

INTERNET-DRAFT
[draft-keyur-prefixlimit-orf-00.txt](#)

Keyur Patel
Chandra Appanna
John Scudder
April 2004

Expires: October 2004

Prefix Limit Based Outbound Route Filter for BGP-4
<[draft-keyur-prefixlimit-orf-00.txt](#)>

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [Section 10 of RFC2026](#).

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."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>

This document is an individual submission. Comments are solicited and should be addressed to the author(s).

Copyright Notice

Copyright (C) The Internet Society (2004). All Rights Reserved.

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.

1. Introduction

The Cooperative Outbound Route Filtering Capability defined in [BGP-ORF] 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.

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 [BCP 14](#), [RFC 2119](#) [[RFC2119](#)].

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 [[BGP-ORF](#)].

INTERNET-DRAFT

Expires: October 2004

April 2004

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 [[BGP-ORF](#)].

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 speaker.

[2.1](#). 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 [[BGP-ORF](#)], and the remaining field of the entry is encoded in the "Type specific part" as follows.

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

[3](#). Capability

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

[4](#). Rules for PrefixLimit 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.

[4.1.](#) 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.

[4.1.1.](#) Enforcing The PrefixLimit

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).

[4.2.](#) 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 [[BGP-ORF](#)].

[5.](#) Acknowledgments

Shyam Suri, David Ward, David Meyer and Robert Raszuk provided valuable comments.

This draft was inspired in part by an earlier draft by Srikanth

[6.](#) Security Considerations

This specification does not change BGP's principal underlying security considerations. However, it does suggest a mechanism by which certain denial of service risks may be reduced.

[7.](#) IANA Considerations

This specification requests a new Cooperative Route Filter [[BGP-ORF](#)] type code.

[8.](#) References

[8.1.](#) Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels" [BCP 14](#), [RFC 2119](#), March 1997.
- [BGP-4] Rekhter, Y., and T. Li, "A Border Gateway Protocol 4 (BGP-4)", [RFC 1771](#), March 1995.
- [BGP-DRAFT] Rekhter, Y., T. Li and S. Hares, "'A Border Gateway Protocol 4 (BGP-4)", work in progress, [draft-ietf-idr-bgp4-23.txt](#), November 2003.
- [BGP-ORF] Chen, E., and Rekhter, Y., "Cooperative Route Filtering Capability for BGP-4", work in progress, [draft-ietf-idr-route-filter-10.txt](#), March 2004.

9. Authors' Addresses

Keyur Patel
Cisco Systems
email: keyupate@cisco.com

Chandra Appanna
Cisco Systems
email: achandra@cisco.com

John Scudder
Cisco Systems
email: jgs@cisco.com

10. Full Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#) and except as set forth therein, the authors retain all their rights.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

INTERNET-DRAFT

Expires: October 2004

April 2004

[11](#). Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

[12](#). Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.

