

# BGP Persistence a.k.a. Long-Lived Graceful Restart

John Scudder

IDR, IETF-87, August 1, 2013

# List of Authors

- Jim Uttaro
- Enke Chen
- Bruno Decraene
- John Scudder
- Clarence Filsfils
- Pradosh Mohapatra
- Yakov Rekhter
- Rob Shakir
- Adam Simpson

# In a Nutshell

- When BGP session goes down,
  - allow relevant routes to “persist” (remain installed, but stale) for a long period of time.
  - Routes are “depreferenced” (only selected as a last resort)
- Intended use
  - “dinosaur killer” rare-but-severe control plane outages
  - Restricted/carefully considered AFI/SAFI and/or topologies

# History

- 01 requested IDR adoption in mid-2012, to strong debate (love, hatred) but no clear consensus.
- Strongest objection was, if used for Internet AFI/SAFIs, possibility of leakage to the Internet At Large.
- 02 is a major revision intended to address this
  - Also analysis, clarity, terminology, {code, spec} reuse

# Regular vs. Long-Lived GR

- Normal GR: don't react to session outage
  - Routes kept, no signaling to rest of network
  - Prioritizes network stability. Assumption is short duration with reversion to previous state.
- LLGR: do react
  - Routes kept but depreferenced: signaling required, network state may change
  - Stale routes are a last resort. Assumption is long duration, use up-to-date state whenever possible.

# High-Level Description

- Many semantics of GR useful for Persistence
  - ... so rather than reinvent, reference.
  - Implementation – minimize new/divergent code
- So what's new/different?
  - Routes can be stale for up to  $2^{24}-1$  seconds
  - Capability to signal support and constrain propagation
    - Stale routes may only be advertised to supporting peers, and are marked as “LLGR\_STALE”
    - Hack for partial deployment, using NO\_EXPORT
  - “NO\_LLGR” community to suppress LLGR treatment

# Operational

- Default off
  - Enable per AFI/SAFI after consideration
  - Generally: avoid if very dynamic, topological diversity. Consider if “semi-static”, topologically boring
- Probably usually scope to a single AS
  - But anyway, limit scope of LLGR routes to “consenting adults”

# To Do

- Multicast VPN requires special consideration
  - Emerging strategy is to never use stale routes in making a new determination of Upstream PE or Upstream Multicast Hop
    - Effectively, a more draconian version of “depreference”
  - Placeholder in -02, detailed language for -03
- Note other option: don’t use LLGR for M-VPN
  - When in doubt, leave it off. Default is off.



# Other issues from 01 debate

- Multi-fault scenario unlikely, poor network design
  - There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy
- Depreferencing may be wrong strategy in face of supernets
  - In some cases yes, in some, no. In main use cases, no.
- Problem too marginal to justify using IDR time
  - Prefer to standardize properly rather than publishing as Informational or Individual Submission
- Solution isn't perfect
  - Perfect is the enemy of good

# Next Steps

- Several implementations underway
- (Re-) Requesting WG adoption