

November 1997

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

[draft-ietf-idr-bgp4-ipv6-00.txt](#)

### Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

To learn the current status of any Internet-Draft, please check the "1id-abstracts.txt" listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), nic.nordu.net (Europe), munnari.oz.au (Pacific Rim), ds.internic.net (US East Coast), or ftp.isi.edu (US West Coast).

### 1. 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.

## **2. 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.

## **3. IPv6 Address Scopes**

IPv6 defines 3 unicast address scopes [[ADDR-ARCH](#)]: global, site-local and link-local. Site-local addresses are non-link-local address 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 restrictions 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.

#### **4. 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.

#### **5. 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.

## **6. Capability Negotiation**

BGP-4 speakers using Multiprotocol Extensions to carry IPv6 NLRI information should use the Capabilities Optional Parameter as defined in [BGP-CAP]. The MP\_EXT Capability Code, as defined in [CAP-MP], is used to negotiate the (AFI, SAFI) pairs available on a particular connection.

## **7. Security Considerations**

Routing protocols such as the one discussed in this document provide control information for the network infrastructure and are, as such, sensible in security terms.

The authors recommend that authentication mechanisms such as IPSEC AH are used to guaranty the integrity and origin of the information.

## **8. Acknowledgments**

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

## **9. References**

- [ADDR-ARCH] "IP Version 6 Addressing Architecture",  
S. Deering and R. Hiden, Internet Draft, November 1997.  
<[draft-ietf-ipngwg-addr-arch-v2-04.txt](#)>
- [BGP-4] "A Border Gateway Protocol 4 (BGP-4)",  
Y. Rekhter and T. Li, [RFC1771](#), March 1995.
- [BGP-CAP] "Capabilities Negotiation with BGP-4",  
R. Chandra and J. Scudder, Internet Draft, August 1997.  
<[draft-ietf-idr-bgp4-cap-neg-00.txt](#)>



- [BGP-MP] "Multiprotocol Extensions for BGP-4",  
T. Bates, R. Chandra, D. Katz, and Y. Rekhter,  
Internet Draft, September 1997.  
<[draft-ietf-idr-bgp4-multiprotocol-01.txt](#)>
- [CAP-MP] "BGP-4 Capabilities Negotiation for BGP Multiprotocol  
Extensions", P. Marques, November 1997.  
<[draft-marques-bgp4-cap-mp-00.txt](#)>
- [IPv6] "Internet Protocol, Version 6 (IPv6) Specification",  
S. Deering and R. Hiden, [RFC1883](#), December 1995.
- [ND] "Neighbor Discovery for IP Version 6 (IPv6)",  
T. Narten, E. Nordmark and W. Simpson,  
Internet Draft, July 1997.  
<[draft-ietf-ipngwg-discovery-v2-00.txt](#)>
- [RIP] "RIPng for IPv6",  
G. Malkin and R. Minnear, [RFC 2080](#), January 1997.

## **10. Author Information**

Pedro R. Marques  
cisco Systems, Inc.  
170 W. Tasman Dr.  
San Jose, CA 95134  
  
email: [roque@cisco.com](mailto:roque@cisco.com)

Francis Dupont  
INRIA Rocquencourt  
Domaine de Voluceau  
BP 105  
78153 Le Chesnay CEDEX  
France

email: [Francis.Dupont@inria.fr](mailto:Francis.Dupont@inria.fr)



