Operational Requirements for Enhanced BGP Error Handling in BGP-4

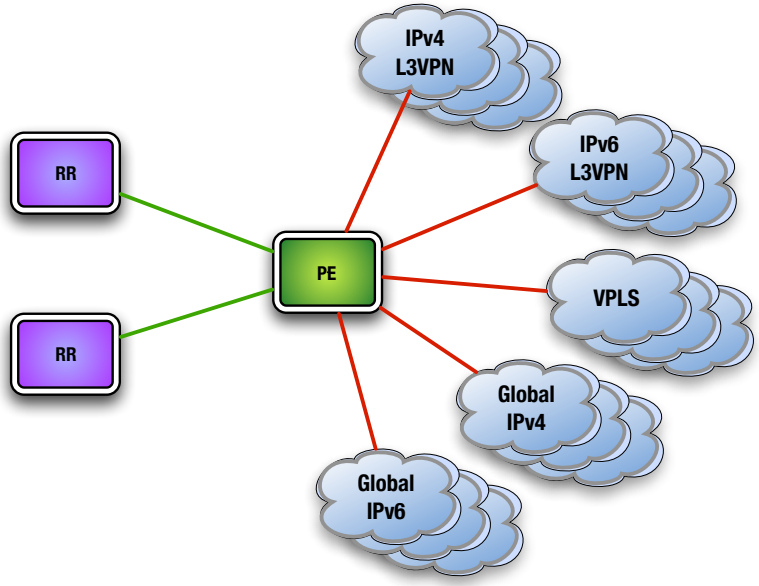
draft-shakir-idr-ops-reqs-for-bgp-error-handling

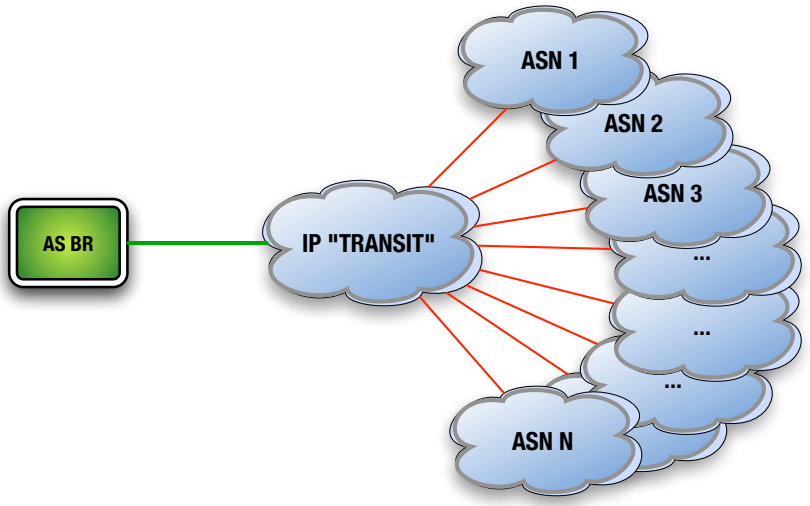
Rob Shakir <rob.shakir@cw.com>

IETF 80 – Prague, CZ.

Problem Statement.

- **NOTIFICATION based on errors in BGP-4 UPDATE messages cause disproportionate failures in Service Provider Networks.**





iBGP

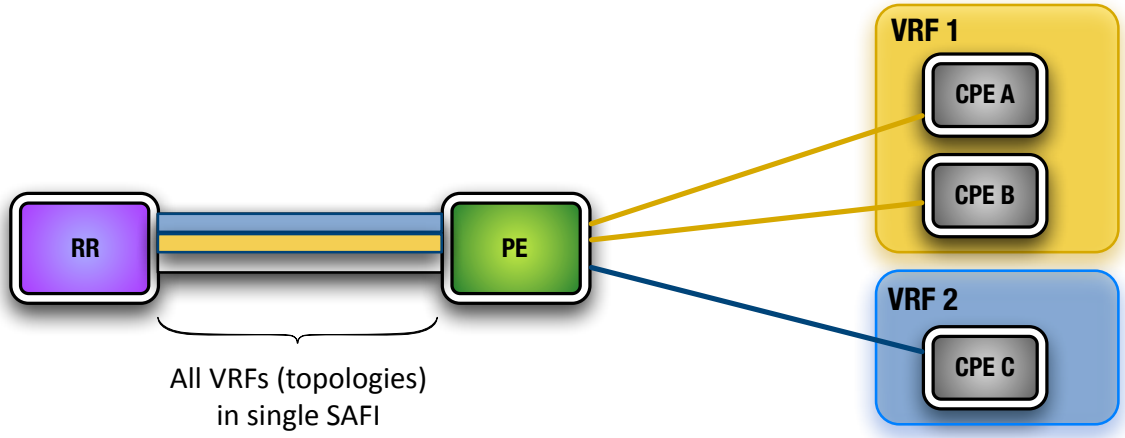
- Multiple AFIs (services) affected.
- Discrete routing topologies affected (e.g. different L3VPNs)

eBGP

- Paths to all NLRI affected despite error in single UPDATE.

Avoiding sending NOTIFICATION.

- **Operator's deployments mean compromises to protocol correctness resulting in invalid routing may be acceptable.**
  - Particularly with multiple AFI – some carrying many discrete topologies.



- **Requirement is to avoid sending NOTIFICATION where possible.**
  - Do not send for erroneous UPDATEs (and hence avoid teardown).
  - Session failure affects all NLRI, where negative impact affects a subset.
  - Required for both eBGP and iBGP.

Recover RIB Consistency.

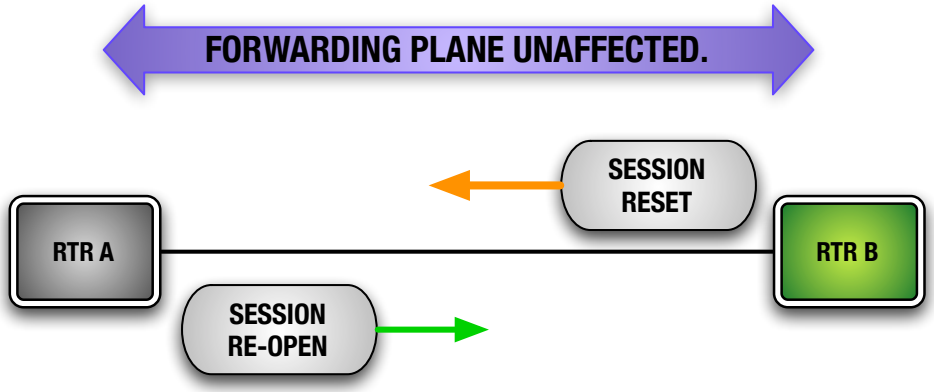
- **Inconsistent RIB (by treating UPDATE as withdraw) compromises protocol correctness.**
  - The resulting RIB inconsistency may have resulted in forwarding loops or black-holes.
  - BGP speaker is aware of this case, if using “treat-as-withdraw”.

- **Whilst such inconsistencies are acceptable, they are clearly sub-optimal.**
  - Mechanism required to recover consistency of the RIB, and remove invalid routing.

- **Whole RIB or specific RIB subset?**
  - ROUTE REFRESH is inefficient where a BGP speaker knows the NLRI transmitted in the invalid UPDATE.
  - Requirement for mechanism(s) to request specific RIB subsets – reduce control-plane load.
  - Allow for such requests to be automatically or manually generated.

Session Reset whilst Maintaining RIB/FIB.

- **Currently NOTIFICATION and session reset is the reaction to an error.**
  - Deals with resetting state that may have resulted in erroneous UPDATE.
  - Major operational issue is the forwarding disruption caused.



- **Benefits of resetting all session state whilst allowing forwarding to continue.**
  - Identical recovery mechanism as is implemented currently, with lower impact to operation of the network.

Monitoring.

- **Additional complexity in the protocol requires further operational visibility.**
  - Let our NOCs know about BGP-4 errors, and respond.
  - Previously NOTIFICATION/tear-down was very visible due to forwarding outages.

- **Enhance monitoring toolset.**
  - Capability to transmit error information between BGP neighbours.
  - Further visibility to determine where errors have occurred, and what they are.

Caveats of Requirements.

- **React to errors (and recover) within available control-plane resource.**
  - Ensure that we do not reach looped scenarios where automatic recovery is available.
- **Exponential (?) Back-Off for RIB recovery requests.**
  - Don’t overload neighbour and/or local BGP speaker with recovery requests.
- **Avoid constant session restarts.**
  - Identify a point at which a session is “bad” if using automatic mechanisms to recover.

Draft Progression.

- **Draft has been presented and discussed at a number of operational forums.**
  - NANOG, UKNOF, LINX.
  - Well supported as a set of requirements for operators (see GROW and IDR mailing lists).
- **Would like WG adoption.**
  - Provides a framework to which IDR/GROW work items can be tied.
  - Intends to avoid “partial solutions” that do not meet the toolset required by operators.
- **Thoughts as to which WG is most suitable?**