Evolution Towards Global Routing Scalability

Team [eFIT → APT → Evolution]

Presented by Lixia Zhang

RRG @ IETF75

Lets Face It

- We are chartered to come up with a solution to scalable routing
- Internet is big
- Internet has no boss
- Any new change need clearly identifiable returns
  - Cost and incentive aliagment

→Need an evolutionary path *towards* scalable routing

Presence 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
  - We try to identify the *landmark* for future directions

→Need an evolutionary path towards scalable routing

  - Relatively more confident about today's problems and feasible solutions
  - See less clearly for 10 years down the road

What is an evolution path: Looking Back

- The Internet routing architecture has gone through several stages of changes
  - Each stage focused on an immediate problem that warrants a change
  - Each stage found a solution with reasonable deployment cost
  - Solutions were taken by individual ASes as/when they felt needed
- The routing system has not closely followed any given prescription envisioned
- The system evolves itself to converge towards desired direction

Evolution –vs– Incremental Deployment

- New architectural solutions (like LISP, APT) can potentially bring big benefits
  - after being deployed by majority of ISPs and edge sites
- "Incremental deployment" of a new design often means that an ISP adapting the new design can inter-operate with legacy ISPs, but
  - Cost associated with new deployment can be high
  - Immediate gain can be low
- An evolutionary path solves specific problems with enough incentives at each step
- Future state is determined by economic forces
- Architecture/protocol designs need to
  - Steer the system towards promising directions at each step
  - Facilitate future changes (that we may not see clear today)

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 RIB, FIB, and update volume under control
  - Show that the first step can move toward a global optimum without getting stuck in local minimum

Internet Is Big

- Different parts feel different degrees of growing pains
- Most Stub ASes don't carry full table internally
  - But many do
- Some ISPs can afford to upgrade routers
  - But some cannot
- Within an AS some routers experience problems more severely than others
  - FIB size
  - Update processing/routing computation

Internet Routing Scalability: a problem?

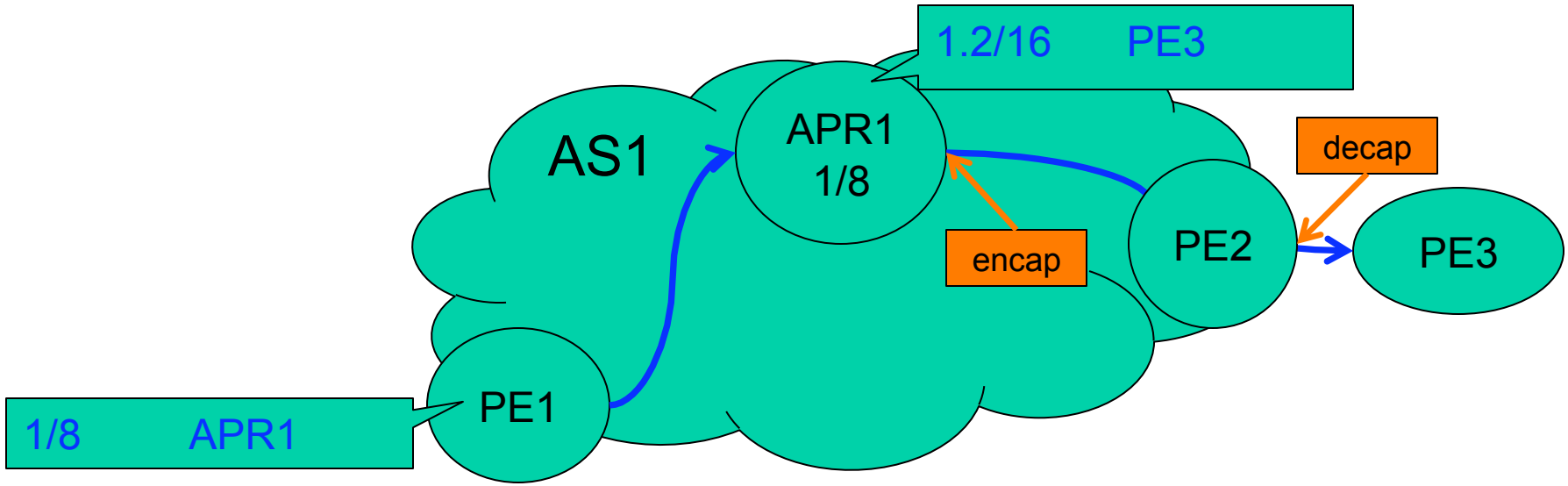
- DFZ routing tables have been growing in a largely uncontrolled way
- Expect fast growth in coming years
  - IPv4 address exhaustion → further fragmentation
  - IPv6 rollout
- Routing table growth brings the following to routers:
  - RIB size growth
  - FIB size growth
  - BGP update growth
    - Going up with RIB size
    - Going up with the network size: large networks inherently have less-well managed parts

First Step: Controlling FIB Size

- Virtual Aggregation
- Deployable by individual ISP
  - Don’t need coordination with anyone else
- No impact upon operations of neighbor networks
- Can bring immediate FIB reduction

Virtual Aggregation is poorman's Map-Encap

- APR holds the Map of all specific prefixes to the exit routers
- Packets first forwarded to APR, then to exit PE
  - ≈ APT/LISP within an AS, concerning FIB



Benefit and Cost of VA

- Bad news first
  - Path stretch
  - With sensible APR placement, preliminary measurement shows the results not too bad (`draft-ietf-grow-va-perf-00`)
- Good news: Shrinking FIB by an order or more
  - Can fit into those resource constrained places
  - Can reduce FIB download delay
    - hence speed up convergence, improve data plane performance
- A silent fact:
  - A smaller number of routers, APRs gain more control power than others

Next Step: RIB Size Reduction

- VA 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
- VA can also reduce RIB size with little impact on neighbor ASes
  - APRs must hold the full table anyway
  - Let APRs peer with downstream neighbors via multihop BGP sessions
    - PLEASE DON'T JUMP UP: yes some issues need to be nailed out here, but nothing seems fatal

≈ APT/LISP within an AS (FIB & RIB)

Gains and Cost of RIB Reduction

- Bad news first: have to make multihop BGP peer sessions work well
- Good news: Like VA, this is decision by individual ASes, pay a cost for some gains
  - Non-APR routers now have small FIB and small RIB
  - In addition: reduced BGP updates as a result of reduced RIB
    - Updates for suppressed prefixes stop at APRs
- A silent fact: APRs gain more control power
  - Since all routing goes through APRs: a good place to support SIDR solution?

What's Next After RIB Reduction?

- The crystal ball looks cloudier when one attempts to look into further future
- Imagine possibilities:
  - Inter-AS mapping exchange?
    - Inter-AS VA [Xiaohu Xu's talk @ IETF74 RTGW]
  - If this happens, the world moves further towards APT, LISP design

    ≈ APT/LISP with an AS cloud
  - The real question: how much is the gain? (to balance out the cost)

How Do We Know We Are Heading to the Right Direction?

- Routing scalability possible through aggregation
- We are enabling aggregation
- We leave decisions of deployment to individual ASes
  - Thinking about all the changes over last 10 years: which one was a simultaneous, joint action by multiple ASes?

Step Up A Level

- There may not be a global mapping table as many people have envisioned
- Individual ISPs are dealing with their own routing table size problems
  - There have been attempts to voluntarily stop routing propagation
  - With 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
  - Stay flexible for extension to meet the need for tomorrow

Evolution –vs– architecture

- In the process of reducing routing/forwarding table size of majority of routers, a minority set takes on more control responsibility
- A promising routing architectural direction: separating control plane from data plane
- What about separating out IP addresses from identifiers, or IPv6 transitions
  - Not aim to solve multiple problems by one solution
  - Aim at a coherent architecture, which facilitates best engineerig solutions for individual problems
- Of course all above is open for debate!

Relation with Other Proposed Solutions

- Complement those solutions starting from "edge" (clean slate design of separating edges from core)
- Paul: "if/when LISP (ILNP) succeeds one day, we no longer need all this stuff (FIB, RIB reduction)"
  - VA provide solutions to meet indivual ASes' problems today while waiting for longer term solutions rollout
- Impose no changes to current practice at edges/ applications while ISPs evolve their own routing structure
  - New developments such as MPTCP, HIP, etc. can proceed in parallel

Thank You

Questions? Comments?