                 Autonomous System Confederations for BGP

Status of this Memo

Abstract

   Border Gateway Protocol [1] is an inter-autonomous system routing
   protocol designed for TCP/IP networks.

   This document describes an extension to BGP which may be used to
   create a confederation of autonomous systems which is represented as
   one single autonomous system to BGP peers external to the
   confederation.

   The intention of this extension is to aid in policy administration
   and reduce the management complexity of maintaining a large
   autonomous system.

   The extension this document describes is widely deployed in the
   Internet today.

Introduction

   It may be useful to subdivide autonomous systems with a very large
   number of BGP speakers into smaller domains for purposes of
   controlling routing policy via information contained in the BGP
   AS_PATH attribute. For example, one may chose to consider all BGP
   speakers in a geographic region as a single entity.

   In addition to improvements in routing policy control, current
   techniques for deploying BGP among speakers in the same autonomous
   system establish a full mesh of TCP connections among all speakers
   for the purpose of exchanging exterior routing information. In
   autonomous systems the number of intra-domain connections that need
   to be maintained by each border router can become significant.

   Subdividing a large autonomous system allows a significant reduction
   in the total number of intra-domain BGP connections, as the

connectivity requirements simplify to the model used for inter-domain
connections.

Unfortunately subdividing an autonomous system may increase the
complexity of policy routing based on AS_PATH information for all
members of the Internet. Additionally, this division increases the
maintenance overhead of coordinating external peering when the
internal topology of this collection of autonomous systems is
modified.

Finally, dividing a large AS may unnecessarily increase the length of
the sequence portions of the AS_PATH attribute. Several common BGP
implementations can use the number of "hops" required to reach a
given destination as part of the path selection criteria. While this
is not an optimal method of determining route preference, given the
lack of other in-band information, it provides a reasonable default
behavior which is widely used across the Internet. Therefore,
division of an autonomous system into separate systems may adversely
affect optimal routing of packets through the Internet.

However, there is usually no need to expose the internal topology of
this divided autonomous system, which means it is possible to regard
a collection of autonomous systems under a common administration as a
single entity or autonomous system when viewed from outside the
confines of the confederation of autonomous systems itself.

Terms and Definitions

    AS Confederation
      A collection of autonomous systems advertised as a single AS
      number to BGP speakers that are not members of the confederation.

    AS Confederation Identifier
      An externally visible autonomous system number that identifies the
      confederation as a whole.

    Member-AS
      An autonomous system that is contained in a given AS
      confederation.

Overview

    IDRP[2] has the concept of a routing domain confederation. An IDRP
    routing domain confederation appears to IDRP speakers external to the
    confederation as a single administrative entity. This extension is
    based upon that work.

In IDRP, routing domain confederations may be nested within each
other or disjoint portions of still larger confederations. The
algorithm BGP defines for additions to the AS_PATH attribute imposes
an additional restriction that AS confederations must be strictly
hierarchical in nature.

AS_CONFED segment type extension

   Currently, BGP specifies that the AS_PATH attribute is a well-known
   mandatory attribute that is composed of a sequence of AS path
   segments. Each AS path segment is represented by a type/length/value
   triple.

   In [1], the path segment type is a 1-octet long field with the two
   following values defined:

              Value   Segment Type

              1   AS_SET: unordered set of ASs a route in the
                              UPDATE message has traversed

              2   AS_SEQUENCE: ordered set of ASs a route in
                              the UPDATE message has traversed

   This document reserves two additional segment types:

              3   AS_CONFED_SET: unordered set of ASs in the local
                              confederation that the UPDATE message
                              has traversed

              4   AS_CONFED_SEQUENCE: ordered set of ASs in the
                              local confederation that the UPDATE
                              message has traversed

Operation

   A member of a BGP confederation will use its confederation identifier
   in all transactions with peers that are not members of its
   confederation. This confederation identifier is considered to be the
   "externally visible" AS number and this number is used in OPEN
   messages and advertised in the AS_PATH attribute.

   A member of a BGP confederation will use its routing domain
   identifier (the internally visible AS number) in all transactions
   with peers that are members of the same confederation as the given
   router.

A BGP speaker receiving an AS_PATH attribute containing a
confederation ID matching its own confederation shall treat the path
in the same fashion as if it had received a path containing its own
AS number.

AS_PATH modification rules

Section 5.1.2 of [1] is replaced with the following text.

When a BGP speaker propagates a route which it has learned from
another BGP speaker's UPDATE message, it shall modify the route's
AS_PATH attribute based on the location of the BGP speaker to which
the route will be sent:

   a) When a given BGP speaker advertises the route to another BGP
   speaker located in its own autonomous system, the advertising
   speaker shall not modify the AS_PATH attribute associated with
   the route.

   b) When a given BGP speaker advertises the route to a BGP
   speaker located in a neighboring autonomous system that is a
   member of the local autonomous system confederation, then the
   advertising speaker shall update the AS_PATH attribute as
   follows:

      1) if the first path segment of the AS_PATH is of type
      AS_CONFED_SEQUENCE, the local system shall prepend its own AS
      number as the last element of the sequence (put it in the
      leftmost position).

      2) if the first path segment of the AS_PATH is not of type
      AS_CONFED_SEQUENCE the local system shall prepend a new path
      segment of type AS_CONFED_SEQUENCE to the AS_PATH, including
      its own confederation identifier in that segment.

   c) When a given BGP speaker advertises the route to a BGP
   speaker located in a neighboring autonomous system that is not a
   member of the current routing domain confederation, then the
   advertising speaker shall update the AS_PATH attribute as
   follows:

      1) if the first path segment of the AS_PATH is of type
      AS_CONFED_SEQUENCE, that segment and any immediately
      following segments of the type AS_CONFED_SET are removed from
      the AS_PATH attribute, leaving the sanitized AS_PATH
      attribute to be operated on by steps 2, or 3.

2) if the first path segment of the remaining AS_PATH is of
type AS_SEQUENCE, the local system shall prepend its own
confederation identifier as the last element of the sequence
(put it in the leftmost position).

3) if there are no path segments following the removal of the
first AS_CONFED_SET/AS_CONFED_SEQUENCE segments, or if the
first path segment of the remaining AS_PATH is of type AS_SET
the local system shall prepend a new path segment of type
AS_SEQUENCE to the AS_PATH, including its own confederation
identifier in that segment.

When a BGP speaker originates a route:

a) the originating speaker shall include an empty AS_PATH
attribute in all UPDATE messages sent to BGP speakers located in
its own autonomous system. (An empty AS_PATH attribute is one
whose length field contains the value zero).

b) the originating speaker shall include its own AS number in an
AS_CONFED_SEQUENCE segment of the AS_PATH attribute of all
UPDATE messages sent to BGP speakers located in neighboring
autonomous systems that are members of the local confederation.
(In this case, the AS number of the originating speaker's member
autonomous system number will be the only entry in the AS_PATH
attribute).

c) the originating speaker shall include its own confederation
identifier in a AS_SEQUENCE segment of the AS_PATH attribute of
all UPDATE messages sent to BGP speakers located in neighboring
autonomous systems that are not members of the local
confederation. (In this case, the confederation identifier of
the originating speaker's member confederation will be the only
entry in the AS_PATH attribute).

Common Administration Issues

It is reasonable for member ASs of a confederation to share a common
administration and IGP information for the entire confederation.

It shall be legal for a BGP speaker to advertise an unchanged
NEXT_HOP and MULTI_EXIT_DISCRIMINATOR attribute to peers in a
neighboring AS within the same confederation. In addition, the
restriction against sending the LOCAL_PREFERENCE attribute to peers
in a neighboring AS within the same confederation is removed. Path
selection criteria for information received from members inside a
confederation may follow the same rules used for information received
from members inside the same autonomous system.

Compatibility

   All BGP speakers participating in a confederation must recognize the
   AS_CONFED_SET and AS_CONFED_SEQUENCE segment type extensions to the
   AS_PATH attribute.

   Any BGP speaker not supporting these extensions will generate a
   notification message specifying an "UPDATE Message Error" and a sub-
   code of "Malformed AS_PATH".

   This compatibility issue implies that all BGP speakers participating
   in a confederation must support BGP confederations, however BGP
   speakers outside the confederation need not support these extensions.

Compatibility Discussion

   We considered the use of a distinct, optional, transitive attribute
   to carry AS confederation information as opposed to specifying new
   types in the existing AS path attribute. This would relax the
   requirement that all BGP speakers participating in a confederation to
   allow the use of legacy units provided they have no external (i.e.
   neither inter-AS nor intra-confederation) connectivity.

   At the time of this writing, an implementation of this extension as
   documented is widely deployed throughout the Internet, therefore the
   value of any change that is incompatible with this document must be
   weighed against the benefit gained from a relaxation of this
   restriction.

References

   [1] Rekhter, Y., and T. Li, "A Border Gateway Protocol 4 (BGP-4)",
       RFC 1771, March 1995.

   [2] Kunzinger, C. Editor, "Inter-Domain Routing Protocol", ISO/IEC
       10747, October 1993.

Security Considerations

   Security issues are not discussed in this memo.

Acknowledgments

   Ravi Chandra and Yakov Rekhter reviewed this document and provided
   constructive and valuable comments.

Author's Address

    Paul Traina
    cisco Systems, Inc.
    170 W. Tasman Dr.
    San Jose, CA 95134

    EMail: pst@cisco.com