IANA IPV6 Workshop
A View From the Trenches
(Some thoughts on IPV6 deployment in enterprise and ISP networks)

Joel Jaeggli
Nokia

This talk is focused on the issues faced by network operators in deploying IPV6

Most of these observations are opinions not facts.

Please debate me!

Agenda

- Understanding the current environment
  - Scary public policy issues
  - Liabilities
- Transition (or lack thereof)
- Deployment (making your network safe for IPV6)
- Deployment Conclusions
- Surprises

Understanding The Current Situation

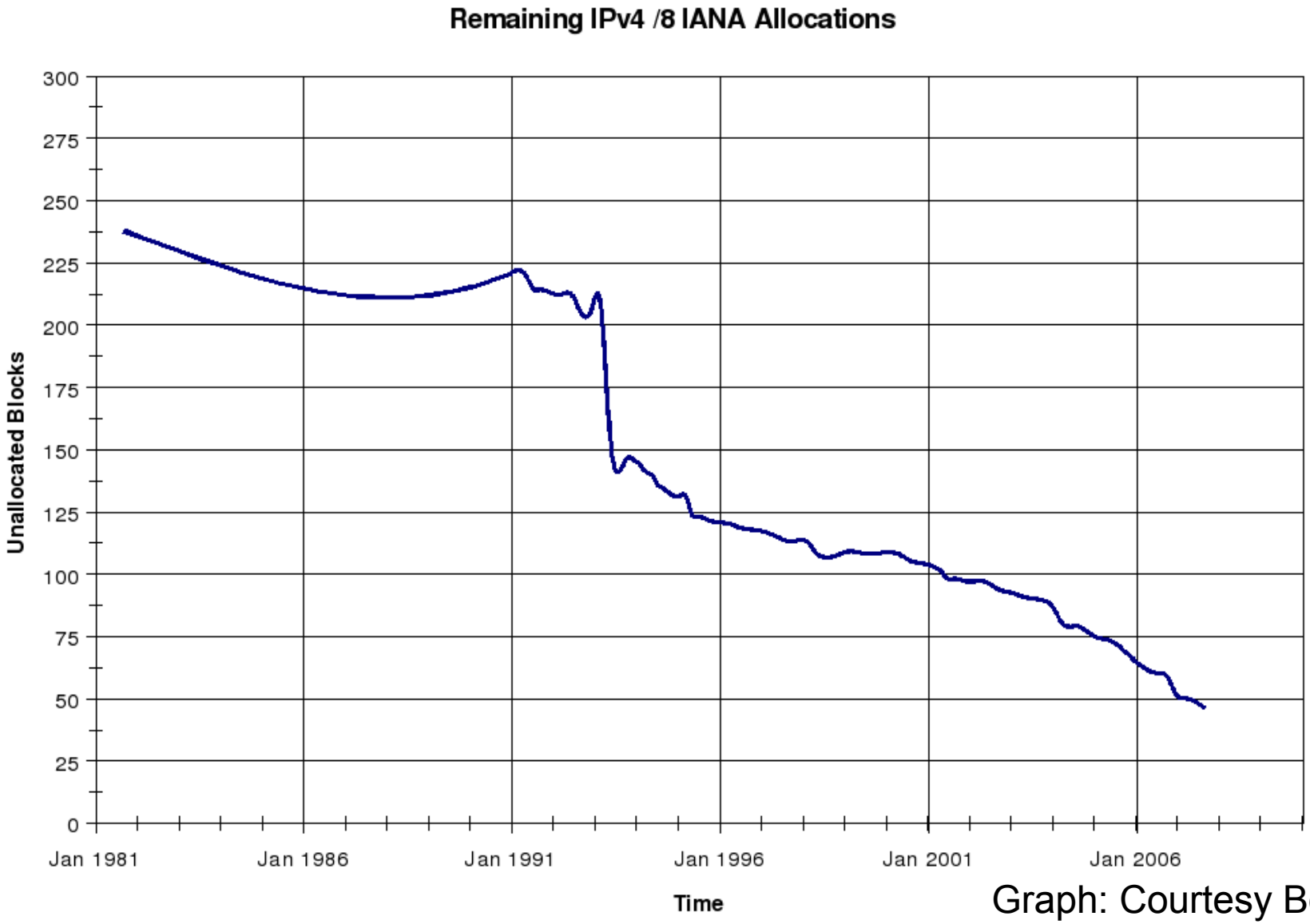
- Predicting the moment of IPV4 exhaustion has become a popular spectator sport.
- Doesn't much matter when, if or how much. Any business that consumes IP addresses in the process of growing has a problem.
- Shareholders and customers will likely be unhappy when told that this “new” liability affects the ability to grow the business.
- IPV6 is not the only solution...

One Alternative

Image attributed to Richard Edden

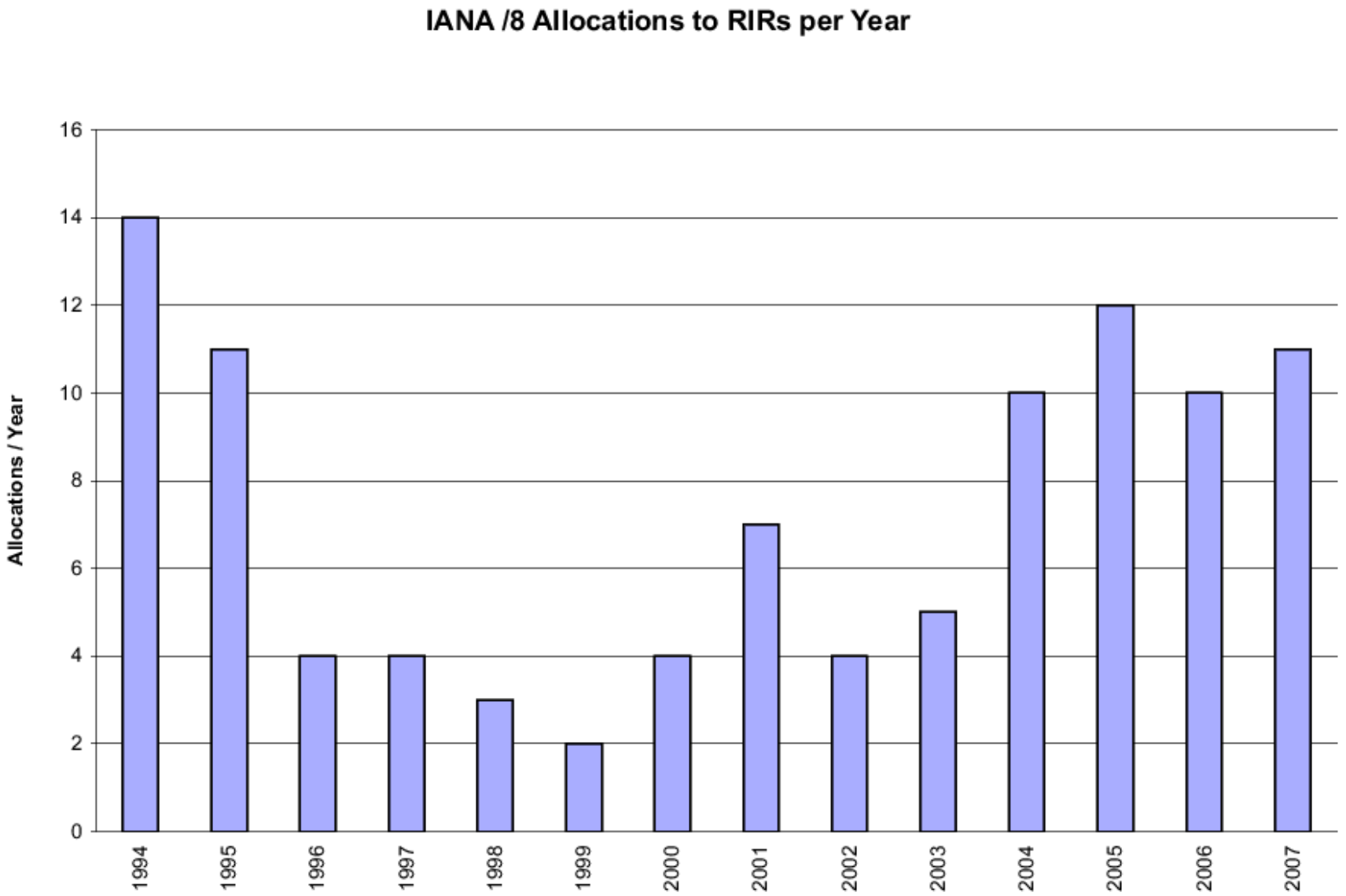
The end of the (ipv4) world is nigher!

- Geoff Huston July 2007



Graph: Courtesy Bob Hinden

The RIR's are doing their job



Graph: Courtesy Bob Hinden

The provocative but boring (and irrelevant) statement.

We actually ran out in 1992!

Scary Public Policy Debate...

- If you're new to this situation, the best place to learn about is not the various RIR public policy mailing lists.
  - Acrimonious debates are producing more heat than light.
  - The fighting, is over the remains of the corpse.
  - Unable to divine future direction if any from content of the mailing lists.

So...

- Don't throw hands up in disgust because there are 70 messages titled “`Re: [ppml] [address-policy-wg] Those pesky ULAs again`” or “`Re: [ppml] Policy Proposal: Global Policy for the Allocation of the Remaining IPv4 Address Space`”
- Focus on the things you can do.
- Secure the resources you need for the business to grow and prosper.

Liabilities

- Inability to secure additional ipv4 addresses due to exhaustion.
- Changes to RIR policy pushing the date at which securing new addresses becomes harder closer to the present day.
- Widespread IPV6 deployment never occurs and IPV4 is what we're stuck with.

Transition (or lack thereof)

- Dual-Stack deployment is not going to slow the Consumption of IPV4 Addresses.
  - The fact of the matter is that more devices will have to share proportionally fewer ipv4 addresses.
  - That means NAT, multiple NAT layers, NAT boxes with the same addresses in use on both sides.
- The “killer app” for IPV6 is 96 more bits.

Killer App

- Increased use of NAT and shorter leases seems inevitable in growing IPV4 networks.
  - Green field deployments get more challenging when large amounts of V4 cannot be acquired.
- IPV6 addresses are available in sufficient quantity to produce stable bindings for as long as a host needs a particular address.
- Peer to Peer applications (not just file sharing mind you) benefit from end-to-end connectivity.
  - IPV6 can preserve, if desired end-to-end reachability.

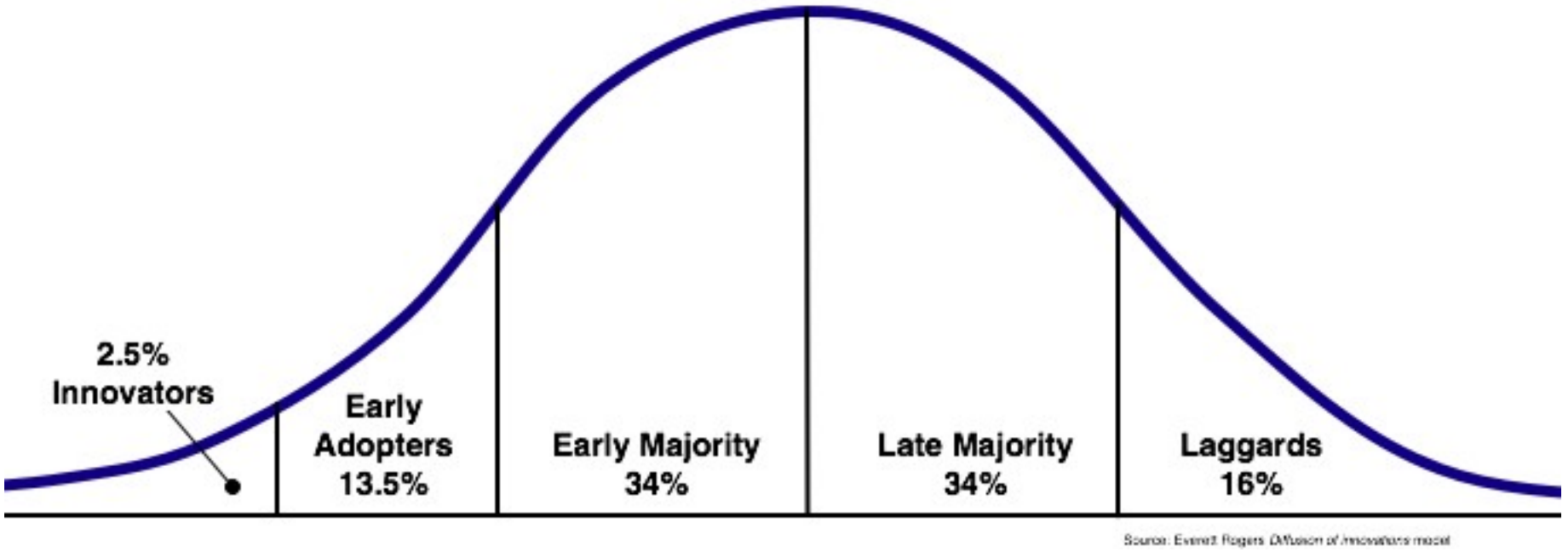
First Order of Business

- Continue to fly the airplane...
  - For ISP's and Enterprise Operators that are growing their operations that means maintaining a supply of IPV4 addresses based on RIR guidelines.
  - It is widely presumed, though not inevitable that address allocation policy will change sometime between now and the last allocation of a /8 from IANA to the RIR's

Making the Network Safe for IPV6

- A substantial amount of the histrionics the specter of V6 deployment is engendering on Operator and standards working group lists is the product of experience gained through operation of the 6BONE and early IPV6 networks.
- There are people willing to suggest throwing all the V6 work out and starting over from scratch.
- I see few reasons other than sullen inertia that this foot-dragging should continue.
- Everett Rogers diffusion of adoption model is applicable.

Everett Rodgers Diffusion of Innovation



Major Complaint
Consider the following Case
(DNS lookup leads you down the garden path)

- Host A wants to connect to host B
- Host A's resolver looks up B and gets back two resource records
  - AAAA
  - A
- Host A believing itself to have IPV6 connectivity attempts to contact host B using IPV6

Garden Path 2

- No IPV6 path between A and B exists
- A waits for the IPV6 connection attempt to fail before trying IPV4
- For some reason this is considered disastrous.
  - Reflects the early nature of existing IPV6 deployments
  - RFC 4943 - IPv6 Neighbor Discovery On-Link Assumption Considered Harmful

IPV4 Case

- Host A looks up B
- Gets:
  - A record
- If host A cannot reach host B?
  - We assume, networking connectivity issue, host down, administrative boundary (SPI Firewall), private address space leakage, etc.

Garden Path Lesson...

- IPV6 (like IPV4) requires a path between to hosts (obvious)
- Network and service design must ASSUME that if proffered, IPV6 only hosts (when they exist) and dual stack hosts will use V6. IPV4 only host will use IPV4.
  - Do not put AAAA records in the DNS for hosts not providing IPV6 services
- Network Operators (that's us) focus on the delivery of reliable transport, do that and this problem fades into the background.

Safe and Sane Deployment Plan

- You can experiment all you want, but experimentation only gets you so far.
- The Plan:
  - Secure resources
  - Focus on the core
  - Transit and peers
  - Eat your own dogfood
  - Early services
  - Deployment to the edge
  - Applications

Secure Resources

- Most large enterprises, large content providers and virtually all ISP networks are going to qualify for address space under existing RIR rules.
- In the ARIN case for example, either:
  - As an IPV6 LIR (6.5.1)
  - Under existing ipv4 number policy (ARIN 6.5.8.1)
- No reason to experiment with PA space if it's not going to be suitable for deployment.

Focus on The Core

- Vendor support for dual stack varies. If you made it a requirement in your last upgrade cycle you're probably well enough off for now.
- Some deployments have chosen alternate approaches example 6PE over an MPLS
  - Not a good excuse in itself to convert you core to MPLS.
- Congruence of IPV4 and IPV6 deployment is desirable. Was not possible in some early deployments.
  - Keeps backbone engineers and IGP sane, though neither is a strict requirement.

Focus on The Core 2

- I really like /64s for point to point links... Fewer typos because all your subnets are the same size. You can use smaller for example /126 or /112, but to what end?
- Get the IPV6 network into the NMS.
  - I've made the mistake more than once of assuming V6 is fine because the routers are up and talking on the V4 side which was already monitored.

Focus on The Core 2

- I really like /64s for point to point links... Fewer typos because all your subnets are the same size. You can use smaller for example /126 or /112, but to what end?
- Get the IPV6 network into the NMS.
  - I've made the mistake more than once of assuming V6 is fine because the routers are up and talking on the V4 side which was already monitored.

Focus on the Core 3

- Come up with a rational address allocation plan.
- if you have a /32 some large portion of it should be reserved for future use.
  - In the nokia.net deployment two /36 address blocks are in use presently

Transit and Peers

- Tunnel brokers are fine for experimentation but not for real work.
- Treat IPV6 like V4.
- Most large transit providers offer it even if sales rep isn't familiar with the offering.
- Typically delivered as part of the same service, buying transit applies equally to V4 and V6

Transit and Peers 2

- Most providers surveyed in North America filter on ARIN guidelines so a PI /48 should work.
  - Announcing /48s out of your /32 is likely to be frowned on by your peers.
- Multihome!
- Put PTR records for IPV6 routers in the DNS Immediately. IPV6 trace-routes with only addresses are hard to debug.

Eat your own dogfood

- Put your engineering staff on dual stack enabled hosts on dual stack enabled networks.
- Enable IPV6 on your management networks, recognizing that only dual stack supporting devices should get RRs.
- Facilities that people use every day become familiar and it becomes obvious when there are problems.

Eat your own dogfood 2

- Early adopter customers (non-retail) are likely to be more tolerant of outages.
  - Some of them can be turned up at this stage.
  - Becomes a value add, especially for customers that are already looking for IPV6 transit.

Early Services

- Once the network is a functional transport for IPV6 It's time to look at deploying services that the leverage and support the deployment and build operational experience.
  - DHCP V6
    - Name server discovery
  - Resolvers that answer queries over ipv6
  - NTP
  - SMTP relays (retail ISP)
    - Remember the PTR records!

Deployment to the Edge

- Situational Dependence. Experience will be different for
  - Wholesale transit and business ISP
  - Consumer broadband deployment
  - Enterprise network deployment

Wholesale transit and commercial service provider

- For Customers with a single upstream and PA (from you) space. /64 point to point link and a /48 or possibly less for some applications example /56 as next hop.
- Modern commercial and even relatively old CPE, including small Cisco and Junipers is capable.

Wholesale and commercial part 2

- Multihomed customers using PA space should be encouraged to consider securing PI V6 space assuming they qualify under ARIN rules.
- Multi-homed customers using PI space should secure the resources that they need.
- Size of the IPV6 routing table is not likely to motivate upgrades of CPE for these customers.

Consumer Broadband

- Cable CMTS line cards will require upgrade.
- Many DSL aggregation routers used in traditional US DSL deployment don't support V6 without forklift upgrade.
- CPE can be replaced incrementally
  - The era when a customer has one IP address and the rest of the devices in the home are behind NAT should be over

Consumer Broadband 2

- Some debate of the size of prefixes to hand to residential customers.
  - ARIN guidelines suggest “/64 when it is known that one and only one subnet is needed”
  - This virtually guarantees the use of NAT due to consumer devices needing a layer 3 boundary to insure congruent V4 and V6 policy (see “apple airport extreme could expose macs via IPV6” discussion for example)

Consumer Broadband 3

  - Residential customers probably need more than a /64. BCP hasn't gelled on this yet /60 /56 /48 have been kicked around.
- DHCP-PD can be used to assign prefixes to customers (part of RFC 3633).
  - requires support in CPE
  - probably want a stable assignment mechanism rather than dynamic, so that customers aren't constantly renumbering their internet networks.
  - Triple play providers have their own ideas since they have multiple devices to manage for the customer

Enterprise Deployment

- Enterprises make extensive use of private address space for a number of reasons.
  - likely want to preserve that functionality
- Enterprises especially large one's have proportionally, lots of devices to manage.
- Stable bindings for devices make a heck of a lot of sense in this context.
- Enterprise mobility applications create bindings for devices outside the company in the Enterprise network (VPNS)

Enterprise Deployment 2

- No point in using private address space in a large enterprise when you don't have to.
  - Announce covering prefix. Null-route the prefixes you're only using internally on your border.
    - More unique than ULA-L
    - If you leak it someplace you should end-up black holing your own traffic (A good thing in this context).
- When you have to, ULA-L is a sufficient replacement for RFC 1918.

Enterprise Deployment 3

- Enterprise may wish to disguise topology when exposing hosts and services to the outside world (this is a perceived benefit of NAT)
  - RFC 4864 details some scenarios for how local network topology can be obscured without breaking end to end reachability (assuming maintaining it is the goal)
    - I wouldn't recommend deploying MIP6 just to support this scenario.
    - l2tp or IP-in-IP is probably a better choice.

Applications

- Large content providers are probably the last people to put AAAA records up for their services. The possible consequences of losing customers with incomplete IPV6 network connectivity is too high. That's ok. In general their address space needs are finite.
- Peer-to-Peer applications, for example VOIP would love to be able to assume that target can be reached without the use of a proxy which is know to have a public IP address.

Applications 2

- The enterprise case has different issues. New applications will likely be deployed IPV6 Enabled.
- Some applications will never migrate (before they are turned off) consider the case with IPX, Decnet, SNA transport on enterprise networks
  - Doesn't matter it's their sandbox.
- In some cases ALG's will be crafted, in other's applications will migrate.

Conclusions

- IPV6 rollout is not a transition.
- IPV6 addresses can be used to support end-to-end reachability. The same property continues to erode in IPV4.
- Scarcity in IPV4 (which as been present for virtually all of the life of the commercial Internet) affects our perception of how we should Allocate and use IPV6.

Conclusions 2

- If IPV6 lasts as long or longer than IPV4 then we should consider it a resounding success.
- It's presumptuous to assume we solved the address space issue for all time. Solutions were designed by humans and inevitably entail compromise.
- We didn't address routing scalability at all except as a consequence of better aggregation.

NAT-PT

- It's back!
- It's likely at some point that IPV6 hosts will be speaking to IPV4 hosts using such a facility at some point.
- RFC 2766 (original)
- RFC 4966 (moved NAT-PT to historic)

Surprise!

- NAT
  - It's not going away.
  - If a mechanism isn't provided to delegate a prefix to end devices, they will NAT V6 in order preserve the layer-3 boundary that they have in IPV4
  - The ability (and willingness) to hand out and route /64s to devices with /128s would side-step this issue. The hardware hasn't be built yet (for the most part) so there's still time to put DHCP-PD in it.

Surprise 2

- Teredo!
  - It'll tunnel IPV6 to your customers even if you aren't providing them services
  - Enterprises will just knock it down (port 3544 UDP)
  - Should shut itself off when IPV6 becomes available locally
  - Yay Windows Live Services...

Surprise 3

- ULA – Unique local addresses.
  - ULA RFC 4193
  - ULA-L – Like RFC 1918 only much larger.
    - FC00::/8
    - 1,099,511,627,776 /48s, 118 per person on the planet in 2004
    - Collisions not assured
- ULA-C (FC01::/8) is currently the subject of some debate.

Thanks!

joel.jaeggli@nokia.com

Bibliography

- http://www.sixxs.net/main/ - IPV6 tunnel broker pointers to supporting ISPs etc.
- http://www.civil-tongue.net/clusterf/ - IPV6 transition wiki, submissions and participants appreciated