BGP Convergence

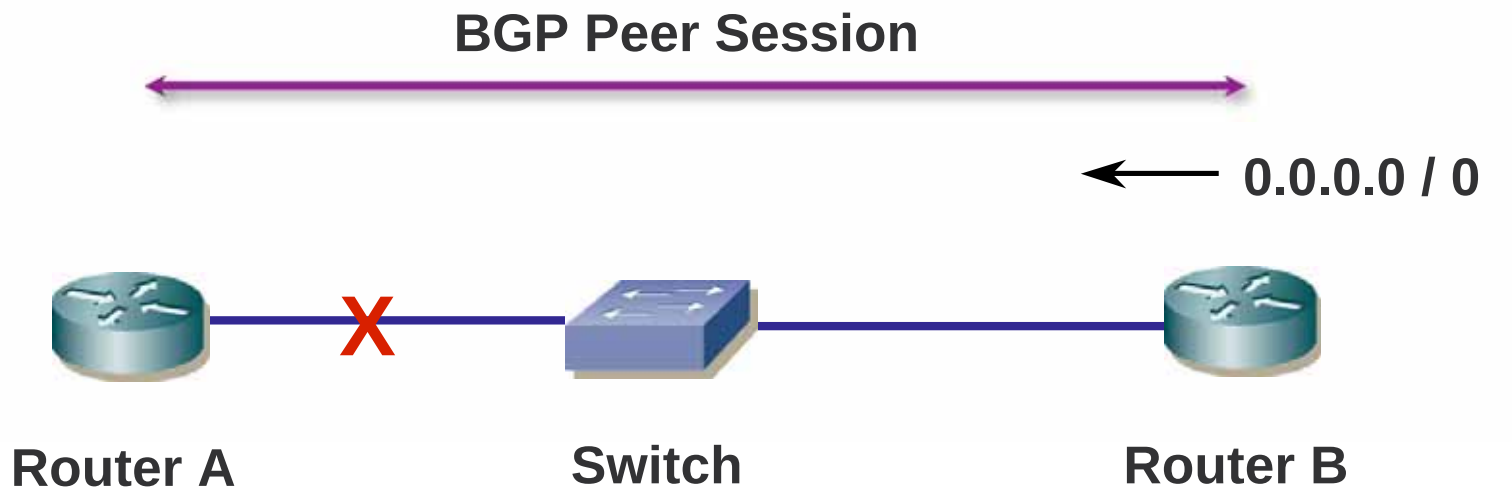
Better handling of silent peer failures

David J. Hughes

Overview

- Outline the impact of silent peer failure
- Review emerging technologies that solve the problem
- Outline things that can help today
- Present operational experience

Silent Peer Failures



Default route will blackhole until BGP timers expire

BGP Timers

There are 2 timers of interest

- Hold Timer
  - Max number of seconds between an UPDATE or KEEPALIVE message
  - Negotiated between peers at session establishment
- Keep-alive Timer
  - Number of seconds between transmission of keepalive packets
  - Usually implemented as HOLD TIME / 3

Common Implementation

Cisco

- **Keepalive Timer 60 Seconds**
- **Hold Timer 180 Seconds**
- **Traffic will be blackholed for up to 179 Seconds !!!**

Juniper

- **Keepalive Timer 30 Seconds**
- **Hold Timer 90 Seconds**
- **Maximum blackhole duration still 89 Seconds**

99.999% reliability is less than 316 seconds of downtime per year !

You cannot offer five nine's SLA's if you use Cisco's default timers.

BGP Next Hop Tracking

Solution A - Next Hop Tracking for BGP

- Event driven removal of prefixes from RIB
- Triggers when Next Hop address dropped from IGP
- No need to wait for Hold timer etc
- Great solution for iBGP peers

BGP Next Hop Tracking

But ……

- **No good for eBGP as next hop isn't in your IGP**
- **Doesn't appear to be supported on Juniper**
- **Very limited support on Cisco**
  - **12.0 S**
  - **12.3 T**
  - **12.4**
  - **Nothing for Cat6500 or 7304**

Bidirectional Forwarding Detection (BFD)

Solution B - Bidirectional Forwarding Detection

- BFD is basically a Hello protocol
- Checks connectivity between the forwarding-planes of peer devices
- Works over direct connections, VC's, Tunnels, MPLS
- Intervals specified in microseconds
- Can provide very fast detection of path failure
- IETF draft supported by Cisco and Juniper
- Great solution for eBGP peers

Bidirectional Forwarding Detection (BFD)

But ….

- **Your upstream must support it**
- **Interaction with routing protocols is specific to the protocol**
  - IETF drafts exist for BGP, OSPF, IS-IS
- **Very limited support for BFD for BGP**
  - Juniper appears to support BFD for IGP's only at present
  - Cisco IOS on 6500's supports BFD for IGP's only
  - Cisco IOS supports BFD for IGP's and BGP in 12.4T

A Solution for Today !

- **BGP peers negotiate the lowest Hold timer at session establishment**
- **IOS supports configurable BGP timers**
- **Lowering the Hold timer can help dramatically**
- **Configure non-default timers in BGP config stanza**
  - `neighbor a.b.c.d timers Keepalive Hold`

  `eg`
  - `neighbor 10.0.0.1 timers 30 90`

Potential Problems

- **Peer device may reject BGP session if Hold timer too low**
  - IOS introduced support for minimum Hold time in a recent 12.0S release. No sign in other releases or JunOS yet.
- **Aggressive timers may fail when router is busy (BGP scanner activity etc)**
- **Increased keepalive activity may impact performance of aggregation routers**

Operational Experience

- Non-default timers implemented with all upstreams
- Upstreams include all Tier-1 providers in Australia
- Currently configured timers
  - 5 second Keepalive timer
  - 15 second Hold timer
- Worst case scenario is 14 seconds.
- Plans to trial 1 second Keepalive for iBGP sessions until Next-Hop tracking available

Summary

- Light at the end of the tunnel
  - BFD for eBGP
  - Next-Hop tracking for iBGP
- Reducing BGP timers is an interim solution
- Cisco default values should be avoided
- 3 * 5 second Keep-alives working well in production