Network Working Group
Internet-Draft
Expires June 1997

Dimitry Haskin
Bay Networks
John W. Stewart, III
ISI

Multiprotocol Extensions for BGP

draft-stewart-bgp-multiprotocol-00.txt

1. 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).

2. Abstract

This document outlines a proposal for a BGP Version 5 (BGP5) which has the ability to carry routing information for a variety of network protocols. The proposed BGP modifications are intended to be simple enough to be easily added to the existing BGP implementations and, at the same time, provide for a longer than 16-bit AS number space.

This document only describes BGP5 support for IPv4 and IPv6 networks.

3. Overview

BGP [BGP4-RFC] is widely deployed as an inter-domain routing protocol for IPv4 and there is a large operational experience with it. Though currently BGP can only carry routing information for IPv4, the existence of the mature BGP implementation base provides a strong motivation for extending it to support the Internet Protocol version 6 (IPv6) and possibly other network protocols.

This document outlines a proposal for creating a BGP Version 5 (BGP5) which has the ability to carry routing information for a variety

of network protocols. Adding this multi-protocol support to BGP significantly increases its lifetime, so care should be taken for BGP5 to not be under-engineered. For this reason, BGP5 supports longer than 2-byte ASs. At the same time, in the authors' view, the proposed modifications are simple enough to be easily added to the existing BGP implementations as not to impede BGP5 deployment. As a matter of fact no changes to the BGP4 data structures are required to support IPv4 routing in BGP5.

To ensure a smooth transition from BGP4 to BGP5 on IPv4 networks it is recommended that upgraded BGP implementations continue BGP4 support at least until BGP5 is universally deployed. The BGP5 proposal ensures that IPv4 routing information that is advertised with BGP5 is fully re-advertisable via BGP4 and vice versa. In this sense BGP4 and BGP5 are fully compatible as far as IPv4 routing is concerned.

This paper presents an overview of the protocol messages used by BGP5. Where details of the protocol (e.g., the state machine and fields of protocol messages) have the same meaning as BGP4, those details are left out of this outline for the sake of brevity.

4. Autonomous System Number Space

This document proposes to adopt a 32-bit Autonomous System number space for IPv6 to accommodate for the aggressive Internet growth. The 32-bit number space is large enough to ensure that no future IPv6 transition to a larger number space will be necessary.

In order to minimize implementation changes that are needed to support IPv4 routing, 16-bit AS number space will continue to be supported with BGP5.

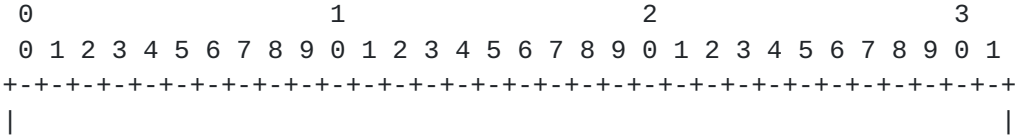
In addition BGP5 will provide a mechanism to gradually introduce the 32-bit AS number space to IPv4 routing.

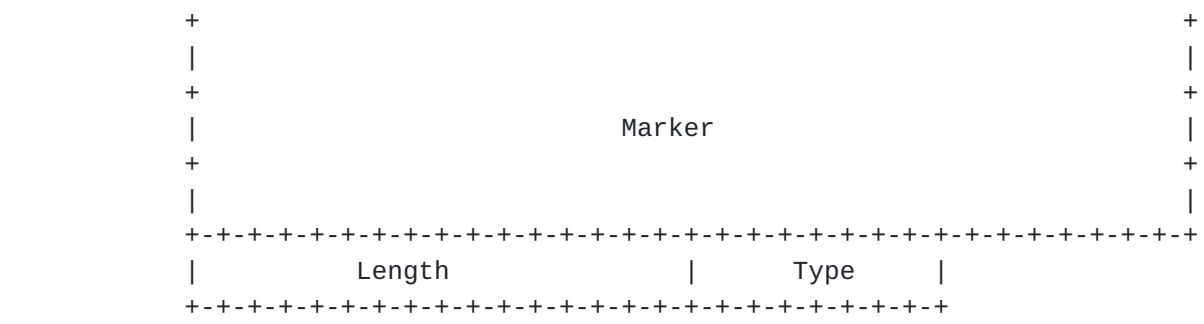
5. Message Formats

This section describes message formats used by BGP.

5.1 Message Header Format

The fixed header for BGP5 is the same as BGP4, namely:

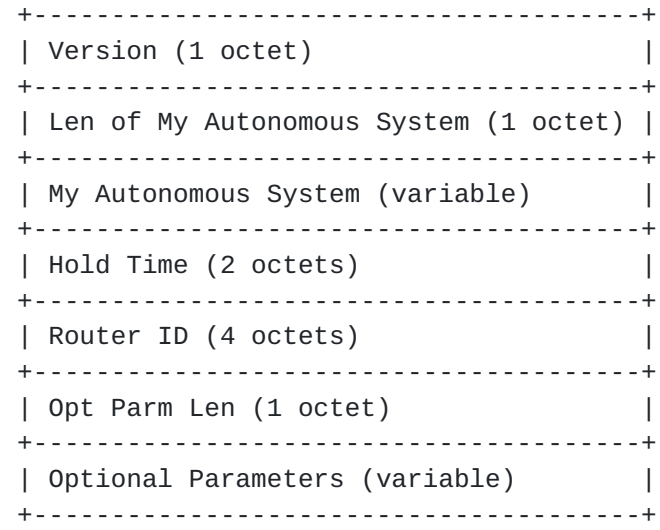




This is as described in BGP4.

5.2 OPEN Message Format

The OPEN message has the following format:



"Version" is a one-octet unsigned integer specifying the protocol version. This paper describes BGP version 5.

"Len of My Autonomous System" is a one-octet unsigned integer specifying the length in octets of the "My Autonomous System" field. To carry IPv4 ASs this field should be 2; to carry IPv6 ASs this field should be 4. The presence of this field permits routing domain identifiers of any length, though this paper only discusses the use of 2-byte ASs and 4-byte [fixed-length] ASs.

"My Autonomous System" identifies the AS of which the sender is a member.

"Hold Time" is as described in BGP4.

"Router ID" is a 32-bit unsigned integer identifying the sender. Because the size of the Router ID is smaller than an IPv6 address, it cannot be set to one of the router's IPv6 addresses (as is commonly done for IPv4). Possible Router ID assignment procedures

for IPv6 include: a) assign the IPv6 Router ID as one of the router's IPv4 addresses or b) assign IPv6 Router IDs through some local administrative procedure (similar to procedures used by manufacturers to assign product serial numbers). The Router ID of 0.0.0.0 is reserved, and should not be used.

"Opt Parm Len" and "Optional Parameters" are as described in BGP4.

5.3 UPDATE Message Format

The UPDATE message always includes the fixed-size BGP header that is followed by one-octet Format Type field which indicates the format and content of the rest of the UPDATE message.

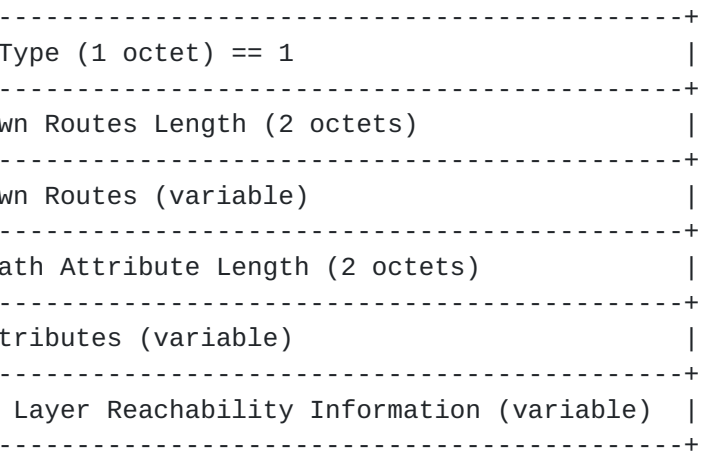
The following UPDATE Format Type values are defined in this document:

1 - BGP4 compatible update with IPv4 routing information
2 - IPv6 routing update
3 - IPv4 routing update with the 4-octet AS numbers

Implementations are not required to support all possible update formats. To allow for future enhancements, implementations are required to ignore UPDATES with format types that they do not support.

5.3.1 BGP4 Compatible Update

This type of UPDATE message has the following format:

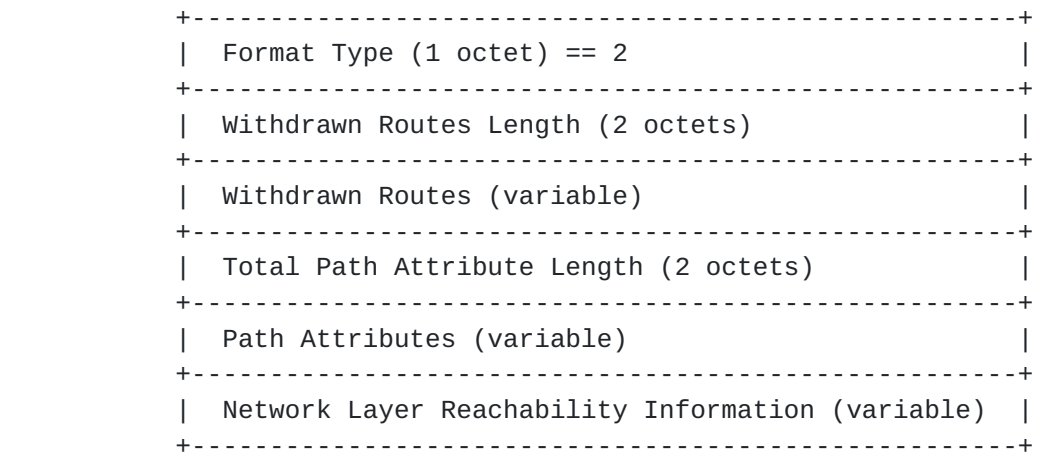


"Format Type" is a value indicating the format of the UPDATE message. It is 1 for the BGP4 compatible format type.

All other fields are as described in BGP4.

5.3.2 IPv6 Routing Update

This type of UPDATE message has the following format:



"Format Type" is a value indicating the format of
the UPDATE message. It is 2 for UPDATE messages that carry
IPv6 routes.

"Withdrawn Routes Length" is a two-octet unsigned integer
specifying the length in octets of the "Withdrawn Routes" field.

"Withdrawn Routes" is a variable-length field containing a list
of address prefixes being withdrawn from service.
Each address prefix is encoded as described in BGP4, except that
the "Length" component of the tuple isn't limited to what would
make sense for IPv4 addresses.

"Total Path Attribute Length" is as described in BGP4.

"Path Attributes" field is as described in BGP4 except:

- The NEXT-HOP attribute is interpreted as a 16-byte
  IPv6 global or site scope address.

- The AGGREGATOR attribute contains the AS number that formed
  the aggregate route (encoded as 4 octets), followed by the
  16-byte IPv6 global or site scope address of the BGP speaker that
  formed the aggregate route.

- The AS-PATH attribute contains four-byte ASs. Two-byte ASs
  can be mapped to four-byte ASs with zero-padding to the left.

- A new LINK-LOCAL-NEXT-HOP attribute.

  LINK-LOCAL-NEXT-HOP is a well-known discretionary attribute
  that contains the link-local address of the router interface
  associated with the global-scope IPv6 address specified in
  the NEXT-HOP attribute of the same UPDATE message.

  LINK-LOCAL-NEXT-HOP shall be included in all IPv6 UPDATE

messages that a given BGP speaker sends to other BGP speakers located in a neighboring autonomous system. Such a link local address should be used as the immediate next hop for forwarding packets to the destinations listed in the UPDATE message.

A BGP speaker shall not include this attribute in IPv6 UPDATE messages that it sends to BGP speakers located in its own autonomous system. If it is contained in an UPDATE message that is received from a BGP speaker which is located in the same AS as the receiving speaker, then this attribute shall be ignored by the receiving speaker. For the routes received from a BGP speaker in the same AS as the receiving speaker, the link-local address of the immediate next hop should be based on the IGP route to the global-scope address provided in the NEXT-HOP attribute.

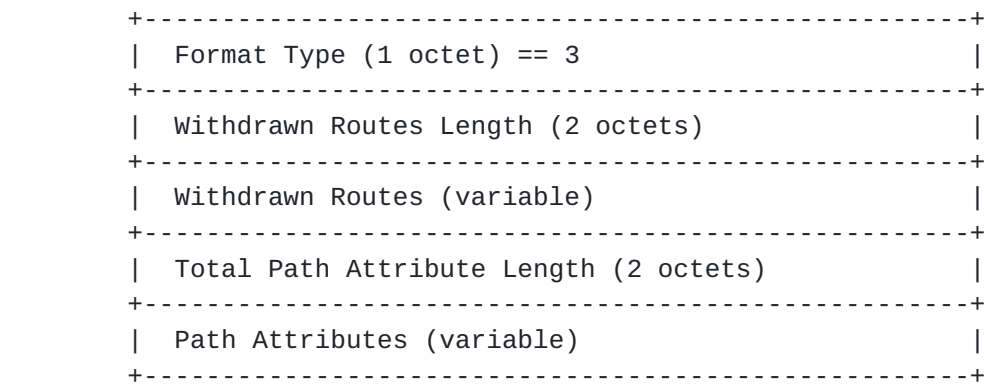
"Network Layer Reachability Information" is as described in BGP4, except that the "Length" component of the tuple isn't limited to what would make sense for IPv4 addresses.

5.3.3 IPv4 Routing Update With the 32-bit AS Numbers

This type of UPDATE message provides support for a larger than 16-bit AS number space that is currently supported in BGP4. Introducing this type of UPDATE well in advance of the 16-bit AS number space exhaustion will allow for smoother IPv4 transition to a larger AS number space.

16-bit AS numbers can be mapped into 32-bit AS numbers by with zero-padding to the left. Therefore it is expected that for an extended period, until BGP5 which supports this type of UPDATE message is universally deployed, bacwards compatibility will be achieved by simply mapping 16-bit AS numbers to zero-extended 32-bit AS numbers and vice versa.

This type of UPDATE message has the following format:



```
   |  Network Layer Reachability Information (variable)  |
   +-----------------------------------------------------+
```

"Format Type" is a value indicating the format of
the UPDATE message. It is 3 for this format type.

"Withdrawn Routes Length" is as described in BGP4.

"Withdrawn Routes" is as described in BGP4

"Total Path Attribute Length" is as described in BGP4.

"Path Attributes" field is as described in BGP4 except:

- The AGGREGATOR attribute contains the 32-bit AS number that formed the aggregate route (encoded as 4 octets), followed by the IPv4 address of the BGP speaker that formed the aggregate route.

- The AS-PATH attribute contains 32-bit ASs.

6. Acknowledgements

[TBA]

7. Security Considerations

[TBA]

8. References

[BGP4-RFC] -- RFC1771

[ASSIGNED-NUMBERS-RFC] -- RFC1700

9. Authors' Addresses

John W. Stewart, III
USC/ISI
4350 North Fairfax Drive
Suite 620
Arlington, VA 22203
+1 703 812 3704
email: jstewart@isi.edu

Dimitry Haskin
Bay Networks
2 Federal St.
Billerica, MA

+1 508 916 8124
email: dhaskin@baynetworks.com