IIJ
Internet Initiative Japan

# IPv6 Transition & Operational Reality

## AfNOG / Rabat

2008.06.03

Randy Bush <randy@psg.com>

http://rip.psg.com/~randy/080603.afnog-op-reality.pdf

# The Bottom Line

I am far less interested in what particular service is available via IPv4, IPv6, or both than I am in ensuring that it makes no difference to the users.

# Just Deploy IPv6!

- If we were designing it again we might do it differently
- But it is too late given IPv4 Free Pool run-out
- We have no choice, deploy IPv6 or break the internet

# Content

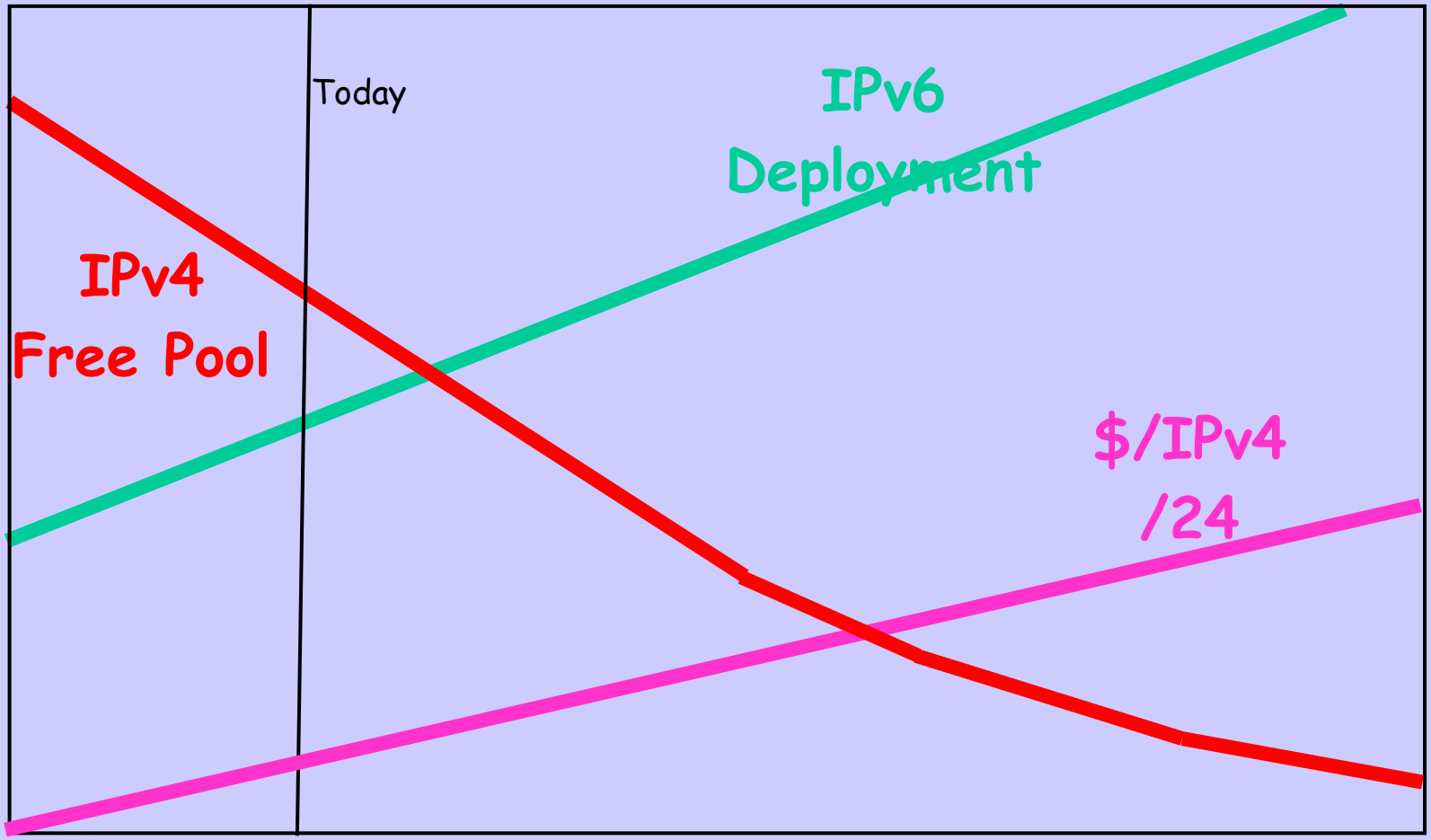
- Some Hard Realities
- Too Many IPv6 Myths
- Strategic Problems in Transition
- Tactical Transition Needs

There is no accompanying paper, so the slides are dense and meant as reference

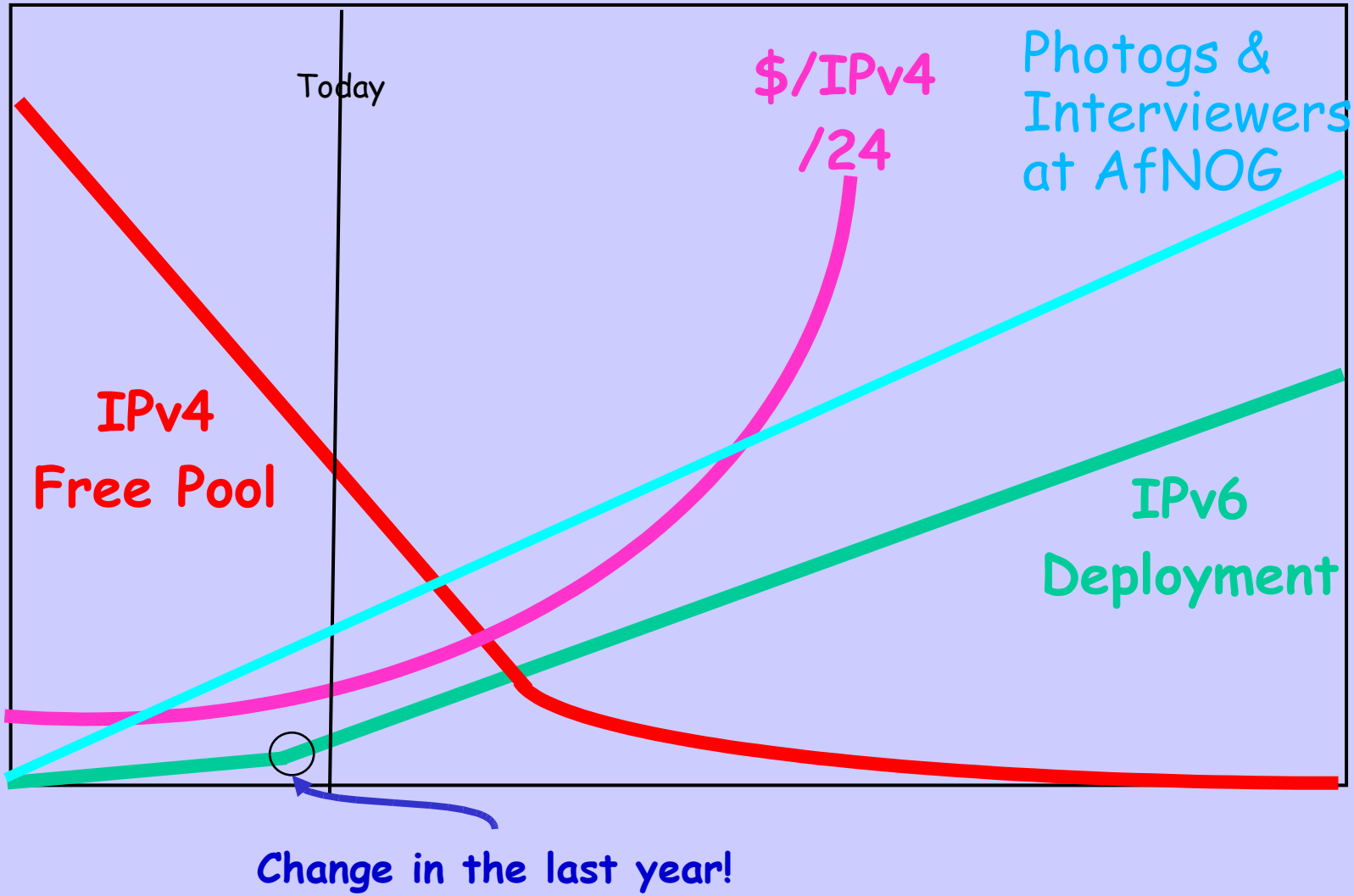
# Reality Therapy

- Massive IPv4 NAT or dual stack with NAT-PT or other NATs, get over it
- IPv6 is a technically better more scalable choice, get over it
- The issues are when and how
- Marketing fantasies are not helping us actually deploy
- Take off the 'rose colored glasses' so we can see what reality is so we can actually make deployment decisions

# What Should Have Happened



# What Is Happening?



# Why Is This Happening?

No transition plan

Declared victory before hard part started

No real long term plan

No realistic estimation of costs

No support for the folk on the front lines

Victory will be next month

This Describes:

a - The US invasion of Iraq

b - IPv6

c - DNSSec

# IPv6 is Incompatible with IPv4 on the Wire!

The Insanity and Short-Sighted Arrogance of this is Utterly Amazing

# Let's Dispel Some Myths

# Myth: IPv4 is Running Out

- IPv4 Free Pool run-out in a few years
- This is in line with the graphs of Frank Solensky over ten years ago
- IPv4 will go to a Trading Model
- Registries will become Title Agents, not allocators, of IPv4 space
- RIRs developing full multi-RIR/LIR open source software to certify and verify title to IPv4 and IPv6 resources

# Myth: IPv6 Transition is Easy

- IPv6 was designed with no serious thought to operational transition
- IPv6 is **on-the-wire incompatible** with IPv4
- Might have been avoided, e.g. if IPv6 had variable length addressing, IPv4 might have become the 32 bit variant
- There are no simple, useful, scalable translation or transition mechanisms

# Myth: IPv6 Eliminates NATs

- An IPv6-**only** site can not reach the IPv4 Internet because it can not source packets from an IPv4 address
- There will be significant IPv4-only Internet for a decade or more
- All IPv6 sites will need IPv4 space and will have NATs with ALGs
- IPv6 **increases** NAT use in short and medium term, i.e. a decade or more

# Myth: IPv6 Reduces Routing Load

- Multi-homing in IPv6 is the same as in IPv4, there is no new routing model
- Traffic engineering in IPv6 is the same as in IPv4, no new TE model
- Enterprises will slice and dice their IPv6 /32s to handle branches etc.
- The routing table will fragment more and more over time

# Myth: IPv6 Space is Infinite

- 64 bits goes to every LAN
- This leaves half the bits gone!
- Some folk use /64 for Point-to-Point!
- RIRs are giving away /32s
- In 15 years we will think of these as we now think of legacy /8s in IPv4 space
- We once thought 32 bits was enough

# Myth: IPv6 is more Secure

- IPv6 does nothing IPv4 does not, though it promised to
- IPSec is the recipe in either case
- IPSec does not work well in a mixed IPv4/IPv6 environment (think VPN from an IPv4-only hotel room)
- It is true that address space scanning will be somewhat harder
- Ha Ha, think botnet scanning and a black market in hot space

# Myth: Incremental Deployment

- For an enterprise, the entire chain, from database back end, through applications, through firewalls, to the border router must all support v6 or the enterprise can not deploy
- For ISP, provisioning systems, monitoring, measurement, billing, …
- And everyone needs support from all their vendors

# Myth: Routers Fully Support IPv6

- But not 100% in hardware
- Especially not if you add ACLs
- Folk do not know this because there is no good IPv6 traffic test equipment
- And all vendors are not spinning the ASICs to solve this
- Not all v4 features are supported over IPv6: MIBs, SNMP over v6, …

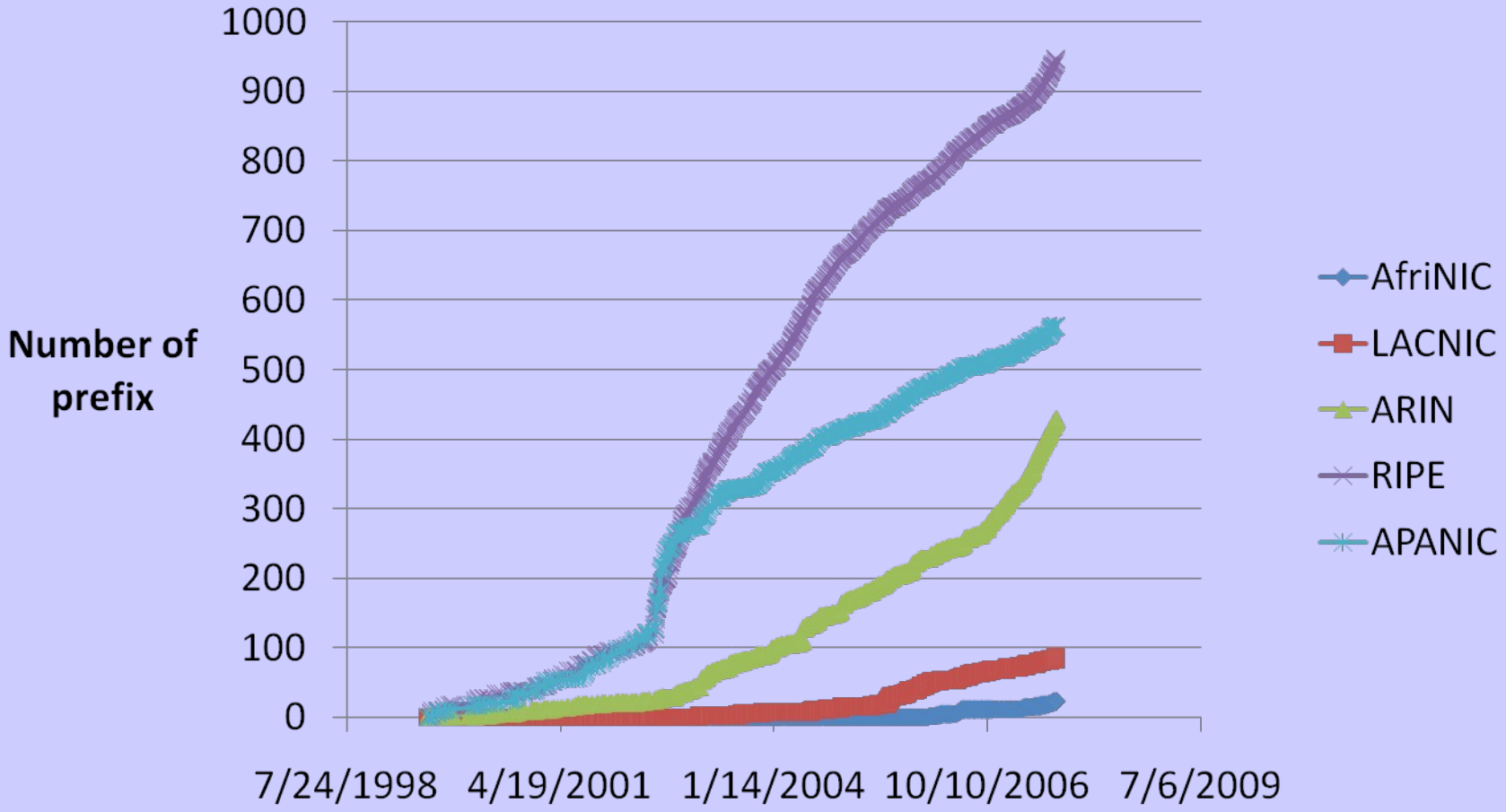
# Myth: No Static Numbering

- IPv6 Auto Configuration is not widely used in enterprise as security policy prefers known (i.e. DHCP) addresses
- Similarly, ISP backbone addresses and customer addresses must be known for logging, audit, CALEA, …

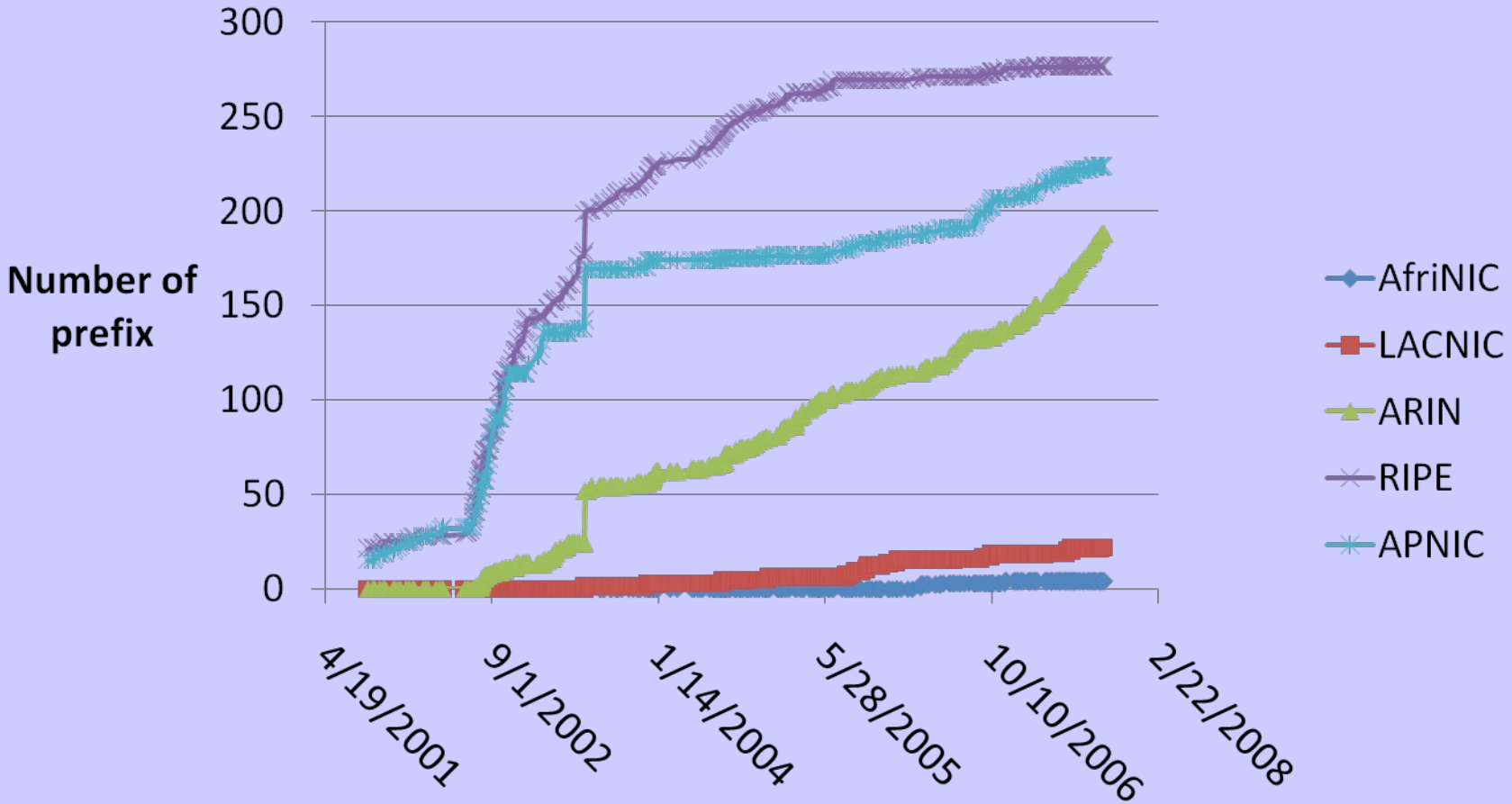
# Myth: IPv6 is Deployed

- Pioneers are still moving cautiously
- Early adopters are just starting to enter the game
- Actual measured traffic is very small (so it makes routers look as if they can handle the traffic)
- But there are success stories

# Prefix Allocation Distribution



# BGP Prefix Announcements



# Myth: IPv6 Ubiquitous in Japan

- NTT.com and IIJ do have dual-stack on leased line services. And NTT.com has dual-stack on ADSL.
- The main high-bandwidth layer-2 service is NTT-East/West B-Flets, for ETTH
- Neither NTT nor IIJ provide native or dual-stack over B-Flets
- NTT-East/West do have a VoIP and TV-to-STB closed IPv6 universe over B-Flets. No Internet.

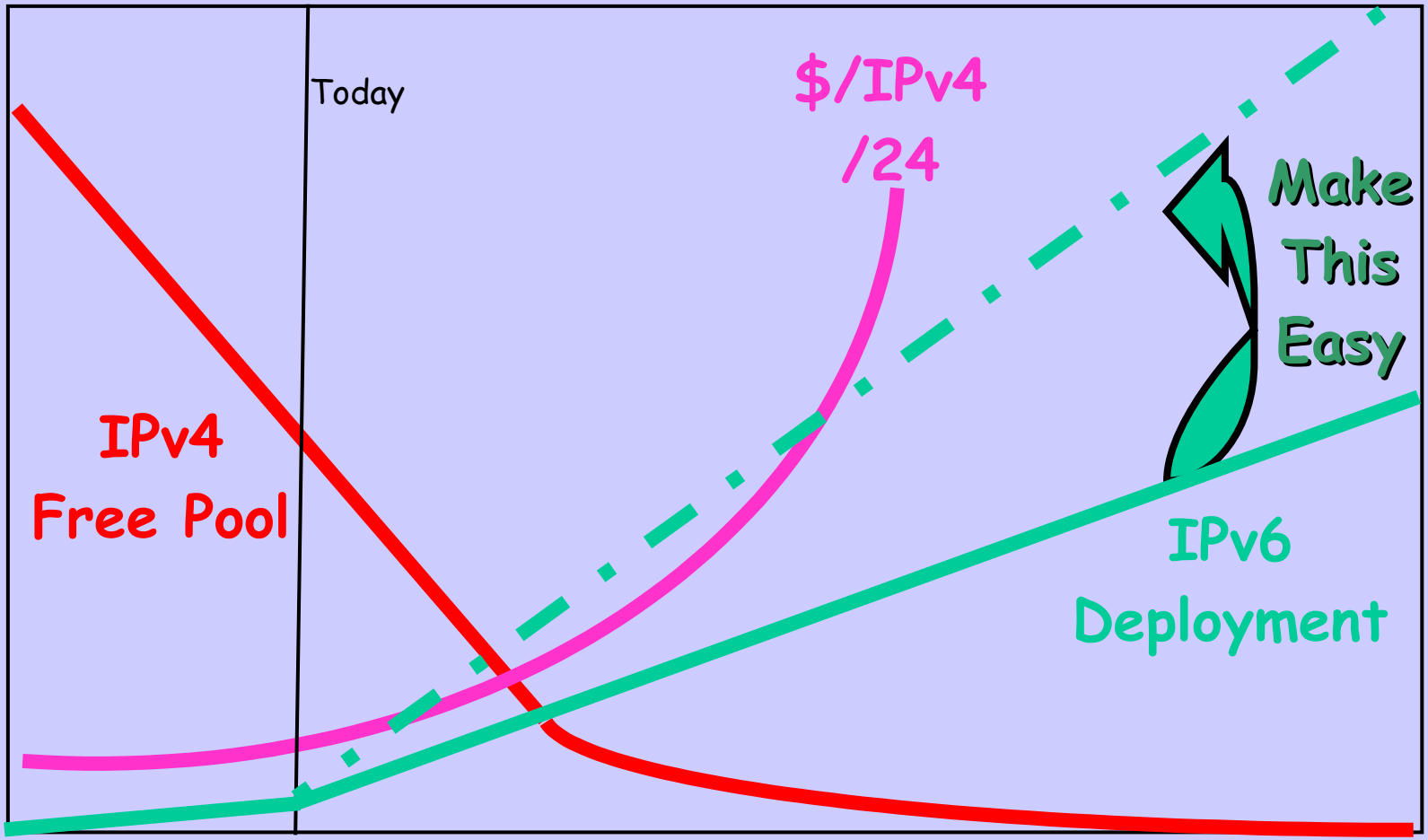
# Myth: IPv6 Will Replace IPv4

- Not given current lack of **universal** vendor support from back end to border router
- It is far easier to use NAT and IPv4
- IPv4 with NATs requires no new expense, conversion, training, …
- This is architecturally horrible, it is just financial reality

# The Reality

- "96 more bits, no magic"
  -- Gaurab Raj Upadhaya
- But we definitely need more bits!
- The key questions are how to use them?
- How to transition without losing anyone or anything?

# What Can We Do?



# How?

- Identify current transition problems
- See that they are fixed
- Ask the IETF to fix the outstanding protocol issues
- Ask IETF to stop embellishing so that we can deploy something stable
- Push vendors to support IPv6 and the tools for us to transition

# What We Should Not Do

- Pretend that there are no transition problems. It just makes things harder.
- Give away IPv6 space in strange ways to "promote" IPv6. IPv4 run-out will promote IPv6 for us.
- Make messes we will have to live with forever.

# Areas of Concern/Study

- Global Issues
- Administrative Infrastructure
- Layers 1 and 2
- Backbone Engineering
- Last Mile/Kilometer
- Consumer/SOHO Self-Installed CPE
- Enterprise
- Server Farm
- Campus
- Exchange Points
- Applications
- Telephony
- More?

# Layers 1 and 2

- DOCSIS 3.0 for Cable
  - CMTS support lacking
  - Massive installed base of DOCSIS 2 modems
- 802
  - All media protocols support IPv6
  - While the protocols support IPv6, this does not at all mean that implementations do

# Backbone Engineering

- Core Routing – conversion to dual stack is slow
- Provisioning, Address Assignment, DNS, …
- DHCPv6 and DNS Integration
- Monitoring and Measurement over v6?
- New line cards are often required!

# Last Kilometer

- Authentication and session setup
- Conversion to IPoE, DHCP expensive
- Provisioning, back-end database, ...
- "How to scale the routing/provisioning combo to deal with million of customers using stable prefix delegation?"

# Consumer Self-Installed CPE

- $50 DSL Modems do not support v6
- $50 Firewalls do not support v6
- Teredo does not really scale
- shim6 is does not solve enterprise or large site, and is not deployable due to security and routing model issues

# Security Devices

- Dave Piscitello made a presentation "IPv6 Support Among Commercial Firewalls" at the last ARIN meeting
  <http://www.arin.net/meetings/minutes/ARIN_XX/PDF/thursday/Firewalls_Piscitello.pdf>
- Less than 1/3 had IPv6 Transport
- 25% supported IPv6 Routing
- And it gets Worse from there

# Enterprise

- Databases, PeopleSoft, Siebold, Business Applications, ...
- Firewalls, VPNs, Access, ...
- Millions of lines of in-house code
- NFS Appliances, unknown
- Load Balancers

# Applications

- Where is the web page with matrix of application by platform showing which are v6 capable and clickable link on how to turn it on?
- http://www.deepspace6.net/docs/ipv6_status_page_apps.html out of date
- Many applications which support v6 have sufficiently poor performance that early adopters are being told to turn v6 off

# SMTP: An Example

- Email/SMTP is a mandatory application
- Everyone needs to be able to send email to arbitrary recipients, i.e. everyone else
- But, due to SPAM, no one can run an open SMTP relay
- So all IPv6 sites need to have the ability to SMTP to arbitrary IPv4 sites
- Therefore everyone needs private dual stack relay until the world is all dual stack SMTP

  [ example by Jeffrey Streifling ]

# Why is Japan in Better Shape?

- Folk with vision (i.e. Murai) convinced the government that early movement to IPv6 was wise for Japan
- Government $upport$ IPv6 research
- Government $upport$ IPv6 development by industry, vendors, …
- Government give$ tax incentive$ to enterprises which become v6 compatible

# What Can We Do?

# Summary

- No More Bull*&^(*
- No More Excuses
- Shut up and Spend the Money

-- Lucy, in a stressed moment

# Principle: One Internet

- Under no circumstances can we allow the Internet to fragment
- During transition, everybody still needs to talk to everyone else at will
- And it would be good if the End to End principle could be kept as much as possible

# Principle: Dual Stack

- The core needs to be dual IPv4/IPv6 during all of transition or kludges will escalate horrifyingly
- The further dual stack goes toward the edge (enterprise, net services, consumer, …) the easier it will be
- Configuration, Management, and Measurement need to be simplified

# Five Phases

- Denial, from both 'sides':
  - We can ignore brain-dead IPv6
  - IPv6 is perfect and those greedy fools just have to deploy it
- Dual stack with IPv4 Dominant
- Dual stack with both widely used
- Dual stack with IPv6 Dominant
- The IPv6 Internet (getting ready for IPv10 transition:)

# Pressure on Routing

- IPv4 address space price escalation and the consequential NATs will put serious new pressure on routing
- If it takes a $10m router to deal with 2m routes and churn then 96% of ISPs die and enterprises can not be DFZ multi-homed
- So **all sized routers**, from enterprise border to ISP core, need to handle **>2m routes with churn**

# Do Not Hack!

- Do not accept hacks around the routing scaling problem such as tunneling from enterprise border to some $10m 'core' router
- Think TLA/NLA and be fearful ☺
- Think ten Monopoly ISPs and be very very fearful

# Forwarding is Behind

- Because of lack of market, it will be five years before **all** major router vendors support dual stack at line rate with ACLs
- Some vendors are not even spinning the ASICs for all platforms and line cards
- Needs to be **all vendors** because ISPs can not be vendor-locked by transition
- So we are not interested in "We can do it, they can't" marketing wars.

# Good IPv6 Test Equipment

- Router/Switch vendors claim wonderful performance
- But you can not test it because there is a serious lack of good test/exercise equipment

# Stop Adding Features

- Stop trying to market IPv6 through more and more kinky features
- IPv4 free pool run-out will either sell IPv6 or there will be an IPv4 NAT world
- Adding features just gives vendors and operators reasons to delay
- Freeze the damned thing and give us a chance to deploy it!

# ULA: A Bad Example

- Because ULA is address-based
- 'Borders' need to filter packets
- To not leak and not accept leaks, needs both source and destination filters
- Do not make special address space
  - Remember 240/4 and that clean-up
  - IPv6 space is supposed to be infinite!
- Give them real IPv6 space and tell them to just not announce it to the DFZ

# Principle: NATs

- End to End Principle is **very** desirable
- But IPv6 on the wire is incompatible with IPv4
- During transition there will be NATs
- Get over it
- But we need to make it so they can fade away and not be there forever

# NAT-PT

- At the edge, the enterprise, consumer, etc. need to run IPv6 but need to talk to both IPv4 and IPv6 services
- When IPv6 becomes dominant, the IPv4 sites will still need to talk to the then predominantly IPv6 Internet
- The IETF needs to standardize 4/6 NAT for ICMP, UDP, TCP, DNS, SMTP, HTTP, SIP, RTP, and maybe an API of how ALGs plug in

# IETF and Reality

- In July 2007, the IETF published RFC 4966 "Reasons to Move the Network Address Translator – Protocol Translator (NAT-PT) to Historic Status"
- This tells you a lot about the IETF, their level of operational clue, and how much they care about religion as opposed to IPv6 deployabilty

# Windows XP

- XP can move payload over IPv6
- But does **not** do DNS over IPv6 transport (**bad bug**! "Buy Vista")
- So, the LAN has 1918 IPv4 space to carry DNS, but no exit for IPv4
- You get an IPv6 and IPv4 Address but should use the IPv6 for all real transport as IPv4 has no default

# NAT-PT & Security

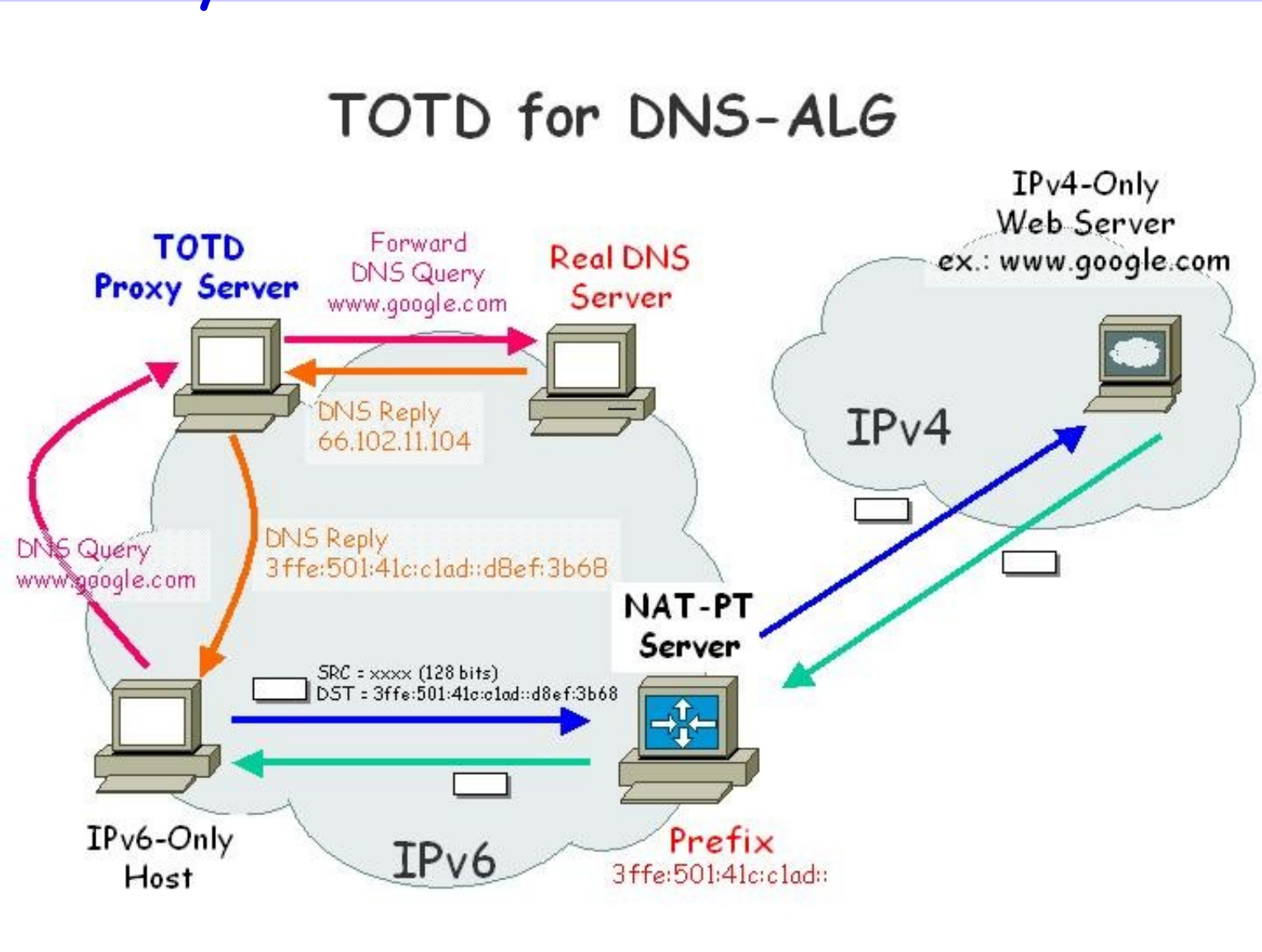
- DNSsec has to terminate on the NAT if translating and use ALG
- IPSec can transit NAT-PT
- DNS, SMTP, HTTP, SIP, RTP ALGs will be critical
- IPsec **must** be made easy for users to configure

# The DNS Hack

- On a pure IPv6 network, if I get an A record, what do I do?
- Panic, you can't use an A record
- So the local DNS Cache has a hack, totd, which takes an A, embeds it within a hacked IPv6 prefix, and synthesizes an AAAA
- NAT-PT knows the hack prefix, and strips it back to IPv4 to dual-stack

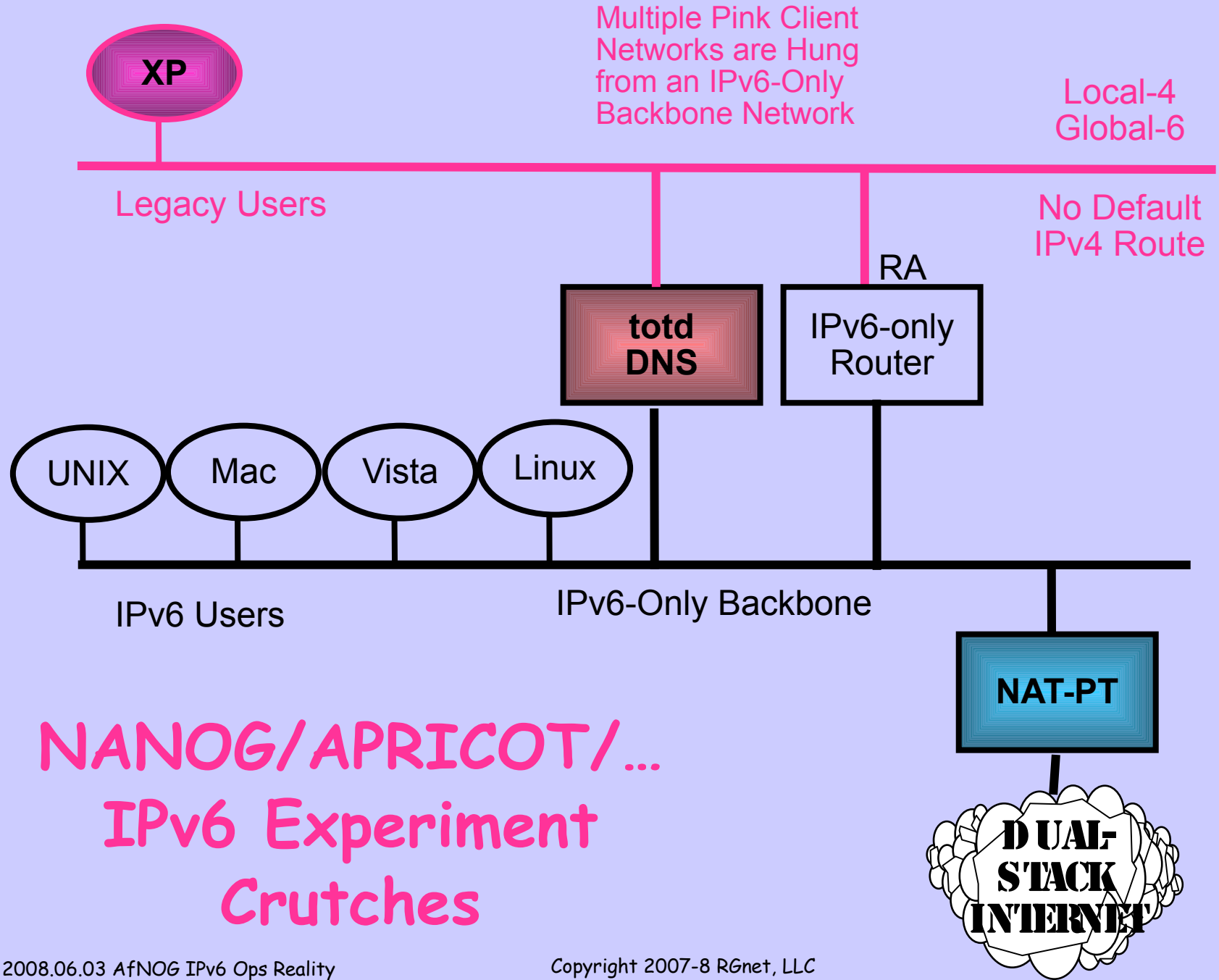
# totd - DNS Trans Proxy

**Picture by Vasaka Visoottiviseth - mahidol.ac.th**



# NANOG Experiment

- 2008.02.17-20 in San Jose, CA
- 550 attendees, all on laptops
- 45mb Exit to Net
- Wireless with multiple SSIDs
- Dual Stack, V6-only, Hacked V6 for XP



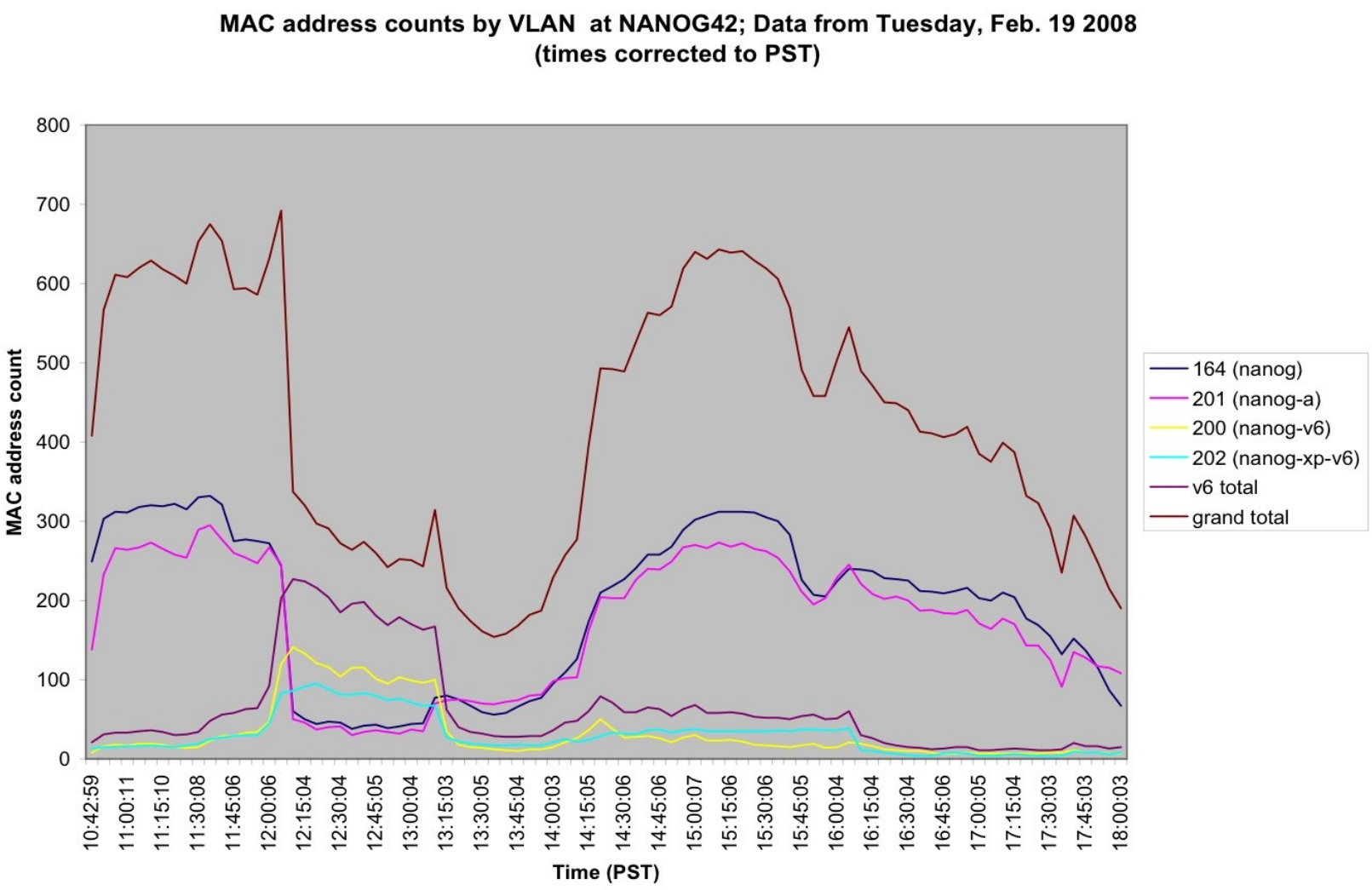

# NANOG/APRICOT/… IPv6 Experiment Crutches

# NANOG Experiment

- All SSIDs were up from start
  - **nanog** and **nanog-a** fully dual-stack
  - **nanog-v6** with NAT-PT & tots
  - **nanog-v6-xp** with 1918
- Tuesday, for an hour or so, nanog and nanog-a went away
- We gathered stats from NAT-PT, net use, AP binding, traffic, …
- Reported on Wednesday

# What We Learned

- ¾ of users said they could get on net but it was actually more like half
- Every component except UNIXes had bugs: NAT-PT, Vista, MacOS, …
- The prize to MacOS which dropped capital A from DNS server entry
- We hope vendors now working to fix


MAC address counts by VLAN at NANOG42; Data from Tuesday, Feb. 19 2008
(times corrected to PST)
MAC address count
800
700
600
500
400
300
200
100
0
10:42:59
11:00:11
11:15:10
11:30:08
11:45:06
12:00:06
12:15:04
12:30:04
12:45:05
13:00:04
13:15:03
13:30:05
13:45:04
14:00:03
14:15:05
14:30:06
14:45:06
15:00:07
15:15:06
15:30:06
15:45:05
16:00:04
16:15:04
16:30:04
16:45:06
17:00:05
17:15:04
17:30:03
17:45:03
18:00:03
Time (PST)
164 (nanog)
201 (nanog-a)
200 (nanog-v6)
202 (nanog-xp-v6)
v6 total
grand total

# Ongoing Experiments

- Similar, but not identical, experiments held at
  - APRICOT/APNIC last week
  - IETF next week, different goals
  - AfNOG/AfriNIC, LACNIC, …
  - ICANN, …

# Summary

- IETF
  - NAT-PT
  - No More 'Features' or hacks, e.g. ULA
- Vendors
  - Dual Stack on the Fast Path with ACLs
  - 2+m Routes with churn on all routers
  - Test Equipment
- ISPs
  - Dual Stack to the Customer Edge
- Governments: incent, don't regulate

# You Can Help!

Search

Login | Settings | Help/Guide | About Trac

Wiki | Timeline | Roadmap | Browse Source | View Tickets | New Ticket | Search

Start Page | Index by Title | Index by Date | Last Change

## IPv4 / IPv6 Operational Transition Information Collection

**ARIN Announcement** http://www.arin.net/announcements/20070521.html



### Goal Statement

### Information Gathering

- Call for Input
- Contributors
- Organizers

### Areas of Investigation

- Global Issues
- Administrative Infrastructure
- Layers 1 and 2
- Backbone Engineering
- Last Mile
- Consumer/SOHO Self-Installed CPE
- Enterprise
- Server Farm
- Campus
- Exchange Point
- Applications
- Telephony

### Resources

- http://www.internet2.edu/presentations/jt2007feb/20070213-broersma.ppt
- http://ipv6.internet2.edu/merit/
- http://ipv6samurais.com/ipv6samurais/demystified/
- http://www.ipv6-to-standard.org/

### Stories of Pioneers

### Other

- Players and roles
- Transition mechanisms

# How You Can Help

http://www.civil-tongue.net/6and4/

write to randy@psg.com

if you can contribute

**Please!**

# Thanks To

ISOC

Internet Initiative Japan

Lucy Lynch, Rob Austein,
Russ Housley, Steve Bellovin &
Jari Arko