BGP policy interactions can be difficult to debug

(BGP Wedgies, RFC 4264)

Complexity can arise in unexpected ways

Timothy G. Griffin
University of Cambridge
tgg22@cam.ac.uk

IETF 83 --- Paris
Network Complexity Research Group (ncrg)
March 26 2012
Multiple Stable States

R1 prefers this path to d
R2 prefers this path to d
¾ Wedgie Example

- AS 1 implements backup link by sending AS 2 a “depref me” community.
- AS 2 implements this community so that the resulting local pref is below that of routes from its upstream provider (AS 3 routes)
Getting wedged...

Happy, happy, joy, joy

Backups are good!

OH NO, I’M WEDGED!

Primary fails

Primary comes back up!
And the Routings are...

**Intended Routing**

Note: this would be the ONLY routing if AS2 translated its “depref me” community to a “depref me” community of AS 3

**Unintended Routing**

Note: This is easy to reach from the intended routing just by “bouncing” the BGP session on the primary link.
Recovery

- Requires manual intervention
- Can be done in AS 1 or AS 2
What is a BGP Wedgie?

- BGP policies make sense locally
- Interaction of local policies allows multiple stable routings
- Some routings are consistent with intended policies, and some are not
  - If an unintended routing is installed (BGP is “wedged”), then manual intervention is needed to change to an intended routing
- When an unintended routing is installed, no single group of network operators has enough knowledge to debug the problem
Load Balancing Example

AS 3 peer peer AS 4
provider provider

customer customer

AS 2

primary link for prefix P2
backup link for prefix P1

AS 1

primary link for prefix P1
backup link for prefix P2

AS 5

Simple session reset my not work!!
Can’t un-wedge with session resets!

Note that when bringing all up we could actually land the system in any one of the 4 stable states --- depends on message order....
Who among us could figure this one out?  
When 1—2 is in New York and 1—5 is in Tokyo?
Full Wedgie Example

AS 1

AS 2

AS 3

peer

peer

peer

AS 4

provider

customer

customer

AS 5

provider

customer

customer

backup links

primary link
And the Routings are...

Intended Routing

Unintended Routing 1
And the Routings are...

Intended Routing

Unintended Routing 2
Resetting 1—2 may not help!!
A lot of **non-local knowledge** is required to arrive at this recovery strategy!

Try to convince AS 5 that their session has be reset (or filtered) **even though it is not associated with an active route!**
Same problems can arise with “traffic engineering” across regional networks.
Look familiar?

Intended Routing for some prefixes in AU....
What is going on?

• There is no guarantee that a BGP configuration has a unique routing solution.
  – When multiple solutions exist, the (unpredictable) order of updates will determine which one is wins.

• Complex policies (weights, communities setting preferences, and so on) increase chances of routing anomalies.
  – ... yet this is the current trend!
How to manage?

- Study the interactions in the wild
  - Several research groups working on this ...
  - Could more edu/research nets follow example of Internet2 and publish configs?
- Guidelines for configuration?
  - This may be as simple as translate depreff me communities in a consistent way
  - Or it may be more complicated, depending on what ISPs are actually doing ...
- Unsolved research problem: autonomy vs global sanity