                      Prefix Limit Based ORF for BGP-4
                   draft-keyur-idr-bgp-prefix-limit-orf-01.txt

Abstract

   The BGP specification allows for "the ability to impose an (locally
   configured) upper bound on the number of address prefixes the speaker
   is willing to accept from a neighbor". In this specification, we
   define a new Outbound Route Filter type for BGP, termed "Prefix Limit
   Outbound Route Filter", which the speaker can use to communicate that
   upper bound to its peer. The peer is then required to abide by the
   limit. This is expected to have benefits in terms of resource
   consumption and more importantly, transparency of operation.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The Cooperative Outbound Route Filtering Capability defined in
[RFC5291] provides a mechanism for a BGP speaker to send to its BGP
neighbor a set of Outbound Route Filters (ORFs) that can be used by
its neighbor to filter its outbound routing updates to the speaker.

This documents defines a new ORF-type for BGP, termed "Prefix Limit
Outbound Route Filter (PrefixLimit ORF)", that can be used to perform
Prefix Limit based route filtering. This filtering mechanism imposes
a limit on a the number of unique prefixes that the BGP speaker can
advertise to its neighbor.

2. Prefix Limit ORF-Type

The Prefix Limit ORF-Type allows a BGP speaker to inform its neighbor
of its prefix limits. That is, it provides a mechanism through which
a BGP speaker can request its neighbor to limit the number of unique
prefixes that neighbor will advertise to the BGP speaker.

Conceptually a Prefix Limit ORF entry consists of the fields "Action, Match, Reserved, Prefix-Limit".

   Action is a two-bit field. The definition and use of the Action field is described in [RFC5291].

   Match is a one-bit field. The value of this field is 0 for PERMIT and 1 for DENY. In the context of the Prefix Limit ORF-Type, DENY indicates that the BGP speaker sending the ORF will terminate the connection in the event that the Prefix Limit is exceeded.

   Reserved is a 5-bit field. The definition and use of the Reserved field is described in [RFC5291].

   The "Prefix-Limit" is a four byte unsigned integer. It states the maximum number of unique prefixes that the ORF sending BGP speaker is willing to accept from the ORF receiving BGP spea

3. Prefix Limit ORF encoding

   The value of the ORF-Type for the Prefix Limit ORF-Type is [TBD].

   A Prefix Limit ORF entry is encoded as follows. The "Action", "Match", and the "Reserved" field of the entry is encoded in the common part [RFC5291], and the remaining field of the entry is encoded in the "Type specific part" as follows.

```
               +---------------------------------------------------+
               |          Prefix-Limit  (4 octets)                 |
               +---------------------------------------------------+
```

4. Capability Specification for Cooperative route filtering with Prefix Limit

   A BGP speaker signals its compliance with this specification by listing the PrefixLimit ORF type in the Cooperative Route Filtering Capability as defined in [RFC5291].

5. Rules for Prefix Limit ORF

   We describe the rules for PrefixLimit primarily in terms of the rules for the router which sends a PrefixLimit ORF to its peer, which we term the "sending speaker", and for the router which receives a PrefixLimit ORF from its peer, which we term the "receiving speaker". Note that a given router may be either a sending or receiving speaker, or both, with respect to any given peering session.

A router which supports PrefixLimit ORF MUST keep track of the number
of prefixes it has advertised to its peer -- when a new prefix is
advertised, the count is incremented, and when a prefix is withdrawn,
the count is decremented. A modification to the route for an
already-advertised prefix does not change the count. We refer to
this count as the "advertised prefix count" for the session. In
effect, the advertised prefix count is equivalent to the size of the
Adj-RIB-Out for the session.

A router which supports PrefixLimit ORF MAY maintain a received
prefix count for its peer, which tracks the number of prefixes it has
accepted from the peer. In effect, the received route count is
equivalent to the size of the Adj-RIB-In for the session. The use of
such a count is elaborated in the following section.

5.1. Rules for Sending Speaker

If a BGP speaker (the sending speaker) is configured to bound the
number of prefixes it is willing to accept from its neighbor, it MAY
advertise the value of that upper bound to that neighbor using
PrefixLimit ORF. In this section and its subsection, when we refer
to "the PrefixLimit" we are referring to the PrefixLimit value most
recently advertised by the sending speaker to the receiving speaker.

If the sending speaker does not maintain a received prefix count, it
is implicitly relying on its peer to correctly abide by this
specification and no further action is required. If the sending
speaker does maintain a received prefix count, it MAY locally enforce
the PrefixLimit, according to the following rules.

5.1.1. Enforcing the Prefix Limit

When the sending speaker sends a PrefixLimit ORF which is less than
its current received prefix count, it SHOULD wait for some interval
before enforcing the new PrefixLimit. The interval to be used is a
matter of local policy. Also, even if the PrefixLimit ORF is greater
than or equal to the current received prefix count, the router may
wish to wait for some interval before enforcing the new limit in
order to allow for UPDATEs which may have been in flight prior to the
receipt of the PrefixLimit ORF by the peer. Subsequent to any such
waiting period, the remaining rules in this section SHALL apply.

If the PrefixLimit is exceeded (either because of a route announced
by the peer or because the peer failed to timely withdraw routes
after the PrefixLimit is revised downward), the peer is in violation,
and the sending speaker MAY take corrective action. The router MAY
also allow the received prefix count to exceed the PrefixLimit by
some amount as a matter of local policy.

Corrective actions MAY include dropping the BGP session or refusing
to accept new prefixes in excess of the PrefixLimit.

If the former option -- dropping the BGP session -- is chosen, the
router MUST indicate this in advance by advertising its PrefixLimit
ORF with the Match flag set to DENY. Also, by default it SHOULD NOT
automatically reestablish the session.

If the latter option -- refusing to accept new prefixes -- is chosen,
the router MUST accept modifications to already-accepted prefixes,
and it MUST accept withdrawals of already-accepted prefixes. If
prefixes are withdrawn, the received prefix count will drop below the
announced PrefixLimit and new prefixes SHOULD be accepted, again up
to but not exceeding the limit. Prefixes which are refused SHOULD
NOT contribute to the received prefix count.

We note that the option of refusing to accept new prefixes will
likely lead to desynchronization of the BGP session and is a flawed
solution at best; operator intervention will be required in order to
restore synchronization (for example, through correction of routing
policies and a subsequent route-refresh).

5.2. Rules for Receiving Speaker

When a PrefixLimit ORF is received, the new Prefix Limit value in the
ORF is considered to be the new maximum Prefix Limit for the
neighbor. In this section, when we refer to "the PrefixLimit" we are
referring to the PrefixLimit value most recently received from the
sending speaker by the receiving speaker.

The receiving speaker MUST NOT advertise a prefix to its peer if
doing so would cause its advertised prefix count to exceed the
PrefixLimit.

The receiving speaker MAY take local action when its advertised
prefix count approaches the PrefixLimit. The nature of the action
(logging, etc) is a matter of local policy, as is the threshold at
which the action occurs.

When the receiving speaker receives a PrefixLimit ORF with When-to-
Refresh set to DEFER, it need not take any additional action unless
its current advertised prefix count exceeds the new PrefixLimit. In
that case, it MUST take immediate steps to correct the violation.

Such steps MAY include withdrawing already-advertised prefixes so as
to reduce the advertised prefix count to be less than or equal to the
PrefixLimit. The selection of which prefixes to withdraw is a matter
of local policy. Another option to correct the violation would be to

drop the session; in this case the session SHOULD NOT be
automatically reestablished.

When the receiving speaker receives a PrefixLimit ORF with When-to-
Refresh set to IMMEDIATE, it behaves as given for DEFER but in
addition advertises its Adj-RIB-Out as specified in [RFC5291].

6. Error handling

ORFs provide information that guides future sending, but any
malformed ORF is simply missed filtering information. If Prefix
Limit ORF is malformed, then the Refresh messages shall simply be
discarded.

7. Security Considerations

This extension to BGP does not change the underlying security issues.
However, it does suggest a mechanism by which certain denial of
service risks may be reduced.

8. Acknowledgements

The authors would like to thank ... for their valuable comments.

9. IANA Considerations

This specification requests a new Cooperative Route Filter [RFC5291]
type code.

10. Normative References

[RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
            Requirement Levels", BCP 14, RFC 2119,
            DOI 10.17487/RFC2119, March 1997,
            <http://www.rfc-editor.org/info/rfc2119>.

[RFC3392]   Chandra, R. and J. Scudder, "Capabilities Advertisement
            with BGP-4", RFC 3392, DOI 10.17487/RFC3392, November
            2002, <http://www.rfc-editor.org/info/rfc3392>.

[RFC4271]   Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A
            Border Gateway Protocol 4 (BGP-4)", RFC 4271,
            DOI 10.17487/RFC4271, January 2006,
            <http://www.rfc-editor.org/info/rfc4271>.

[RFC5291]   Chen, E. and Y. Rekhter, "Outbound Route Filtering
            Capability for BGP-4", RFC 5291, DOI 10.17487/RFC5291,
            August 2008, <http://www.rfc-editor.org/info/rfc5291>.

   [RFC5292]   Chen, E. and S. Sangli, "Address-Prefix-Based Outbound
               Route Filter for BGP-4", RFC 5292, DOI 10.17487/RFC5292,
               August 2008, <http://www.rfc-editor.org/info/rfc5292>.

Authors' Addresses

   Keyur Patel
   Arrcus, Inc

   Email: keyur@arrcus.com

   John Scudder
   Juniper Networks

   Email: jgs@juniper.net

   Marco Marzetti
   PCCW Global
   450 Spring Park Place, Suite 100
   Herndon , VA 20170

   Email: mmarzetti@pccwglobal.com

   Jakob Heitz
   Cisco
   170 W. Tasman Drive
   San Jose, CA 95124

   Email: jheitz@cisco.com