

idr  
Internet-Draft  
Expires: February 22, 2009

B. Dickson  
Afilias Canada, Inc  
August 21, 2008

A New BGP Standards Action Community, LAST\_RESORT  
draft-dickson-idr-last-resort-05

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

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/ietf/lid-abstracts.txt>.

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

This Internet-Draft will expire on February 22, 2009.

Copyright Notice

Copyright (C) The IETF Trust (2008).

Abstract

This Internet Draft describes a new Standards Action BGP community, LAST\_RESORT.

This community provides a simple and easily deployable solution to a certain class of BGP "wedgies".

Initial deployment is expected to be achieved by voluntary use in the network operator community-at-large.

Internet-Draft

BGP Community: LAST\_RESORT

August 2008

Long-term adoption via software enforcement of the community, will improve global behavior, and simplify router configurations.

The Standards Action range of communities (previously limited to the "well-known" communities) ensures that the expectation of (eventual) router support is reasonable.

#### Author's Note

This Internet Draft is intended to result in this draft or a related draft(s) being placed on the Standards Track for idr.

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 [\[5\]](#).

Intended Status: Proposed Standard.

## [1.](#) Background

Even when all the best current practises are observed, operational problems may be experienced when running a BGP network.

One particularly thorny problem is BGP "wedgies" [\[1\]](#).

While not often articulated, the common problem is the use of local policy settings within (and in particular, at) AS boundaries, which often override the original intent of "backup" BGP announcements.

It is somewhat ironic that such local policies are often achieved by use of BGP Communities, and that the lack of a common community, is one source of the problem.

### [1.1.](#) The Local Policy Problem

The BGP "wedgie" problem occurs when a number of mechanisms in path selection interact across multiple routers, where the resulting state is stable, but is not deterministic, but rather is the result of a race.

Generally speaking, the "wedgie" problem is induced by the application of non-default values to prefixes, by way of local policy

setting, such as modification of AS-path, setting of Multi-Exit Discriminators (MEDs), local preference setting, or similar mechanisms.

One particular class of "wedgies" is the result of the lack of an

explicit global mechanism for expressing de-preferring announcements via "back-up" providers. In essence, the applied local policies that interfere with such "back-up" announcements could be described as "not well informed".

## 2. Proposed Change: A New BGP Standards Action Community

To solve this particular class of problem, a new BGP Standards Action Community, LAST\_RESORT, is proposed. This is a standard, not an extended, community. This is a new value to be assigned by IANA. In particular, if this instant Draft is adopted as a WG Draft, an [RFC 4020](#) Early Assignment is requested.

## 3. Changes to BGP Behavior

A BGP speaker receiving a prefix from an EBGp neighbor with LAST\_RESORT MUST assign the lowest-possible preference value of LOCAL\_PREFERENCE. This LOCAL\_PREFERENCE MUST be assigned AFTER the application of local policy mechanisms, and MUST NOT be able to be over-ridden without first removing the LAST\_RESORT community from the prefix. A BGP speaker receiving a prefix from an IBGP neighbor with LAST\_RESORT MUST ignore the LAST\_RESORT community.

The distinction between EBGp and IBGP ensures backwards compatibility, particularly when a heterogeneous set of routers in an AS doing IBGP exists.

The restriction preventing overriding LAST\_RESORT ensures that local policy mechanisms do not interfere with LAST\_RESORT. This also facilitates ease in deployment, as most router configurations would not require modification.

Note Well, that this order of evaluation, with local policy before LAST\_RESORT, means that the only way to apply local policy to routes

with LAST\_RESORT, is to strip LAST\_RESORT on ingress, then apply local policy, then set LAST\_RESORT again on egress.

This restriction is by design, so that AS-wide policy remains consistent even in a heterogenous environment of routers that may or may not understand LAST\_RESORT.

#### [4.](#) LAST\_RESORT - Basic Method

The main reason for establishing the LAST\_RESORT Community is to permit the global implementation of "backup only" announcements,

Dickson

Expires February 22, 2009

[Page 3]

---

Internet-Draft

BGP Community: LAST\_RESORT

August 2008

whose purpose and intent are clear and unambiguous. It is not to facilitate change of policies, or to circumvent local policies. Rather, it is to make possible the implementation of policies where those have been negotiated by two or more parties.

Currently, there are several documented scenarios in the "Wedgies" RFC [\[1\]](#) where the mutually desired policy is either unable to be implemented, or does not deterministically reach the desired state.

Application of the LAST\_RESORT Community on announcements sent to a backup provider, permits these problem classes to be resolved.

The same prefix is announced to both the primary and backup provider. When announced to the primary provider, the LAST\_RESORT Community is NOT set. When announced to the backup provider, the LAST\_RESORT Community IS set.

The propagation of the LAST\_RESORT instance will be limited by the availability of paths, and inhibited by the existence of paths which do not have LAST\_RESORT applied to them.

In Figure 1 (of [Appendix A](#)), the LAST\_RESORT instance will be seen by the backup provider, and be passed with LAST\_RESORT to the backup provider's transit provider. The latter will prefer any other instance without LAST\_RESORT, even if it has policy for applying a LOCAL\_PREFERENCE to the received prefix instances. Should the other instance be withdrawn, the LAST\_RESORT will be selected and subsequently propagated.

## [5.](#) Security Considerations

No additional security considerations beyond those already present in BGP are introduced.

## [6.](#) IANA Considerations

IANA will need to assign a new code point from BGP Standards Action Communities for LAST\_RESORT.

## [7.](#) Acknowledgements

The author wishes to acknowledge the helpful guidance of Joe Abley and Tony Li. The author also wishes to acknowledge the assistance and suggestions of Joel M. Halpern in simplifying the original "Backup-only" concept to that of a BGP community, and of Olivier

Dickson

Expires February 22, 2009

[Page 4]

---

Internet-Draft

BGP Community: LAST\_RESORT

August 2008

Bonaventure in clarifying the LOCAL\_PREFERENCE mechanisms.

## [8.](#) References

### [8.1.](#) Normative References

- [1] Griffin, T. and G. Huston, "BGP Wedgies", [RFC 4264](#), November 2005.
- [2] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), January 2006.
- [3] Chandrasekeran, R., Traina, P., and T. Li, "BGP Communities Attribute", [RFC 1997](#), August 1996.
- [4] Kompella, K. and A. Zinin, "Early IANA Allocation of Standards Track Code Points", [BCP 100](#), [RFC 4020](#), February 2005.

### [8.2.](#) Informative References

- [5] Bradner, S., "Key words for use in RFCs to Indicate Requirement

Levels", [BCP 14](#), [RFC 2119](#), March 1997.

## [Appendix A](#). BGP Wedgie Examples

The following examples from [RFC 4264](#) [1] show the effects of the proposed changes, in resolving "wedgie" issues.

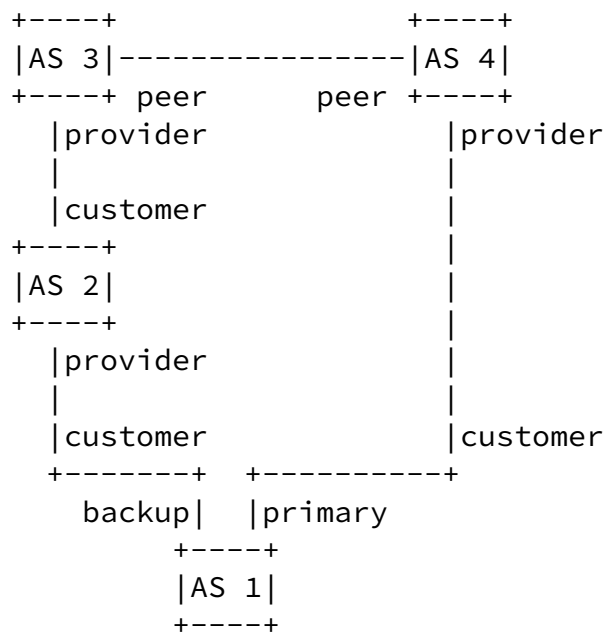


Figure 1

In Figure 1 above, the announcement via the backup link is sent with LAST\_RESORT.

- o AS 4 sends AS\_PATH "4 1" to AS 3.
- o AS 2 receives the LAST\_RESORT path from AS 1, and sends AS\_PATH "2 1" to AS 3, also with LAST\_RESORT.
- o AS 3 and AS 4 exchange their respective "best" paths.
- o AS 3 prefers the path "4 1" over "2 1" because "2 1" is LAST\_RESORT.
- o AS 3 sends a withdrawal of the LAST\_RESORT path to AS 4.
- o AS 3 sends its "best", AS\_PATH "3 4 1" to AS 2.

This state will be reached regardless of sequence of disconnects and reconnects.

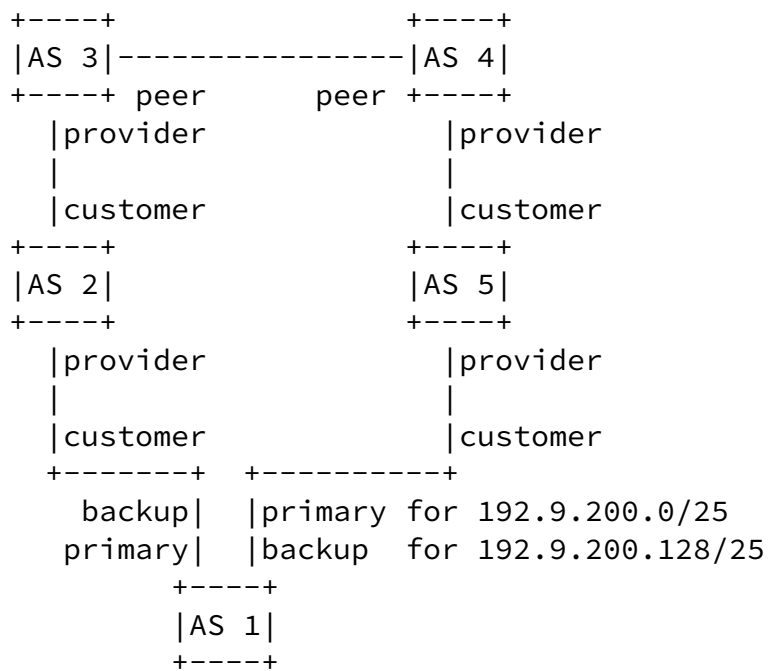
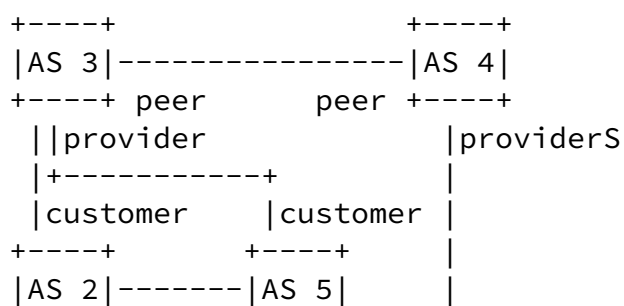


Figure 2

In Figure 2 above, the announcements via the backup links will work the same as in Example 1.



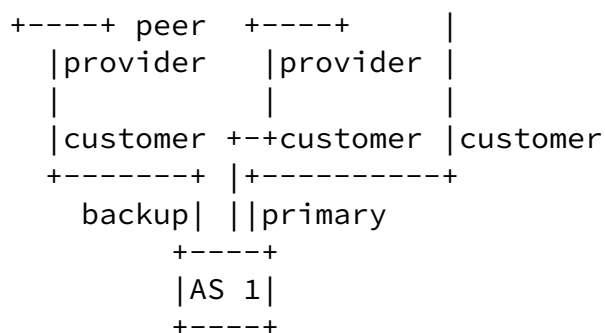


Figure 3

In Figure 3 above, the announcements via both backup links will result in:

- o AS 2 selecting its best path via "3 4 1" (the only path it hears from AS 3)
- o AS 2 hearing one of two paths from AS 5:
  - \* "5 3 4 1"
  - \* LAST\_RESORT with path "5 1"
- o AS 2 hearing a LAST\_RESORT directly from AS 1

Any announcement that AS 3 hears from AS 2 will always be marked LAST\_RESORT. (The same will be true of AS 5.) Thus, any combination of break/restore on any links in any order, will always result in the desired state being reached.

#### Author's Address

Brian Dickson  
 Afiliat Canada, Inc  
 4141 Yonge St,  
 Suite 204  
 North York, ON M2P 2A8  
 Canada

Email: [brian.peter.dickson@gmail.com](mailto:brian.peter.dickson@gmail.com)  
 URI: [www.afiliat.info](http://www.afiliat.info)



Copyright (C) The IETF Trust (2008).

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 the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM 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.

## 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](mailto:ietf-ipr@ietf.org).

## Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).