

Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (1999). All Rights Reserved.

Abstract

BGP-4 Multiprotocol Extensions [BGP-MP] defines the format of two BGP attributes (MP_REACH_NLRI and MP_UNREACH_NLRI) that can be used to announce and withdraw the announcement of reachability information. This document defines how compliant systems should make use of those attributes for the purpose of conveying IPv6 routing information.

1. Introduction

The BGP-4 protocol [BGP-4] in particular, and path vector routing protocols in general, are mostly independent of the particular Address Family for which the protocol is being used.

IPv6 falls under the generic category of protocols for which BGP-4 is suitable and, unless stated otherwise in this document, the BGP-4 procedures to apply when using BGP-4 to carry IPv6 reachability information are those defined in [BGP-4] and in subsequent documents that extend or update the BGP-4 specification.

In terms of routing information, the most significant difference between IPv6 and IPv4 (for which BGP was originally designed) is the fact that IPv6 introduces scoped unicast addresses and defines particular situations when a particular address scope must be used. This document concerns itself essentially with the necessary rules to accommodate IPv6 address scope requirements.

2. IPv6 Address Scopes

IPv6 defines 3 unicast address scopes [ADDR-ARCH]: global, site-local and link-local. Site-local addresses are non-link-local addresses which are valid within the scope of a "site" and cannot be exported beyond it. As this document makes no assumption on the characteristics of a particular routing realm where BGP-4 is used, it makes no distinction between global and site-local addresses and refers to both as "global" or "non-link-local". Network administrators must however respect address scope restrictions and should be aware that the concepts of a BGP-4 routing domain and "site" are orthogonal notions and that they may or may not coincide in a given situation.

Companion IPv6 specifications further define that only link-local address can be used when generating ICMP Redirect Messages [ND] and as next hop addresses in some routing protocols (eg. RIPng [RIP]).

This restriction does imply that an IPv6 router must have a link-local next hop address for all directly connected routes (routes for which the given router and the next hop router share a common subnet prefix).

Link-local addresses are not, however, well suited to be used as next hop attributes in BGP-4 given the rules defined for this attribute in the protocol specification [BGP-4].

For the above reasons, when BGP-4 is used to convey IPv6 reachability information it is sometimes necessary to announce a next hop attribute that consists of a global address and a link-local address. The following section describes the rules that should be followed when constructing the Network Address of Next Hop field of an MP_REACH_NLRI attribute.

3. Constructing the Next Hop field

A BGP speaker shall advertise to its peer in the Network Address of Next Hop field the global IPv6 address of the next hop, potentially followed by the link-local IPv6 address of the next hop.

The value of the Length of Next Hop Network Address field on a MP_REACH_NLRI attribute shall be set to 16, when only a global address is present, or 32 if a link-local address is also included in the Next Hop field.

The link-local address shall be included in the Next Hop field if and only if the BGP speaker shares a common subnet with the entity identified by the global IPv6 address carried in the Network Address of Next Hop field and the peer the route is being advertised to.

In all other cases a BGP speaker shall advertise to its peer in the Network Address field only the global IPv6 address of the next hop (the value of the Length of Network Address of Next Hop field shall be set to 16).

As a consequence, a BGP speaker that advertises a route to an internal peer may modify the Network Address of Next Hop field by removing the link-local IPv6 address of the next hop.

4. Transport

TCP connections, on top of which BGP-4 messages are exchanged, can be established either over IPv4 or IPv6. While BGP-4 itself is independent of the particular transport used it derives implicit configuration information from the address used to establish the peering session. This information (the network address of a peer) is taken in account in the route dissemination procedure. Thus, when using TCP over IPv4 as a transport for IPv6 reachability information, additional explicit configuration of the peer's network address is required.

Note that the information referred above is distinct from the BGP Identifier used in the BGP-4 tie breaking procedure. The BGP Identifier is a 32 bit unsigned integer exchanged between two peers at session establishment time, within an OPEN message. This is a system wide value determined at startup which must be unique in the network and should be derived from an IPv4 address regardless of the network protocol(s) a particular BGP-4 instance is configured to convey at a given moment.

The use of TCP over IPv6 as transport protocol for IPv6 reachability information also has the advantage of providing explicit confirmation of IPv6 network reachability between two peers.

5. Security Considerations

The extensions defined in this document allow BGP to propagate reachability information about IPv6 routes. As such, no new security issues are raised beyond those that already exist in BGP-4 and its use with IPv4.

6. Acknowledgments

This document derives from the work on BGP-4 Multiprotocol Extensions by Tony Bates, Ravi Chandra, Dave Katz and Yakov Rekhter.

7. References

- [ADDR-ARCH] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 2373, July 1998.
- [BGP-4] Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995.
- [BGP-MP] Bates, T., Chandra, R., Katz, D. and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 2283, February 1998.
- [IPv6] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.
- [ND] Narten, T., Nordmark, E. and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.
- [RIP] Malkin, G. and R. Minnear, "RIPng for IPv6", RFC 2080, January 1997.

8. Author Information

Pedro R. Marques
Cisco Systems, Inc.
170 W. Tasman Dr.
San Jose, CA 95134
USA

Phone: +1 408 527-5202
Fax: +1 408 526-4952
Email: roque@cisco.com

Francis Dupont
GIE DYADE
INRIA Rocquencourt
Domaine de Voluceau
BP 105
78153 Le Chesnay CEDEX
FRANCE

Phone: +33 1 39 63 52 13
Fax: +33 1 39 63 58 66
Email: Francis.Dupont@inria.fr

9. Full Copyright Statement

Copyright (C) The Internet Society (1999). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

