BGP Link Bandwidth Extended Community

draft-rfernando-idr-link-bandwidth-00

Pradosh Mohapatra <pmohapat@cisco.com>,
Rex Fernando <rex@juniper.net>

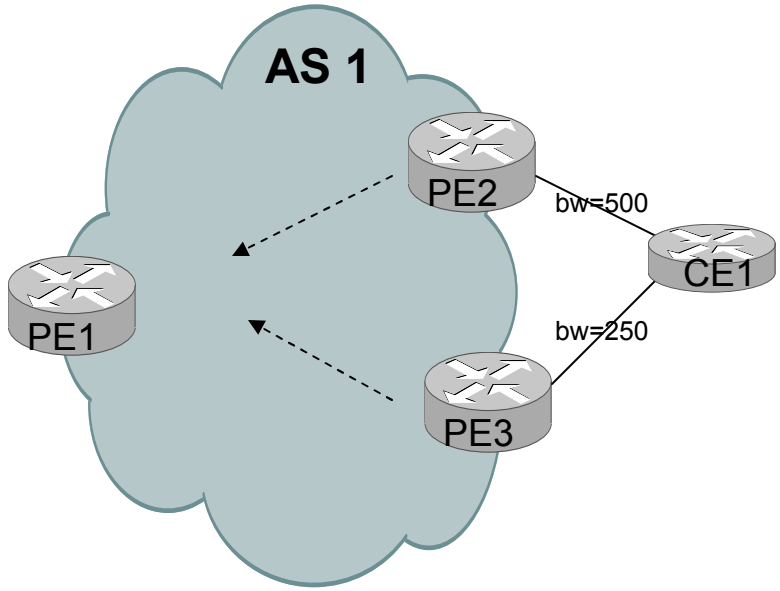
Acknowledgements: Yakov Rekhter, Srihari Sangli, Dan Tappan

Basic Idea

- Define an extended community to carry the DMZ link's bandwidth (or an associated weight) to enable ingress routers to do unequal cost load balancing
- Already implemented by multiple vendors; not standardized (hence potential for interoperability issues)

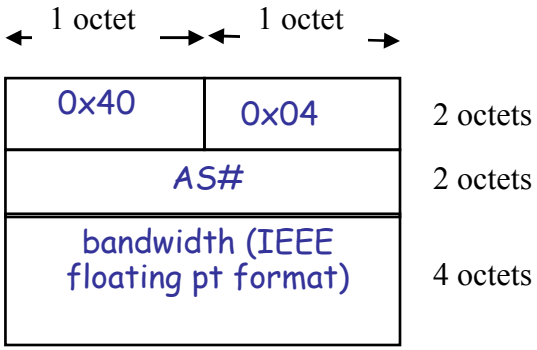
Motivation

- CE1 is one-hop EBGP neighbor dually connected to PE2 and PE3 (links with different bandwidths)
- When PE2 & PE3 send received prefixes from CE1 to IBGP mesh, they include the bandwidth in an extended community
- PE1 does unequal cost load balancing based on the bandwidth values (e.g. 2:1 in this case)



Format & Usage

- Optional non-transitive
- Type high=0x40, low=0x04
- AS# in next two octets (AS_TRANS in case of 4-byte ASes)
- Bandwidth in bytes/sec in next four octets (IEEE floating point format)
- Restricted to cases where IBGP multipath can safely be used
  - IGP cost to nexthop should be same or use of tunneling for unequal cost



Document Status

- Comments are welcome
- We would like the draft to be WG draft