Network Working Group	P. Mohapatra
Internet-Draft	Cisco Systems
Intended status: Standards Track	R. Fernando
Expires: October 24, 2009	Juniper Networks
	April 22, 2009

TOC

BGP Link Bandwidth Extended Community
draft-ietf-idr-link-bandwidth-00.txt

Status of this Memo

Copyright Notice

Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

This document describes an application of BGP extended communities that allows a router to perform unequal cost load balancing.

Table of Contents

---

1. Introduction

TOC

When a BGP speaker receives multiple paths from its internal peers, it could select more than one path to send traffic to. In doing so, it might be useful to provide the speaker with information that would help it distribute the traffic unequally based on the cost of the external (DMZ) link. This document suggests that the external link bandwidth be carried in the network using a new extended community [RFC4360] (Sangli, S., Tappan, D., and Y. Rekhter, “BGP Extended Communities Attribute,” February 2006.) - the link bandwidth extended community.

1.1. Requirements Language

TOC

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 RFC 2119 (Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” March 1997.) [RFC2119].

TOC

2. Link Bandwidth Extended Community

When a BGP speaker receives a route from a directly connected external neighbor (the external neighbor that is one IP hop away) and advertises this route (via IBGP) to internal neighbors, as part of this advertisement the router may carry the bandwidth of the link that connects the router with the external neighbor. The bandwidth of such a link is carried in the Link Bandwidth Community. The community is optional non-transitive. A border router MUST strip the link bandwidth community from a route when it advertises the route to an external neighbor.

It is noteworthy that the bandwidth carried in the Link Bandwidth extended community is the configured bandwidth of the EBGP link. It does not depend on the amount of traffic transiting that link.

The value of the high-order octet of the extended Type Field is 0x40. The value of the low-order octet of the extended type field for this community is 0x04.

The value of the Global Administrator subfield in the Value Field SHOULD represent the Autonomous System of the router that attaches the Link Bandwidth Community. If four octet AS numbering scheme is used [RFC4893] (Vohra, Q. and E. Chen, “BGP Support for Four-octet AS Number Space,” May 2007.), AS_TRANS should be used in the Global Administrator subfield.

The bandwidth of the link is expressed as 4 octets in IEEE floating point format, units being bytes per second. It is carried in the Local Administrator subfield of the Value Field.

---

3. Deployment Considerations

TOC

This document proposes to use the Link Bandwidth extended community for the purpose of load balancing in the following two scenarios. The first scenario is when the candidate paths are identical until and including the IGP distance step in the BGP decision process. The second scenario is when the traffic goes via a tunneled network, in which case the candidate paths are identical for all steps before the IGP distance step in the BGP decision process. Use of this community for other scenarios is outside the scope of this document.

If there are multiple paths to reach a destination and if only some of them have link bandwidth community, the receiver should not perform unequal cost load balancing based on link bandwidths.

TOC

4. Acknowledgments

The authors would like to thank Yakov Rekhter, Srihari Sangli and Dan Tappan for proposing unequal cost load balancing as one possible application of the extended community attribute.

---

5. IANA Considerations

TOC

This document defines a specific application of the two-octet AS specific extended community. IANA is requested to assign a sub- type value of 0x04 for the link bandwidth extended community.

Name	Value
non-transitive Link Bandwidth Ext. Community	0x4004

---

6. Security Considerations

TOC

There are no additional security risks introduced by this design.

---

7. Normative References

TOC

[RFC2119]	Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC4360]	Sangli, S., Tappan, D., and Y. Rekhter, “BGP Extended Communities Attribute,” RFC 4360, February 2006 (TXT).
[RFC4893]	Vohra, Q. and E. Chen, “BGP Support for Four-octet AS Number Space,” RFC 4893, May 2007 (TXT).

Authors' Addresses

TOC

	Pradosh Mohapatra
	Cisco Systems
	170 W. Tasman Drive
	San Jose, CA 95134
	USA

Phone:	
Email:	pmohapat@cisco.com
	Rex Fernando
	Juniper Networks
	1194 N. Mathilda Ave
	Sunnyvale, CA 94089
	USA
Phone:	
Email:	rex@juniper.net