RRG Recommendation Discussion

@ IETF77

March 26, 2010

Identified Problems

- Routing scalability for both v4 and v6
  - Site multihoming
  - Traffic engineering
- Host multihoming and TCP's tie to IP address
- Mobility support

Dimensions of Design Space

Scale by enabling route aggregation

1. Enforcing address aggregatability all the way into end hosts
2. Enforcing address aggregatability in DFZ
3. Enforcing address aggregatability with increasing scope, starting from single router
4. All of the above (?)

Solution Requirements (1)

- Enforcing address aggregatability all the way into end hosts
  - Require changes to all hosts, to DNS
  - Changes to site operations to support multiple prefixes
  - Semi-automated renumbering when changing provider
  - No change to DFZ routing

Solution Requirements (2)

- Enforcing address aggregatability in DFZ
  - No change to hosts or DFZ
  - Require a mapping system to be built
  - Require changes to edge routers (CE or PE)
  - Require packet encapsulation across DFZ

Solution Requirements (3)

- Enforcing address aggregatability with increasing scope, starting from single router
  - No changes to hosts
  - Changes to individual AS to reduce FIB

Proposed solutions to each problem

- Site multihoming
  - Map-and-encap
    - AIS (Aggregation with Increasing Scope)
    - CES (Core-Edge Separation)
  - Host multihoming solutions + renumbering→site multihoming (or CE Elimination)
- Traffic engineering
  - AIS: no change to today's practice
  - CES class of solutions can handle to certain degree
  - CEE-based solution: varies
- Host multihoming: varies
- Mobility: varies

Class 1: Transmogrification

1. NOL: Name overlay
2. ILNP: Identifier-Locator Network Protocol
3. AIS: Aggregation with Increasing Scope (evolution)

Class 2: Map-n-Encap

1. RANGI:Routing Architecture for the Next Generation Internet
2. LISP: Locator Identifier Separation Protocol
3. Ivip
4. HIPv4
5. Global Locator, Local Locator, and Identifier Split (GLI-Split)
6. Tunneled Inter-domain Routing (TIDR)
7. Routing and Addressing in Networks with Global Enterprise Recursion (IRON-RANGER) Aggregation with Increasing Scope

Class 3: Mapping System Designs (for CES)

1. Compact routing in locator identifier mapping system
2. LMS: Layered mapping system
3. 2-phased mapping
4. Enhanced Efficiency of Mapping Distribution Protocols in Map-and-Encap Schemes
5. Accessory: Name-Based Sockets

Class 1: Transmogrification

NOL

- Form of NAT/PAT
  - Hide multi-homing
- External PA aliasing for site services in DNS
- Requires host changes to reach servers behind NTR

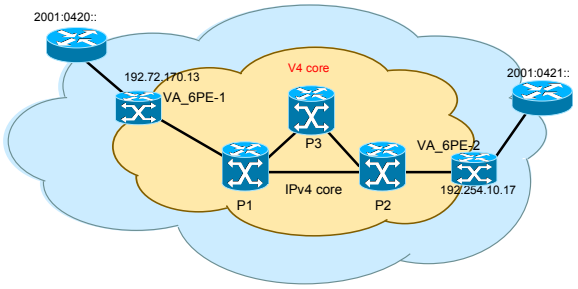
ILNP

- Locators and identifiers are first-class objects
- Splits v6 address in half
- Requires host changes
- Uses DNS as mapping system
- Needs renumbering support

Aggregation with Increasing Scope

- At a router: FIB aggregation
- Within an AS: virtual aggregation
- Can extend aggregation to multiple ASes when/once they all turned on VA
- Deployable by individual parties to control one's own routing table size
- Handles v4-v6 interworking

Carrying IPv6 with VA based network



Class 2: Map-n-Encap

RANGI

- Map-n-encap
  - v6 as transport for v4
  - Similar to HIP, but with structured ID
  - Crypto based
- Use IPv4 addresses in low order 32 bits of IPv6 address as identifier
- Reachability is a concern

LISP

- Map-n-encap edge to edge
- Mapping done by ALT
- IP-UDP encapsulation: packet size increase
- Reachability remains a difficult problem
- Already has a WG

Ivip

- Map-n-encap edge to edge
- Mapping changes are flooded globally and instantly
  - The changes include those due to host mobility
- feasibility is a concern
- Requires all routers be modified

TIDR

- Map-n-encap edge to edge
- Use BGP to distributing the identifier-to-locator mapping
  - Split prefixes into RIB and TIB (Tunnel Info Base, similar to EID in LISP)
- Require changes to all routers

hIPv4

- Map-n-encap
- Two locators: ALOC, ELOC
- Uses a shim to stash unused locator
- Requires host changes, avoids fragmentation issues

GLI-Split

- Map-n-encap
- Need 2 new mapping systems
  - local mapping system maps IDs →LLs
  - global mapping system maps IDs→GLs
- requires host changes and special GLI-gateways
  - Hosts perform heavy lifting of all mapping lookups

IRON-RANGER

- Map-n-encap
- Assumes a hierarchy of recursively-nested networks
  - RLOC addresses in underlying network; EID addresses in overlay
  - More-specific EID prefixes added to router FIBs on-demand, only to routers that need them
- RIB loaded from centrally-managed file; no dynamic routing protocol needed
- has its own tunneling protocol: SEAL

Class 3: Mapping system designs

Compact routing in locator identifier mapping system

- Mapping system only
- Based on compact routing
- Intended for map-n-encap class of solutions

LMS: Layered mapping system

- Hierarchical mapping system
- Administered independently of ISPs
- Concerns about even distribution of mapping load

2-phased mapping

- First phase: prefix →AS numbers
  - M:M mapping
  - Stored in a registry system
- Second phase: AS# → ETR address
- ITR first finds AS#, then finds ETR
- Require changes to all routers

EEMDP

- Enhanced Efficiency of Mapping Distribution Protocols in Map-and-Encap Schemes
- Reduce mapping entries through aggressive aggregation
  - i.e. allowing holes in the aggregation and treating them with special handling

Name-Based Sockets

- Abstract BSD sockets to operate on FQDNs rather than v4 addresses
- Requires application redesign
- Gives OS more flexibility in fulfilling application requests
- Decrease reliance on explicit IP addreses

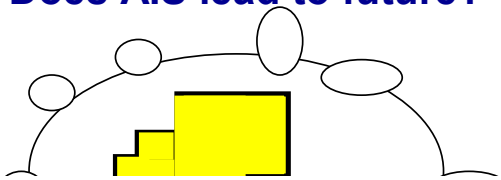
Rationale

- We must have a solution (for IPv6)
- All of the ‘permanent’ solutions require major changes before benefit
- Major changes take time
- Major changes -> make best possible change
- Tactical and strategic changes not incompatible

Recommendation

- AIS
- ILNP
- Renumbering support

Does AIS lead to future?



Trend: Yellow area grows as needed
Yellow area: Gloc; White: Lloc

What about identifier separation from IP address and mobility support? Solved separately (running code exists, see http://www.ietf.org/internet-drafts/draft-zhu-mobileme-doc-01.txt)

From Butler Lampson

- "The test of your architecture is whether you can explain the rules that tell you what your system cannot do.
- "If you claim your system can do everything, then you do not have an architecture; you just have a dream."

— http://www.nets-find.net/Meetings/July09Meeting/July09Meeting.php

"Why The Internet Only Just Works"

— http://www.cs.ucl.ac.uk/staff/m.handley/papers/only-just-works.pdf

- "I believe that this has historically been the natural state of the Internet and it is likely to remain so in future. Unless this is understood, then it's hard to understand which problems are really cause for concern, and which we can safely ignore or put off solving till some later date."
- "Solutions that have actually been deployed in the Internet core seem to have been developed just in time, perhaps because only then is the incentive strong enough. In short, the Internet has at many stages in its evolution only just worked."
- This was never fun or safe. - Tli