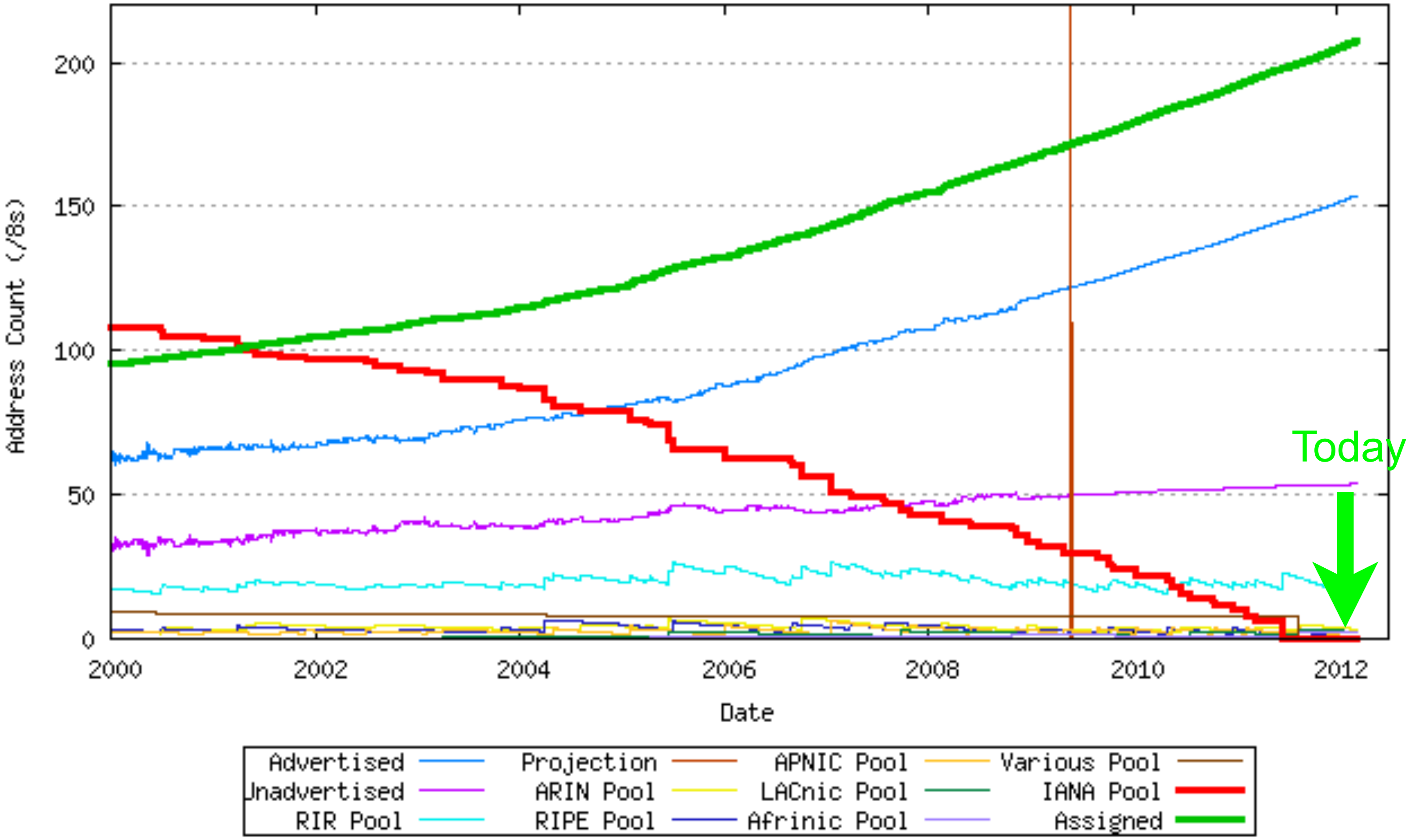
HURRICANE ELECTRIC
INTERNET SERVICES

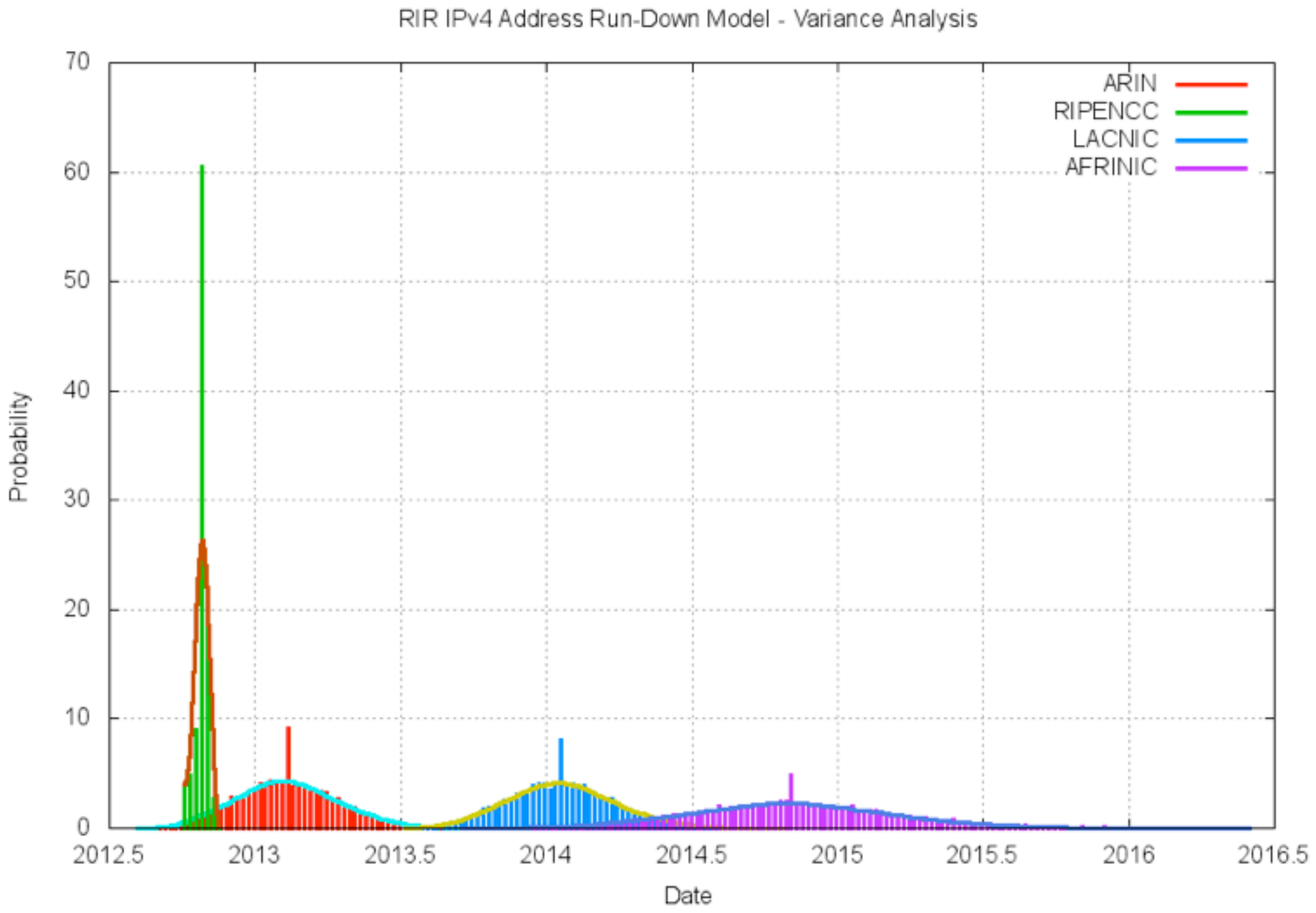
IPv6 -- No longer optional

Owen DeLong
owend@he.net

Why is this important? – Today



RIR Free Pool Projections Geoff Huston's math:



RIR Free Pool Update
My speculation:

RIR	Non-Austerity Free Pool (9/4/2011)	Austerity Date?
ARIN	5.22 /8s	3/2013?
AfriNIC	4.74 /8s	1/2014?
RIPE	1.34 /8s	11/2012?
LACNIC	1.72 /8s	9/2013?
~~APNIC~~	~~0.00 /8s~~	~~OUT 4/15/11~~

IPv4 Runout Process

- IANA runs out first, ~~~2011~~ February 3, 2011
- RIRs start running out probably ~~in 2012~~ around ~~June, 2011~~ APNIC ran out April 15, 2011
- End-User providers start running out shortly after RIR runout. Most likely, the larger ones first (APNIC happening now)
- After ISPs start running out, an increasing number of your customers/users ~~will have~~ are experiencing limited or seriously degraded ability to connect via IPv4, possibly even no ability.

IPv6 Transition -- How ready are we?

- Things that are ready
  - ☑ Backbones
  - ☑ CMTS Systems (DOCSIS 3)
  - ☑ MacOS (10.4+)
  - ☑ Linux (2.6 Kernels)
  - ☑ Windows (7, 2008, XP (limited))
  - ☑ WiMax (specification, head end equipment)
  - ☑ LTE (some)
  - ☑ CPE (very limited)
  - ☑ Early Adopters and some industry experts
  - ☑ Hurricane Electric
  - ☑ Me

IPv6 Transition -- How ready are we?

- Things that are NOT ready
  - PON Systems
  - DSL Systems
  - CMTS Systems (DOCSIS 2)
  - WDS/EVDO/HSPA
  - WIMAX (handsets, providers)
  - Older Windows (XP and earlier)
  - Embedded systems
  - Printers
  - Home entertainment devices
  - CPE (most)
  - Most IT staff and management

Quick survey

How many of you have started planning IPv6 in your organization?

How many of you have IPv6 running in a test environment?

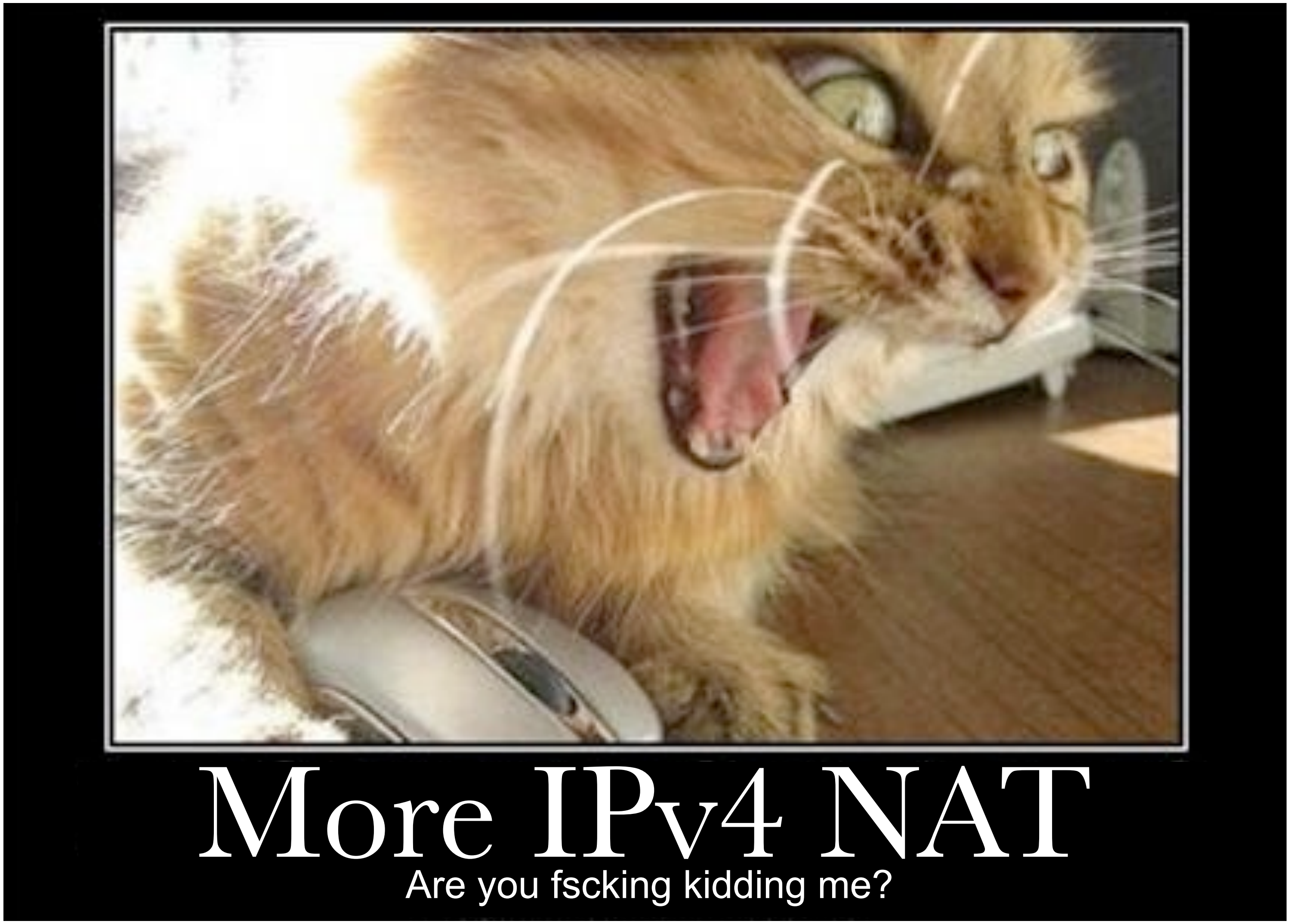
How many of you have started deploying IPv6 to your organization?

How many of you have a fully production dual-stack environment running in your organization?

This is a room full of IPv6 proponents.

- Results from other rooms:
  - Planning? -- average about 5%
  - Test environment? -- average about 2%
  - Deploying? -- Average 1-2 hands
  - Full production? -- Usually just my hand.
- We have to do better!
  - If you’re not planning, why?
  - If you’re deploying, keep moving.
  - Full Production? Help the others!

LoL Kitteh sez:



Shared Network, Shared Fate

- I hear a lot of people say “I don’t need to do IPv6, I have enough IPv4 addresses for years to come.”
  - Are you really on the internet just to talk to your own organization?
  - There simply aren’t enough addresses for everyone that wants/needs to be on the internet in IPv4. If you want to be able to reach new participants, that’s going to require IPv6.
  - Workarounds all come with bad tradeoffs.

The real questions...

How many of you think your organization will be fully IPv6 ready by February, 2012?

What do you plan to do to fix that?

How do you plan to cope with a world where there are no more IPv4 addresses available?

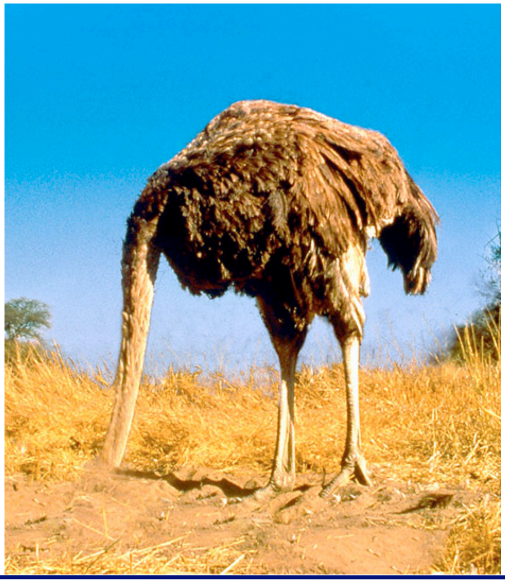
How do you plan to cope with a world where some of your customers have only IPv6 connectivity, or, severely degraded IPv4 connectivity?

The final question...

Which Approach will you take?

IPv4/IPv6 Dual Stack Now

IPv4 is just fine.
We just need MOAR NAT!!



What we’ll cover

- Basics of IPv6
- IPv6 Addressing Methods
  - SLAAC
  - DHCP
  - Static
  - Privacy
- Mikrotik Configuration for Native Dual Stack
- IPv6 without a native backbone available
- Free IPv6?

Some additional topics

- Routing
- Firewalls
- DNS
- Reverse DNS
- Troubleshooting
- Staff Training

Basics: IPv4 vs. IPv6

Property	IPv4 Address	IPv6 Address
Bits	32	128
Total address space	3,758,096,384 unicast 268,435,456 multicast 268,435,456 Experimental/other (Class E, F, G)	42+ Undecilion assignable[1] 297+ Undeciliion IANA reserved[2]
Most prevalent network size	/24 (254 usable hosts)	/64 (18,446,744,073,709,551,616 host addresses)
Notation	Dotted Decimal Octets (192.0.2.239)	Hexidecimal Quads (2001:db8:1234:9fef::1)
Shortening	Suppress leading zeroes per octet	Suppress leading zeroes per quad, longest group of zeroes replaced with ::

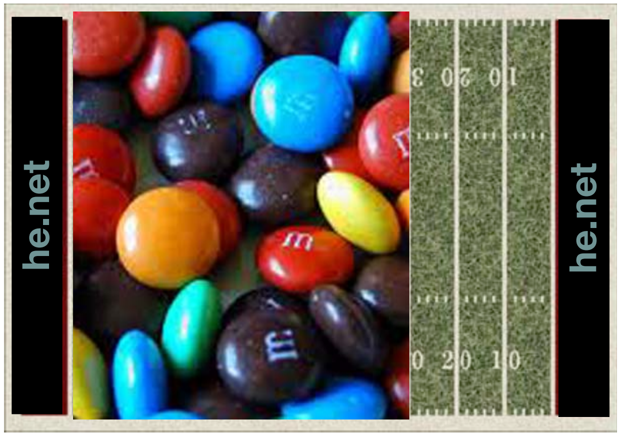
[1] 42,535,295,865,117,307,932,921,825,928,971,026,432 assignable unicast (1/8th of total)

[2] 297,747,071,055,821,155,530,452,781,502,797,185,024 IANA reserved (7/8th of total)

Network Size and Number of networks (The tasty version)

One IPv4 /24 -- 254 M&Ms

One IPv6 /64 -- Enough M&Ms to fill all 5 of the great lakes.



Full Address Space, One M&M per /24 covers 70% of a football field

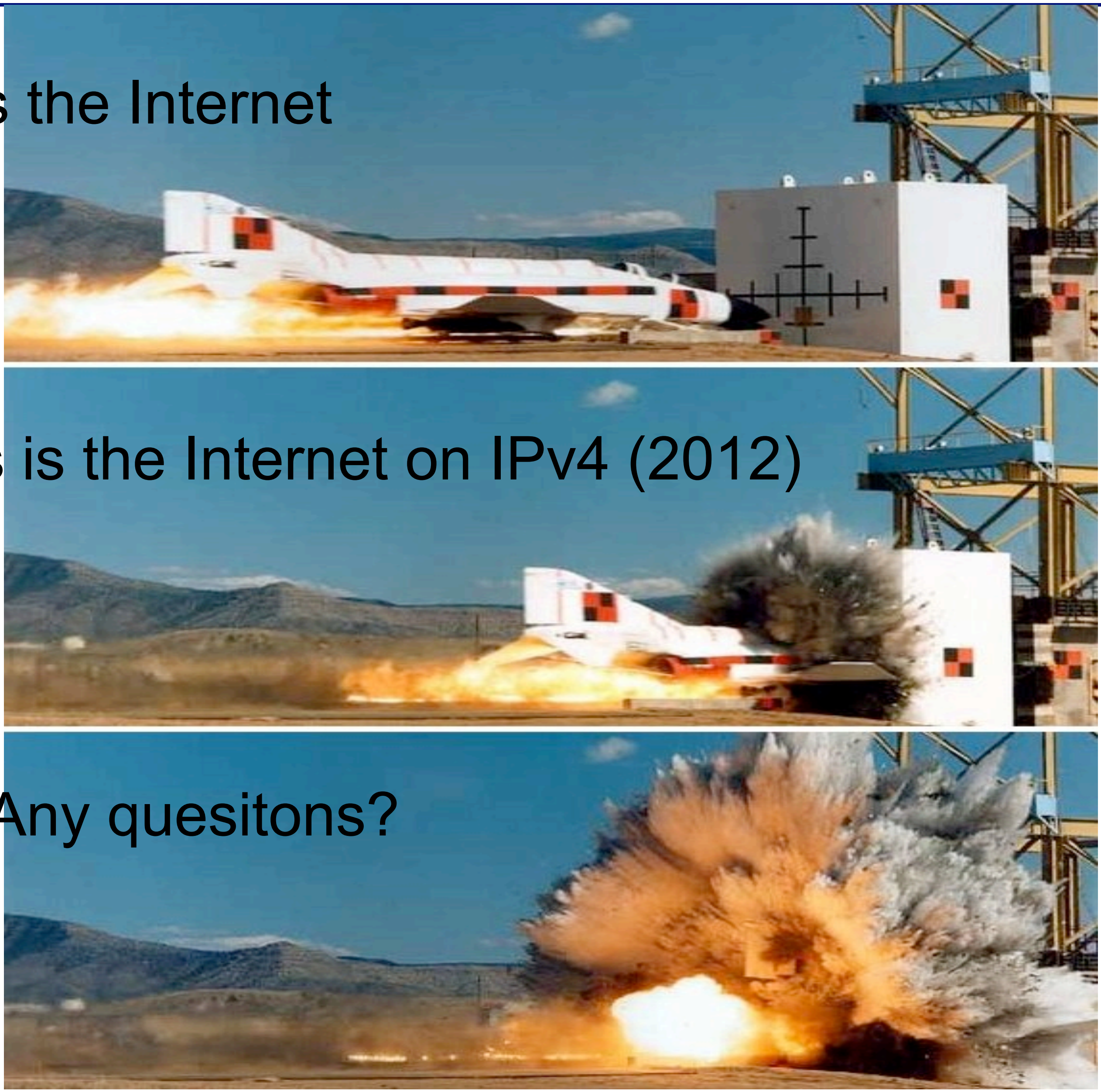
Full Address Space, One M&M per /64 fills all 5 great lakes.

Comparison based on Almond M&Ms, not plain. Caution! Do not attempt to eat a /64 worth of any style of M&Ms.

Basics: IPv4 vs. IPv6 thinking

Thought	IPv4 dogma	IPv6 dogma
Assignment Unit	Address (/32)	Network (/64)
Address Optimization	Tradeoff -- Aggregation, Scarcity	Aggregation (At least for this first 1/8th of the address space)
Address Issue Methodology	Sequential, Slow Start, frequent fragmentation	Bisection (minimize fragmentation), issue large, minimal requests for more, aggregate expansions.
NAT	Necessary for address conservation	Not supported, Not needed -- Breaks more than it solves (other than possible NAT64)
Address Configuration	Static, DHCP	Stateless Autoconf, Static, some DHCP (needs work), DHCP-PD (NEW!!)

This is the Internet



This is the Internet on IPv4 (2012)

Any quesitons?

Basics Address Scopes

- Link Local -- fe80::<UUVV:WW>ff:fe<XX:YYZZ> only valid on directly attached subnet.
- Site Local (deprecated) -- Only valid within site, use ULA or global as substitute.
- Unique Local Addresses (ULA) -- Essentially replaces IPv4 RFC-1918, but, more theoretical uniqueness.
- Global -- Pretty much any other address, currently issued from 2000::/3, globally unique and valid in global routing tables.

Basics: Stateless Autoconfiguration

- Easiest configuration
- No host configuration required
- Provides only Prefix and Router information, no services addresses (DNS, NTP, etc.)
- Assumes that all advertising routers are created equal, rogue RA can be pretty transparent to user (RA guard required on switches to avoid)

RA Guard -- PUSH YOUR VENDORS!!

- RA has a serious vulnerability
  - Compare to rogue DHCP
  - Accidental Rogue RA
    - breaks stuff
    - easy to find
    - easy to mitigate
  - Malicious Rogue RA
    - Virtually undetectable
    - All your packets are belong to us
    - Coffee Shop nightmare

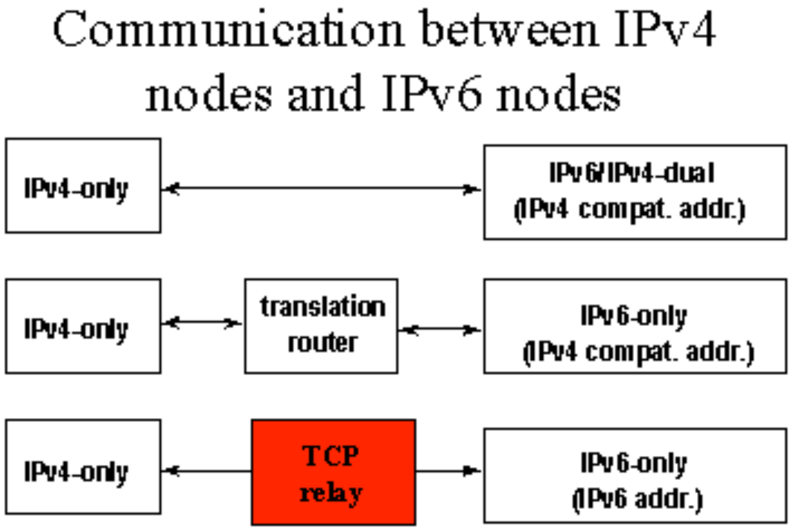
Stateless Autoconfiguration Process

- Host uses MAC address to produce Link Local Address. If MAC is EUI-48, convert to EUI-64 per IEEE process: invert 0x02 bit of first octet, insert 0xFFFE between first 24 bits and last 24 bits fe80::<EUI-64>
- IPv6 shutdown on interface if duplicate detected.
- ICMP6 Router Solicitation sent to All Routers Multicast Group

Stateless Autoconfigration Process (cont.)

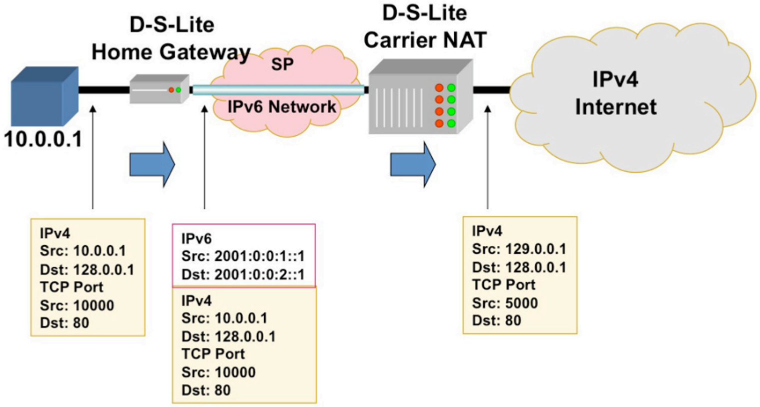
- Routers send ICMP6 Router Advertisement to link local unicast in response. Also sent to All Hosts Multicast group at regular intervals.
- Router Advertisement includes Prefix(es), Preference, Desired Lifetime, Valid Lifetime.
- Host resets applicable Lifetime counters each time valid RA received.
- Address no longer used for new connections after Desired lifetime expires.
- Address removed from interface at end of Valid lifetime.
- Prefix(es)+EUI-64 = Host EUI-64 Global Address, netmask always /64 for SLAAC.

If you think IPv6 is hard, wait until you try any of these.


Communication between IPv4 nodes and IPv6 nodes
IPv4-only
IPv6/IPv4-dual (IPv4 compat. addr.)
IPv4-only
translation router
IPv6-only (IPv4 compat. addr.)
IPv4-only
TCP relay
IPv6-only (IPv6 addr.)

As yet undefined/unimplemented Magic (TCP relay could be SSH tunnel)

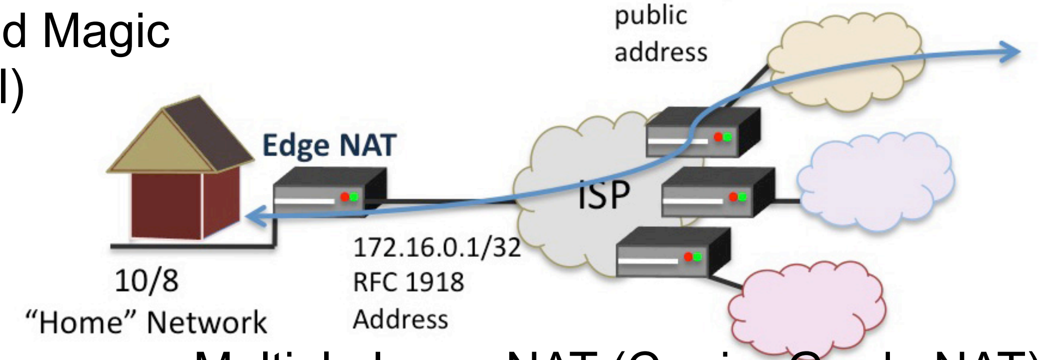
Dual Stack Lite (ISC)



IPv4
Src: 10.0.0.1
Dst: 128.0.0.1
TCP Port
Src: 10000
Dst: 80

IPv6
Src: 2001:0:0:1::1
Dst: 2001:0:0:2::1
IPv4
Src: 10.0.0.1
Dst: 128.0.0.1
TCP Port
Src: 10000
Dst: 80

IPv4
Src: 129.0.0.1
Dst: 128.0.0.1
TCP Port
Src: 5000
Dst: 80



Multiple Layer NAT (Carrier Grade NAT)

DHCPv6

- Can assign prefixes other than /64 -- Ideally larger (/48) prefixes to routers which then delegate various networks automatically downstream, a few limited implementations of this feature.
- Can assign addresses to hosts, cannot provide default router information.
- Can provide additional information about servers (DNS, Bootfile, NTP, etc.)
- Vendor support still lacking in some areas

Static Addressing

- IPv6 can be assigned statically, same as IPv4
- Common to use one of two techniques for IPv4 overlay networks:
  - Prefix::<addr> (first 12 bits of 64 bit <addr> must be 0)
  - Either <addr> is IPv4 last octet(s) expressed as BCD, or <addr> is IPv4 last octet(s) converted to hex.
  - e.g. 192.0.2.154/24 -> 2001:db8:cafe:beef::154/64 (BCD) or 2001:db8:cafe:beef::9a/64 (Hex)
  - These mappings won't conflict with autoconfigured addresses since autoconfigured addresses will never be 000x:xxxx:xxxx:xxxx.

Privacy Addresses

- Essentially a pathological form of Stateless Address Autoconfiguration which uses a new suffix for each flow and obfuscates the MAC address.
- RFC-3041
- Uses MD5 Hash with random component to generate temporary address
- Preferred and Valid lifetimes derived from SLAAC address
- Unfortunate default in Lion and Vista/later

Multiple addresses per interface

- IPv4 has some support for this in most implementations.
- IPv6 has full support for this in all implementations.
- IPv4, multiple addresses/interface are exception.
- IPv6, single address on an interface nearly impossible in useful implementation (link local required, global optional)

IPSEC

- In IPv4, IPSEC is add-on software.
- In IPv6, IPSEC is a required part of any IPv6 implementation
- IPv6 does NOT require IPSEC utilization
- IPSEC is considerably easier to configure in IPv6.
- IPSEC automation may be possible in future IPv6 implementations.

Configuring IPv6 Native on RouterOS

- Important Caveats
  - Use 5.20 or later for IPv6
    - 5.14 has bugs with handling IPv6 nameservers.
    - Always test any software release to ensure suitability before deploying in your production environment.
  - RouterOS doesn't support multiple addresses in the same prefix on an interface
    - Can't do static+EUI-64, for example
- The good news
  - Unlike most other vendors, RouterOS does fully support RDNSS (YAY!).

Configuring IPv6 Native on RouterOS

```
/interface ethernet
set 0 name=ether1 speed=100Mbps

/ipv6 address
add address=2620::930:0:0:0:cafe:beef/64 advertise=no disabled=no \
     eui-64=no interface=ether1

/ipv6 nd
add advertise-dns=yes advertise-mac-address=yes disabled=no interface=ether1 \
     managed-address-configuration=no other-configuration=no ra-lifetime=1h
/ipv6 nd prefix
add autonomous=yes disabled=no interface=ether1 on-link=yes \
     preferred-lifetime=1h prefix=2620::930:0:0:0:0:0/64 \
     valid-lifetime=1d

/ipv6 nd prefix default
set autonomous=yes preferred-lifetime=1w valid-lifetime=4w2d
```

IPv6 without a native connection

- Three options (In order of preference)
  - 6in4 -- Tunnel your IPv6 in an IPv4 GRE Tunnel
  - 6to4 -- Tunnel your IPv6 in an auto-tunnel using an any-casted IPv6 mapping service
  - Teredo -- Tunnel your IPv6 in an auto-tunnel using a multi-server auto-configured process defined by Microsoft.

Why 6in4

- GRE is well understood by most networkers
- Simple and deterministic
- No anycast magic -- Simplifies debugging
- Controlled by two endpoint adminsitrators -- Greatly simplifies debugging
- Disadvantage: Manual config, but, not hard.

Why 6to4

- Automatic configuration
- When it works, it’s pretty clean and relatively self-optimizing.
- May be good option for mobile devices (laptop, cellphone, etc.)
- Hard to troubleshoot when it doesn’t work.
- Disadvantage: Anycast == Non-deterministic debugging process.

Why Teredo?

- Autoconfiguration
- May bypass more firewalls than 6to4
- Enabled by default in Windows (whether you want it or not)
- Meredo available for Linux (client and server)
- Disadvantage: Complicated and tricky to debug if problems occur.

Configuring a GRE tunnel in RouterOS (Dual Stack over IPv4)

```
/interface gre
add comment="GRE Tunnel to delong-l42-rtr-01 (J2320)" disabled=no \
    dscp=0 l2mtu=65535 local-address=76.103.58.47 mtu=1476 \
    name=gre-22 remote-address=192.124.40.130

/ip address
add address=192.124.40.214/30 disabled=no interface=gre-22 network=192.124.40.212

/ipv6 address
add address=2620::930:7005:0:0:0:2/64 advertise=no disabled=no \
    eui-64=no interface=gre-22
```

IPv6 For Free? YES!!

- Several tunnel brokers offer free IPv6.
  - My favorite is the HE Tunnelbroker at www.tunnelbroker.net
- If you or your organization has a presence at an exchange point with Hurricane Electric, we currently offer free IPv6 Transit.

Routing

- Usual suspects
  - OSPF (OSPFv3)
  - BGP (BGP4 Address Family inet6)
  - RA and RADVD
  - Support in Quagga and others

Firewalls

- /ipv6 firewall
  - Just like /ip firewall with bigger addresses

```
/ipv6 firewall address-list
add address=2620::930:0:0:0:400:7/128 disabled=no list=Example
/ipv6 firewall filter
add action=accept chain=example disabled=no dst-address-list=Example dst-port=80 protocol=tcp
src-address=::/0
add action=accept chain=example disabled=no dst-address-list=Example dst-port=53 protocol=udp
add action=accept chain=example disabled=no dst-address-list=Example dst-port=53 protocol=tcp
```

DNS

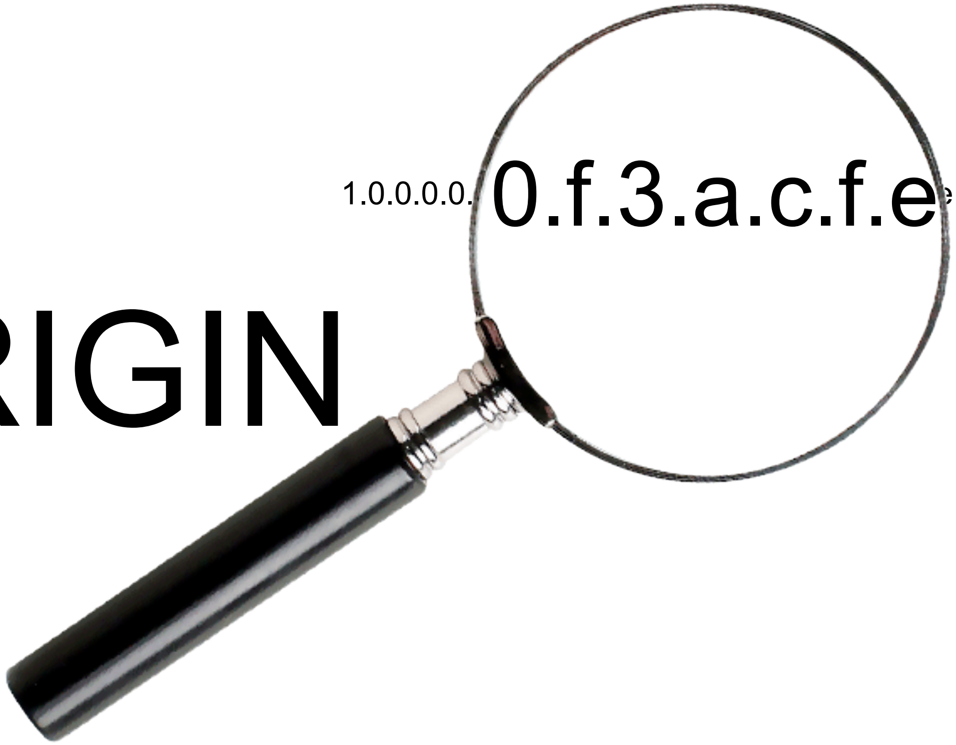
- Forward DNS
  - Instant IPv6 -- Just add AAAA
- Reverse DNS
  - Slightly more complicated
  - ip6.arpa
  - 2620:0:930::200:2 -> 2620:0000:0930:0000:0000:0000:0200:0002
  - 2620:0000:0930:0000:0000:0000:0200:0002 -> 2000:0020:0000:0000:0000:0390:0000:0262
  - 2000:0020:0000:0000:0000:0390:0000:0262 -> 2.0.0.0.0.0.2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.3.9.0.0.0.0.0.0.2.6.2.ip6.arpa

DNS -- BIND Configuration

- Current BIND versions ship with IPv6 template zones (hints, rfc1912, etc.)
- IPv6 addresses valid in ACLs just like IPv4, same rules
- Zone configuration identical except reverse zones for IPv6 ranges called “ip6.arpa”:

```
zone "0.3.9.0.0.0.0.0.2.6.2.ip6.arpa" IN {
      type master;
      file "named.2620:0:930::-48.rev";
};
```

The two best IPv6 Reverse DNS tools:



1.0.0.0.0.0.f.3.a.c.f.e.e.8.b.d.0.1.0.0.2

$ORIGIN

DNS -- BIND Configuration

- In IPv6 Reverse Zone files, $ORIGIN is your friend!
- Forward Zones A for IPv4, AAAA for IPv6, basically what you’re used to:

```
mailhost                    IN       A        192.159.10.2
                            IN       AAAA     2620:0:930::200:2
```

- Reverse Zones PTR records, as described above:

```
$ORIGIN 0.0.2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.3.9.0.0.0.0.0.0.2.6.2.ip6.arpa.
1.0.0.0                     IN         PTR        ns.delong.sj.ca.us.
2.0.0.0                     IN         PTR        owen.delong.sj.ca.us.
4.0.0.0                     IN         PTR        irkutsk.delong.sj.ca.us.
```

DNS -- Reverse DNS Details

- In this example, we see:

```
$ORIGIN 0.0.2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.3.9.0.0.0.0.0.0.2.6.2.ip6.arpa.
1.0.0.0                    IN       PTR      ns.delong.sj.ca.us.
2.0.0.0                    IN       PTR      owen.delong.sj.ca.us.
4.0.0.0                    IN       PTR      irkutsk.delong.sj.ca.us.
```

- $ORIGIN saves us lots of typing for 2620:0:930::200:
- Each entry contains the 4 hex digits for the last quad (0001, 0002, 0004)
- Note each nibble is a zone boundary

DNS -- Common Reverse DNS mistakes

- Not enough zeroes -- 2620:0:930::200:2 is much easier to type, but, remember for reverse DNS you have to expand all those suppressed zeroes before you reverse the address.
- Missing dots (.) -- Every nibble gets one.
  - 2.0.0.0.0.0.2.0.0.0.0.0.0.00.0.0.3.9.0.0.0.0.0.0.2.6.2
  - Do you see the error in the previous line?
- Reversing first then expanding
  - 0.0.0.2.0.2.0.0.0.0.0.0.0.0.0.0.0.3.9.0.0.0.0.0.2.6.2.0

Troubleshooting

- Mostly like troubleshooting IPv4
- Mostly the same kinds of things go wrong
- Just like IPv4, start at L1 and work up the stack until it all works.
- If you are using IPv4 and IPv6 together, may be easier (due to familiarity) to troubleshoot L1-2 on IPv4.

Troubleshooting

- Common problems
  - Cannot ping remote IPv6 address on Tunnel
  - Cannot ping remote IPv6 address on ethernet
  - Cannot ping MY IPv6 address (tunnel or ethernet)
  - Cannot reach IPv6 Internet
    - Long waits for IPv6 enabled websites
    - Long delays in host resolution
  - Why don’t my IPv6 neighbors show up in ARP?

A wee bit about Neighbor Discovery and other tools

- No broadcasts, no ARP
- This is one of the key differences with IPv6.
- Instead an all hosts multicast address is used.
- IPv4: arp 192.0.2.123
- IPv6: ip -f inet6 neigh show 2620:0:930::200:2
- ping -> ping6
- traceroute -> traceroute6
- telnet, ssh, wget, etc. just work

Cool SSH trick

- Special for those that made it through the whole presentation:
- If you have a dual stack host you can SSH to in between an IPv4 only and an IPv6 only host that need to talk TCP, then, you can do this from the client:
- ssh user@dshost -L <lport>:server:<dport>
- Then, from the client, connect to localhost:lport and the SSH tunnel will actually protocol translate the session.

SSH trick example

- myhost -- IPv6-only host 2620:0:930::200:f9
- dshost -- IPv4/v6 dual stack host: 192.159.10.2 and 2620:0:930::200:2
- desthost -- IPv4-only host 192.159.10.100
- On myhost I type:
  - ssh owen@2620:0:930::200:2 -L 8000:192.159.10.100:80
  - Then, I can browse to http://[::1]:8000
- My browser will connect to the ssh tunnel via IPv6, and, the SSH daemon at dshost will pass the contents along via IPv4.

Staff Training

- Hopefully this presentation works towards that.
- You'll need more.
- Plan for it.
- Budget for it.
- Allocate time for it.
- If possible, have the staff being trained leave their pagers/blackberries/iPhones/etc. in the car during training.

Q&A

?

Contact:

Owen DeLong
IPv6 Evangelist
Hurricane Electric
760 Mission Court
Fremont, CA 94539, USA
http://he.net/

owend at he dot net
+1 (408) 890 7992