# IPv6 augmentation for TLDs

bill manning

bmanning@ep.net

# Design and Implementation Objectives and Assumptions

- Do not impact the production systems
- Provide IPv6 capabilities where they are desired/requested
- IPv6 transit capability will not match IPv4 transit capability in either the short or near term

# Relevant Standards

- Dejura
  - The IETF has some guidelines for DNS servers
  - The IETF is developing operational guidelines for IPv6 use, these remain Internet Drafts
    - Several of the authors of these drafts are not and have not been DNS operators
- Defacto
  - The predominant DNS implementation (BIND - Berkley Internet Name Domain) has features not defined through the IETF
  - Behaviors change, depending on what minor version is currently installed

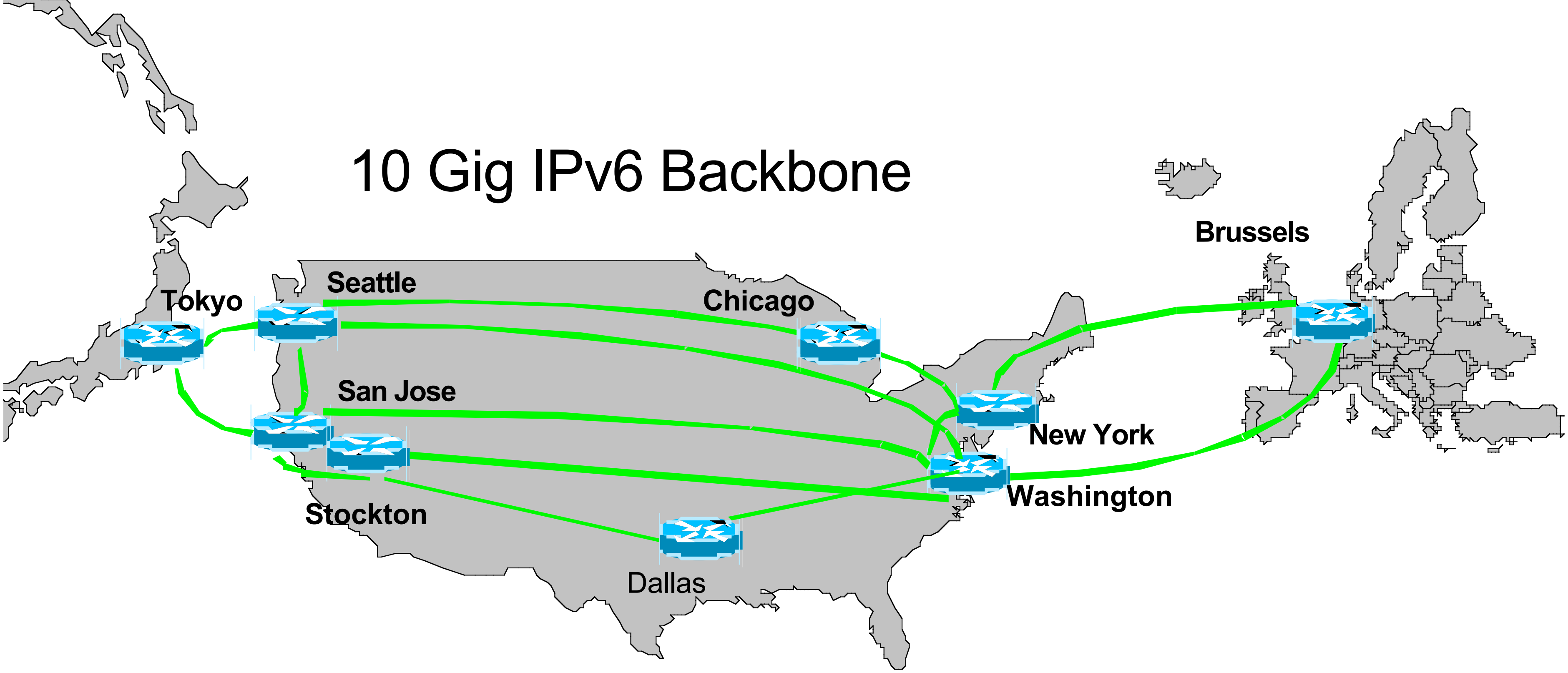
# BIND features

- query/response:
  - BIND servers have started retaining information on the transport protocol used by the query and will attempt to send a response on that same protocol.
- Response sort order:
  - More recent versions of BIND will attempt to sort the data when composing a reply to a query, preferencing the type of data based on the original query.
  - Example: If the query was over IPv6 and the answer has IPv4 and IPv6 data, the IPv6 data is returned first. If the query came in on IPv4 and the answer has IPv4 and IPv6 data, the IPv4 data is returned first.

# Example from Sprint From NANOG-31 meeting

# IPv6 Backbone Today

10 Gig IPv6 Backbone



# SprintLink IPv6 history

- 1997: Obtained 6bone address space (3ffe:2900::/24)
  - Original router under my desk ☺
- 1998: Totaling 15 customers using tunnels to 6bone
- 1999: Totaling 40 customers using tunnels to 6bone
  - Move router out to the network…
- 2000: Obtained ARIN space (2001:440::/35→ /32)
  - EOY 2000 Total: 110 customer using tunnels to 6bone.
- 2001-2002: Added 4 more IPv6 capable PoP's
  - Brussels, Washington DC, San Jose, New York
  - Member of the NY6IX exchange
  - Turning up customers at 2-3 per week.
- 26-May-2004: 300 IPv6 Tunneled Connections; 2 Native
  - Request frequency has slowed considerable (1 every week)

# Sprint IPv6 Offering

- Routers are IPv6 stand-alone boxes
  - No dynamic protocol-level interaction with the IPv4 network (SprintLink; AS1239).
- GRE tunneling is used between IPv6 routers
  - Over SprintLink (IPv4) infrastructure.
- iBGP full-mesh between AS6175
- ISIS runs as IGP
  - Looks eerily similar to SprintLink backbone, but with Hex addresses.

# Why We Did It This Way

- Router OS Dependencies:
  - Either features/code we need to run IPv4 do not support IPv6, or visa versa.
- Customer requirements:
  - Customer: “Yes, we require IPv6, IMMEDIATELY!”
  - Sprint: “Yes Sir! How much will you pay for IPv6?”
  - Customer: “Pay? No, we just *WANT* it. We don’t want to *PAY* for it”
- Overlay model removes Router software dependencies
  - Allows for more experimentation on the IPv6 side of things.
  - Allows us to deploy minimum capital boxes (or depreciated hardware) to support IPv6 for the price-point that customers require (to wit; $0.00/meg).
- Protocol is not ‘fully-cooked’ yet.

# What This Allows Us to Do

- Offering is free to any IPv4 customer of Sprint
  - Or any entity with a static tunnel endpoint (if you are nice to us).
    - This is what customers expect to pay for IPv6 today.
- Goal is to promote the usage of IPv6, within the confines of the current abilities of the protocol
  - We are very careful not to bend the rules (rfc2772)
    - This brings the “issues” with the current rules to light, and forces attention upon them
    - Multi-Homing, Micro-mobility, renumbering, etc…
- Today, Sprint uses 6Bone space for all customer numbering
  - We assume that ‘real’ ipv6 numbering schemas will evolve with the protocol, and do not want to get ourselves stuck

# Why is the IPv6 Network Tunneled??

```
sl-bb1v6-rly#sho int pos 0/0/0 | inc rate
  1 minute input rate 181000 bits/sec, 64 packets/sec
  1 minute output rate 189000 bits/sec, 65 packets/sec
```

```
sl-bb20-nyc#sho int pos 2/0 | inc rate
  1 minute input rate 3001474000 bits/sec, 1057309 packets/sec
  1 minute output rate 5028579000 bits/sec, 1303165 packets/sec
```

Time: 1930 EST, 10-Dec-2003

# Next Steps

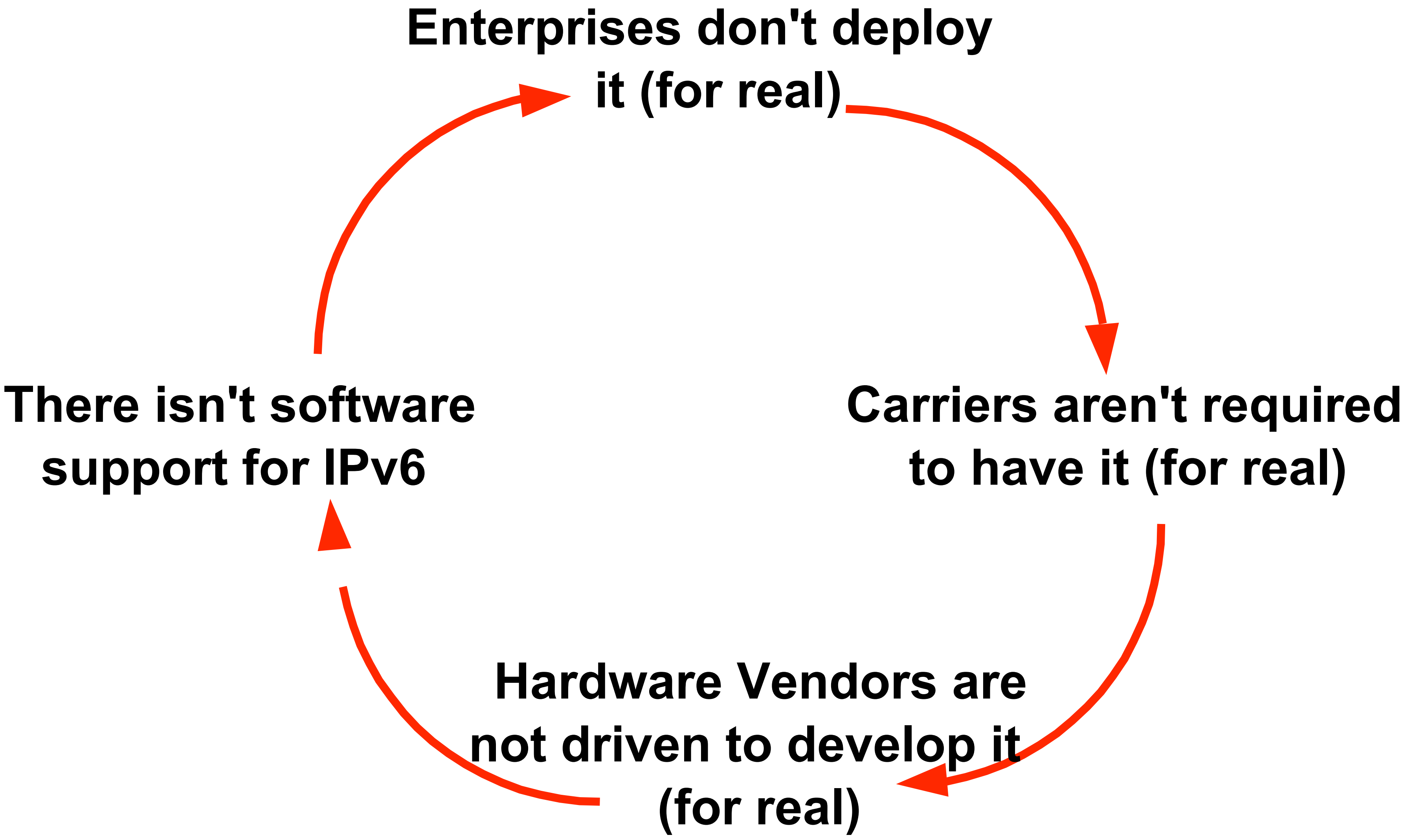
- Build the native transition plan:
  - Somewhat done
- Use Layer 2 protocol ID field to dynamically tunnel IPv6 packets to overlay network
  - Gives customers the native ‘feel’ without the native ‘pain’

# Why Does Sprint Not Lead the Charge to Migrate?

3 Main Reasons:

- We do what brings in money
  - IPv6 is a definite “check-box” on RFP’s, but it is not a stand-alone profit generator for Sprint (yet)
- The Router Vendors do not compel us to do so.
  - Router Vendor Software is not there yet. It has IPv6 forwarding, but misses the things that would entice enterprises
    - Next-headers, IPSec, Mobility extensions, etc..
- Multi-Homing is not solved yet.
  - If you do IPv6 now you might either:
    - Run out of Router memory (at scale)
    - Run out of Addresses (eventually)
  - Something has to give.
  - Solution MIGHT leave IPv6 looking VERY different than it does today.

# Vendors: the vicious circle



Enterprises don't deploy it (for real)

Carriers aren't required to have it (for real)

Hardware Vendors are not driven to develop it (for real)

There isn't software support for IPv6

# Background

- IPv6 data is distinct from IPv6 transport
  - It is possible -NOW- to publish IPv6 data in the DNS. This is potentially useful for environments that run dual-stack services on end systems.
- This is the major problem with deploying a new transport protocol like IPv6. For Example:
  - A resolver, with a single transport, queries for an address of an endsystem it would like to communicate with. The answer (if it gets one) may be an address on a non-supported transport.
  - In a mixed environment, without coordination, it is possible that the resolver is unable to reach any authoritative server. They may all be on the "other" transport.
  - These failure conditions are impediments to IPv6 adoption.
- Recent BIND specific augmentation does much to help mitigate the concerns.

# Issues

- Protocol specifications
- The ARIN example
- Middle Box
- End Systems

# Protocol Specification

- DNS has a defined size limit of 512 bytes.
- UDP fragmentation is operationally -BAD-
  - NAT boxes tend to drop UDP fragments
- The defined limit is 512 bytes !!!!
  - not IPv6 friendly :)
- HOW MANY SERVERS CAN BE DEFINED?
  - …before fragmentation occurs?

# TLD Recommendations

For a TLD and major branches directly defined in that zone

- Presume that the ARIN model (described later) is a “best fit” because:
  - They do not have a clean overlap between IPv4 and IPv6 services
  - Reduced impact on the production services
- Stand up authoritative servers on IPv6 where possible
  - This has as a precondition, the creation of a fully linked IPv6 transport capable mesh connecting the participating DNS services.
- Submit a formal request to ICANN/DoC for adding IPv6 glue for the TLD delegation

# Current Policy Recommendations for TLD operators

- RSSAC to ICANN
  - "Based on empirical testing, please proceed w/ TLD delegations at your earliest" http://www.rssac.org/rssac-v6tldglue
- IETF to TLDs
  - Mind the fragments… And here is a calculator to determine when fragmentation will occur. http://www.ietf.org/internet-drafts/draft-ietf-dnsop-respsize-01.txt

# The RSSAC Recommendation

On Fri, 12 Dec 2003 18:17:26 +0900

Jun Murai <jun@wide.ad.jp> wrote:

Dear ICANN board,

After considering input from experts including reports of relevant lab tests the committee recommends that IANA proceed with adding AAAA glue records to the delegations of those TLDs that request it. The committee does not foresee negative effects to overall DNS operations as a consequence of such additions.

<…>

Jun Murai, as the chairman of RSSAC

# The IETF Guidance

From dnsop-respsize IETF I-D

“With a mandated default minimum maximum message size of 512 octets, the DNS protocol presents some special problems for zones wishing to expose a moderate or high number of authority servers (NS RRs). This document explains the operational issues caused by, or related to this response size limit.”

# Integration "How-To"

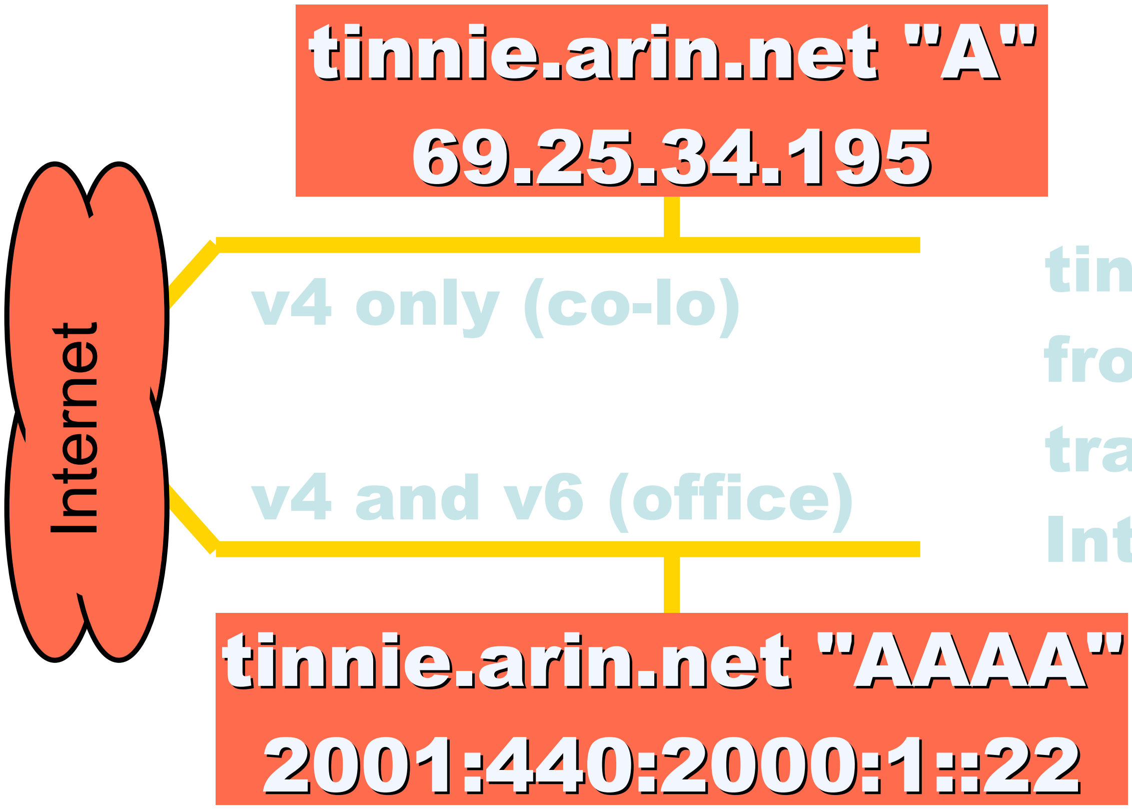
Depends on your current deployment:

- Do Services & Transport overlap in my environment?
- Is there a desire to minimize impact on production services?

# Experiences from ARIN – The ARIN Deployment Model

- The IPv6 transport does not match the IPv4 transport
- IETF recommended dual stack service is not practical
- ARIN has a 24/7/365 production requirement
- IPv6 capability is added to the normal hardware lifecycle

# One Name, Two Machines



tinnie.arin.net "A"
69.25.34.195

v4 only (co-lo)

v4 and v6 (office)

tinnie.arin.net "AAAA"
2001:440:2000:1::22

tinnie "AAAA" slaves from tinnie "A" via v4, transparent to the Internet at large

# Non-dual Stack DNS

- Running non-dual stack servers for a zone on v4 and v6 can be done two ways
  - Having the servers have an A “x” or AAAA record
  - Using one server name on two machines
- BIND seeks A and AAAA for all NS names
  - Recommendation to use "one name, two machines”
    - The production requirements for IPv4 capable service do not allow a single name, single machine
    - IPv6 users should be presented the same environment to the extent possible
    - It is impossible to predict the capability of any given community at a given time - so the names should remain the same

# A Specific Issue

- The "other" v6 service ARIN runs, SSH:
- # ssh tinnie.arin.net
  - AAAA is preferred over A
  - If you wanted to reach tinnie A this would fail
- They once did a "tail -f log" on the wrong host
  - Trying to debug why wasn't an event being logged?
  - Good thing it wasn't an "rm" command
- Otherwise, acceptable but sub-optimal

# The Issue – Generalized

- If the "A" server is running other services that can't be brought to v6
  - Separate the services physically, or
  - Separate the services via domain names
- ARIN separated by purchasing a new server
  - Newer hardware was brought online as part of the lifecycle process

# ARIN Summary

- Modern equipment is IPv6 capable
- IPv6 transport from commercial vendors is sporadic
- IPv6 can be deployed without impacting production services
- IPv6 users do not have to be perceived as marginalized
  - All the services are available to ARIN members, regardless of transport

## Recommendations

- Use latest acceptable versions of software
- Use the same physical media for IPv4 and IPv6
- Get in early, while the bandwidth is easy to handle and grow with it

# Coordination with Others

Background and detail from other sectors

- V6 is important to DOC:: http://www.ntia.doc.gov/ntiahome/press/2004/IPv6_01152004.htm
- Concern about stability:

  " Given the Department's interest in IPv6, and more importantly, in the continued smooth operation and stability of the …<dns>…, we want to see a full-blown technical proposal … that includes … what steps would be taken to protect the smooth operation of … <the dns>…"

  Kathy Smith, NTIA

  Communication to Educause on the application for

  adding IPv6 support to .EDU
- Hence documented procedures for adding IPv6 support in the root zone for TLDs had to be defined. ICANN has that burden.

# ICANN status

- ICANN procedural guidelines for public comment
  http://www.iana.org/procedures/comments.html
- ICANN procedures have been approved as of 13 July 2004 and are being implemented
- The backlog of requests is being processed as they meet the normal criteria that are laid out in their proposals
  - Most should be processed within weeks of being released
- We then move on to native v6 support for the root servers
  - May take another 6-9 months of work

# Technical and Operational Documents

Technical and Operational Documents that support proposals meeting NTIA criteria

- DNS Response Size Issues
  http://www.ietf.org/internet-drafts/draft-ietf-dnsop-respsize-01.txt
- RSSAC recommendation for IANA process on AAAA glue
  http://www.rssac.org/rssac-v6tldglue.html
- Commerce Department Task Force Requests Comments on Benefits and Costs of Transition to New Internet Protocol
  http://www.ntia.doc.gov/ntiahome/press/2004/IPv6_01152004.htm
- DNS IPv6 transport operational guidelines
  http://www.ietf.org/internet-drafts/draft-ietf-dnsop-ipv6-transport-guidelines-02.txt
- IANA Administrative Procedure for Root Zone Name Server Delegation and Glue Data
  http://www.iana.org/procedures/comments.html

# Servers in the Context of the Overall DNS

- DNS service presumes a common namespace across every useable transport protocol
  - The original DNS design presumed a single transport protocol - IPv4
- DNS service is a cooperative engagement between the servers and the end-systems
  - May be impacted by devices and services in the infrastructure
- The servers and end-systems ability to comprehend and adjust to a common namespace in two distinct transport domains in jeopardy without proper planning and execution.

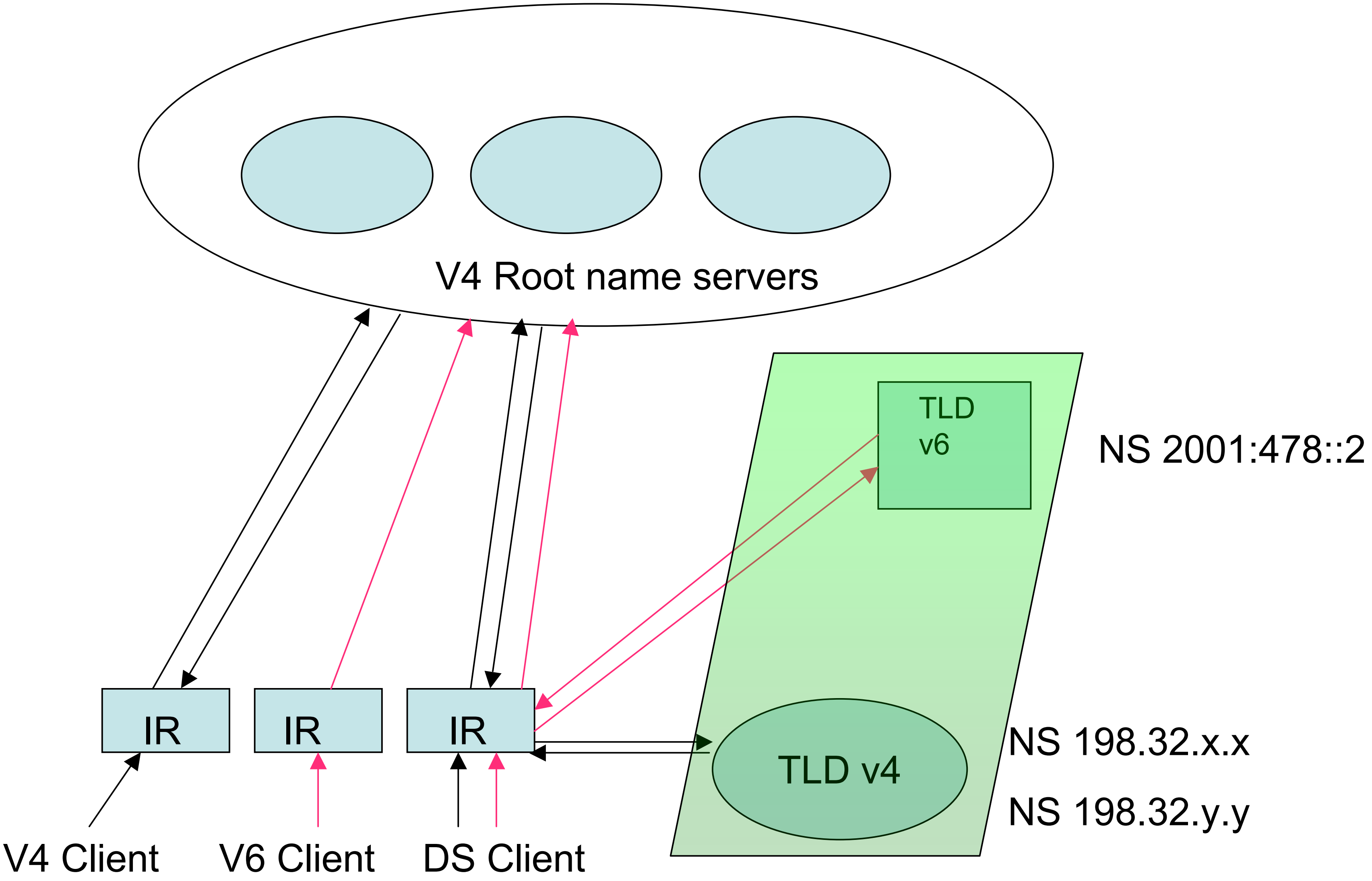
# Other Parts of the System

- The “infrastructure”
  - “Middle-Box”, Proxies, and NATs
  - “Hijacking” the request & response - fabricate something that they think “might” be wanted.
  - Bridging between transports
- The resolver(s)
  - May not be a single “resolver” - some applications have their own
  - Based on OS capabilities
- Lifecycle - what is the replacement cycle for hardware/software/applications?

# Root Server Considerations

- Mostly about namespace fragmentation
  - How to present a consistant namespace over multiple transports
  - What about (future) v6-only areas?
  - Consequences of v6-only servers

# Tying this together


V4 Root name servers
TLD
v6
NS 2001:478::2
IR
IR
IR
TLD v4
NS 198.32.x.x
NS 198.32.y.y
V4 Client
V6 Client
DS Client

# The Generic Recommendation

The best we can do is:

- Have authoritative servers for every zone available over all transports
  - Maintain a single namespace and coherency for endusers
  - Dual stack the service, not the machines
- Make iterative mode  resolvers dual-stack
- Run current software on servers
- Accelerate the lifecycle process to bring onboard new gear that is IPv6 capable as quickly as possible.

# End of Presentation