BGP Persistent Route Oscillation Solution (draft-walton-bgp-route-oscillation-stop-01.txt)

Daniel Walton, Alvaro Retana, Enke Chen
{dwalton, aretana, enkechen}@cisco.com

What is the problem?

- BGP Persistent Route Oscillation Condition (rfc3345)
  - Combination of RR and/or confederations with the use of MED, may cause persistent route oscillations.

- Type I = single-level Route Reflection or AS Confederations
  - Solved by following the *Deployment Considerations* in rfc2796 and/or rfc3065.

- Type II = More than one tier of Route Reflection or Sub-ASs
  - Not all workarounds are operationally feasible.

Is the problem real?

- Yes!!
  - Several ISPs have been experiencing the oscillation (on dual-tier deployments).

- Real world numbers (taken a couple of years ago – average of several measurements):
  - Total number of routes: 112086
  - Oscillation Candidates: 9011
  - Oscillating Routes: 899

Solution (1)

- In a network with a full iBGP mesh, all routers have consistent and equivalent routing information – the MED-type oscillation doesn't occur.

- Solution: allow the advertisement of all available paths in order to recreate the consistent and equivalent view.

  – Results in the same amount of routing information as a full iBGP mesh.

Solution (2)

- A BGP speaker (route reflector or a sub-AS border router) MUST advertise the following to its iBGP or confederation peers:
  - The "neighbor AS based Group Best Path", for each neighbor AS (which includes the best path)

- Notes:
  - "neighbor AS based Group Best Path" = path considered best among all paths from the same neighbor AS
  - Only "neighbor AS based Group Best Paths" for which the tie breaker (when comparing the path to the best path) occurs after the MED is compared need to be propagated.
  - Result = each speaker will advertise at most n "neighbor AS based Group Best Paths" to its non-eBGP peers (one for each neighbor AS), including the best path.

Next Steps

- Adopt this draft as a WG document
  - Addresses a hole in the way BGP is specified today.
  - Solves a real problem in deployed networks.
  - Relatively low impact if additional routes are propagated when the MED oscillation is detected.

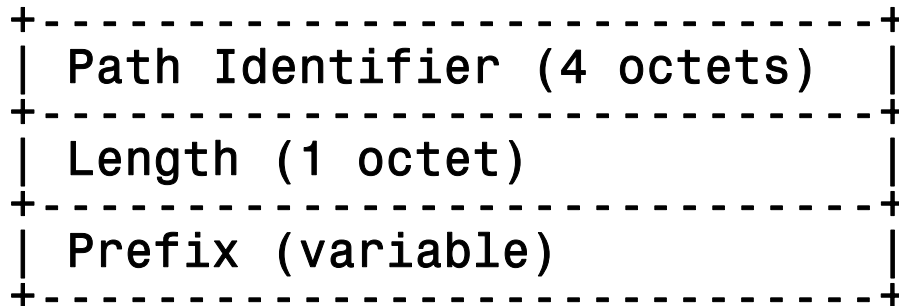# Advertisement of Multiple Paths in BGP (draft-walton-bgp-add-paths-03)

Daniel Walton, Alvaro Retana, Enke Chen

{dwalton, aretana, enkechen}@cisco.com

Proposal

- Mechanism that allows the advertisement of multiple paths for the same prefix without the new paths implicitly replacing any previous ones.
  - Summary: add a *path identifier* to the encoding.
  - The intent of this extension is to be used in a ***controlled fashion*** for applications that require only partial propagation of the routing information, or specific individual recipients.

NLRI Encoding

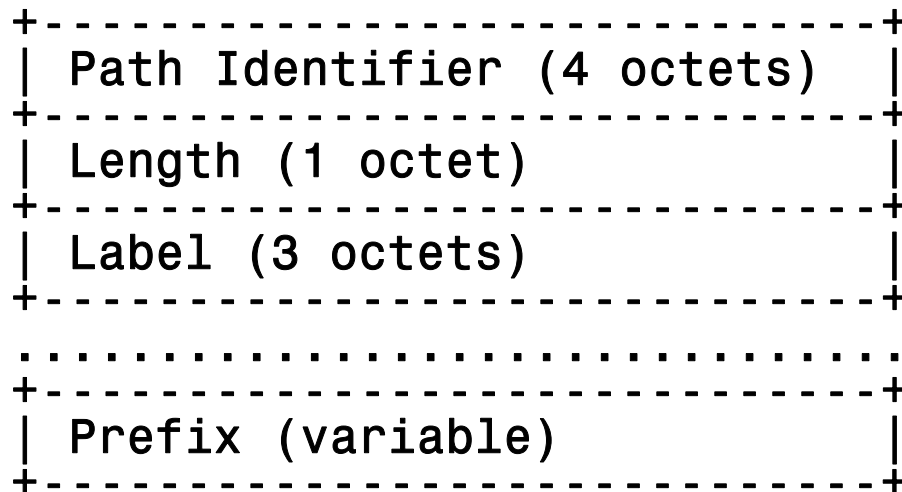- Extension to *Multiprotocol Extensions for BGP-4* (RFC 2858) and the base spec (RFC 1771).
  - The *Path Identifier* field is used to distinguish between different prefixes.

```
+-------------------------------+
| Path Identifier (4 octets)    |
+-------------------------------+
| Length (1 octet)              |
+-------------------------------+
| Prefix (variable)             |
+-------------------------------+
```

NLRI Encoding (Cont.)

- Extension to *Carrying Label Information in BGP-4* (RFC 3107).

```
+------------------------------+
| Path Identifier (4 octets)   |
+------------------------------+
| Length (1 octet)             |
+------------------------------+
| Label (3 octets)             |
+------------------------------+
................................
+------------------------------+
| Prefix (variable)            |
+------------------------------+
```

Capability Advertisement

- New Capability: *ADD-PATH*
  - Code TBD
  - Length is variable..
- Value: 0 or more tuples:

```
+------------------------------------------------------+
| Address Family Identifier (2 octets)                 |
+------------------------------------------------------+
| Subsequent Address Family Identifier (1 octet)       |
+------------------------------------------------------+
| Path Identifier Type (1 octet)                       |
+------------------------------------------------------+
```

Used for display/troubleshooting.

Capability Advertisement

- Update in -04 Version to be published after IETF
- New Capability: *ADD-PATH*
  - Code TBD
  - Length is variable..
- Value: 0 or more tuples:

```
+---------------------------------------------------------+
| Address Family Identifier (2 octets)                    |
+---------------------------------------------------------+
| Subsequent Address Family Identifier (1 octet)          |
+---------------------------------------------------------+
```

Operation

- Advertisement of the capability indicates ability to receive multiple paths for all negotiated AFI/SAFIs.

- Advertisement of specific AFI/SAFI information in the capability indicates the intent to send multiple paths for it.

  – Only in this case must the new encoding be used.

Applications

- MED Oscillation
- Several multipath applications
- Route Server
- Other applications are left for further study.

Next Step

- Adopt this draft as a WG document

- Update rfc3107 to use the proposed encoding for the advertisement of multiple paths and remove the capability code it defined.