Solving the Routing Scalability Problem -- The Hard Parts

Jari Arkko
APRICOT 2007, Bali, Indonesia

Outline

- Where are we on this?
- Some hard bits
- Proposed plan of action

Where Are We on This?

- There is a lot of interest
- People willing to design solutions
- Discussion forums & meetings exist
- Pretty good understanding and agreement about why this problem occurs

Where Are We on This? II

- There are different interpretations of how serious or not the problem is
  - Not everyone believes we have an issue that cannot be addressed by throwing more hardware at it without significant cost impact
- Different opinions with regards to what is most important, e.g. FIB size vs. dynamics
- People working for many different directions and on different time scales
- Have not hit all the hard questions yet

Hard Parts I – The Meta Issue

- Agreeing on how serious the problem is
- Throw hardware or protocols at it?
- But the engineering community should not work only on Internet threatening issues!
- Can we improve the current design?
  - E.g., more users or more provider independence or more multihoming for the users with the same effort
- Sets limits on what kind of solutions can be considered

Hard Parts II – Router Scalability

- Not just about the forwarding decision
  - Also need BGP computation and communication, move data from the RIB to FIB, meaningful management tools for large tables, and so on
- Conversely, router hardware has to do many other things as well
  - Filtering, prioritization, source address validation, tunneling, … (list keeps growing)
- What you see is a sum of different factors
- And commercial issues affect this, too…

Hard Parts II – Deployment

- Deployment and use is what counts
- The hard part is an actual table impact!
- What is the motivation for deployment?
  - Host/router/peer/DNS/...
- If the same organization spends the cost and gets the benefits, we have a good model
- If not, it is questionable what motivates others to deploy something new

Hard Parts II – Deployment Cont'd

- Relatively easy to upgrade some interested set of end hosts

- Very hard or impossible to expect upgrades from everyone

- Its a complete non-starter to require application modifications

Hard Parts III – Applications

- Referrals – how do they work?

- Host stores peer's address in file and attempts to contact it later when the host stack and router have lost the context. Can you find the peer's locator?

- Or, host sends what it thinks is an address to a peer in SIP/SDP. Does the peer know where to send the packet?

- Particularly hard problem when communicating with legacy nodes AND simultaneously reducing DFZ table size

Hard Parts IV – Security

- How do you secure the mapping?

- Are dynamic changes allowed? Can I claim that your identity is now in my computer?

- The solutions that we have seen have wildly different approaches to security

Hard Parts V – Scope

- How ambitious is this effort?
- Routing scalability in the fixed network?
- … with multihoming?
- … with mobility?
- … with secure identifiers (e.g. HITs)
- … with e2e security (e.g. HIP ESP)?
- … with denial-of-service defences (Hi3)?
- … clean slate?

Hard Parts VI – Limits of an IP Solution

- Ease of renumbering is not just a host / router problem – DNS, firewalls, application configs, etc. are involved

- The pressure to keep the same locators may not go away completely

- Solutions that employ identifier space that looks syntactically like an address may get additional pressure to route on identifiers as well

What Can the IETF Do?

- Routing table size growth causes pain
- There is reason to believe we do not have a short term technology problem
  - But hard work and many commercial issues are ahead. Much of this is outside IETF scope, however.
- IETF can help in short term protocol work
  - Such as tuning BGP better for today's challenges
- IETF can also help by looking at architectural changes
  - Takes time to develop (and more to deploy)

Overall Plan

We need to in parallel

- Continue tracking the problem
- Keep educating the operator community
- Encourage implementation improvements
- Start up short-term BGP improvements
- Encourage Id-Loc split experimentation
- Eventually produce an IETF Id-Loc split

Identifier-Locator Split

- Its easy to charter additional work here

- However, lets not forget that deployment is the true change, not a new invention

- Should focus on things that we currently cannot do (such as control from the network)

- Look at both IPv4 and IPv6 -- be backwards compatible

- Not a replay of the 1990's – we know more now

- Will take time!

- IRTF work on clean slate designs, experimental RFCs on candidate ideas, IETF standard work