BGP Route Oscillation Reduction - A Single-path Approach

Enke Chen
(enke@redback.com)

Observation on Route Reflection

- Currently a RR "reflects"only the best path
  - This results in route withdraws by RR (even when it still has a valid route from client/non-client)
  - As shown in Case #1, the route withdraw by RR causes the complete loss of one-cluster's routing information from other clusters
    - Over-reduction of routing info (compared with full-mesh)
- Routing information can be increased if RR is revised to "relay" a path between its clients and non-clients

Proposed Revision to RR

- Toward a client:
  - advertise a route from a non-client
- Toward a non-client:
  - advertise a route from a client
- In both cases:
  - Advertise the overall best if possible
  - Otherwise advertise group best if possible

Revised Approach: Pros and Cons

- Pros
  - Increase available routing information within the paradigm of adverting a single-path: the memory usage is no worse than the closest-exit approach
  - Eliminate most of route oscillations (incl. two cases shown in Sue's slides)
  - No protocol change - revision to the semantics of route advertisement
  - Simple to implement and deploy
- Cons
  - Route oscillation is not eliminated in some cases

BGP Route Oscillation Detection and Deployment Considerations

Enke Chen
(enke@redback.com)

Route Oscillation Detection

- Persistent route oscillation occurs with small number of routes in certain network topologies.
  - Huge problem for the affected customers
- Needs to detect their occurrences and workaround
- Fully automated detection is difficult to get right
- Suggested approach: protocol assistance + human analysis
  - Maintain flap statistics for IBGP routes
    - At least one implementation already (by Redback)
    - “show bgp route flap-statistics internal”
  - Routes with high flapping frequencies are candidates for further analysis

Deployment Considerations

- Enable “deterministic-med” in route selection
  - Inconsistent route selection leads to forwarding loops even w/o RR or confederation
- Consider fully-mesh IBGP speakers
  - Ask your vendor to support 500+ sessions !
- Route Reflection
  - Use only one-level route reflection
    - Why more than one is needed ???
  - Fully-mesh clients within a cluster
    - Clients are not the bottleneck and they do not need to sit idle

Deployment Considerations (cont.)

- Use MED only when needed
- Consider network topology in designing route reflection (or confederation) topology
  - Intra-cluster metric << inter-cluster metric
- Do not transition the best path from one EBGP path to another based on router-id
  - Why churn for something that is random ?
- Monitor route oscillation and workaround
  - Maintain and analyze route flap statistics for IBGP

Acknowledgement

- Thanks to Pedro Marques, John Scudder, Yakov Rekhter and Jenny Yuan for their valuable inputs and discussions on this topic.