An Aggregation-Based Evolutionary Path towards Global Routing Scalability

Team [APT → Evolution]

Presented by Lixia Zhang

RRG @ IETF76
The problem and the solution

• Routing tables continue to grow
  – Multihoming, traffic engineering, (lately) “security”
  – Benefits to those who deaggregate; costs to those who carry full table

• Basic approach to scalable routing:
  route aggregation
Aggregation: what it is

• Aggregation: use a shorter prefix to cover multiple longer prefixes of the same block
  + shrink FIB and/or RIB
  + reduce churn

• Why Aggregation: The farther away you are from a prefix, the less details you need
  - may introduce stretch, just like any other hierarchy structures
    - e.g. routing over autonomous systems introduces stretch compared to globally flat topology
Why we haven’t had a solution deployed

• Internet is big and diverse $\rightarrow$ routing scalability problem is not universal
  – edge sites vs. ISPs; new core routers vs. old PE routers
  – Affordability varies: Some feel that this is a serious long-term problem; others feel it’s not an issue, just pay

• Internet has no boss $\rightarrow$ no universal buy-in, no flag day/year/decade
So,

• Requirement for any solution:
  – Can be deployed by individual parties
  – needs clearly identifiable returns

• Is there anything one can do to reduce local RIB/FIB without relying on others’ good will?
  – Yes

• Will this eventually lead us towards the global routing scalability?
  – Yes, starting from local route aggregation
Present vs. Future

• Applications, requirements, and technology have all been changing over time
• History does not show that we are particularly good in predicting futures with any accuracy
  – We know better about specifics of current time
  – See less clearly for N years down the road
→ Relatively more confident about today’s problems, feasible solutions
  – Identify the landmark for future directions
What is an evolution path

• Each stage
  – focuses on an immediate problem that warrants a change.
  – offers a solution with reasonable deployment cost considering the problem.
  – can be taken by individual ASes as/when they see fit.
• converging towards desired direction
• Notes:
  (1) Some networks may not adopt any solutions
  (2) Different networks may be in different stages, therefore they must be able to co-exist.
Evolutionary Path based on Aggregation

• Basic idea: Apply aggregation with increasing scopes, from local to global

• Evolutionary path
  – *local router* $\rightarrow$ FIB aggregation, no stretch
  – *Intra-AS* virtual aggregation $\rightarrow$ further reduce FIB, but with stretch
     • Also opens possibility to reduce RIBs on non-ARP routers, open issues remain to be resolved
  – *Inter-AS* virtual aggregation using existing BGP sessions $\rightarrow$ reduce stretch
  – APRs peer with neighboring ASBRs over multi-hop BGP sessions $\rightarrow$ reduce RIB size of non-APR routers
Defining Incremental Deployment

• What is “Incremental deployment”? 
• So far: Co-existence of the new and old worlds. 
  – An ISP running new architecture can inter-operate with legacy ISPs. 
  – But the cost associated with new deployment can be high, while the immediate gain can be low. 
  – Routing table reduction may not come until most of the world have converted (e.g. LISP, APT)
Real definition of incremental deployability

• In addition to Co-existence of the new and old worlds
• must also provide enough incentives at each stage.
  – Future state is determined by economic forces.
• Architecture/protocol designs need to
  – Steer the system towards promising direction; *aggregation in our work*
  – facilitate future changes
The Goal of This Discussion

• Show an example of an evolutionary path towards scaling the global routing architecture
  – illustrate feasibility of convergence towards scalable routing

• The particular path mentioned in the example are not meant as a fixed prediction
  – Solutions for today: feel confident
  – Solutions further out: less sure

• The direction: bring FIB, RIB and update volume under control
  – Show that the first step can move toward a global optimum without getting stuck in local minimum
FIB Size Reduction without Stretch

• FIB aggregation by *individual router*
  - If two numerically aggregatable prefixes share the same next-hop, aggregate them into one.
  - Almost no impact on packet forwarding.
  - No impact on routing.
  - Compatible with future solutions.

• Gain: up to *70% reduction* of FIB size, no stretch.

• Cost: CPU cycles, but controllable

• Deployment: a software upgrade at local router.
FIB Size Reduction with Stretch

- Intra-AS Virtual Aggregation (Francis et. al.)
- Aggregate most prefixes into virtual prefixes; leave out popular prefixes
  - Ideally most traffic load does not have stretch
- Deployable by individual ISP, no impact on other networks
- Bring immediate FIB reduction
  - Cost: path stretch; additional complexity
- Preliminary evaluation: the FIB size can be reduced by a factor of 10 or more with minimal stretch
RIB Size Reduction for Non-APR Routers

- Current VA proposals did not touch RIB to avoid impact on neighbor ASes
  - Need to provide full BGP table to downstream neighbors who want it
  - FIB is a local business
- But it’s possible to reduce RIB size of non-APR routers.
  - APRs must hold the full table anyway
  - Let APRs peer with neighbor ASes’ border routers via multi-hop eBGP sessions
  - PLEASE DON’T JUMP UP: yes some issues need to be nailed out here, but nothing seems fatal.
Inter-AS Virtual Aggregation (Francis et. al.)

• If neighboring ASes deployed VA: exchange mapping information
  – exit router address of more specific prefixes under virtual prefixes.
• Deployable by neighboring ISPs
• No impact upon operations of other networks
• Can reduce path stretch
RIB and Update Reduction for APRs

• The crystal ball looks cloudier when one attempts to look into further future

• Possibilities:
  – Routers further away from destination ASes may ignore more specific prefixes
  – Reduced RIB, reduced updates
How Do We Know We Are Heading to the Right Direction?

- Scale Routing through aggregation
- Enable aggregation with increasing scopes.
- We give decisions to individual ASes
  - Thinking about all the changes over last 10 years: which one was a joint action by multiple ASes?
Step Up A Level

• There will **not** be a single global routing hierarchy/a mapping table as many people have envisioned (including ourselves)

• Individual ISPs are dealing with their own routing table size problem.
  – There have been attempts to voluntarily stop routing propagation.
  – With FA and VA: one can send as many routes as one wants to neighbors, the receiving AS will aggregate as much as it needs.
What about “architecture”

• The goal: scalable routing architecture
• From dictionary: building structures; layout, formation, arrangement
• Good routing architecture
  – Fullfill the function needed today
    • Put FIB, RIB, updates under control
  – Must be incrementally deployable!
  – Stay flexible for extension to meet the need for tomorrow
Relation with Other Proposed Solutions

- Our proposal complement those solutions starting from “edge” (more “clean slate”)
- Paul: if/when LISP (ILNP) succeeds one day, we no longer need all this stuff (FIB, RIB reduction)
  - VA provide solutions to meet individual ASes’ problems today while waiting for longer term solutions to rollout.
- Impose no changes to current practice at edges/applications while ISPs evolve their own routing structure
  - New developments such as MPTCP, HIP, etc. proceed in parallel.
Thank You

Questions? Comments?