

Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 17, 2015

A. Retana
Cisco Systems, Inc.
November 13, 2014

Advertisement of Multiple Paths in BGP: Implementation Report
draft-retana-idr-add-paths-implementation-01

Abstract

This document reports the results of an ADD-PATH implementation survey. The survey had 22 questions about implementations' support for advertising multiple paths in BGP. After a brief summary of the results, each response is listed. This document contains responses from six implementers who completed the survey.

The editor did not use external means to verify the accuracy of the information submitted by the respondents. The respondents are considered experts on the products they reported on.

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 <http://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 May 17, 2015.

Copyright Notice

Copyright (c) 2014 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 (<http://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.

Table of Contents

- [1. Introduction](#) [3](#)
- [2. Requirements Language](#) [3](#)
- [3. Results of the Survey](#) [3](#)
 - [3.1. Overview of Differences](#) [3](#)
 - [3.2. Implementation Identification](#) [4](#)
 - [3.3. Implementations and Interoperability](#) [5](#)
- [4. Implementation Report](#) [5](#)
 - [4.1. Section 2: How to Identify a Path](#) [6](#)
 - [4.1.1. Base Behavior](#) [6](#)
 - [4.1.2. Path Identifier Assignment](#) [6](#)
 - [4.1.3. Path Identifier Assignment \(2\)](#) [6](#)
 - [4.1.4. Route Re-advertisement](#) [7](#)
 - [4.1.5. Received Path Identifier](#) [7](#)
 - [4.2. Section 3: Extended NLRI Encodings](#) [8](#)
 - [4.2.1. Base Behavior](#) [8](#)
 - [4.3. Section 4: ADD-PATH Capability](#) [8](#)
 - [4.3.1. Base Behavior](#) [8](#)
 - [4.4. Section 5: Operation](#) [9](#)
 - [4.4.1. Base Behavior](#) [9](#)
 - [4.4.2. Implicit Replacement](#) [9](#)
 - [4.4.3. Silently Ignore](#) [9](#)
 - [4.4.4. Send/Receive Logic](#) [10](#)
 - [4.4.5. Update Procedure](#) [10](#)
 - [4.4.6. Update Generation with Encoding](#) [11](#)
 - [4.4.7. Multiple Address Family Support](#) [11](#)
 - [4.4.8. Multiple Address Family Support \(2\)](#) [12](#)
 - [4.4.9. Bestpath](#) [12](#)
 - [4.4.10. Path Identifier Persistency](#) [13](#)
 - [4.4.11. Graceful Restart](#) [13](#)
 - [4.5. Section 6: Applications](#) [14](#)
 - [4.5.1. Applications](#) [14](#)
 - [4.6. Section 7: Deployment Considerations](#) [15](#)
 - [4.6.1. Deployment Experience](#) [15](#)
- [5. Security Considerations](#) [15](#)
- [6. IANA Considerations](#) [15](#)
- [7. Acknowledgements](#) [15](#)

8. References	15
8.1. Normative References	15
8.2. Informative References	16
Author's Address	16

[1.](#) Introduction

This document reports results from a survey of implementations of the Advertisement of Multiple Paths in BGP [[I-D.ietf-idr-add-paths](#)], where a BGP [[RFC4271](#)] extension that allows the advertisement of multiple paths for the same address prefix without the new paths implicitly replacing any previous ones is defined. The essence of the extension is that each path is identified by a path identifier in addition to the address prefix.

The ADD-PATH implementation survey had 22 detailed questions about compliance with [[I-D.ietf-idr-add-paths](#)]. Six implementers (Cumulus Networks, Cisco Systems, Exa Networks, Juniper Networks, Alcatel-Lucent and CZ.NIC) completed the survey. [Section 3.1](#) provides an overview of the differences between the implementations. [Section 4](#) provides a compilation of the results.

[2.](#) Requirements Language

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.](#) Results of the Survey

The respondents replied "Yes" or "No" to the survey's questions to indicate whether their implementation supports the Functionality/Description of the [[RFC2119](#)] language in [[I-D.ietf-idr-add-paths](#)]. The respondents replied "Other" to indicate an alternate behavior and had the opportunity to provide comments in all cases. Some questions were informative.

[3.1.](#) Overview of Differences

This section provides the reader with a shortcut to the points where the implementations differ.

Two of the implementations work only in receive-mode; they don't implement any advertisement of routes. Obviously, those implementations are not compliant with the sections related to the advertisement of routes. Taking that fact into account, all the responders had consistent and compliant answers to all the sections of the survey.

[3.2.](#) Implementation Identification

3.3.1. Cumulus

Company/Organization Name: Cumulus Networks

Implementation Name/Version: quagga

Date: 11/3/2014

Contact Name: Daniel Walton

Contact e-mail: dwalton@cumulusnetworks.com

3.3.2. Cisco

Company/Organization Name: Cisco Systems

Implementation Name/Version: IOS-XE

Date: 11/03/2014

Contact Name: Mohammed Mirza

Contact e-mail: mohamirz@cisco.com

3.3.3. Exa

Company/Organization Name: Exa Networks

Implementation Name/Version: ExaBGP

Date: 01/11/2014

Contact Name: Thomas Mangin

Contact e-mail: thomas.mangin@exa-networks.co.uk

3.3.4. Juniper

Company/Organization Name: Juniper Networks

Implementation Name/Version: JUNOS 11.3 and later

Date: August 2011

Contact Name: Jeff Haas

Retana

Expires May 17, 2015

[Page 4]

Internet-Draft

ADD-PATH Implementation Report

November 2014

Contact e-mail: jhaas@juniper.net

3.3.5. ALU

Company/Organization Name: Alcatel-Lucent

Implementation Name/Version: SROS

Date: 11/10/2014

Contact Name: Adam Simpson

Contact e-mail: adam.simpson@alcatel-lucent.com

3.3.6. CZ.NIC

Company/Organization Name: CZ.NIC

Implementation Name/Version: BIRD

Date: 2014-11-12

Contact Name: Ondrej Zajicek

Contact e-mail: santiago@crfreenet.org

3.3. Implementations and Interoperability

	Cumulus	Cisco	Exa	Juniper	ALU	CZ.NIC
Cumulus		Yes				Yes
Cisco		Yes				
Exa		Yes				
Juniper						
ALU		Yes				
CZ.NIC						

4. Implementation Report

For every item listed, the respondents indicated whether their implementation supports the Functionality/Description or not (Yes/No) according to the [[RFC2119](#)] language indicated. Any comments are included. If appropriate, the respondents indicated with "Other" the fact that the support is neither Yes/No (an alternate behavior, for example). Refer to the appropriate sections in [[I-D.ietf-idr-add-paths](#)] for additional details.

4.1. [Section 2](#): How to Identify a Path

4.1.1. Base Behavior

Functionality/Description: Is your implementation compatible with the use of the Path Identifier as described in this section?

[[RFC2119](#)]: N/A

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Yes	
Juniper	Yes	

ALU	Yes
CZ.NIC	Yes

4.1.2. Path Identifier Assignment

Functionality/Description: Explain how Path Identifiers are assigned in your implementation.

[RFC2119]: N/A

Implementation Comments

Implementation	Comments
Cumulus	quagga is RX only for now so this is not an issue
Cisco	Each net has unique path-id per paths under it. The path ids that are withdrawn can get assigned to the newer paths.
Exa	By the user
Juniper	Incrementally assign an id based on the N+1 of the max(N) of the path ids already assigned.
ALU	Path IDs are per address family. Every new advertised path uses the next available path ID (in sequential order) for the address family.
CZ.NIC	Each route source (like add_path-unaware BGP peer) has allocated fixed path id.

4.1.3. Path Identifier Assignment (2)

Functionality/Description: "...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."

Can your implementation uniquely identify an advertised path based on the (prefix, path identifier) pair?

[RFC2119]: MUST

Implementation Yes/No/Other Comments

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Other	This is left to the user of the application to do.

Juniper	Yes
ALU	Yes
CZ.NIC	Yes

[4.1.4.](#) Route Re-advertisement

Functionality/Description: "A BGP speaker that re-advertises a route MUST generate its own Path Identifier to be associated with the re-advertised route."

Does your implementation generate a new Path Identifier when re-advertising a route?

[RFC2119]: MUST

Implementation	Yes/No/Other	Comments
Cumulus	Other	Comments quagga does not support TX yet
Cisco	Yes	
Exa	Other	ExaBGP does not re-advertise routes
Juniper	Yes	
ALU	Yes	
CZ.NIC	Other	New path_id is allocated for each unique path_id received through add_path-aware BGP session.

[4.1.5.](#) Received Path Identifier

Functionality/Description: "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."

Does your implementation treat a received Path Identifier as an opaque value?

[RFC2119]: SHOULD NOT

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Yes	
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

[4.2. Section 3](#): Extended NLRI Encodings

[4.2.1.](#) Base Behavior

Functionality/Description: Does your implementation use the encodings specified in this section?

[RFC2119]: N/A

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Yes	
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

[4.3. Section 4](#): ADD-PATH Capability

[4.3.1.](#) Base Behavior

Functionality/Description: Is your implementation able to send and receive the ADD-PATH Capability as described in this section?

[RFC2119]: N/A

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Yes	
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

[4.4. Section 5](#): Operation

[4.4.1.](#) Base Behavior

Functionality/Description: Is your implementation compatible with the operation described in this section?

[RFC2119]: N/A

Implementation	Yes/No/Other	Comments
Cumulus	Other	RX yes, TX not implemented
Cisco	Yes	
Exa	Yes	
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

[4.4.2.](#) Implicit Replacement

Functionality/Description: "...a new advertisement for a given address prefix and a given path identifier replaces a previous advertisement for the same address prefix and path identifier."

Does your implementation replace previous advertisements with the same (prefix, path identifier) pair?

[RFC2119]: N/A

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Other	ExaBGP does not implement a FIB
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

[4.4.3.](#) Silently Ignore

Functionality/Description: "If a BGP speaker receives a message to withdraw a prefix with a path identifier not seen before, it SHOULD silently ignore it."

Does your implementation silently ignore the withdraw of a prefix with a new path identifier?

[RFC2119]: SHOULD

Implementation	Yes/No/Other	Comments
Cumulus		
Cisco	Yes	
Exa	Other	ExaBGP is a "BGP engine", it only convert BGP packet to some JSON that another application can consume (and vice-versa).
Juniper	Yes	
ALU	Yes	
CZ.NIC		

4.4.4. Send/Receive Logic

Functionality/Description: "For a BGP speaker to be able to send multiple paths to its peer, that BGP speaker MUST advertise the ADD-PATH capability with the Send/Receive field set to either 2 or 3, and MUST receive from its peer the ADD-PATH capability with the Send/Receive field set to either 1 or 3, for the corresponding <AFI, SAFI>."

Does your implementation follow the send/receive logic as specified in this section?

[RFC2119]: MUST

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Yes	
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

4.4.5. Update Procedure

Functionality/Description: "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 ability to send multiple paths for the <AFI, SAFI>, and also receives the ADD-PATH Capability from the peer indicating its ability to receive multiple paths for the <AFI, SAFI>..."

Does your implementation follow normal procedures when generating UPDATES if the ADD-PATH capability is not sent and received?

[RFC2119]: MUST

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Yes	
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

4.4.6. Update Generation with Encoding

Functionality/Description: "...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."

If the ADD-PATH capability has been sent and received, does your implementation generate new UPDATES using the (prefix, path identifier) pair and the encodings defined in this document?

[RFC2119]: MUST

Implementation	Yes/No/Other	Comments
Cumulus	Other	TX is not supported yet
Cisco	Yes	
Exa	Yes	
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

[4.4.7.](#) Multiple Address Family Support

Functionality/Description: "The peer SHALL act accordingly in processing an UPDATE message related to a particular <AFI, SAFI>."

Does your implementation support the use of the ADD-PATH capability for multiple <AFI, SAFI> pairs?

[[RFC2119](#)]: SHALL

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Yes	
Juniper	Yes	
ALU	Yes	
CZ.NIC	Other	BIRD currently does not support multiple pairs in one connection, separate connection is used for IPv4 and IPv6 (unicast).

[4.4.8.](#) Multiple Address Family Support (2)

Functionality/Description: Which <AFI, SAFI> pairs does your implementation support when using the ADD-PATH capability?

[[RFC2119](#)]: N/A

Implementation	Comments
Cumulus	IPv4 unicast and IPv6 unicast
Cisco	ipv4 unicast and ipv6 unicast
Exa	1/1 2/1 1/4 2/4
Juniper	IPv4 Unicast, IPv6 Unicast, IPv4 Labeled Unicast, IPv6

	Labeled Unicast
ALU	1/1, 1/4, 1/128, 2/1, 2/4, 2/128
CZ.NIC	IPv4 unicast and IPv6 unicast

[4.4.9.](#) Bestpath

Functionality/Description: "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."

Does your implementation include the bestpath when multiple paths are announced to a neighbor, as described?

[\[RFC2119\]](#): SHOULD

Implementation	Yes/No/Other	Comments
Cumulus	Yes	
Cisco	Yes	
Exa	Other	ExaBGP does not have a FIB, this is user controlled.
Juniper	Yes	
ALU	Yes	
CZ.NIC	Yes	

[4.4.10.](#) Path Identifier Persistency

Functionality/Description: "As the Path Identifiers are locally assigned, and may or may not be persistent across a control plane restart of a BGP speaker..."

Are the path identifiers persistent across control plane restarts in your implementation?

[RFC2119]: N/A

Implementation Yes/No/Other Comments

Cumulus	No	
Cisco	No	XE-BGP-ADD-Paths need to have HA enhancements
Exa	Other	User controlled
Juniper	Other	In the case of the BGP graceful restart feature, path IDs are not persistent. In the case of the JUNOS Non-stop Routing feature, they persist.
ALU	No	With high availability (HA) the path IDs are persistent if there is still one remaining control card after reset/failure of the other control card.
CZ.NIC	No	

[4.4.11](#). Graceful Restart

Functionality/Description: "...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."

Please explain how your implementation addresses Graceful Restart.

[RFC2119]: SHOULD

Implementation Comments

Cumulus	Quagga has partial GR support (it is GR aware for other restarting nodes) but does not maintain the forwarding plane during a restart.
Cisco	XE-BGP-ADD-Paths need to have HA enhancements
Exa	No FIB, not relevant
Juniper	During BGP graceful restart procedures, the receiving speaker ignores the path-id for purposes of identifying a matching route. Once a refreshed route has been correlated to a previous path, the path-id is

ALU	updated. Graceful restart is supported for the receiving router role so by definition graceful restart does not affect the forwarding plane.
CZ.NIC	FIB is not modified until initial graceful restart phase is finished.

4.5. [Section 6](#): Applications

4.5.1. Applications

Functionality/Description: Please list or explain which applications that require the propagation of multiple paths are supported by your implementation.

[[RFC2119](#)]: N/A

Implementation Comments

Cumulus	None yet....RX onlys
Cisco	1. RR client to RR use cases for ipv4 and ipv6. 2. RR to RR clients(could be ASBRs) use cases for ipv4 and ipv6.
Exa	N/A
Juniper	Persistent route flap damping suppression. Distribution of additional destinations or BGP nexthops for multi-path purposes.
ALU	Add-Paths ion IBGP sessions allows for better load-sharing (more ECMP paths), advertisement of potential backup paths, reduced routing churn.
CZ.NIC	(iBGP) route reflector / RR client, (eBGP) route server / RS client, use cases where paths are distributed for other purposes than filling FIBs (like topology-aware CDNs).

4.6. [Section 7](#): Deployment Considerations

4.6.1. Deployment Experience

Functionality/Description: Please comment on deployment experience with your implementation.

[RFC2119]: N/A

Implementation Comments

Cumulus

Cisco

Exa

Cisco routers exporting ADD-PATH routes to ExaBGP, routes are then stored in a distributed Database. A complex best path selection (including latency) is performed on the stored routes, and the best routes are then re-injected in the core via ExaBGP.

Juniper

ALU

CZ.NIC

5. Security Considerations

This document reports the results of an ADD-PATH implementation survey. As such, it does not introduce any security risks.

6. IANA Considerations

This document has no IANA actions.

7. Acknowledgements

The editor would like to thank Daniel Walton, Mohammed Mirza, Thomas Mangin, Jeff Haas, Adam Simpson and Ondrej Zajicek.

8. References

8.1. Normative References

[I-D.ietf-idr-add-paths]

Walton, D., Retana, A., Chen, E., and J. Scudder, "Advertisement of Multiple Paths in BGP", [draft-ietf-idr-add-paths-10](#) (work in progress), October 2014.

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

[8.2.](#) Informative References

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

Author's Address

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

Email: aretana@cisco.com

