

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 18, 2014

A. Retana
Cisco Systems, Inc.
R. White
J. Heitz
Ericsson
June 16, 2014

**BGP Based Generic TransPort (bgpBGP)
draft-white-bgpbgp-00**

Abstract

A wide array of information is being carried (or proposed to be carried) through BGP peering sessions. While this information is necessary and valid, BGP was not designed to carry blobs of information, but rather network layer reachability and information related directly to forwarding traffic from peer to peer. This document proposes a new BGP message type with a well defined structure to use BGP peering sessions for information passed from provider to provider along edge peering points, the BGP based Generic transPort (bgpBGP). This message type is designed to allow any pair of BGP speakers to transfer information within an existing session, or for BGP peering semantics to be used with multihop sessions between "information exchange speakers" within an autonomous system. The new message type is designed to provide flexibility by allowing the encoding of virtually any information within a BGP session through the use of type-length-vector formatting.

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 December 18, 2014.

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.	Requirements notation	2
2.	Description of the Problem Space	2
3.	The Generic Transport Message	3
4.	Negotiating BGP Based Generic Transport	6
5.	Peer Restart/Database Refresh	6
6.	The Generic Transport Database Descriptor and Request	7
7.	The Certificate Generic Transport Block	8
8.	IANA Considerations	8
9.	Security Considerations	9
10.	References	9
	10.1. Normative References	9
	10.2. Informative References	9
	Authors' Addresses	10

[1.](#) Requirements notation

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.](#) Description of the Problem Space

In many ways, BGP has become the "default pigeon carrier" of the Internet. In order to carry routing information, two autonomous systems must already have functioning BGP peering sessions configured, and BGP itself is easily extensible through address families, extended communities, and other mechanisms. The existence of working connections and easy extensibility, however, creates a situation where BGP is being asked to carry data that reaches far beyond network layer reachability information. For instance, Inter-Domain SLA Exchange [[I-D.ietf-idr-sla-exchange](#)] provides a mechanism

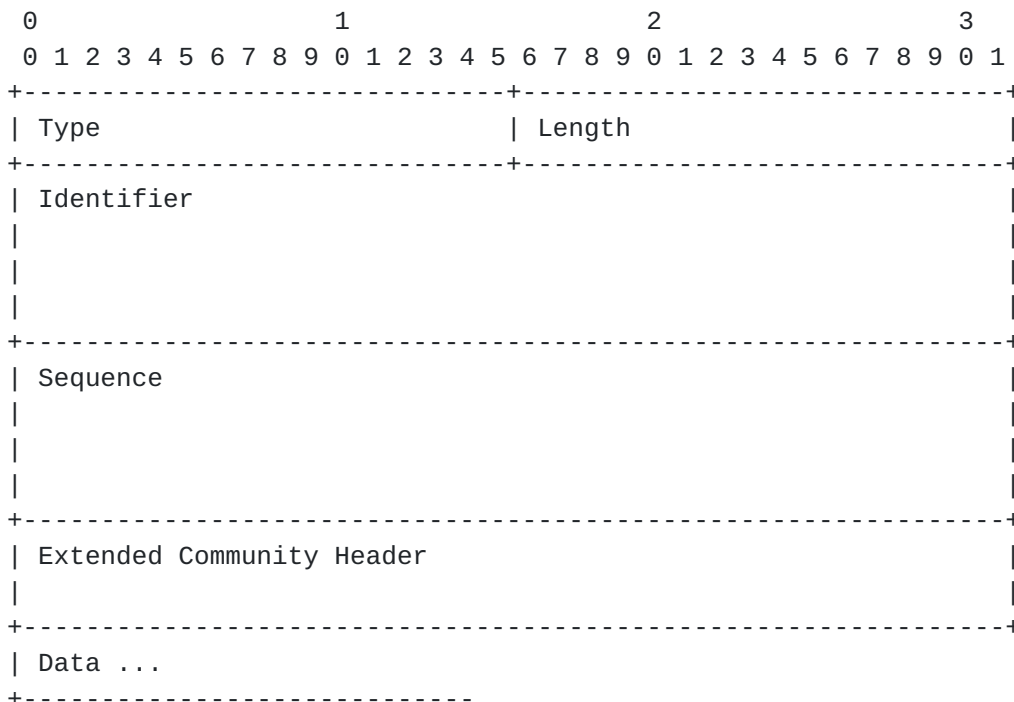
whereby service level information can be exchanged through a BGP peering session between speakers located in different autonomous systems. Another example is North-Bound Distribution of Link-State and TE Information using BGP [[I-D.ietf-idr-ls-distribution](#)], which distributes link state information through a special encoding through BGP.

Both of these use cases, and many others, such as transporting X.509 certificates, could be served better by carrying information outside the standard NLRI formatting of BGP. This draft defines a new message type that would applications to carry information through a BGP session by encoding them in a new message type, rather than embedding them in what appears to be standard BGP routing information. [[RFC4721](#)] This new message type is configured between two peers through a standard capabilities exchange process, and carries type-length-vector encoded data.

Two uses for BGP Generic Transport are described in this document: carrying X.509 certificates and a generic transport database description. These use cases provide justification and the needed framework within which to place this new message type.

3. The Generic Transport Message

The Generic Transport Message (GTM) header is formatted as described in BGP [[RFC4721](#)] [section 4](#), with a type code of [TBD]. Within the GTM message is a series of TLVs, Generic Transport Blocks (GTBs), formatted as:



The fields above are:

- o Type: A two octet unsigned integer describing the GTB type
- o Length: A two octet unsigned integer describing the length of the GTB in octets
- o Identifier: A sixteen octet field which can be used by the application to uniquely identify the information carried in this GTB
- o Sequence: A sixteen octet field which can be used by the application to indicate the relative ordering of information of the same type and identity carried in this GTB
- o Extended Community Header: An eight octet field formatted as described in [[RFC4360](#)]
- o Data: The data, as described in the specific GTB format

GTB types SHOULD be set to a code allocated by IANA for any GTB format which passes through the IETF process (including experimental or individual contribution). Experimental GTB type codes are intended for local experimental use when developing applications using GTB to transfer data through a BGP session, or applications using this technique wholly within a single administrative domain.

The identifier can be used by the application to indicate some unique property of the information carried in the data field. This allows the BGP speaker to determine if this information is unique, or information already contained in a local database.

The sequence field can be used by the application to indicate the "freshness" of the information carried in the data field. The most obvious use of this field would be a monotonically increasing number indicating a newer version of some information advertised previously.

The extended community header is intended to allow for the use of extended communities to control the distribution of the data included in this GTB through route targets or other means. BGP speakers MUST control the flooding, distribution, and processing of the GTB as indicated in the information carried in this field.

Processing received GTBs will proceed as follows:

- o If a route target, community value, or other field in the extended community header indicates the GTB should not be accepted by the local router (through inbound filtering), the GTB is silently discarded
- o If the type and identifier fields do not match any locally stored GTB, the received GTB is silently discarded
- o If the type, identifier, and sequence fields match a GTB stored locally, follow the matching id/sequence number procedure outlined below
- o If the type and identifier fields match a GTB stored locally, and the sequence number is less than the sequence number of the locally stored GTB, the received GTB is silently discarded
- o If the type and identifier fields match a GTB stored locally, and the sequence number is set to 0, follow the withdraw procedure outlined below
- o If the type and identifier fields match, and the sequence number is higher than the sequence number of the locally stored GTB, follow the update procedure outlined below

Matching ID/Sequence procedure:

- o If the remainder of the GTB matches an existing GTB, silently discard the received GTB

- o If the remained of the GTB does not match an existing GTB, log an error and discard the received GTB

Withdraw procedure:

- o The received GTB with a sequence number set to 0 is forwarded to all peers
- o The GTB is removed from local storage

Update procedure:

- o The received GTB is forward to all peers, limited by the route target or other filtering options indicated in the extended community header
- o The received GTB replaces the existing GTB in local storage

4. Negotiating BGP Based Generic Transport

The BGP Based Generic Transport 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 0.

By advertising the BGP Based Generic Transport Capability to a peer, a BGP speaker conveys to the peer that the speaker is capable of receiving and properly handling BGP Based Generic Transport messages.

5. Peer Restart/Database Refresh

In certain situations, one speaker in a peering session may restarted (as outlined in [[RFC4724](#)]), a local filtering change may require reprocessing of the entire GTB database (or some subset, similar to the situation outlined in [[RFC2918](#)]), or the GTB database may need to be refreshed for some other reason. While route refresh and graceful restart assume BGP has some knowledge of the information being exchanged between peers, BGP generic transport assumes BGP is carrying essentially opaque data. Therefore the procedure relies on a database description and database request process, rather than end of table markers, as follows:

Restarting speakers MAY hold a copy of the GTB database in nonvolatile storage to protect information carried through GTB across restarts. If a local copy of the GTBs is available when a resynchronization begins, the originating speaker MUST mark the existing entries as "dirty," so nonexistent entries can be removed at the end of the resynchronization process.

The speaker which wishes to resynchronize (either due to a restart, change in filtering policy, or for any other reason), sends its peer a Database Descriptor (GTBDD) Request with the type field set to 0x01, the sequence number set to 0x0, and the identifier set to the local router id. Future GTBDDs will have a monotonically increasing sequence number.

The speaker receiving this request will format a series of GTBDDs (type 0x02), describing the local database of GTBs, as described below. These descriptor lists carry the type, identifier, and sequence number of each GTB in the sender's database. The sending speaker ends the list with a GTBDD type 0x04, which indicates the entire database has been transmitted.

The resynchronizing speaker compares this list of GTBs to local storage. For each GTB, the resynchronizing speaker determines if it has an up to date local copy of the GTB, or not. For GTBs with a matching local copy, the "dirty" marker set above MUST be cleared.

The resynchronizing speaker sends a GTBDD request (type 0x03) for each GTB which it is missing in its local storage, finishing with a GTBDD end of database marker (type 0x04).

The speaker receiving this request will transmit the requested GTBs, which will be processed according to normal rules (as outlined above).

At the end of processing, all GTBs which were marked as "dirty" MUST be removed from the local database.

6. The Generic Transport Database Descriptor and Request

The generic transport database descriptor can be used to describe the current state of the local GTB database in a BGP speaker. For this GTP, the following fields will be used:

- o For a database descriptor request, the type field will be set to 0x01
- o For a database description listing, the type field will be set to 0x02
- o For a GTB request, the type field will be set to 0x03
- o For an end of database marker, the type field will be set to 0x04

- o The length field will be set to the length of the database descriptor
- o The I bit will be set indicating this is an IANA defined type code
- o The T bit will be clear indicating this is a non-transitive GTB
- o The identifier field will be set to router ID of the BGP speaker sending the GTB
- o The sequence number will be set to a monotonically increasing number set by the originator indicating the block of entry descriptors in the current GTB

The data field will consist of a list of GTB headers. If this GTB is a descriptor, these GTB headers will be taken as a list of GTBs contained in the local database of the transmitter. If this GTB is a request, these GTB headers will be taken as a list of GTBs the speaker would like the receiver to transmit to the originator of the GTB.

7. The Certificate Generic Transport Block

The certificate GTP carries a Route Origin Authentication (ROA) X.509 certificate as described in [[RFC6482](#)]. For this GTP, the following fields will be used:

- o The type field will be set to 0x05
- o The length field will be set to the length of the ROA
- o The I bit within the extended community header will be set indicating this is an IANA defined type code
- o The T bit within the extended community header will be set indicating this is a transitive GTB
- o The identifier field will be set to the prefix contained in the ROA
- o The sequence number will be set to the expiration time of the certificate, encoded as described in [[RFC6482](#)]

8. IANA Considerations

This document introduces the BGP Based Generic Transport Capability. The capability code needs to be assigned by IANA per [[RFC5492](#)].

This document introduces a new BGP message type, BGPBGP. The type code needs to be assigned by IANA.

This document introduces a new GTB type for describing the type and format of information carried in a GTB TLV within a GTM. This is a new 32 bit number space that needs to be registered with the IANA.

9. Security Considerations

None.

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2918] Chen, E., "Route Refresh Capability for BGP-4", [RFC 2918](#), September 2000.
- [RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", [RFC 4360](#), February 2006.
- [RFC4721] Perkins, C., Calhoun, P., and J. Bharatia, "Mobile IPv4 Challenge/Response Extensions (Revised)", [RFC 4721](#), January 2007.
- [RFC4724] Sangli, S., Chen, E., Fernando, R., Scudder, J., and Y. Rekhter, "Graceful Restart Mechanism for BGP", [RFC 4724](#), January 2007.
- [RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", [RFC 5492](#), February 2009.
- [RFC6482] Lepinski, M., Kent, S., and D. Kong, "A Profile for Route Origin Authorizations (ROAs)", [RFC 6482](#), February 2012.

10.2. Informative References

- [I-D.ietf-idr-ls-distribution]
Gredler, H., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and TE Information using BGP", [draft-ietf-idr-ls-distribution-05](#) (work in progress), May 2014.

[I-D.ietf-idr-sla-exchange]

Shah, S., Patel, K., Bajaj, S., Tomotaki, L., and M.
Boucadair, "Inter-domain SLA Exchange", [draft-ietf-idr-sla-exchange-03](#) (work in progress), April 2014.

Authors' Addresses

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

Email: aretana@cisco.com

Russ White
Ericsson

Email: russw@riw.us

Jakob Heitz
Ericsson

Email: jakob.heitz@ericsson.com

