). 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].

2. 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.

3. 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.

4. ADD-PATH Capability

The ADD-PATH Capability is a new BGP capability [RFC5492]. 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>.

5. 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 [RFC5492].

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>.

A BGP speaker SHOULD include the bestpath when more than one path are
advertised to a neighbor unless the bestpath is a path received from
that neighbor.

When deployed as a provider edge router or a peering router that
interacts with external neighbors, a BGP speaker usually advertises
at most one path to the internal neighbors in a network. In the case
the speaker is configured to advertise multiple paths to the internal
neighbors, it should include the Edge_Discriminator attribute defined
in [FAST-CONV] in order to make the route selection consistent inside
the network.

As the Path Identifiers are locally assigned, and may or may not be
persistent across a control plane restart of a BGP speaker, an
implementation SHOULD take special care so that the underlying

forwarding plane of a "Receiving Speaker" as described in [RFC4724]
is not affected during the graceful restart of a BGP session.

6. 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.

7. 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.

8. IANA Considerations

IANA has assigned capability number 69 for the ADD-PATH Capability
described in this document. This registration is in the BGP
Capability Codes registry.

9. Security Considerations

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

10. Acknowledgments

We would like to thank David Cook and Naiming Shen for their
contributions to the design and development of the extension.

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

11. References

11.1. Normative References

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

   [RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement
   with BGP-4", RFC 5492, February 2009.

   [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.

   [RFC4724] Sangli, S., E. Chen, R. Fernando, J. Scudder, and Y.
   Rekhter, "Graceful Restart Mechanism for BGP", RFC 4724, January
   2007.

   [FAST-CONV] Mohapatra, P., R. Fernando, C. Filsfils, R. Raszuk, "Fast
   Connectivity Restoration Using BGP Add-path", Work in Progress, March
   2011.

11.2. Informative References

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

12. Authors' Addresses

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

   Email: dwalton@cisco.com

   Alvaro Retana
   Hewlett-Packard Co.

2610 Wycliff Road
Raleigh, NC   27607

Email: alvaro.retana@hp.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