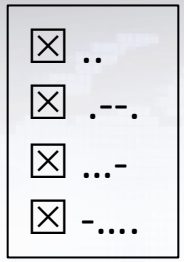
Successfully Deploying IPv6

Presented by Scott Hogg

CTO GTRI, Chair Emeritus RMv6TF, Infoblox IPv6 COE

NANOG On The Road 11, Denver, CO, May 10th, 2016

Successful IPv6 Planning



- Build a cross-functional IPv6 deployment team
  - Multidisciplinary, Collaborative, Cooperative
- Organizations need to treat IPv6 as a “Program” not just like a typical smaller IT “Project”.
  - IPv6 transition is made up of many projects that will span multiple years and cross the entire enterprise.
- Regular & frequent meetings are key to maintaining pace.
- Just like anything, executive buy-in and support is essential.
- Enterprise IPv6 Deployment Guidelines (RFC 7381) provides a good roadmap for all organizations

Training for Success

- Assume your IT organization has not taken the initiative to immerse themselves in learning IPv6.
- People need to be trained early in the process, but not too early that they forget what they learned.
  - Train “just in time”, not years before an IPv6 address is actually configured on a production device.
- Training specific for different IT skillsets
  - Appdev, sysadmin, network engineer, security admin, DevNetSecOps, helpdesk, desktop support, PMs, …
- Much of your IPv4 experience is applicable to IPv6.
- Overcome fear the larger addresses – “Think in Hex”.

IPv6 Addressing

- IPv4-Think is dangerous when planning IPv6 addressing
  - Don’t use decimal #s, don’t embed VLAN #, don’t IPv4 address converted to hex, and then put into IPv6 address
- Perform addressing for simplicity and ease of use and management
  - Don’t be concerned about lots of reserved space
- There is no scarcity of IPv6 addresses, so there can be no waste
  - Don’t try to assign only the minimum-needed prefix length
  - Plan for the number of subnets, not the number of hosts

IPv6 Addressing: “Wasting” a /64

- Some people can’t seem to get the sparseness of nodes within a /64 and considered it “wasteful”
- A single /64 can have 18 quintillion possible nodes
- When the human population reaches 10B, each person could receive 1,844,674,407 nodes on the /64
- If you had a network with 1M nodes, your efficiency is infinitesimally small $n=x/1.8 \text{ x } 1019 = 1.0 \text{ x } 106/1.8 \text{ x } 1019 = 0.00000000000005421011$

- Whether you have 2 or 2M nodes, you are using only a small fraction of the whole /64

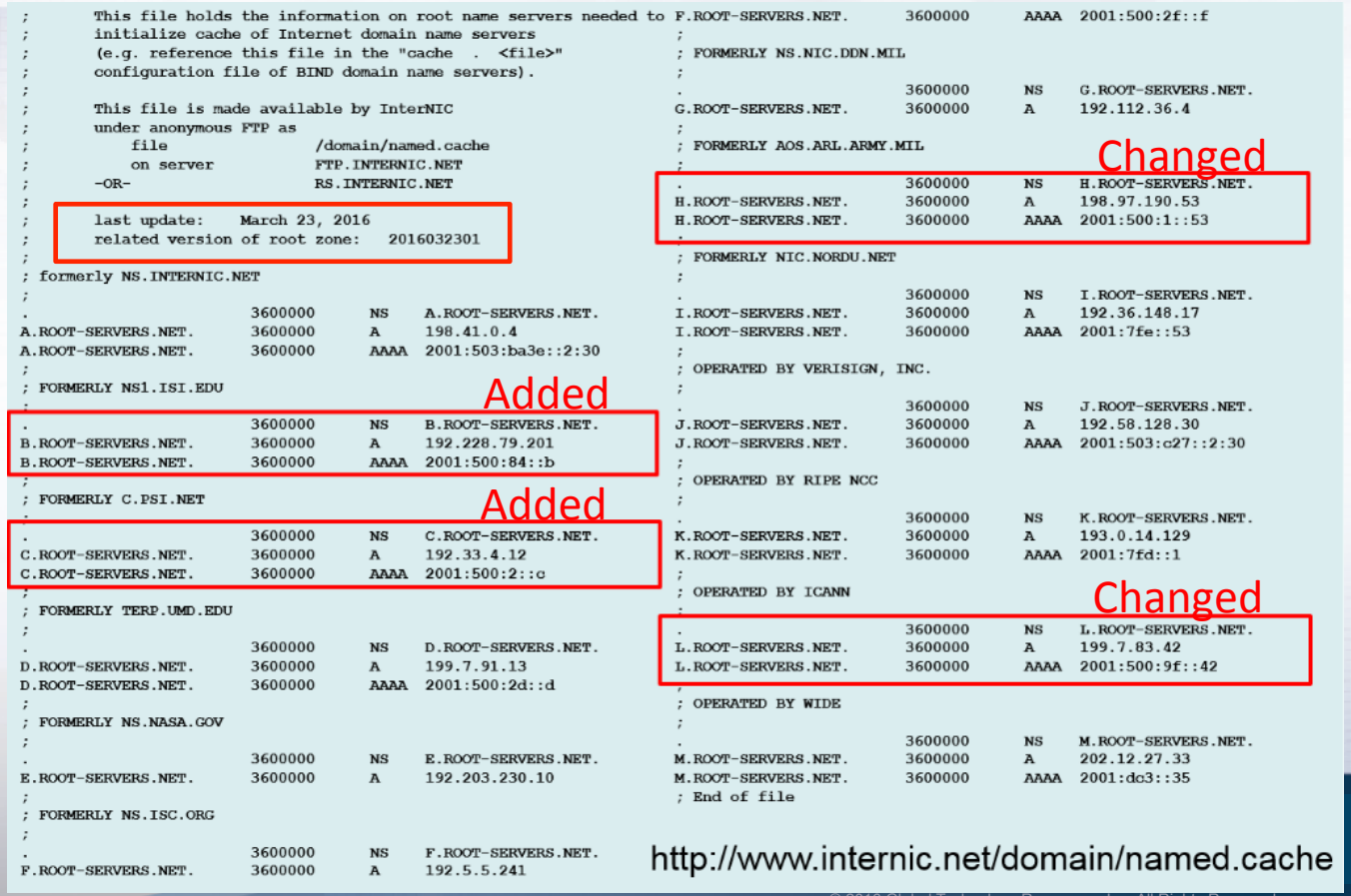
https://community.infoblox.com/t5/IPv6-Center-of-Excellence/How-Many-IPv6-Nodes-Can-You-Have-on-a-LAN/ba-p/6092

IPv6 Addressing

- Don’t force levels of hierarchy that are not needed.
- Use standard prefix lengths: /48, /56, /64
- Use nibble-boundary – Don’t use /50, /57, /65, …
- Consistency between sites can increase operational efficiency, however, not every site needs the same addressing plan.
  - Branches need a different plan than a data center “site”.
- Stick with Global Unicast Addresses (GUA) 2000::/3
  - Use these everywhere, you don’t need NAT66 (Read RFC 4864)
- Avoid Unique Local Addresses (ULA) FC00::/7 (FD00::/8)

https://community.infoblox.com/t5/IPv6-Center-of-Excellence/3-Ways-to-Ruin-Your-Future-Network-with-IPv6-Unique-Local/ba-p/6177

Dual-Protocol DNS: IPv6 Root Zone (named.root)



```
;       This file holds the information on root name servers needed to
;       initialize cache of Internet domain name servers
;       (e.g. reference this file in the "cache  .  <file>"
;       configuration file of BIND domain name servers).
;
;       This file is made available by InterNIC
;       under anonymous FTP as
;           file                /domain/named.cache
;           on server           FTP.INTERNIC.NET
;       -OR-                    RS.INTERNIC.NET
;
;       last update:    March 23, 2016
;       related version of root zone:   2016032301
;
; formerly NS.INTERNIC.NET
;
.                        3600000      NS    A.ROOT-SERVERS.NET.
A.ROOT-SERVERS.NET.      3600000      A     198.41.0.4
A.ROOT-SERVERS.NET.      3600000      AAAA  2001:503:ba3e::2:30
;
; FORMERLY NS1.ISI.EDU
;
.                        3600000      NS    B.ROOT-SERVERS.NET.
B.ROOT-SERVERS.NET.      3600000      A     192.228.79.201
B.ROOT-SERVERS.NET.      3600000      AAAA  2001:500:84::b
;
; FORMERLY C.PSI.NET
;
.                        3600000      NS    C.ROOT-SERVERS.NET.
C.ROOT-SERVERS.NET.      3600000      A     192.33.4.12
C.ROOT-SERVERS.NET.      3600000      AAAA  2001:500:2::c
;
; FORMERLY TERP.UMD.EDU
;
.                        3600000      NS    D.ROOT-SERVERS.NET.
D.ROOT-SERVERS.NET.      3600000      A     199.7.91.13
D.ROOT-SERVERS.NET.      3600000      AAAA  2001:500:2d::d
;
; FORMERLY NS.NASA.GOV
;
.                        3600000      NS    E.ROOT-SERVERS.NET.
E.ROOT-SERVERS.NET.      3600000      A     192.203.230.10
;
; FORMERLY NS.ISC.ORG
;
.                        3600000      NS    F.ROOT-SERVERS.NET.
F.ROOT-SERVERS.NET.      3600000      A     192.5.5.241
F.ROOT-SERVERS.NET.      3600000      AAAA  2001:500:2f::f
;
; FORMERLY NS.NIC.DDN.MIL
;
.                        3600000      NS    G.ROOT-SERVERS.NET.
G.ROOT-SERVERS.NET.      3600000      A     192.112.36.4
;
; FORMERLY AOS.ARL.ARMY.MIL
;
.                        3600000      NS    H.ROOT-SERVERS.NET.
H.ROOT-SERVERS.NET.      3600000      A     198.97.190.53
H.ROOT-SERVERS.NET.      3600000      AAAA  2001:500:1::53
;
; FORMERLY NIC.NORDU.NET
;
.                        3600000      NS    I.ROOT-SERVERS.NET.
I.ROOT-SERVERS.NET.      3600000      A     192.36.148.17
I.ROOT-SERVERS.NET.      3600000      AAAA  2001:7fe::53
;
; OPERATED BY VERISIGN, INC.
;
.                        3600000      NS    J.ROOT-SERVERS.NET.
J.ROOT-SERVERS.NET.      3600000      A     192.58.128.30
J.ROOT-SERVERS.NET.      3600000      AAAA  2001:503:c27::2:30
;
; OPERATED BY RIPE NCC
;
.                        3600000      NS    K.ROOT-SERVERS.NET.
K.ROOT-SERVERS.NET.      3600000      A     193.0.14.129
K.ROOT-SERVERS.NET.      3600000      AAAA  2001:7fd::1
;
; OPERATED BY ICANN
;
.                        3600000      NS    L.ROOT-SERVERS.NET.
L.ROOT-SERVERS.NET.      3600000      A     199.7.83.42
L.ROOT-SERVERS.NET.      3600000      AAAA  2001:500:9f::42
;
; OPERATED BY WIDE
;
.                        3600000      NS    M.ROOT-SERVERS.NET.
M.ROOT-SERVERS.NET.      3600000      A     202.12.27.33
M.ROOT-SERVERS.NET.      3600000      AAAA  2001:dc3::35
; End of file
```

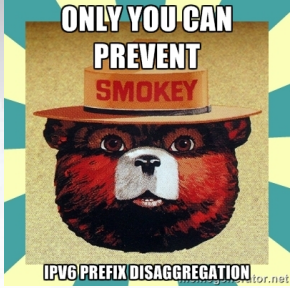
Disabling Router Advertisements in a Data Center

- As soon as IPv6 address is added to router's interface, it sends out an ICMPv6 Type 134 RA with A=1, L=1
- Sending an RA activates the IPv6 stack in all hosts on that network – that may not be what you intend or want
- In a data center environment, servers are statically addressed
  - They don't need the RA to learn their first-hop router
- Suppressing RSs and RAs allows you to turn on IPv6 one server at a time and deliberately deploy IPv6
- Consider sending RA with A=0, L=1, M=0, O=0

IPv6 Routing

- IP addressing and routing go hand-in-hand.
- All IP routing protocols now have IPv6 capabilities.
- Separating control plane for two data planes is preferred.
  - Establish BGP peer over IPv4 TCP 179 for sharing IPv4 routes
  - Establish BGP peer over IPv6 TCP 179 for sharing IPv6 routes
- Peering using global IPv6 addresses is preferred
- Don’t forget to use a 32-bit RID in the IPv6 routing process.
- Bidirectional Forwarding Detection (BFD) is now available in may routing protocols on many platforms
- Consider using locally-administered link-local addresses.
  - fe80::cccc:0001, fe80::dddd:0002, …

IPv6 Prefix Disaggregation

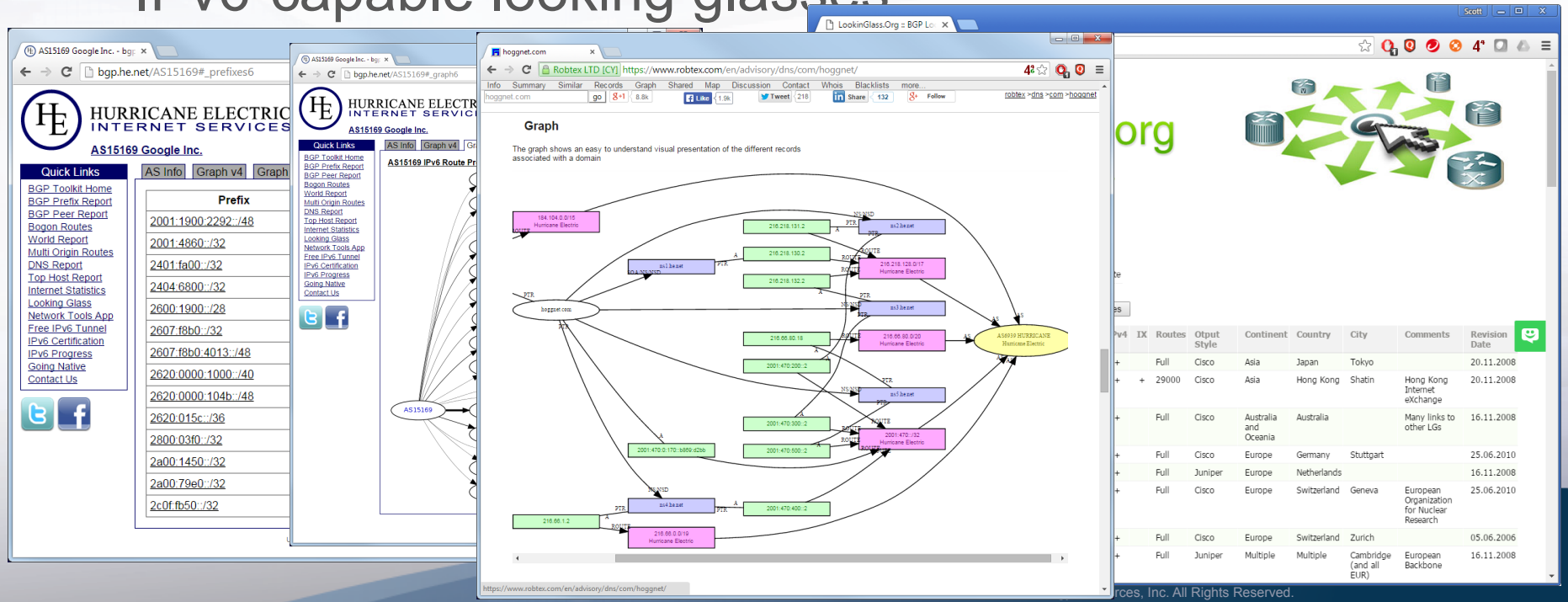


- If you have been allocated a /32, /36, /40, /44, congratulations, now advertise it with BGP
- Do not disaggregate that prefix into many /48s
- There is little reason to advertise more-specific /48s
  - Anycast (e.g. root name servers)
  - Testing, Monitoring
- Disaggregation is bad form and shows your lack of expertise and coolness

https://community.infoblox.com/t5/IPv6-Center-of-Excellence/Only-You-Can-Prevent-IPv6-Prefix-Disaggregation/ba-p/4201

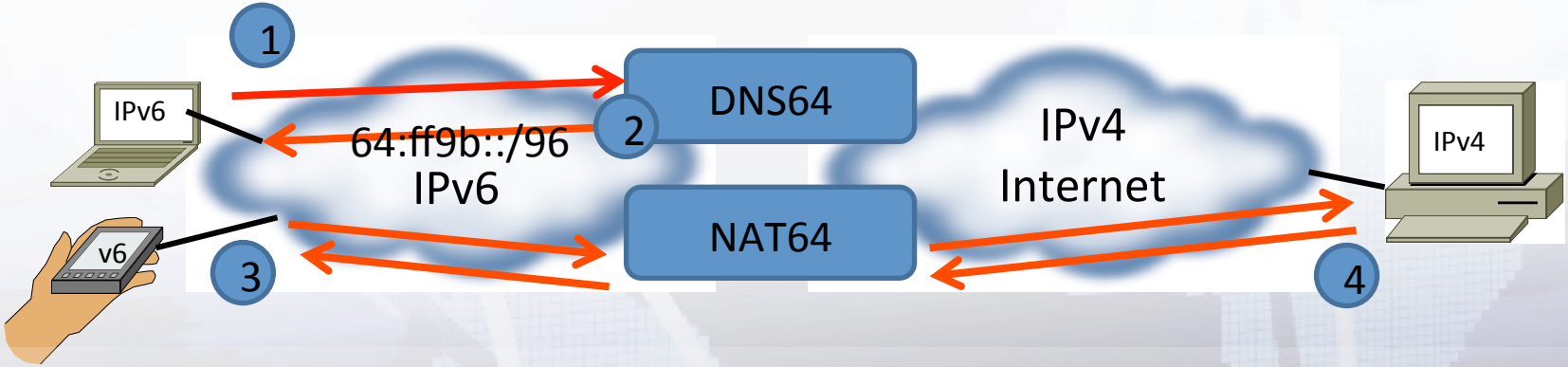
IPv6 BGP Looking Glasses

- View yourself from the Internet-perspective using IPv6-capable looking glasses



DNS64 & NAT64 Operations

- Translation technique that uses NAT and DNS in combination
- Uses the IPv6 prefix 64:ff9b::/96 to map the 32-bit IPv4 DNS responses into an IPv6 address that an IPv6-only client can use
- DNS64 (RFC 6147) and NAT64 (RFC 6146) “appears” to share state information about connections but they do not directly share state information



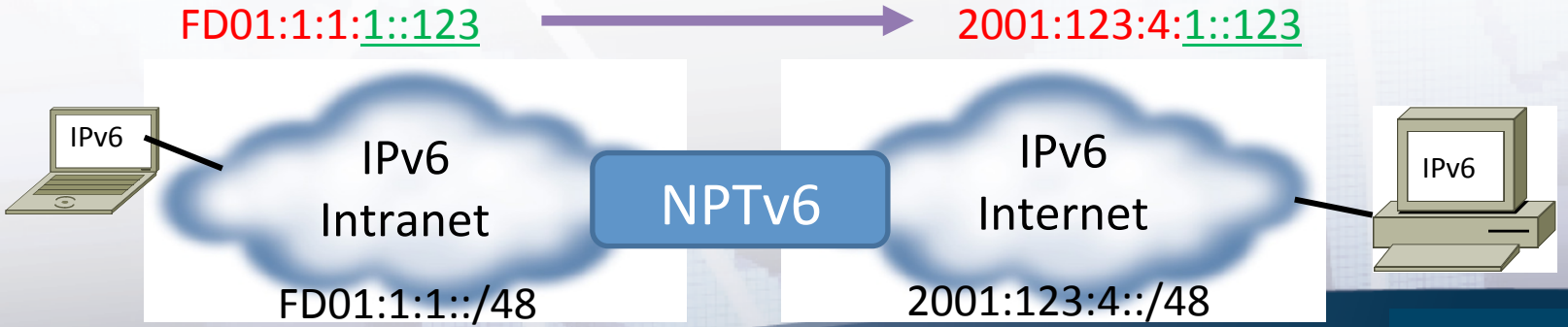
http://www.networkworld.com/article/2231256/cisco-subnet/testing-nat64-and-dns64.html

DNS64 & NAT64 Drawbacks

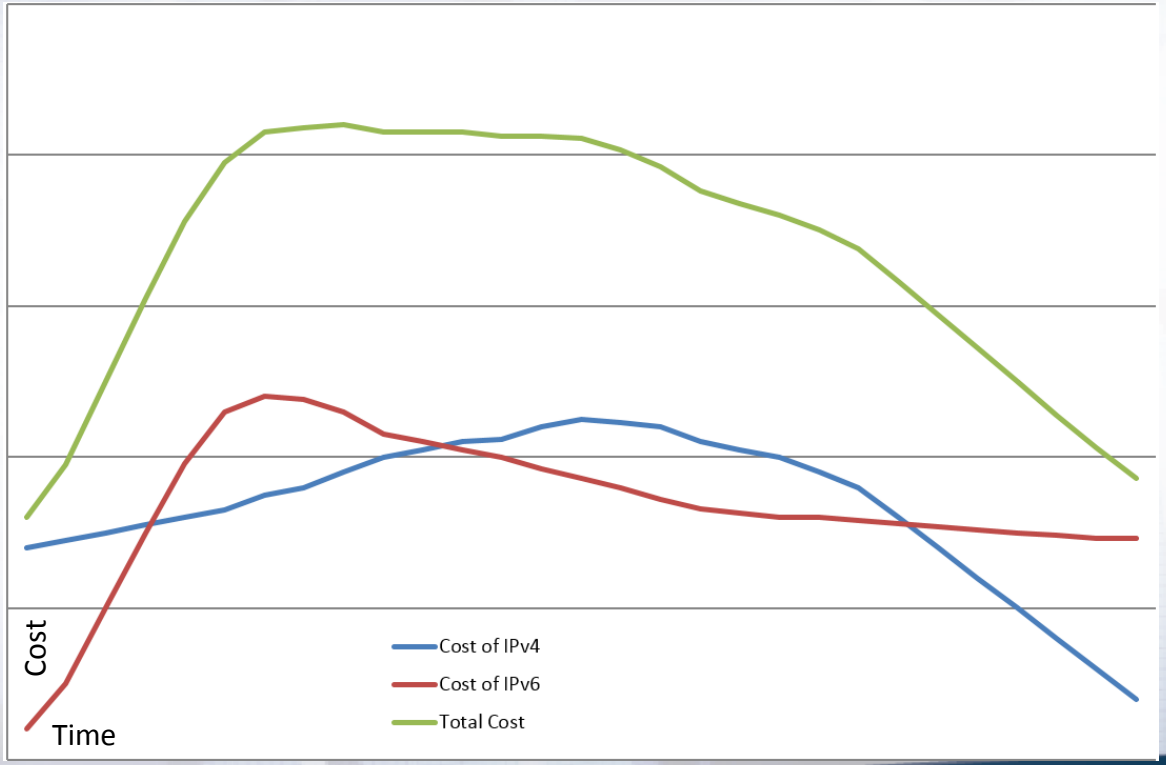
- NAT64 only works for applications that start with a DNS lookup
- Breaks with applications that embed an IPv4 address in the URL (embedded literals) (e.g. http://192.168.1.1/)
- Breaks for domains that have IPv4-only DNSSEC signed domains
- ~15% of applications break with IPv6 native or break with NAT64
  - Spotify, WhatsApp, Skype, SIP, RTSP, H.323, XMPP peer to peer
  - http://tinyurl.com/nat64-breakage

NAPT/NPT/NPTv6 (NAT66)

- Network Prefix Translation (NPTv6) (RFC 6296)
- Translates source address from one prefix to another prefix
- Remainder of IPv6 address remains the same
- 1:1 mapping – stateless
- Some networking/perimeter products support this today



Dual Stack Increases OPEX Costs



http://www.networkworld.com/article/2222870/cisco-subnet/dual-stack-will-increase-operating-expenses.html

Benefits of IPv6-only

- Reduced OPEX costs by running only a single IP protocol
- IPv6 addressing (operations) is simpler
  - No NAT makes everything better
- Reduced dependence on increasingly expensive IPv4 addresses
  - If you know you are going to need more IPv4, buy it now
  - Sell your public IPv4 at the peak price
- No need to purchase and maintain CGN/LSN systems
- In some cases IPv6 performs better than IPv4

https://community.infoblox.com/t5/IPv6-Center-of-Excellence/IPv4-Address-Trading-for-Fun-and-Profit/ba-p/3496

Is IPv6 Faster Than IPv4?

- There are now several studies analyzing if IPv6 is faster than IPv4.
  - Google's 2010 paper titled "Evaluating IPv6 Adoption in the Internet"
- Geoff Huston of APNIC at NANOG 66
  - 6to4 and Teredo are responsible for most of the connection failures
  - He concluded that native IPv6 can be as-fast as IPv4
- Paul Saab at Facebook has shows data from Mobile Proxygen that shows IPv6 is faster for them.
  - "Facebook says it has seen users' News Feeds loading 20 percent to 40 percent faster on mobile devices using IPv6".
- Hurricane Electric (HE) Global IPv6 Deployment Progress Report
  - "Percentage of IPv6 rDNS Nameservers where IPv6 is as fast or faster than IPv4 (within 1ms): 74.9%"

Dual-Protocol Applications

- Assessing current code for IPv6-capability
  - Most applications do not create socket-level connections.
  - Most applications use higher-level APIs or rely on lower-level web services for connectivity.
- Create code that is Address-Family (AF) independent.
- Presentation-to-Numeric (p2n) & Numeric-to-Presentation (n2p)
  - Robustness principle: Be conservative in what you send, be liberal in what you accept.
- Be careful of data structures for storing 128-bit addresses.
- Create code that performs dual-protocol DNS resolution and incorporates Happy Eyeballs (RFC 6555).
- Write code that properly handles Path MTU Discovery (PMTUD).

IPv6 Security Considerations

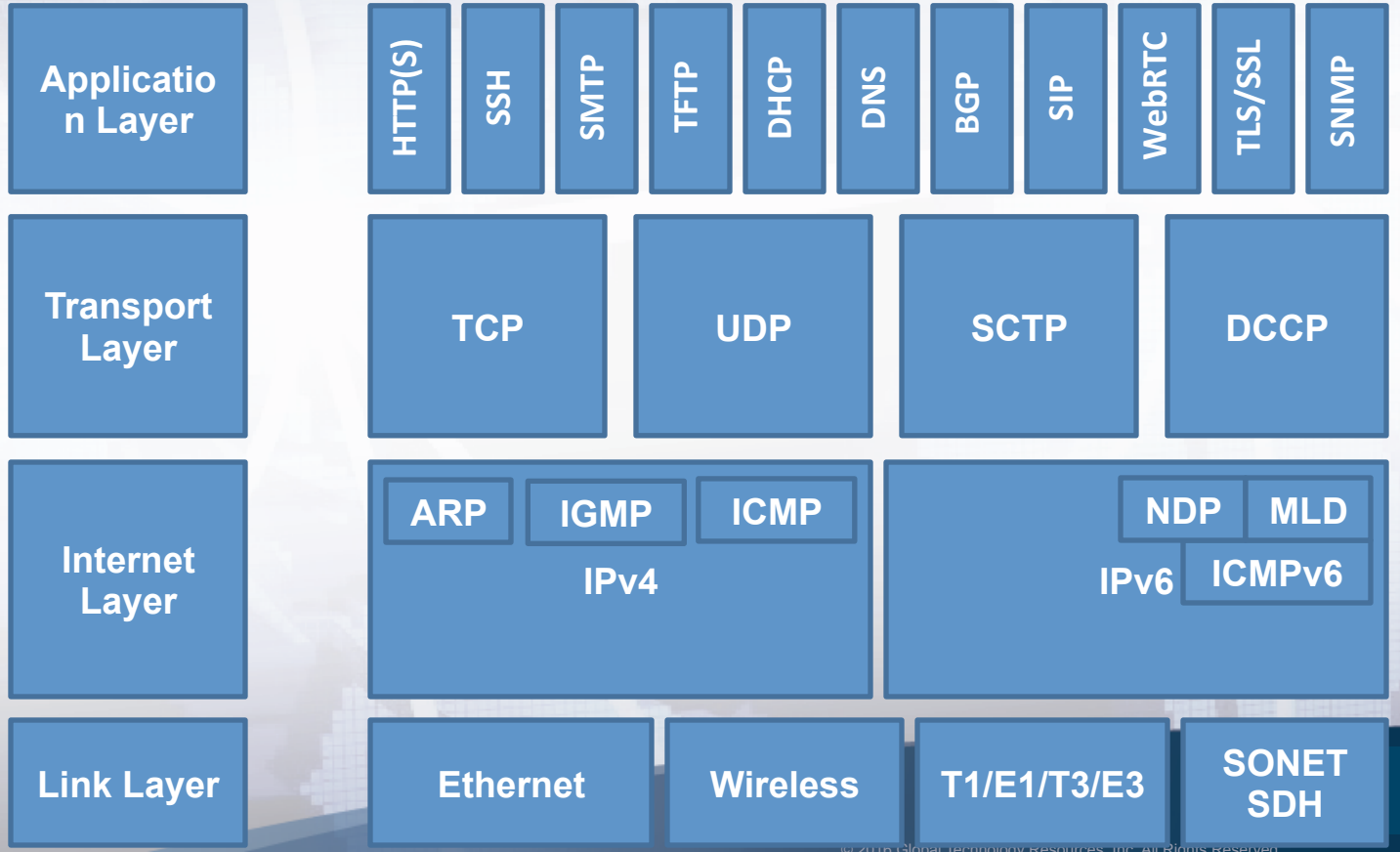
- Understand how IPv4 and IPv6 are different in terms of networking (NDP, extension headers, dynamic tunnels)
- Learn how IPv6 NDP, link-local, and multicast packets are already on your network
- Don't deploy IPv6 if you lack the products to secure the protocol properly.
- Don't be overly worried about IPv6 NDP security weaknesses.
  - You likely haven't secured ARP on your IPv4 LANs either.

https://community.infoblox.com/blogs/2015/02/10/holding-ipv6-neighbor-discovery-higher-standard-security

Troubleshooting Dual Protocol Networks

- Even if you do not deploy IPv6, there could still be IPv6-related issues that you must deal with.
- You now have IPv6-enabled nodes in your environment.
- Using a disciplined troubleshooting methodology will pay dividends when dealing with multi-part problems.
- Keep your IT infrastructure documentation constantly updated as you deploy IPv6.
  - Readily available docs will reduce your MTTR.
- Troubleshoot IPv6 in segments (LAN1, WAN, LAN2).

Troubleshooting Dual Protocol Networks



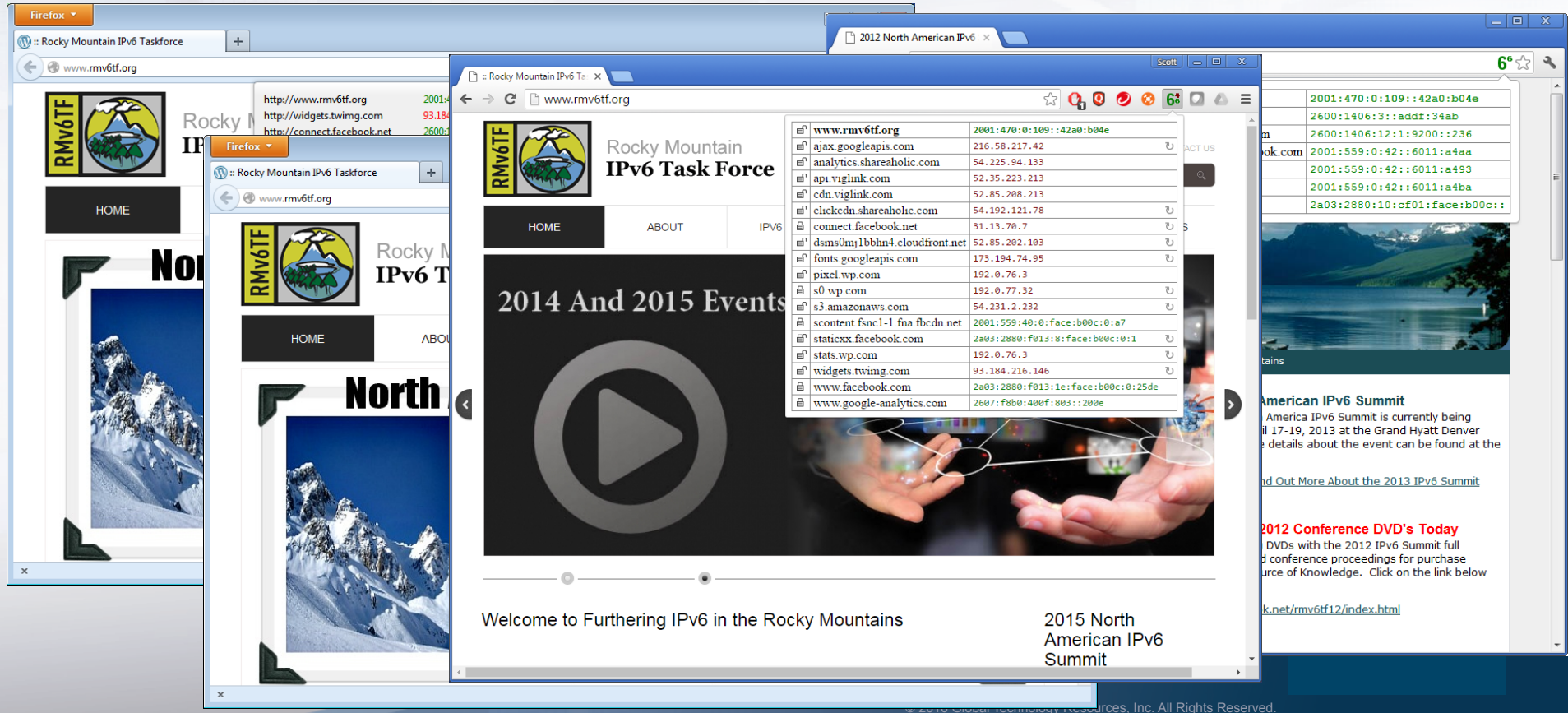
Troubleshooting Dual Protocol Networks

- Our network management and operations systems must be dual-protocol capable and give us visibility to IPv6.
- Troubleshooting NDP requires a magnifying lens.
  - You may need to break out the protocol analyzer.
  - Looking for an IPv6 needle in a haystack of IPv4.
- We want test systems to automatically check both protocols in parallel.

Troubleshooting Dual Protocol Networks

- Different applications and different OSs create dual-protocol connections in different methods.
  - Happy Eyeballs, RFC 6555, Microsoft NCSI, Apple Mac OS X & iOS
- Some connections could use IPv4 and/or IPv6.
  - Web pages could be delivered over a combination of protocols. How do you know which protocol was used?
  - IPv6 Browser add-ons, plug-ins can be helpful
    - IPvFox (Firefox), SixOrNot, IPvFoo (Chrome)

IPv6-Troubleshooting Browser Additions



IPv6 Roses and Thorns

- What we are doing right?
  - IPv6 support now exists in most products
  - Mobile and residential subscribers now using IPv6
  - IPv6 adoption doubling every year, hockey-stick graphs
- What could we do better?
  - Corporate enterprise internal access networks
  - SMB dependence on PA IPv6 addresses without NPTv6/NAT66
  - Content providers – we need more dual-protocol web sites
  - Cloud services need IPv6 (AWS, MS Azure, Google)
  - Better Geolocation data for IPv6
  - Security Reputation data for IPv6

Thank You!

Scott Hogg, CTO GTRI

shogg at gtri.com, @scotthogg