Agenda

- Goal & use case.
- Persistence & Graceful Restart: complementary use cases.
- Enabling both Persistence & Graceful Restart.
- Stale routes are less preferred.
- Deployment considerations.
Goal & Use case

- BGP persistence targets catastrophic failure …
  - degraded routing is better than nothing.

- … in a controlled network / environment.
  - In scope: BGP/MPLS VPNs, routes internal to an AS.
  - Out of scope: Internet inter-domain routing.

- Typical use case is the failure of both redundant Route Reflector.
  - including eBGP multi-hop for inter AS option C
  - i.e. BGP control plane only equipments
Persistence & Graceful Restart: complementary use cases.

1. Introduction .............................................. 4
1.1. BGP Graceful Restart and BGP persistence targets
different use cases ........................................... 4
http://tools.ietf.org/html/draft-uttaro-idr-bgp-persistence-01#section-1.1

GR: control plane restart

- Assumption: control plane to go back quickly, all protocols possibly affected, “certain” that forwarding is not affected.
- attempt local recovery: keep using routes, quickly falls back if peer is dead.

Persistence: catastrophic BGP failure

- Assumption: BGP only failure, possibly large scope & long duration, no certainty on route validity.
- degraded mode: use live path if available, otherwise stale path is better than nothing,
  relies on others protocol (IGP, BFD, link layer…) to check BGP Next Hop liveliness.

Largely independent usages

- One could enable GR or Persistence or both or none.

Translate into different solutions

- GR: short timer, no attribute change, no route advertisement, negotiate capability with peer.
- Persistence: long timer, lower preference of stale routes, re-advertise, no capability negotiation.
Persistence and Graceful Restart interactions

- Persistence and GR can be enabled independently
  - GR only \(\rightarrow\) RFC 4724
  - Persistence \(\rightarrow\) draft bgp-persistence

- If both are enabled on a BGP session, the principle is to start first with Graceful Restart
  a. If GR recovers \(\rightarrow\) GR ends, back to normal BGP \(\rightarrow\) Persistence never used
  b. else (GR fails) \(\rightarrow\) Persistence starts

- Both cases detailed in the draft.

Stale routes are less preferred

- **Route preference requirement:**
  a. A stale path is less preferred than a live path.
  b. Between stale paths, (pre-stale) relatives preference are kept.

- **Mechanism investigated:**
  - Cost Community, Local Pref, well known community, BGP attribute
  - both for eBGP & iBGP
Proposed way to lower the preference

4.1.2. Lower route preference .............................................. 8

iBGP (within AS)

• LOCAL_PREF decreased by a configured value
  • Pro: Available now, incremental deployment
  • Con: Some limitations (e.g. interwork w/ existing LP values)

• Optionally (long term): BGP cost community
  • Pro: flexible
  • Con: feature availability / no incremental deployment

eBGP (between ASes)

• well known STALE community
  • to be translated in iBGP as per above
Deployment considerations

- If BGP cost community used, all routers needs to be compliant with I-D.ietf-idr-custom-decision
  - otherwise, forwarding loops may form.

- BGP persistence requires a way to validate BGP Next Hop reachability / liveliness
  - ... since BGP KEEPALIVE can't be used anymore
  - e.g. for iBGP: IGP, LDP ordered mode
  - e.g. for eBGP: BFD, link layer, physical layer
thank you
Annex: main changes introduced in v1

Changes -01

- PERSIST community removed

- Use of local_pref or cost_community to lower the preference of the path within an AS. Between AS, the STALE community is used to convey the information.

- Deployment considerations section enhanced.

- Introduction explains why GR and persistence are different and target different needs.

- Security section refer to RFC RFC 4781.

- New section describing interaction between GR and Persistence.

Persistence & GR interactions (a)

Case a: GR succeed and Persistence never kicks in:

1. BGP session failure --> GR behavior applies.
   * Route marked as stale.
   * Route are kept unchanged (hence not re-advertised).

2. BGP session is re-established before GR timer expires --> GR succeed, GR behavior applies

   1. Route are refreshed.
   2. When End-of-RIB is received, route still marked as stale are removed.
   3. If routes have changed, routes are updated in the FIB and re-advertised to peer as per regular BGP.
Persistence & GR interactions (b)

Case b: GR fails and Persistence kicks in:

1. BGP session failure --> GR behavior applies
   * Route marked as stale.
   * Route are kept unchanged (hence not re-advertised).

2. Expiry of GR restart-time-expiry timer --> GR behavior ends, Persistent behavior applies.
   1. GR stale routes are marked as Persistence stale and their preference is lowered.
   2. As a result, regular BGP best path computation runs and possibly select alternate routes.
      + If routes have changed, routes are updated in the FIB.
      + Updated routes are advertised to peer as needed.

3. Session now runs in persistence mode as defined in this document