IPv6 Deployment - Facts
Keys to Deploying IPv6 Successfully

- **Facts:**
  - – **Millions of nodes are running IPv4 today**
  - – **Some nodes will never upgrade to IPv6**
    - » **Large investment in IPv4 applications**
- **Consequences:**
  - – **IPv4 and IPv6 will coexist for an extended period**
  - – **Transition should prevent isolation of IPv4 nodes**
  - – **Transition must consider security ramifications.**

- **No disruption - no Flag Day**
  - – **IPv6 and IPv4 routers and hosts can interoperate**
- **No Dependencies - Incremental upgrade and deployment**
  - – **IPv6 routers and hosts can be deployed in a highly diffused and incremental fashion**
- **Low start-up costs**
  - – **Make transition as easy as possible for end-users, system administrators, and network operators**

Transition Mechanisms

- **Many solutions to deliver IPv6 packets**
  - **One size does not fit all**
- **Basic transition tools**
  - **Tunnels**
  - **Translation**
- **IPv4 and IPv6 can share same physical infrastructure**
  - **Coexist in the box and on the wire**

IPv6 transition - transit

- **Too many “transition” methods**
  - **6to4**
  - **teredo**
  - **isatap**
  - **et.al.**
- **Should select a subset**
  - **For DNS services - a single choice will ensure the smoothest transition**
  - **some transition methods will redirect packets over non-optimal paths**
  - **transition methods left in place will be perceived as an impediment to native IPv6 adoption**
- **None of these directly impact DNS from the server perspective, they do impact user perception of DNS availability**

2.0 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) is 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.**

2.1 Prior Work

- **Test beds**
  - **EP.NET - http://www.rs.net/**
  - **WIDE/ISI/ISC work**
- **Early Adopters**
  - **TLDs from all regions**

2.1.3 WIDE/ISI/ISC work on IPv6 impact at the resolver.

- **Akira Kato wrote the following internet draft while working as a graduate student at ISI and then at ISC.**
  - **www.ietf.org/internet-drafts/draft-ietf-dnsop-respsize-01.txt**
- **1.1. The DNS standard (see [RFC1035 4.2.1]) limits message size to 512 octets. Even though this limitation was due to the required minimum UDP reassembly limit for IPv4, it is a hard DNS protocol limit and is not implicitly relaxed by changes in transport, for example to IPv6.**
- **1.2. The EDNS0 standard (see [RFC2671 2.3, 4.5]) permits larger responses by mutual agreement of the requestor and responder. However, deployment of EDNS0 cannot be expected to reach every Internet resolver in the short or medium term. The 512 octet message size limit remains in practical effect at this time.**

2.1.4 Early adopters

- **participants in the RS test bed**
  - **JP, KR, CN, NL, DE, SE, MIL, INT**
- **some on the ICANN adopters for TLD use**
  - **JP, FR, KR**
  - **All are running production DNS service for their TLDs on IPv6 transport**

3.0 Issues

- **Protocol specifications**
  - **DNS**
  - **IPv6**
- **Server considerations**
  - **OS implementation of IPv6**
- **The ARIN example**
- **Middle Box**
- **End Systems**

3.1 Protocol

- **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 I DEFINE?**
  - **…before fragmentation occurs?**

Does Size Matter?

will these “glue” records fit in a single 512byte UDP packet?

what happens when AAAA are added?

```
lump.              in     ns                  foo.
                                              bar.
                                              joy.
                                              zen.
                                              delta.somewhere.else.
                                              pdc.
foo.          in a            192.0.2.53
                              192.168.2.53
                              10.10.0.42
bar.          in a            172.16.7.77
                              192.0.2.42
                              192.168.255.8
                              172.17.17.3
joy.          in a            300.44.44.44
zen.          in a            299.5.5.53
              in a            298.7.6.5
              in a            297.6.5.4
delta.somewhere.else. in a  410.9.8.7
                              192.168.33.33
                              127.53.6.6
pdc.          in a            198.32.64.12
```

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

"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."

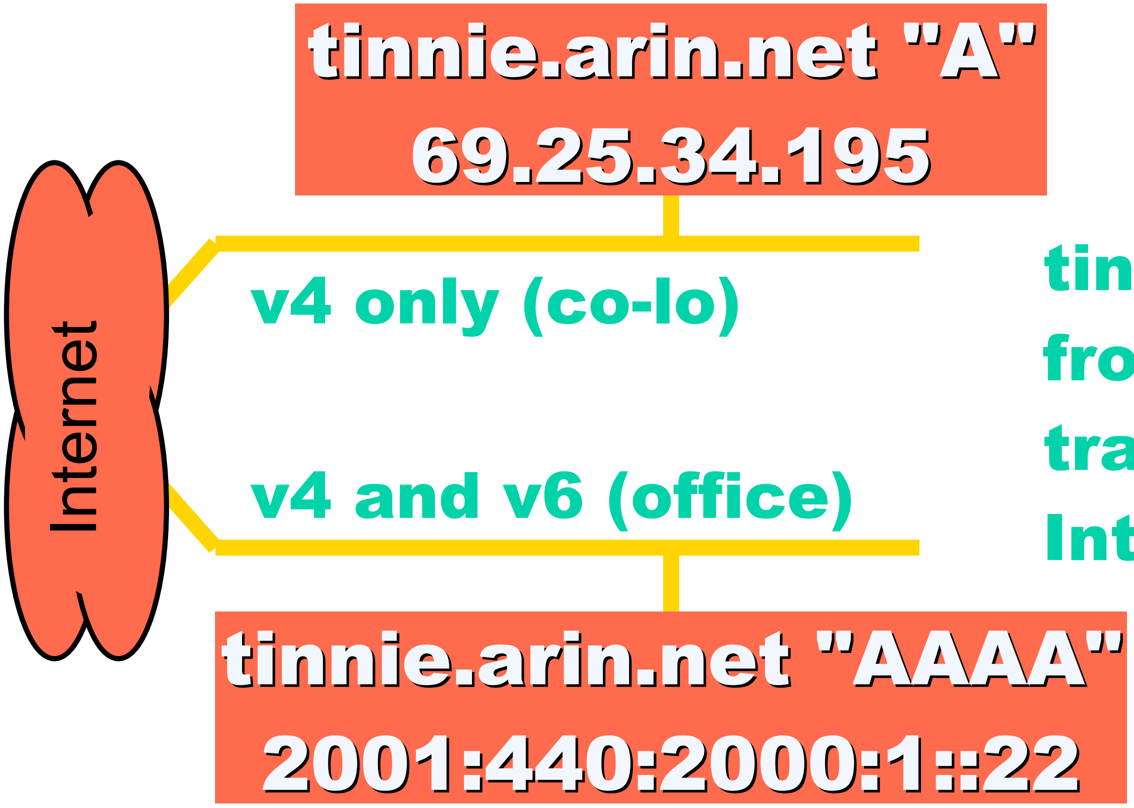
4.2 Integration “How-To”

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

4.2.0 Experiences from ARIN - The ARIN deployment model.

- **The IPv6 transport does not match the IPv4 transport.**
- **IETF recommended dual stack service is not practical.**
- **They have a 24/7/365 production requirements**
- **IPv6 capability is added to the normal hardware lifecycle**

4.2.1 One name, two machines



Internet

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

4.2.2 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

4.2.3 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

4.2.4 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**

4.2.5 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**

4.2.6 Other Implementation choices

- Ø **Dual stack**
  - Ø **presumes that all systems and services have working IPv4 and IPv6 protocol stacks.**
  - Ø **the assumption that there will be consistent, homogeneous dual-stack support throughout an organization**
  - Ø **Untested application interaction when presented with dual-stack operating systems.**
    - Ø **See the ARIN example above, with the distinction between SSH and Bind behavior**
  - Ø **places unwanted pressure on production systems to adopt IPv6 capability.**

4.3 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**

- **Concerned 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.**

4.3.1 ICANN status

- ICANN procedural guidelines for public comment. http://www.iana.org/procedures/comments.html
- ICANN procedures have been approved as of 13july2004 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.

5.0 Technical and operational documents that support proposals meeting NTIA criteria

- http://www.ietf.org/internet-drafts/draft-ietf-dnsop-respsize-01.txt
- http://www.rssac.org/rssac-v6tldglue
- http://www.ntia.doc.gov/ntiahome/press/2004/IPv6_01152004.htm
- http://www.ietf.org/internet-drafts /draft-ietf-dnsop-ipv6-transport-guidelines-02.txt
- http://www.iana.org/procedures/comments.html

6.0 Servers in the context of the overall DNS

- **DNS service presumes a common namespace across every useable transport protocol**
  - **The original DNS design presumes 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.**

6.1 Other parts of the System

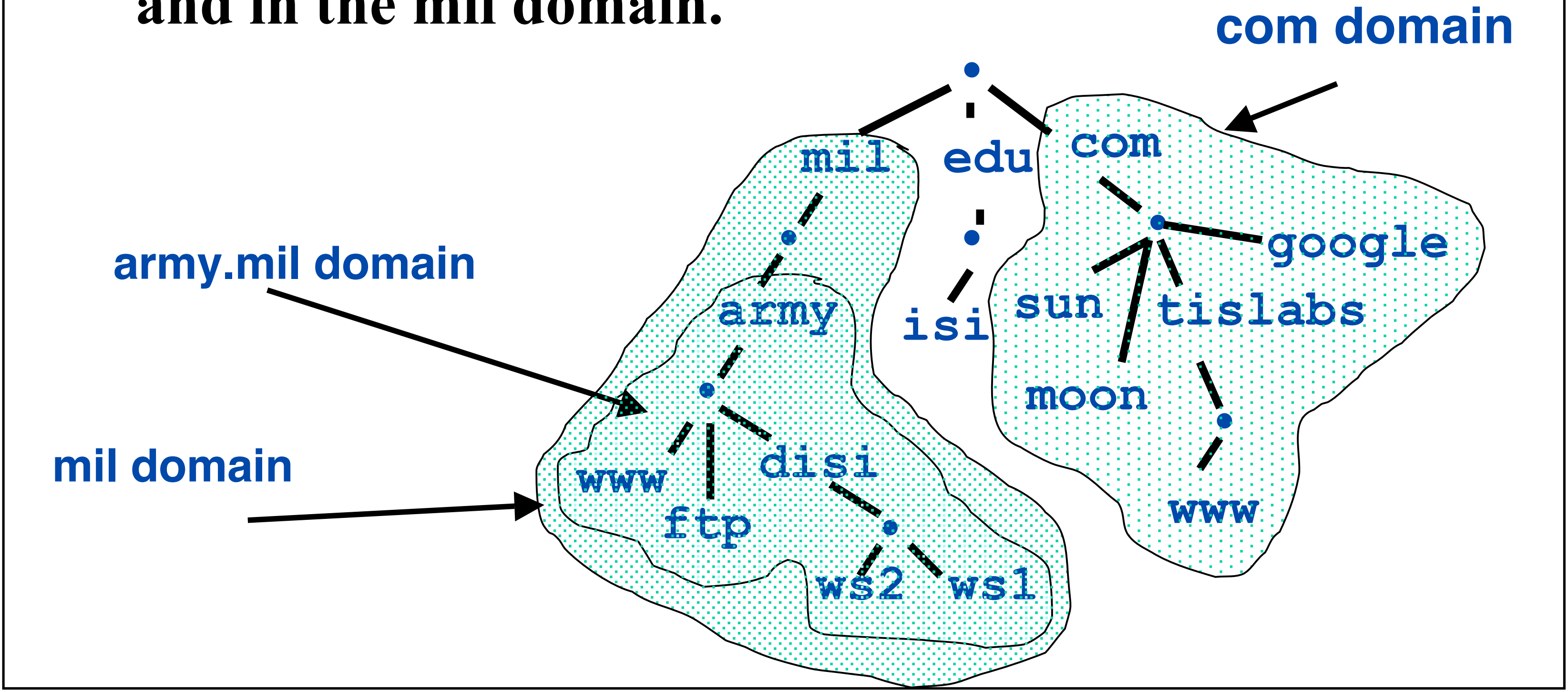
- **The “infrastructure”**
  - **“Middle-Box” , Proxies, and NATs**
  - **“hijacking” the request & response - fabricate something that they think “might” be wanted.**
  - **Bridging between transports**
- **The resolver(s)**
  - **may not be a single “resolver” - some applications have their own**
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

6.2 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**

7.0 Namespace & Delegations

- **Domains are “namespaces”**
- **Everything below .com is in the com domain.**
- **Everything below army.mil is in the army.mil domain and in the mil domain.**



7.0 The DNS database

- **”DNS is a distributed, autonomous, coherent, reliable database.” - Paul Vixie**
- **This defines the dns namespace. The presumption is that the namespace is visable over the transport protocol(s) in use**
- **The autonomy follows from the process of delegation. The parent gives up** responsibility **for the contents of the subdomain at the time of delegation.**
  - **among the responsibilities that accrue to the delegee is their selection of servers**
    - **this includes the selection of available transport capabilities**
- **These attributes are applicable at each delegation point from the root down.**

7.1 ”Distributed”

- **At the time of delegation, the subdomain manager is required to select a suite of servers to act as athoritative sources for the delegation.**

- **Most of these servers need to be visable to the rest of the Internet, so that anyone may resolve names in the delegated space.**

7.2 ”Coherency”

- **Coherence require that the answer to a query** must **be the same regardless of how the answer was obtained.**
  - **The Internet has a long tradition of being able to cope with outages, failures and brokenness.**
  - **It is** much **harder to try to cope with lying.**

7.3 ”Autonomy”

- **We will have to live with whatever solution we decide on for several decades.**
  - **If during that time the zone owners don't realize (or agree) to deploy their zones on IPv6 the namespace fragmentation problems will persist.**
  - **If during that time, application developers do not upgrade to IPv6 aware resolvers, we will have deployed DNS over IPv6 for no purpose.**
  - **Ultimately** each zone manager and application developer **are responsible for native IPv6 support.**

7.4 Transport & namespace visablity

- **Lookups for v6 data works today, since the entire tree is available in v4 and all resolvers knows v4.**
- **Lookups would continue to work tomorrow if only all the servers and resolvers stayed ”virtually” dual-stack.**
  - **I'm only talking about** full service **resolvers, not stub resolvers, nor forwarding caches.**
  - **This will not help v4 if zones migrate to v6 only.**

7.4.1 Balkanization / Bridging

- **When a zone administrator makes a determination that certain DNS data will only be made visable or available over a single transport protocol.**
  - **Such data ”disapears” from the visable namespace in other transports**
- **Bridging is an intermediary that attempts to map/remap queries from one tranport to another.**
  - **inability to discriminate when mapping is -NOT- required.**
  - **may be deployed anywhere a v4/v6 transit border crossing exists**
    - **» tracking transport bridge deployment is roughly identical to tracking NAT deployment.**

7.4.2 Balkanization vs NAT/bridging

- **”Balkanization” (i.e. uneven zone availability) is mostly unknown territory.**
  - **it may create new failure modes for apps and services**
  - **it may drive deployment of zones over v6 transport**
  - **it will cause confusion**
- **Bridging is a known evil. It has a number of drawbacks and will only work to a certain extent.**
  - **but to that extent it does keep the namespace together**

7.4.5 Disadvantages to a "working" bridging solution

- Without bridging, deployment of zones over v6 becomes necessary. We believe this is the desired goal.
- With bridging, deployment may never happen and bridging will have to stay forever (working progressively worse as the v6 part grows).

One way solutions?

- Bridging from a v6 client to a v4 server is easier than the opposite way around.
- This could be taken as an argument that by keeping the entire namespace available over v4 transport everything is fine.
  - That is a false argument, since at some (remote) point the assumption of v4 prevalence will be false.
  - An assumption that everything is always available over v4 transport will eventually be false.

Where to solve problems with a single namespace and multiple transports?

- **Close to the resolver (i.e. client)?**
- **Close to the server?**
- **On the v4/v6 border (as a network wide "service")?**
- **The cost of maintaining bridging should somehow (but how?) end up at the doorstep of the zone owners (to create incentive to deploy).**
  - **This does not happen, creating chokepoints that function as attractive targets.**

The generic recommendation

- **The best we can do is:**
  - **have authoritative servers for every zone available over all transports**
    - » **maintain a single namespace - coherency for endusers**
  - **make full service resolvers v irtually dual-stack**
- **run current software on servers**
- **accelerate the lifecycle process to bring onboard new gear that is IPv6 capable as quickly as possible.**
- **This limits support scenarios to easy FAQs**
  - **rather than dependancy on transport bridging.**

FIN