IPv6-only Deployment in Broadband and Cellular Networks IPv4 as-a-Service

RIPE77

October, 2018

Amsterdam, Netherlands

@JordiPalet

(jordi.palet@theipv6company.com)

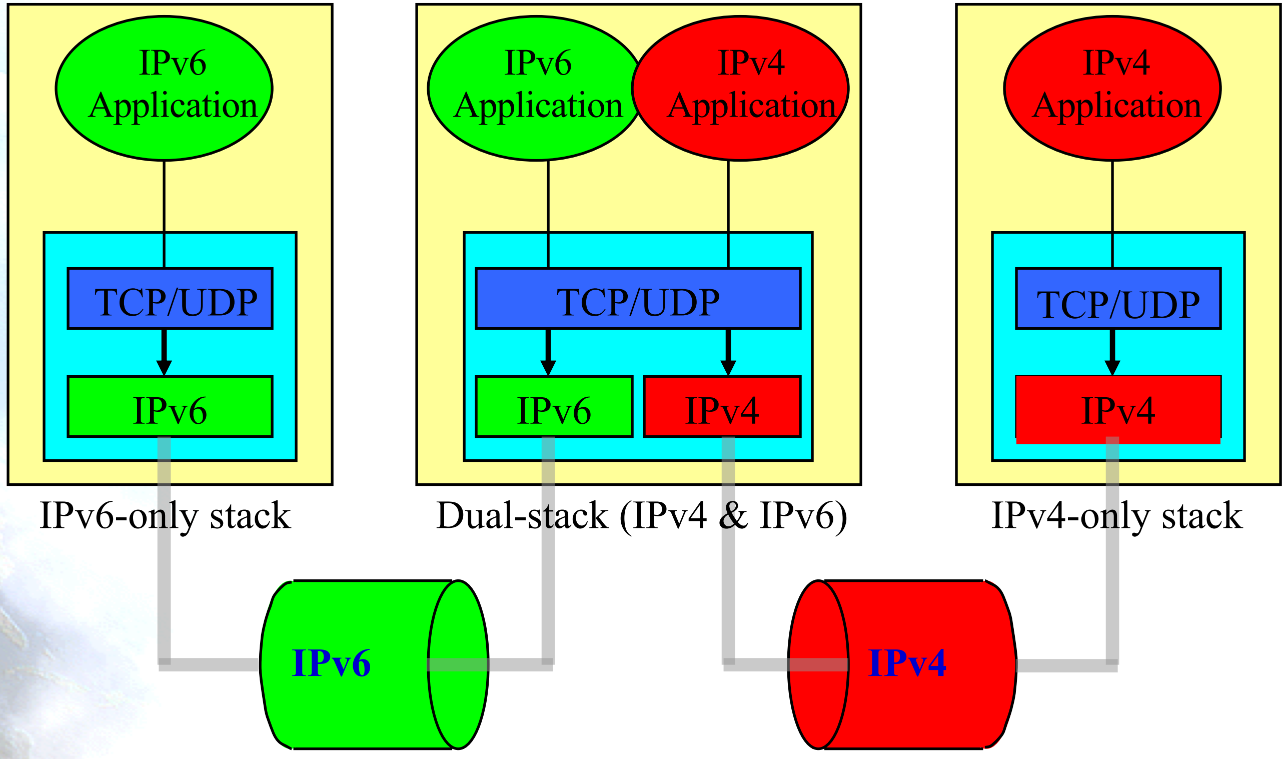
Transition / Co-Existence Techniques

- IPv6 has been designed for easing the transition and coexistence with IPv4
- Several strategies have been designed and implemented for coexisting with IPv4 hosts, grouped in three categories:
  - Dual stack: Simultaneous support for both IPv4 and IPv6 stacks
  - Tunnels: IPv6 packets encapsulated in IPv4 ones
    - This has been the commonest choice
    - Today expect IPv4 packets in IPv6 ones!
  - Translation: Communication of IPv4-only and IPv6-only. Initially discouraged and only “last resort” (imperfect). Today no other choice!
- Expect to use them in combination!

Dual-Stack Approach

- When adding IPv6 to a system, do not delete IPv4
  - This multi-protocol approach is familiar and well-understood (e.g., for AppleTalk, IPX, etc.)
  - In the majority of the cases, IPv6 is be bundled with all the OS release, not an extra-cost add-on
- Applications (or libraries) choose IP version to use
  - when initiating, based on DNS response:
    - if (dest has AAAA record) use IPv6, else use IPv4
  - when responding, based on version of initiating packet
- This allows indefinite co-existence of IPv4 and IPv6, and gradual app-by-app upgrades to IPv6 usage
- A6 record is experimental

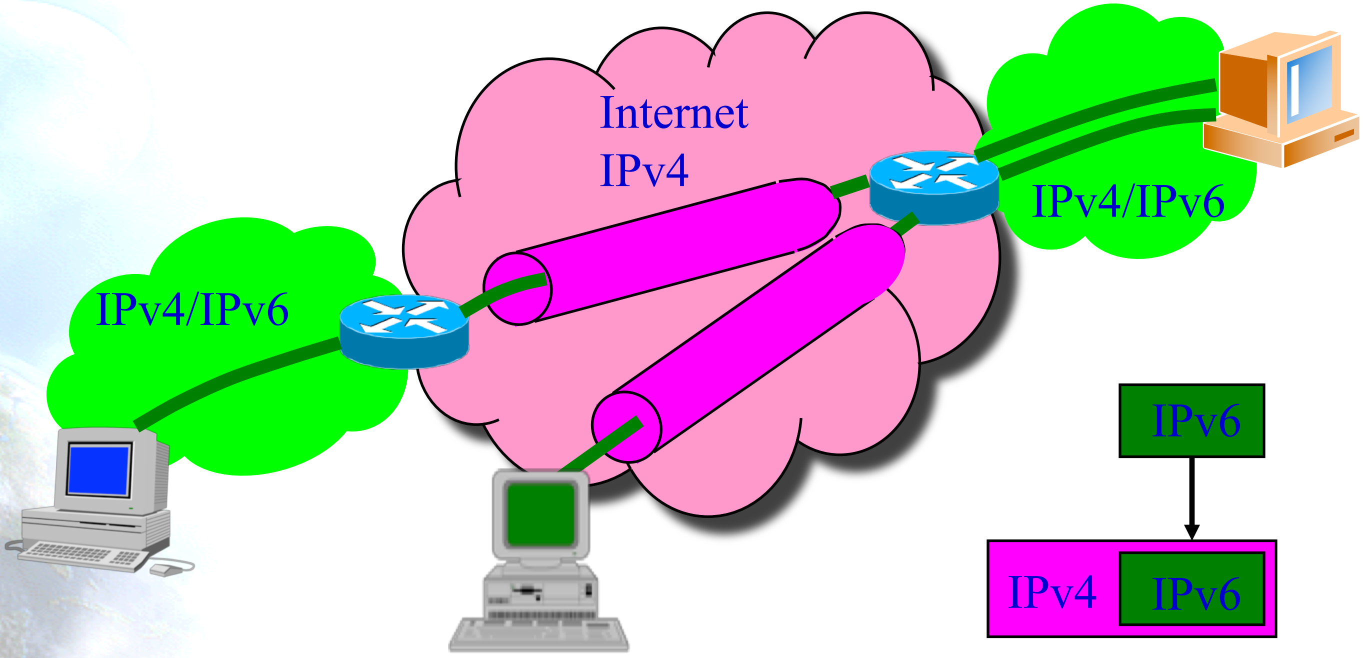
Dual-Stack Approach

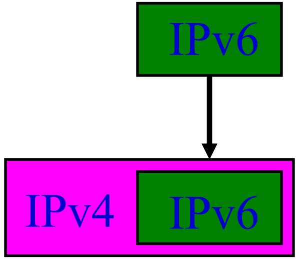


Tunnels to Get Through IPv6-Ignorant Routers

- Encapsulate IPv6 packets inside IPv4 packets (or MPLS frames) in order to provide IPv6 connectivity through IPv4-only networks
- Many methods exist for establishing tunnels:
  - manual configuration
  - “tunnel brokers” (using web-based service to create a tunnel)
  - “6over4” (intra-domain, using IPv4 multicast as virtual LAN)
  - “6to4” (inter-domain, using IPv4 addr as IPv6 site prefix)
- Can view this as:
  - IPv6 using IPv4 as a virtual link-layer, or
  - an IPv6 VPN (virtual public network), over the IPv4 Internet (becoming “less virtual” over time, we hope)

Tunnels IPv6 in IPv4





Translation IPv4/IPv6

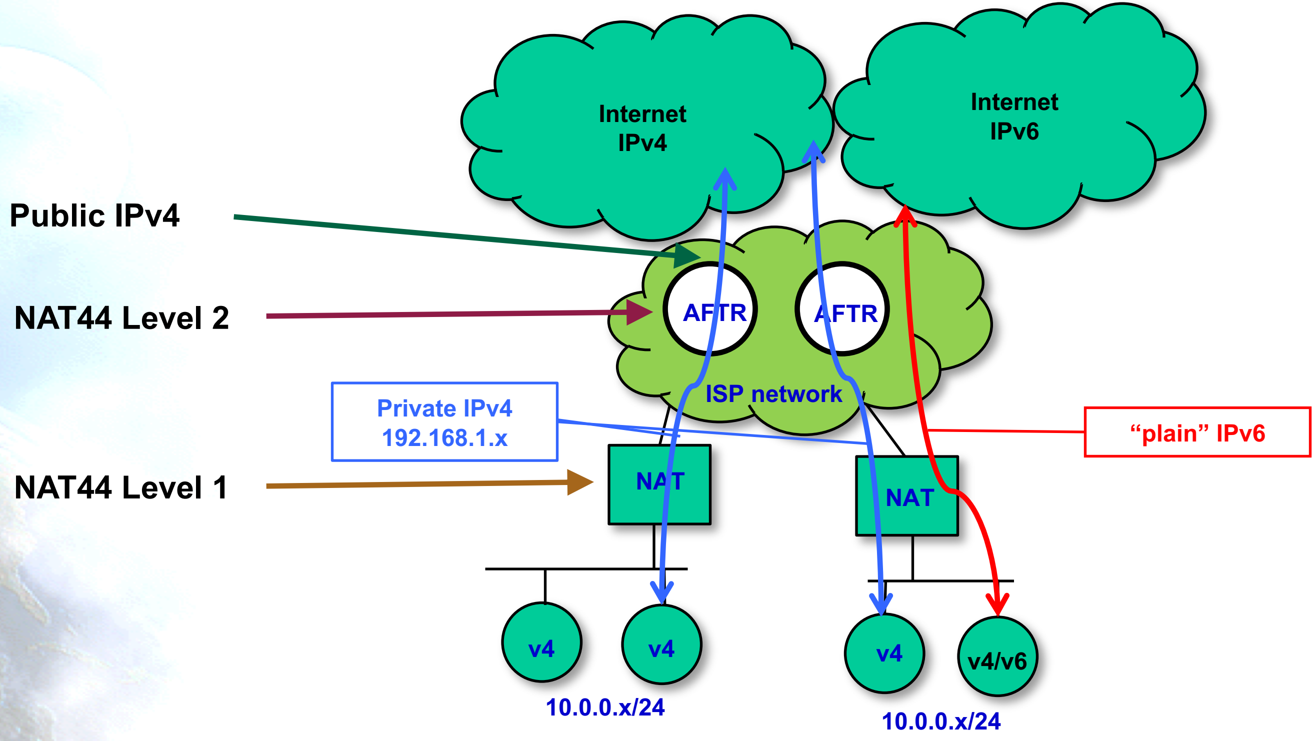
- May prefer to use IPv6-IPv4 protocol translation for:
  - new kinds of Internet devices (e.g., cell phones, cars, appliances)
  - benefits of shedding IPv4 stack (e.g., serverless autoconfig)

- This is a simple extension to NAT techniques, to translate header format as well as addresses
  - IPv6 nodes behind a translator get full IPv6 functionality when talking to other IPv6 nodes located anywhere
  - they get the normal (i.e., degraded) NAT functionality when talking to IPv4 devices
  - methods used to improve NAT functionality (e.g, RSIP) can be used equally to improve IPv6-IPv4 functionality

IPv6 Transition Mechanisms

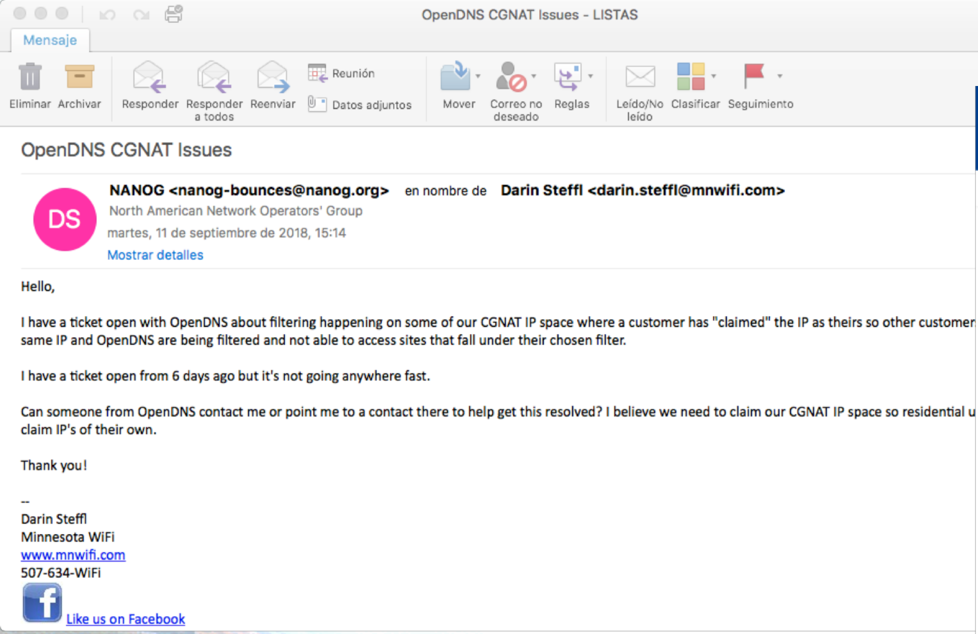
- Some transition mechanism based on tunnels and/or translation:
  - 6in4 [6in4]
  - TB [TB]
  - TSP [TSP]
  - 6to4 [6to4]
  - Teredo [TEREDO], [TEREDOC]
  - Túneles automáticos [TunAut]
  - …
  - ISATAP [ISATAP]
  - 6over4 [6over4]
  - Softwires
  - 6RD
  - NAT64
  - DS-Lite
  - lw4o6
  - 464XLAT
  - MAP E/T
  - …

NAT444


Internet IPv4
Internet IPv6
Public IPv4
NAT44 Level 2
AFTR
AFTR
ISP network
Private IPv4
192.168.1.x
"plain" IPv6
NAT44 Level 1
NAT
NAT
v4
v4
v4
v4/v6
10.0.0.x/24
10.0.0.x/24

CGN breaks …

- UPnP-IGD (Universal Plug & Play - Internet Gateway Device protocol)
- NAT-PMP (NAT Port Mapping Protocol)
- Other NAT Traversal mechs
- Security
- AJAX (Asyncronous Javascript And XML)
- FTP (big files)
- BitTorrent/Limewire (seeding – uploading)
- On-line gaming
- Video streaming (Netflix, Hulu, …)
- IP cameras
- Tunnels, VPN, IPsec, ...
- VoIP
- Port forwarding
- ...



OpenDNS CGNAT Issues

DS **NANOG <nanog-bounces@nanog.org>** en nombre de **Darin Steffl <darin.steffl@mnwifi.com>**
North American Network Operators' Group
martes, 11 de septiembre de 2018, 15:14
Mostrar detalles

Hello,

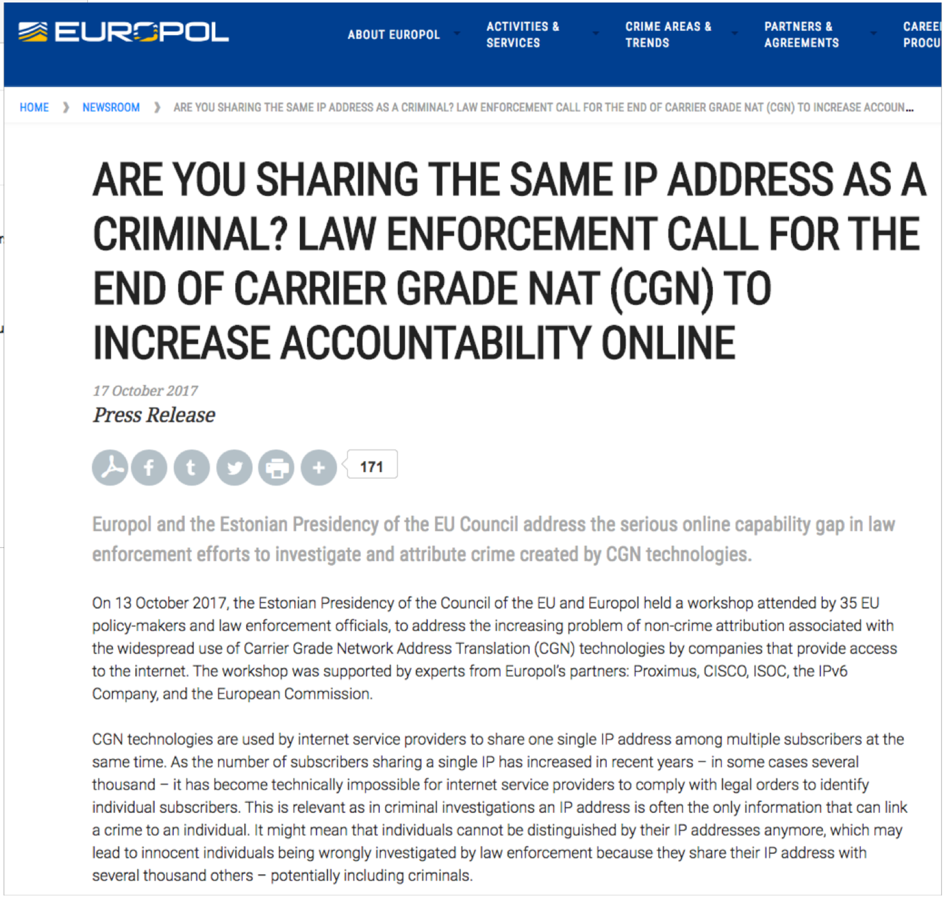
I have a ticket open with OpenDNS about filtering happening on some of our CGNAT IP space where a customer has "claimed" the IP as theirs so other customer
same IP and OpenDNS are being filtered and not able to access sites that fall under their chosen filter.

I have a ticket open from 6 days ago but it's not going anywhere fast.

Can someone from OpenDNS contact me or point me to a contact there to help get this resolved? I believe we need to claim our CGNAT IP space so residential u
claim IP's of their own.

Thank you!

--
Darin Steffl
Minnesota WiFi
www.mnwifi.com
507-634-WiFi
Like us on Facebook


EUROPOL
ABOUT EUROPOL
ACTIVITIES & SERVICES
CRIME AREAS & TRENDS
PARTNERS & AGREEMENTS
CAREE PROCU
HOME › NEWSROOM › ARE YOU SHARING THE SAME IP ADDRESS AS A CRIMINAL? LAW ENFORCEMENT CALL FOR THE END OF CARRIER GRADE NAT (CGN) TO INCREASE ACCOUN...
171

ARE YOU SHARING THE SAME IP ADDRESS AS A CRIMINAL? LAW ENFORCEMENT CALL FOR THE END OF CARRIER GRADE NAT (CGN) TO INCREASE ACCOUNTABILITY ONLINE

17 October 2017

Press Release

Europol and the Estonian Presidency of the EU Council address the serious online capability gap in law enforcement efforts to investigate and attribute crime created by CGN technologies.

On 13 October 2017, the Estonian Presidency of the Council of the EU and Europol held a workshop attended by 35 EU policy-makers and law enforcement officials, to address the increasing problem of non-crime attribution associated with the widespread use of Carrier Grade Network Address Translation (CGN) technologies by companies that provide access to the internet. The workshop was supported by experts from Europol's partners: Proximus, CISCO, ISOC, the IPv6 Company, and the European Commission.

CGN technologies are used by internet service providers to share one single IP address among multiple subscribers at the same time. As the number of subscribers sharing a single IP has increased in recent years – in some cases several thousand – it has become technically impossible for internet service providers to comply with legal orders to identify individual subscribers. This is relevant as in criminal investigations an IP address is often the only information that can link a crime to an individual. It might mean that individuals cannot be distinguished by their IP addresses anymore, which may lead to innocent individuals being wrongly investigated by law enforcement because they share their IP address with several thousand others – potentially including criminals.

We don’t have IPv4 …

- **IPv4 exhaustion avoids**
  - **Assigning IPv4 to end-users**
  - **Assigning IPv4 even in public networks**
  - **Keep scalable interoperability with IPv4-only networks**
- **Consequence: In many cases, we need to deploy IPv6-only networks**
  - OpEx
  - No IPv4 resources (CapEx if you buy them)
  - Performance
  - Efficiency
  - RFCs
  - Other issues …

Dual Stack Lite (DS-Lite)

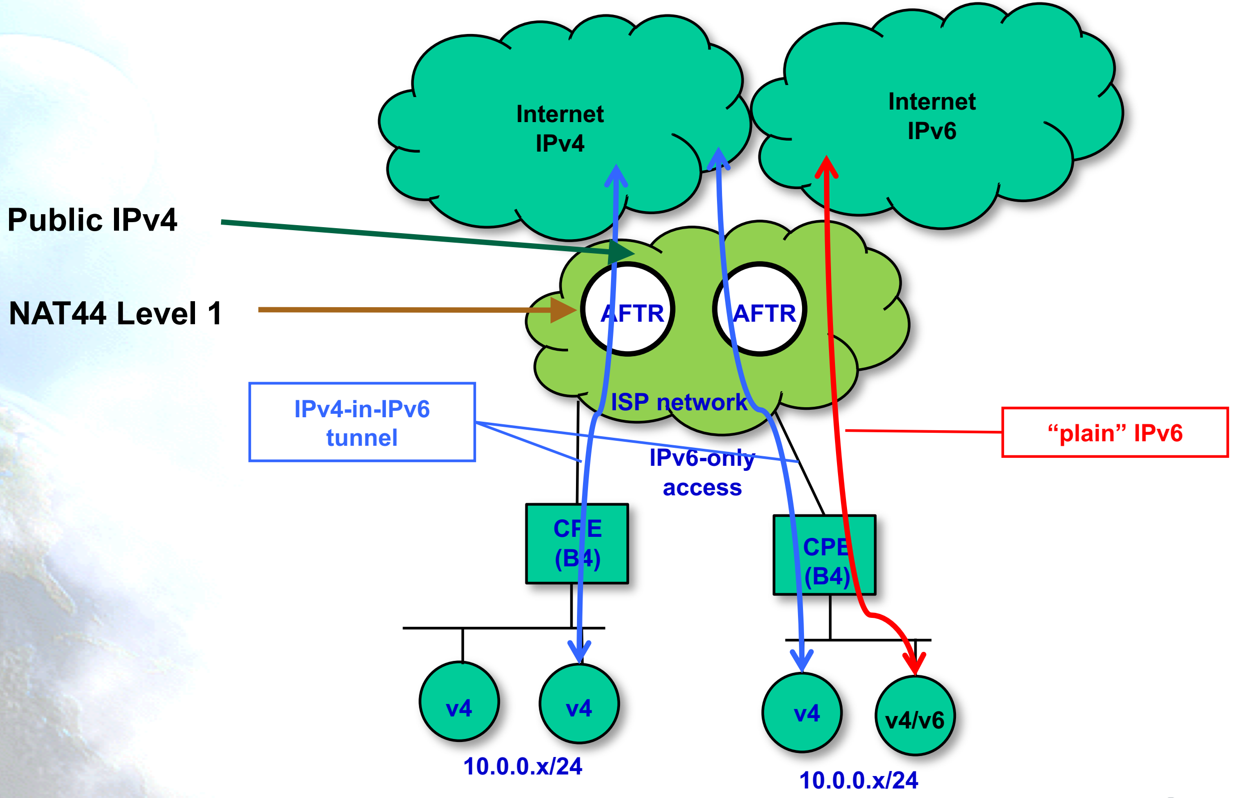
- To cope with the IPv4 exhaustion problem.

- Sharing (same) IPv4 addresses among customers by combining:
  - Tunneling
  - NAT

- No need for multiple levels of NAT.

- Two elements:
  - DS-Lite Basic Bridging BroadBand (B4)
  - DS-Lite Address Family Transition Router (AFTR)
    - Also called CGN (Carrier Grade NAT) or LSN (Large Scale NAT)

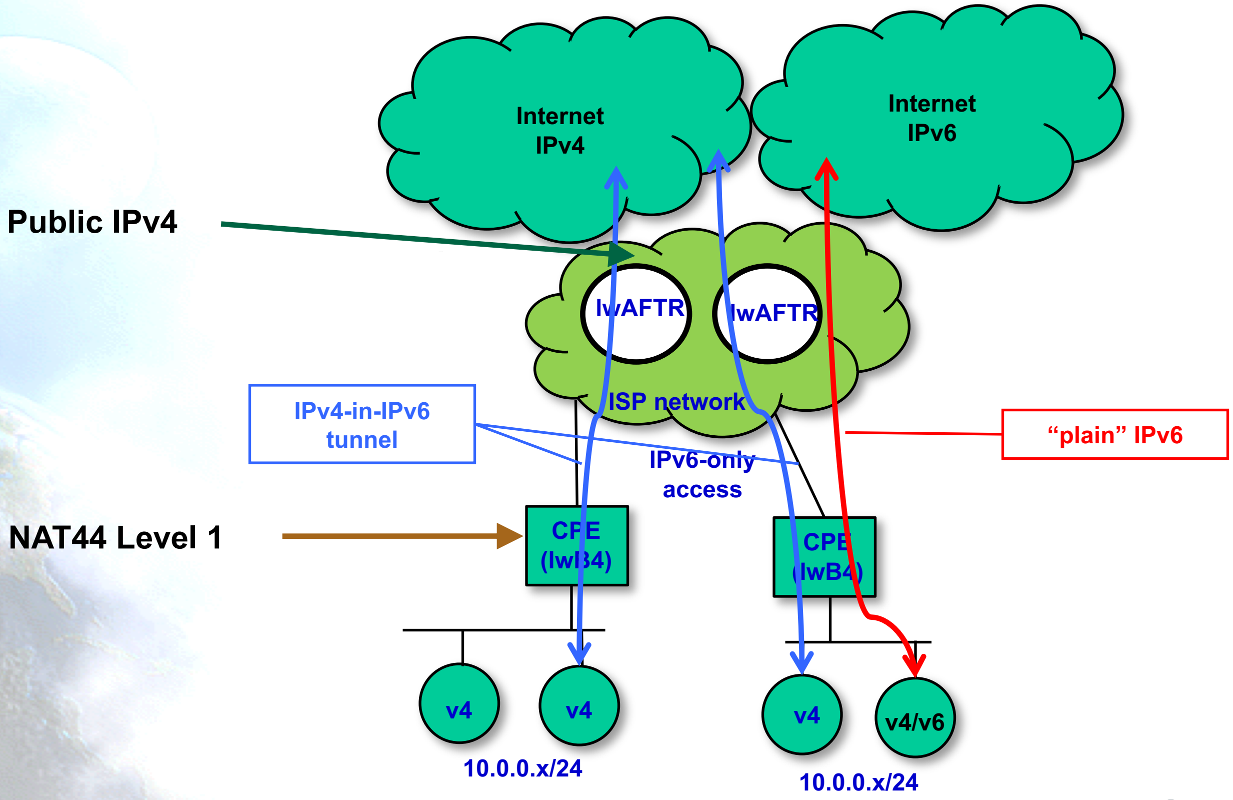
DS-Lite


Internet
IPv4
Internet
IPv6
Public IPv4
NAT44 Level 1
AFTR
AFTR
ISP network
IPv4-in-IPv6
tunnel
IPv6-only
access
"plain" IPv6
CPE
(B4)
CPE
(B4)
v4
v4
v4
v4/v6
10.0.0.x/24
10.0.0.x/24

Lightweight 4over6 (lw4o6)

- Similar to DS-Lite -> Changes NAT location
  - – Better scalability
  - – Reduces logging
- Sharing SAME IPv4 addresses among several customers, combining:
  - Tunneling
  - NAT
- No need for multiple levels of NAT
- Two elements:
  - Lw Basic Bridging BroadBand (lwB4) - CPE
  - Lw Address Family Transition Router (lwAFTR)

lw4o6


Internet
IPv4
Internet
IPv6
Public IPv4
lwAFTR
lwAFTR
ISP network
IPv4-in-IPv6
tunnel
"plain" IPv6
IPv6-only
access
NAT44 Level 1
CPE
(lwB4)
CPE
(lwB4)
v4
v4
v4
v4/v6
10.0.0.x/24
10.0.0.x/24

NAT64 (1)

- When ISPs only provide IPv6 connectivity, or devices are IPv6-only (cellular phones)
- But still some IPv4-only boxes are on the Internet
- Similar idea as NAT-PT, but working correctly
- Optional element, but decoupled, DNS64
- Good solution if IPv4 is not required at the client
  - Client is IPv6-only
- Some apps don’t work (Skype …)
  - Peer-to-peer using IPv4 “references”
  - Literal addresses
  - Socket APIs

NAT64 (2)

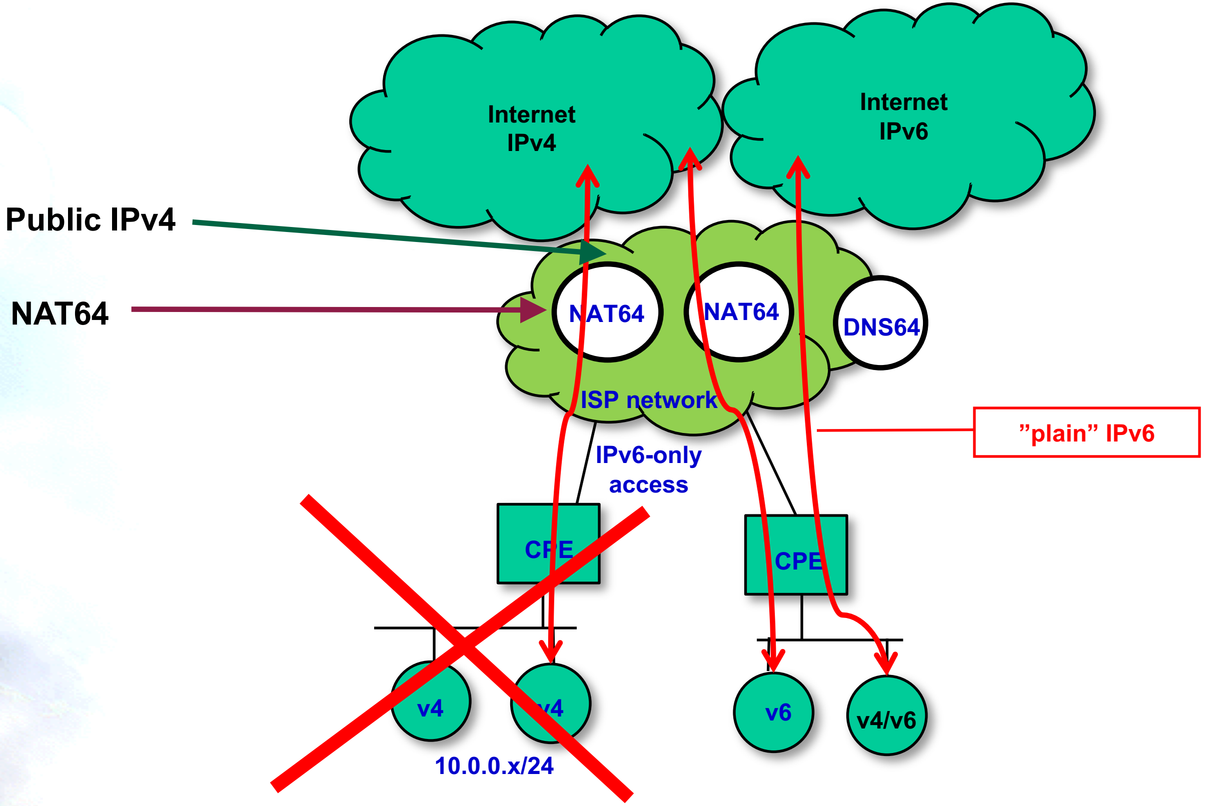
- Stateful NAT64 is a mechanism for translating IPv6 packets to IPv4 packets and vice-versa
  - The translation is done by translating the packet headers according to the IP/ICMP Translation Algorithm.
  - The IPv4 addresses of IPv4 hosts are algorithmically translated to and from IPv6 addresses by using a specific algorithm.
  - The current specification only defines how stateful NAT64 translates unicast packets carrying TCP, UDP and ICMP traffic.
  - DNS64 is a mechanism for synthesizing AAAA resource records (RR) from A RR. The IPv6 address contained in the synthetic AAAA RR is algorithmically generated from the IPv4 address and the IPv6 prefix assigned to a NAT64 device

- NAT64 allows multiple IPv6-only nodes to share an IPv4 address to access the IPv4 Internet

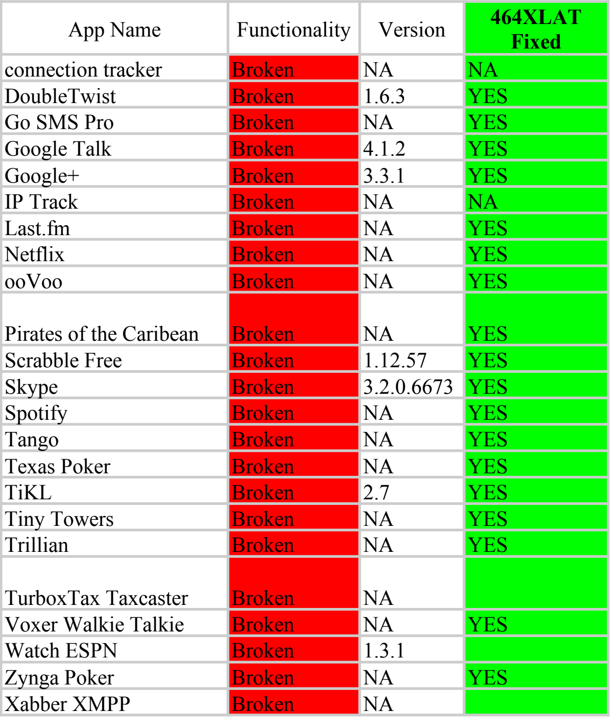
NAT64 (3)

- It’s known that there are things that doesn’t work:
  - Everything out of TCP,UDP, or ICMP: Multicast, Stream Control Transmission Protocol (SCTP), the Datagram Congestion Control Protocol (DCCP), and IPSEC
  - Applications that carry layer 3 information in the application layer: FTP [RFC6384], SIP/H323
  - Some apps: online gaming, skype, etc.
    - Peer-to-peer using IPv4 “references”
  - Literal addresses
  - Socket APIs

NAT64 (4)


Internet
IPv4
Internet
IPv6
Public IPv4
NAT64
NAT64
NAT64
DNS64
ISP network
IPv6-only
access
”plain” IPv6
CPE
CPE
v4
v4
10.0.0.x/24
v6
v4/v6

NAT64 breaks …



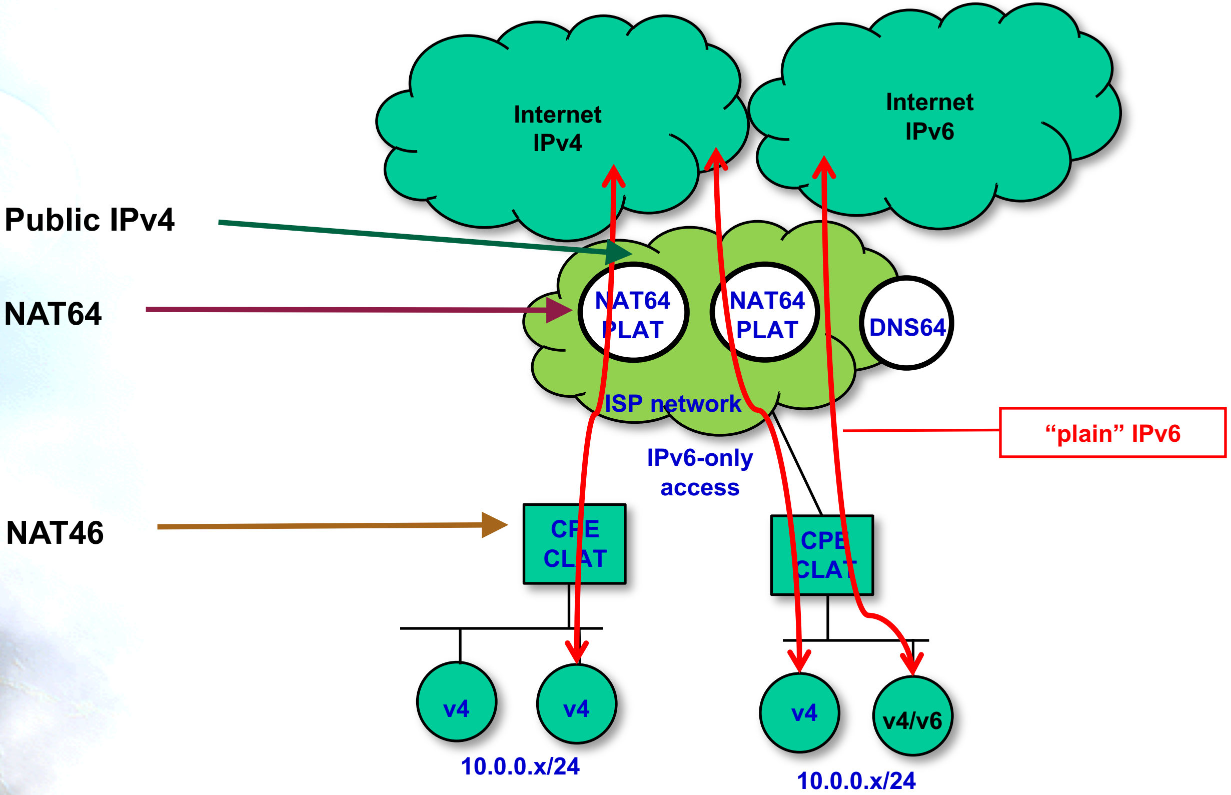
App Name	Functionality	Version	**464XLAT Fixed**
connection tracker	Broken	NA	NA
DoubleTwist	Broken	1.6.3	YES
Go SMS Pro	Broken	NA	YES
Google Talk	Broken	4.1.2	YES
Google+	Broken	3.3.1	YES
IP Track	Broken	NA	NA
Last.fm	Broken	NA	YES
Netflix	Broken	NA	YES
ooVoo	Broken	NA	YES
Pirates of the Caribean	Broken	NA	YES
Scrabble Free	Broken	1.12.57	YES
Skype	Broken	3.2.0.6673	YES
Spotify	Broken	NA	YES
Tango	Broken	NA	YES
Texas Poker	Broken	NA	YES
TiKL	Broken	2.7	YES
Tiny Towers	Broken	NA	YES
Trillian	Broken	NA	YES
TurboTax Taxcaster	Broken	NA	
Voxer Walkie Talkie	Broken	NA	YES
Watch ESPN	Broken	1.3.1	
Zynga Poker	Broken	NA	YES
Xabber XMPP	Broken	NA	

***T-Mobile**

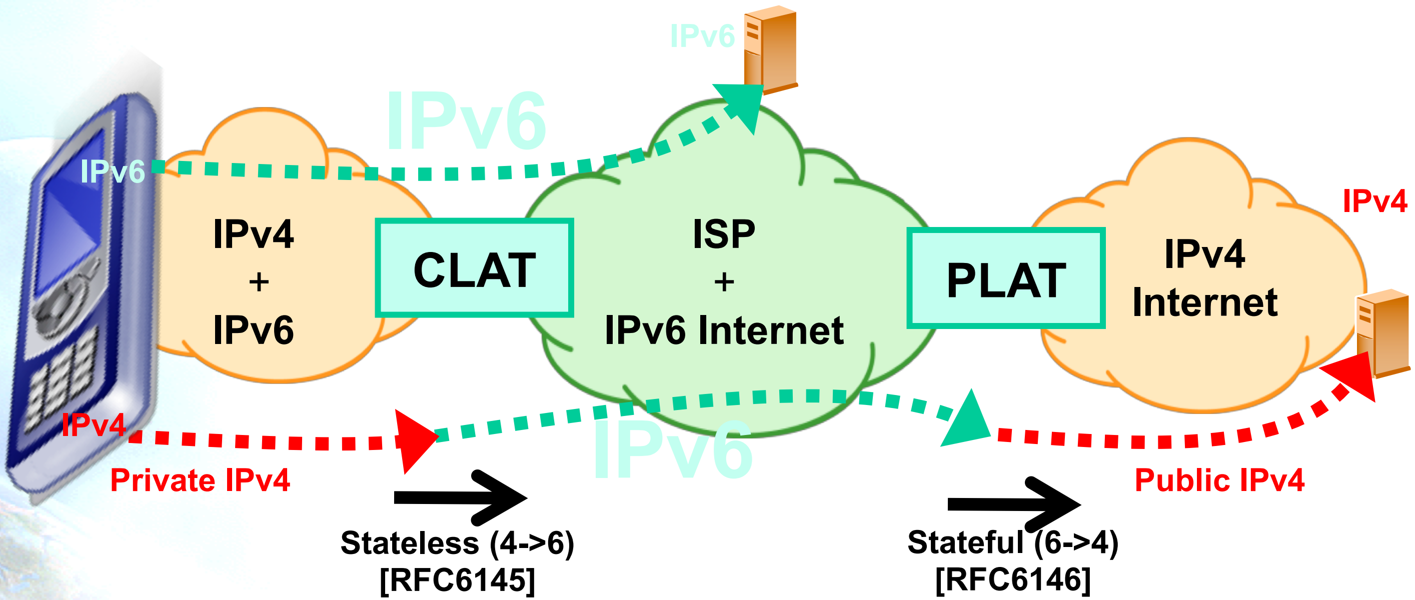
464XLAT

- 464XLAT (RFC6877): RFC6145 + RFC6146
- Very efficient use of scarce IPv4 resources
  - N*65.535 flows per each IPv4 address
  - Network growth not tied to IPv4 availability
- IPv4 basic service to customers over an-IPv6 only infrastructure
  - **WORKS** with applications that use socket APIs and literal IPv4 addresses (Skype, etc.)
- Allows traffic engineering
  - Without deep packet inspection
- Easy to deploy and available
  - Commercial solutions and open source

464XLAT


Internet
IPv4
Internet
IPv6
Public IPv4
NAT64
NAT64
PLAT
NAT64
PLAT
DNS64
ISP network
"plain" IPv6
IPv6-only
access
NAT46
CPE
CLAT
CPE
CLAT
v4
v4
v4
v4/v6
10.0.0.x/24
10.0.0.x/24

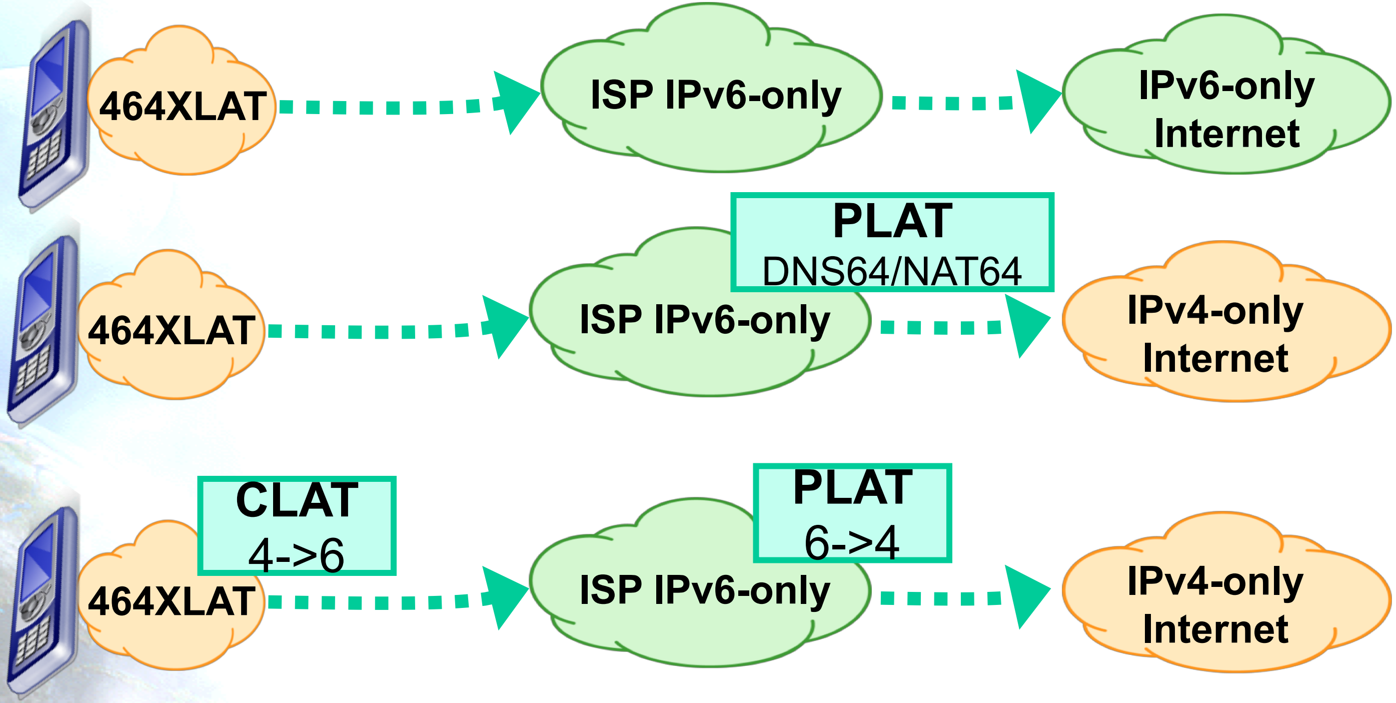
How 464XLAT works?



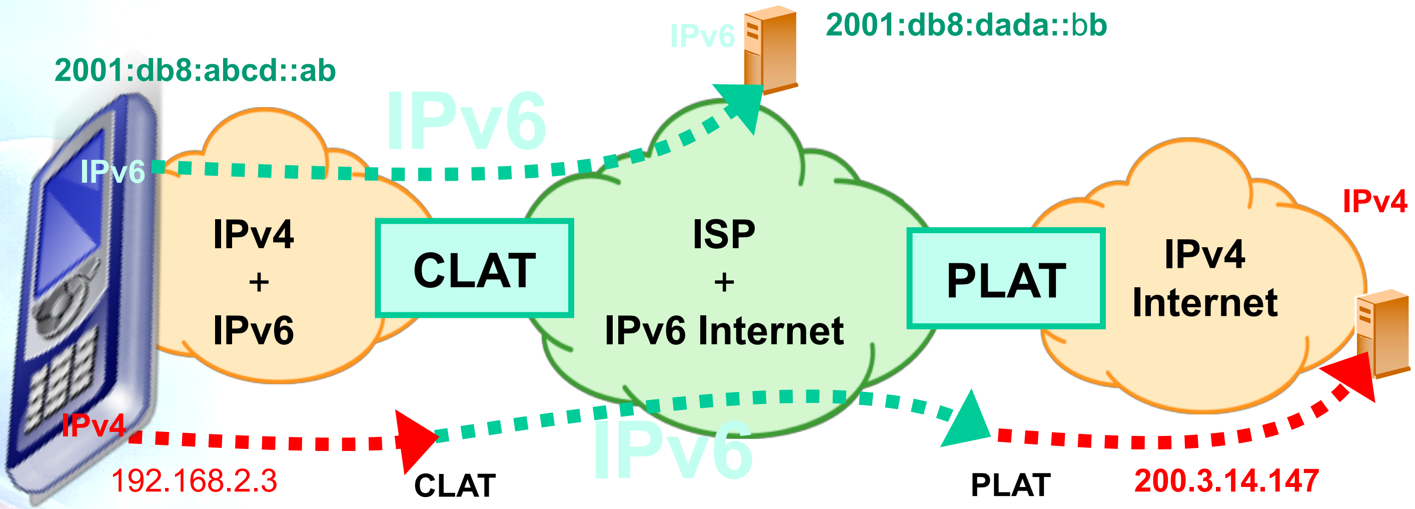
CLAT: Customer side translator (XLAT)

PLAT: Provider side translator (XLAT)

Possible “app” cases



464XLAT Addressing



CLAT
XLATE SRC prefix
[2001:db8:abcd::/96]
XLATE DST prefix
[2001:db8:1234::/96]

PLAT
IPv4 pool
(192.1.0.1 – 192.1.0.250)
XLATE DST prefix
[2001:db8:1234::/96]

IPv4 SRC
192.168.2.3
IPv4 DST
200.3.14.147

→ **Stateless** XLATE **[RFC6145]**

IPv6 SRC
2001:db8:abcd::192.168.2.3
IPv6 DST
2001:db8:1234::200.3.14.147

→ **Stateful XLATE [RFC6146]**

IPv4 SRC
192.1.0.1
IPv4 DST
200.3.14.147

Availability and Deployment

- NAT64:
  - A10
  - Cisco
  - F5
  - Juniper
  - NEC
  - Huawei
  - Jool, Tayga, Ecdsys, Linux, OpenBSD, …
- CLAT
  - Android (since 4.3)
  - Nokia
  - Windows
  - NEC
  - Linux
  - Jool
  - OpenWRT
  - Apple (sort-of, is Bump-in-the-Host [RFC6535] implemented in Happy Eyeballs v2) - IPv6-only since iOS 10.2
- Commercial deployments:
  - T-Mobile US: +68 Millions of users
  - Orange
  - Telstra
  - SK Telecom
  - …
  - Big trials in several ISPs

MAP Encapsulation (MAP-E)

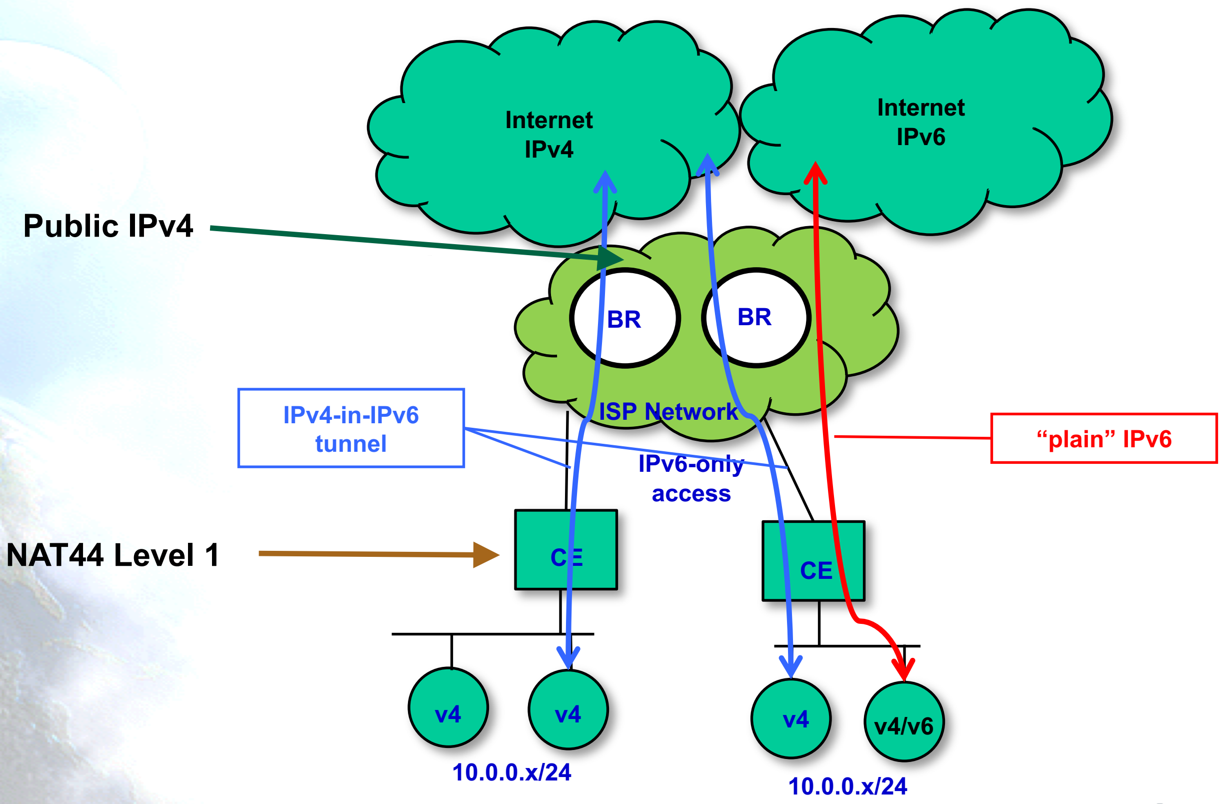
- Mapping of Address and Port with Encapsulation

- Is a “stateless” DS-Lite
  - Provision of an IPv4 prefix, address or “shared” address
  - Algorithmic mapping between IPv4 and an IPv6 address
  - Extends CIDR to 48 bits (32 IP + 16 port)

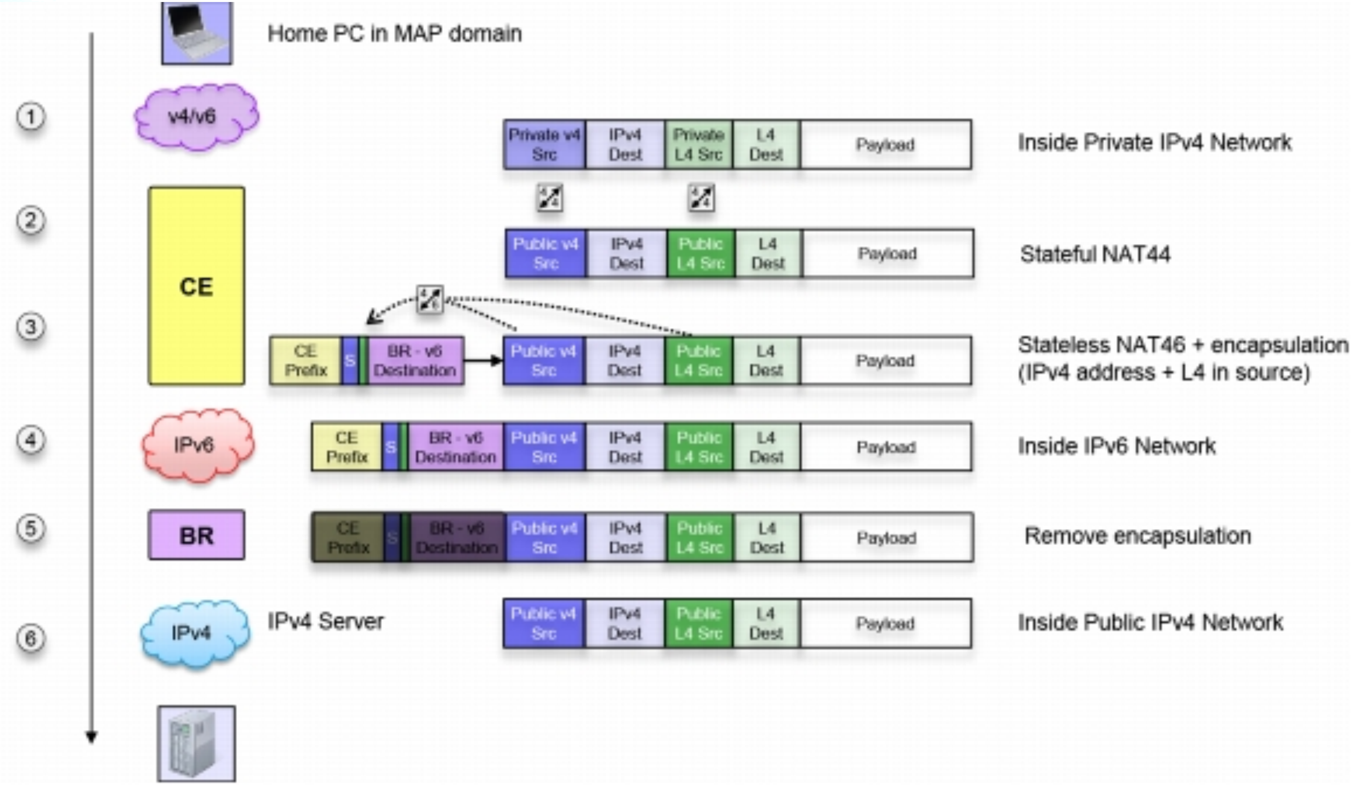
- Allows encapsulating IPv4 in IPv6 for both mesh and hubs&spoke topologies, including mapping-independent IPv4 and IPv6

- Two elements:
  - MAP Customer Edge (CE)
  - MAP Border Relay (BR)

MAP-E


Internet
IPv4
Internet
IPv6
Public IPv4
BR
BR
ISP Network
IPv4-in-IPv6
tunnel
IPv6-only
access
"plain" IPv6
NAT44 Level 1
CE
CE
v4
v4
v4
v4/v6
10.0.0.x/24
10.0.0.x/24

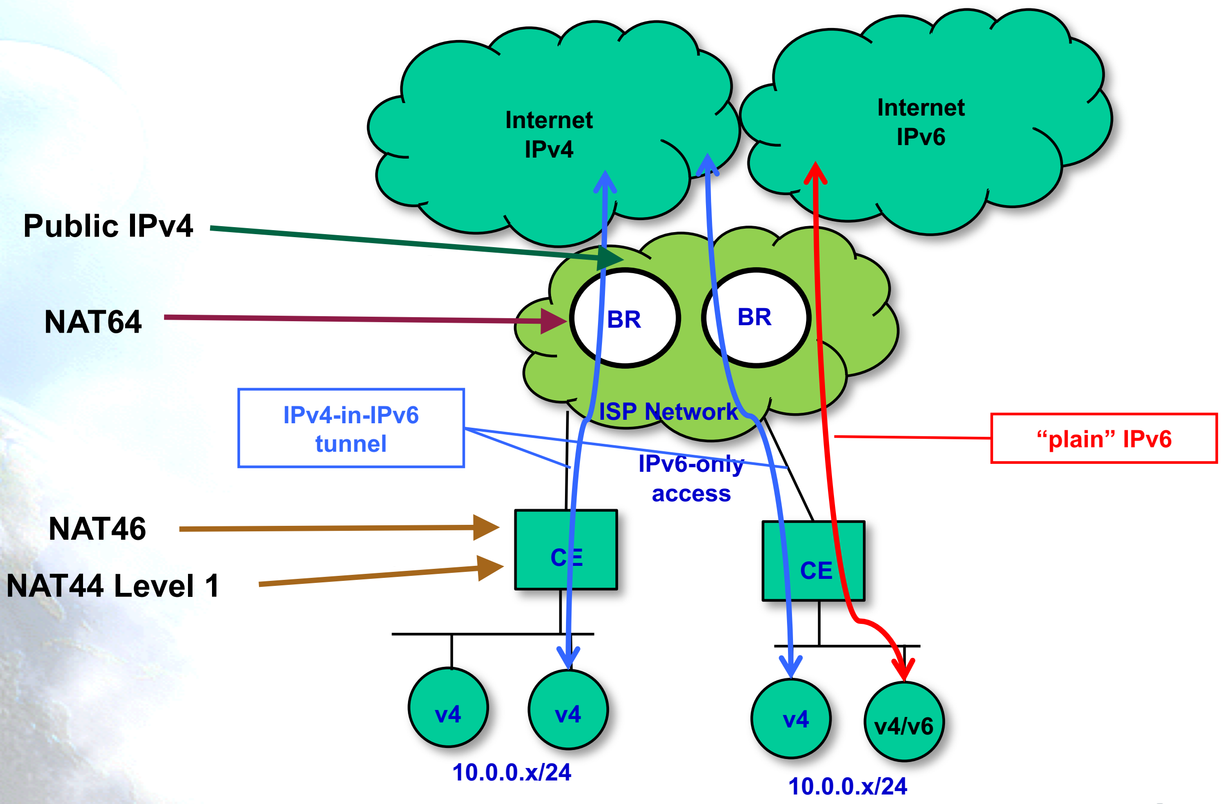
MAP-E Packet Path



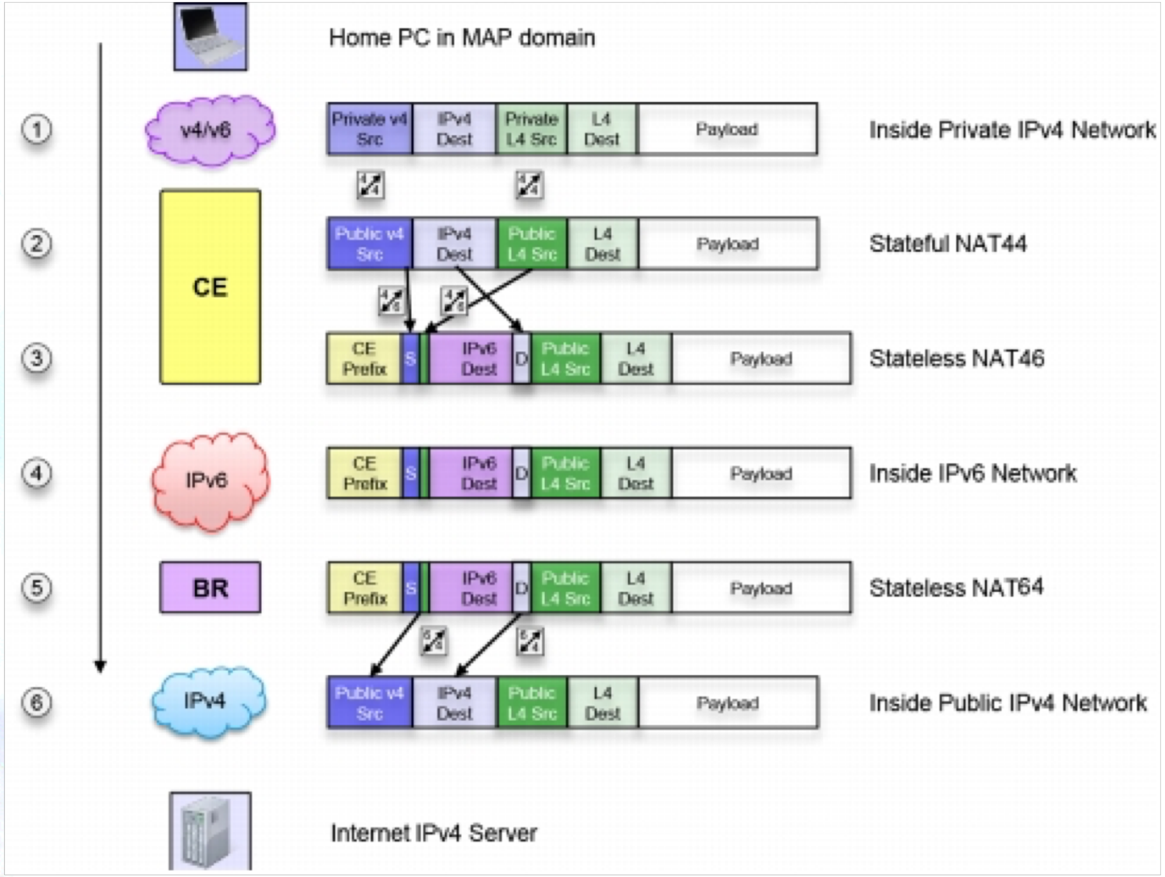
MAP Translation (MAP-T)

- Mapping of Address and Port using Translation
- Similar to MAP-E
- Similar to 464XLAT in the sense of the double translation NAT46 (CLAT) and NAT64 (PLAT)

MAP-T

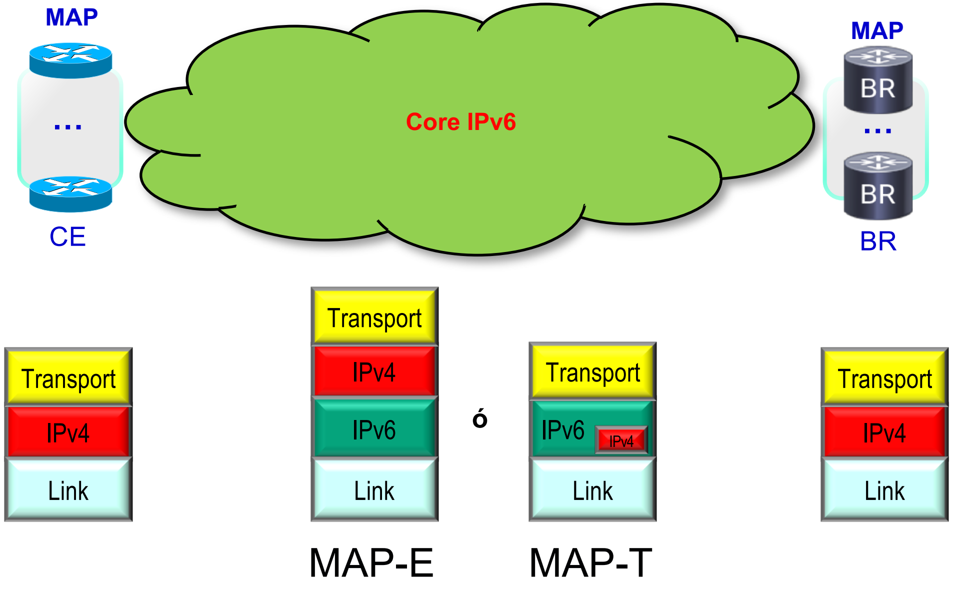

Internet
IPv4
Internet
IPv6
Public IPv4
NAT64
BR
BR
ISP Network
IPv4-in-IPv6
tunnel
IPv6-only
access
"plain" IPv6
NAT46
NAT44 Level 1
CE
CE
v4
v4
v4
v4/v6
10.0.0.x/24
10.0.0.x/24

MAP-T Packet Path

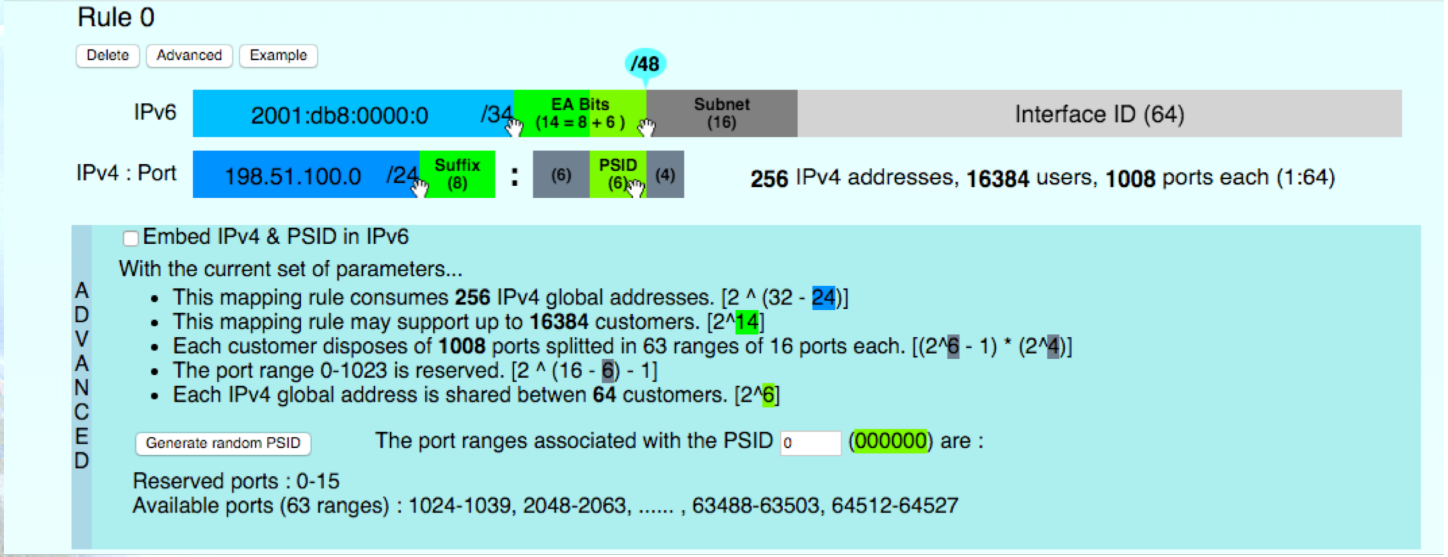


MAP-E vs MAP-T

- MAP-E uses extra 20 bytes for the encapsulation (IPv4-in-IPv6 tunnel).



MAP Addressing



Rule 0

/48

IPv6 | 2001:db8:0000:0 /34 | EA Bits (14 = 8 + 6) | Subnet (16) | Interface ID (64)

IPv4 : Port | 198.51.100.0 /24 | Suffix (8) | : | (6) | PSID (6) | (4)

256 IPv4 addresses, **16384** users, **1008** ports each (1:64)

☐ Embed IPv4 & PSID in IPv6

With the current set of parameters...

- This mapping rule consumes **256** IPv4 global addresses. [2 ^ (32 - 24)]
- This mapping rule may support up to **16384** customers. [2^14]
- Each customer disposes of **1008** ports splitted in 63 ranges of 16 ports each. [(2^6 - 1) * (2^4)]
- The port range 0-1023 is reserved. [2 ^ (16 - 6) - 1]
- Each IPv4 global address is shared betwen **64** customers. [2^6]

The port ranges associated with the PSID 0 (000000) are :

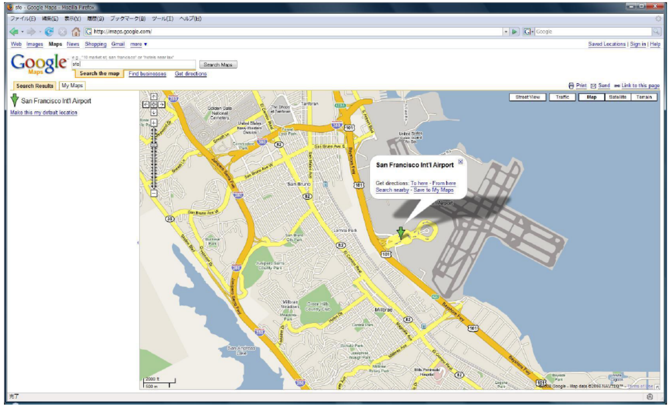
Reserved ports : 0-15
Available ports (63 ranges) : 1024-1039, 2048-2063, ... , 63488-63503, 64512-64527

Comparing Transition …

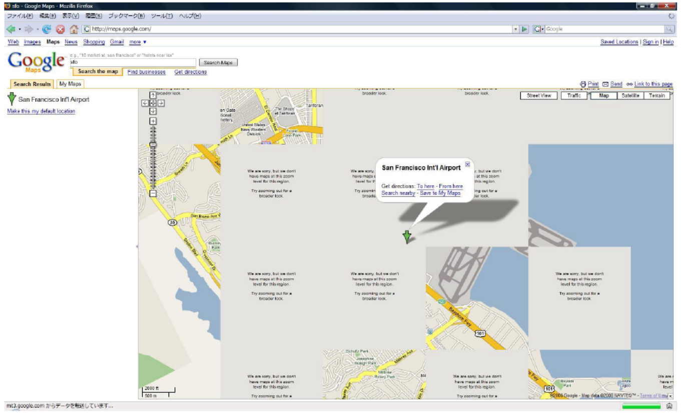
	6RD	Softwires v2	NAT444	DS-Lite	Lw4o6	NAT64	464XLAT	MAP-E	MAP-T
Tunel/Translation (X)	T 6in4	T 6in4	X	T 4in6	T 4in6	X	X	T 4in6	X
Dual-stack LAN	YES	YES	optional	YES	YES	NO	YES	YES	YES
IPv4 Multicast	YES	YES	YES	NO	NO	NO	NO	NO	NO
Access Network	IPv4	IPv4	IPv4 /dual	IPv6	IPv6	IPv6	IPv6	IPv6	IPv6
Overhead	20 bytes	40 bytes	-	40 bytes	40 bytes	20 bytes	20 bytes	40 bytes	20 bytes
Impact in IPv6 addressing plan	YES	NO	NO	NO	NO	NO	NO	YES	YES
CPE Update	YES	YES	optional	YES	YES	YES	YES	YES	YES
NAT44/NAPT	CPE	CPE	CPE + CGN	CGN	CPE	CPE	CPE	CPE	CPE
46/64 Translation	-	-	-	-	-	ISP	ISP +/or CPE	-	CPE + ISP
Translation at ISP with or w/o state	-	-	with	-	-	with	with	w/o	w/o
Scalability	High	Medium	Medium	Medium	High	High	High	High	High
Performance	High	Low	Low	Low	High	Medium	High	High	High
ALGs	NO	NO	YES	YES	NO	YES	YES	YES	YES
Any Protocol or only-TCP/UDP/ICMP	YES	YES	YES	YES	YES	NO	NO	NO	NO
Sharing IPv4 Ports	NO	NO	YES	YES	YES	NO	NO	YES	YES
IPv6 Aggregation	NO	NO	optional	YES	YES	YES	YES	YES	YES
IPv4 Mesh	YES	YES	YES	NO	NO	NO	NO	YES	YES
IPv6 Mesh	YES	NO	optional	YES	YES	YES	YES	YES	YES
Impacts on logging	NO	NO	YES	YES	NO	YES	YES	NO	NO
HA simplicity	High	Low	Low	Low	High	Medium	High	High	High
DPI simplicity	Low	Low	High	Low	Low	High	High	Low	High
Support in cellular	NO	NO	YES	NO	NO	YES	YES	NO	NO
Support in CPEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
	15.5	12.5	10.5	9.5	15	11.5	14	13	13.5

How many ports per user?

Max 30 Connections

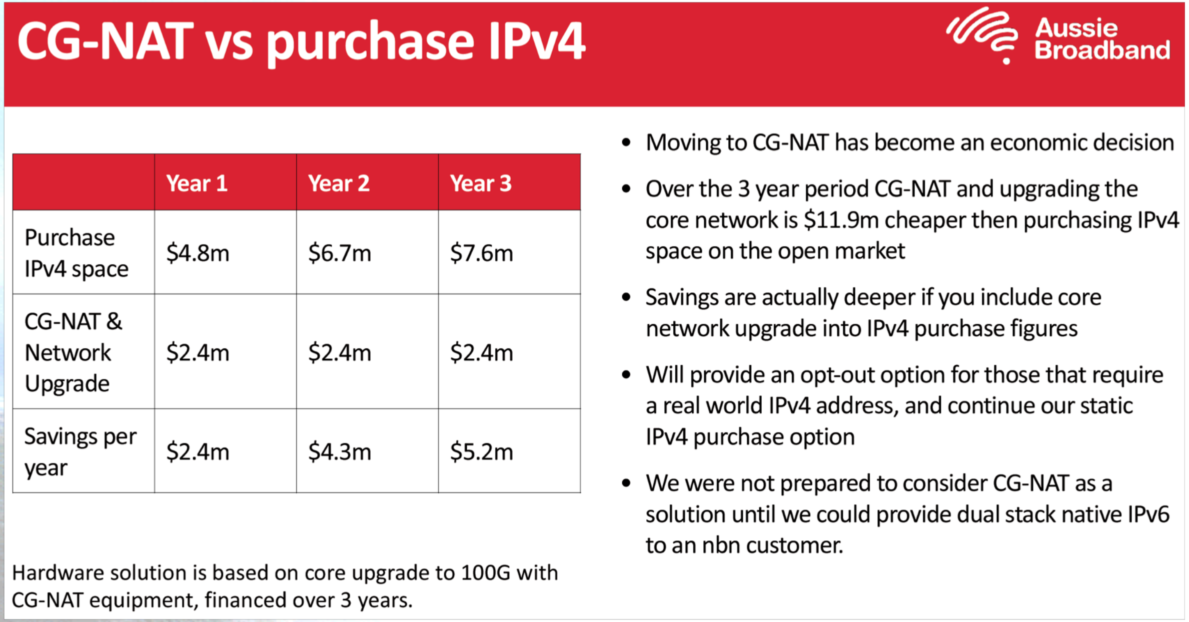

Google Maps
Search Results
San Francisco Int'l Airport
Make this my default location
San Francisco Int'l Airport
Get directions: To here - From here
Search nearby - Save to My Maps

Max 15 Connections


Google Maps
Search Results
San Francisco Int'l Airport
Make this my default location
San Francisco Int'l Airport
Get directions: To here - From here
Search nearby - Save to My Maps
We are sorry, but we don't have maps at this zoom level for this region.
Try zooming out for a broader look.

- **Possibly a minimum of 300 per user behind each CPE**
  - **More as AJAX/similar technologies usage increase**
  - **Times average number of users behind each NAT**
  - **And going up**
- **Be aware of IP/port sharing implications …**

Buying CGNs or IPv4 Addresses



CG-NAT vs purchase IPv4

Aussie Broadband

	Year 1	Year 2	Year 3
Purchase IPv4 space	$4.8m	$6.7m	$7.6m
CG-NAT & Network Upgrade	$2.4m	$2.4m	$2.4m
Savings per year	$2.4m	$4.3m	$5.2m

Hardware solution is based on core upgrade to 100G with CG-NAT equipment, financed over 3 years.

- Moving to CG-NAT has become an economic decision
- Over the 3 year period CG-NAT and upgrading the core network is $11.9m cheaper then purchasing IPv4 space on the open market
- Savings are actually deeper if you include core network upgrade into IPv4 purchase figures
- Will provide an opt-out option for those that require a real world IPv4 address, and continue our static IPv4 purchase option
- We were not prepared to consider CG-NAT as a solution until we could provide dual stack native IPv6 to an nbn customer.

https://www.ausnog.net/sites/default/files/ausnog-2018/presentations/2.6_Phil_Britt_AusNOG2018.pdf

- You buy CGNs instead of IPv4 addresses
  - You start rotating the IPv4 pools at the CGNs because they get blocked after some time
  - Then you discover a couple of years after, that all your IPv4 addresses
  - Then you buy new addresses …
- Why not buying the addresses (now that are cheaper and available) instead of buying the CGNs?

Recommended Reading

- Basic Requirements for IPv6 Customer Edge Routers (RFC7084)
  - Originally include support only for 6RD and DS-LITE
  - Being updated to include support for 464XLAT, MAP T/E, lw4o6, …
  - https://datatracker.ietf.org/doc/draft-ietf-v6ops-transition-ipv4aas/
- BCOP RIPE690:
  - https://www.ripe.net/publications/docs/ripe-690
- Point-to-point links:
  - https://datatracker.ietf.org/doc/draft-palet-v6ops-p2p-links/
- NAT64 deployment guidelines:
  - https://datatracker.ietf.org/doc/draft-ietf-v6ops-nat64-deployment/

DNSSEC Considerations

- DNS64 modifies DNS answers and DNSSEC is designed to detect such modifications, DNS64 can break DNSSEC

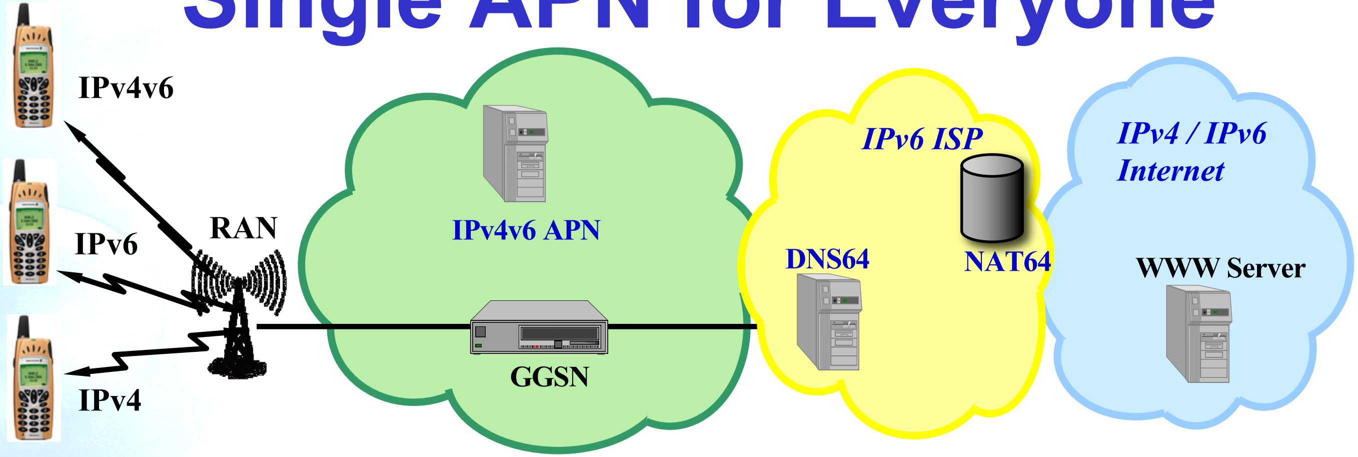
- In general, DNS servers with DNS64 function, by default, will not synthesize AAAA responses if the DNSSEC OK (DO) flag was set in the query. In this case, as only an A record is available, it means that the CLAT will take the responsibility, as in the case of literal IPv4 addresses, to keep that traffic flow end-to-end as IPv4, so DNSSEC is not broken

- Today no apps in cellular that use DNSSEC, but you should be ready for that
  - Consider apps used by means of tethering
  - Very relevant for non-cellular networks

Cellular: 464XLAT or Other Transition?

- 6RD
- DS-Lite
- MAP-E or MAP-T
- …
- No way!
  - Not implemented in smartphones
  - Require using lots of IPv4 addresses
  - Heavy setup and network overhead, require DHCP
  - Take less advantage of “multiplexing” IPv4 addresses & ports, than stateful NAT64

Single APN for Everyone



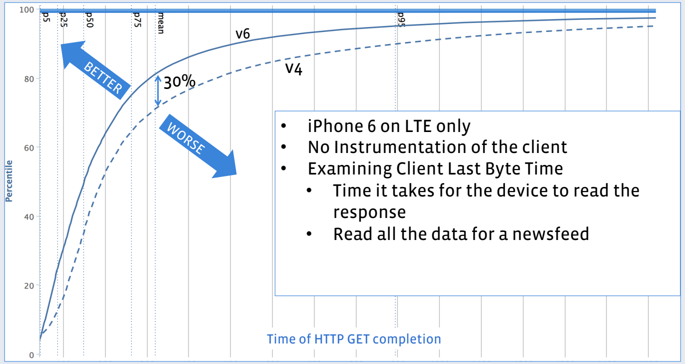
- Single APN
  - Supporting Dual-Stack and Single-Stack
  - Cellular IPv6 deployment is easy because the network supports whatever the UE ask.
  - Progressive deployment, as slow or fast as you want
    - One new phone, all new phones, then OTA old ones
- DNS supporting RFC7050
  - Discovery of the IPv6 Prefix Used for IPv6 Address Synthesis

Tethering

- RFC7278
  - Extending an IPv6 /64 Prefix from a Third Generation Partnership Project (3GPP) Mobile Interface to a LAN Link
  - The UE is switched from an IPv6 host mode to an IPv6 router-and-host mode
- If the UE is also a CLAT, it provides IPv4 service with private addresses to the “tethered” devices

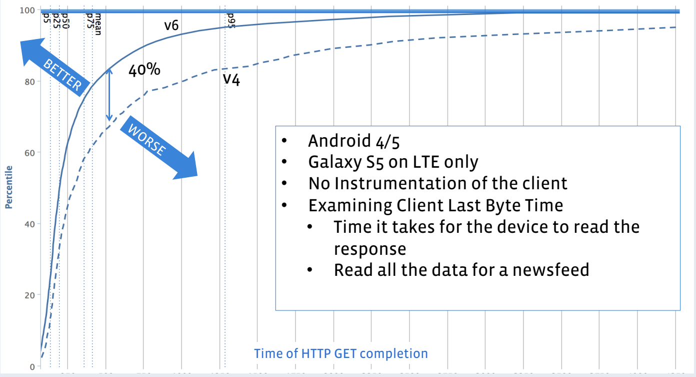
Performance

US Mobile Performance – Dual Stack Provider iOS



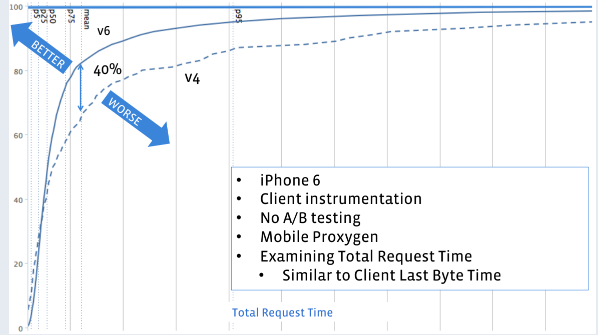
- iPhone 6 on LTE only
- No Instrumentation of the client
- Examining Client Last Byte Time
  - Time it takes for the device to read the response
  - Read all the data for a newsfeed

US Mobile Performance – Dual Stack Provider Android



- Android 4/5
- Galaxy S5 on LTE only
- No Instrumentation of the client
- Examining Client Last Byte Time
  - Time it takes for the device to read the response
  - Read all the data for a newsfeed

US Mobile Performance – Dual Stack Provider iOS

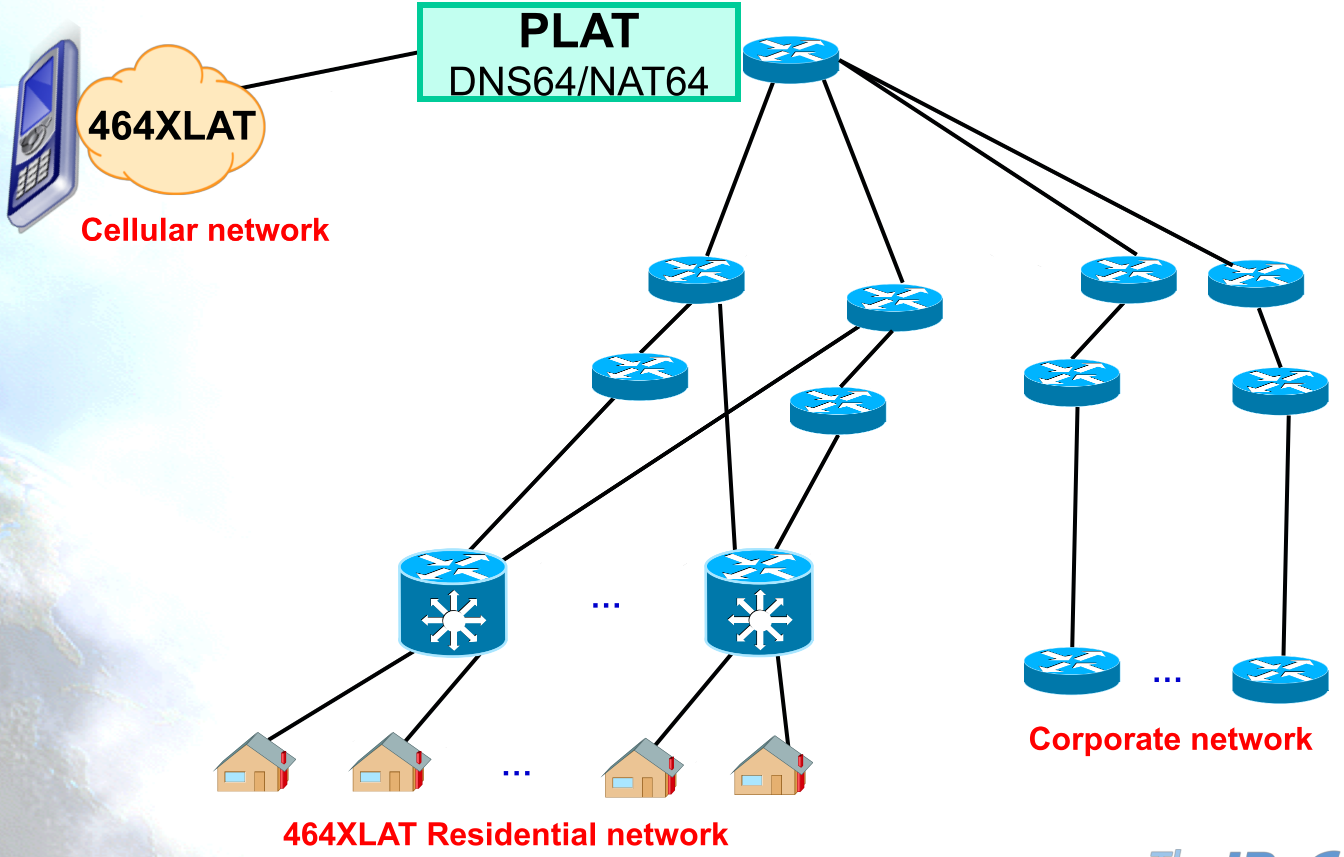


- iPhone 6
- Client instrumentation
- No A/B testing
- Mobile Proxygen
- Examining Total Request Time
  - Similar to Client Last Byte Time

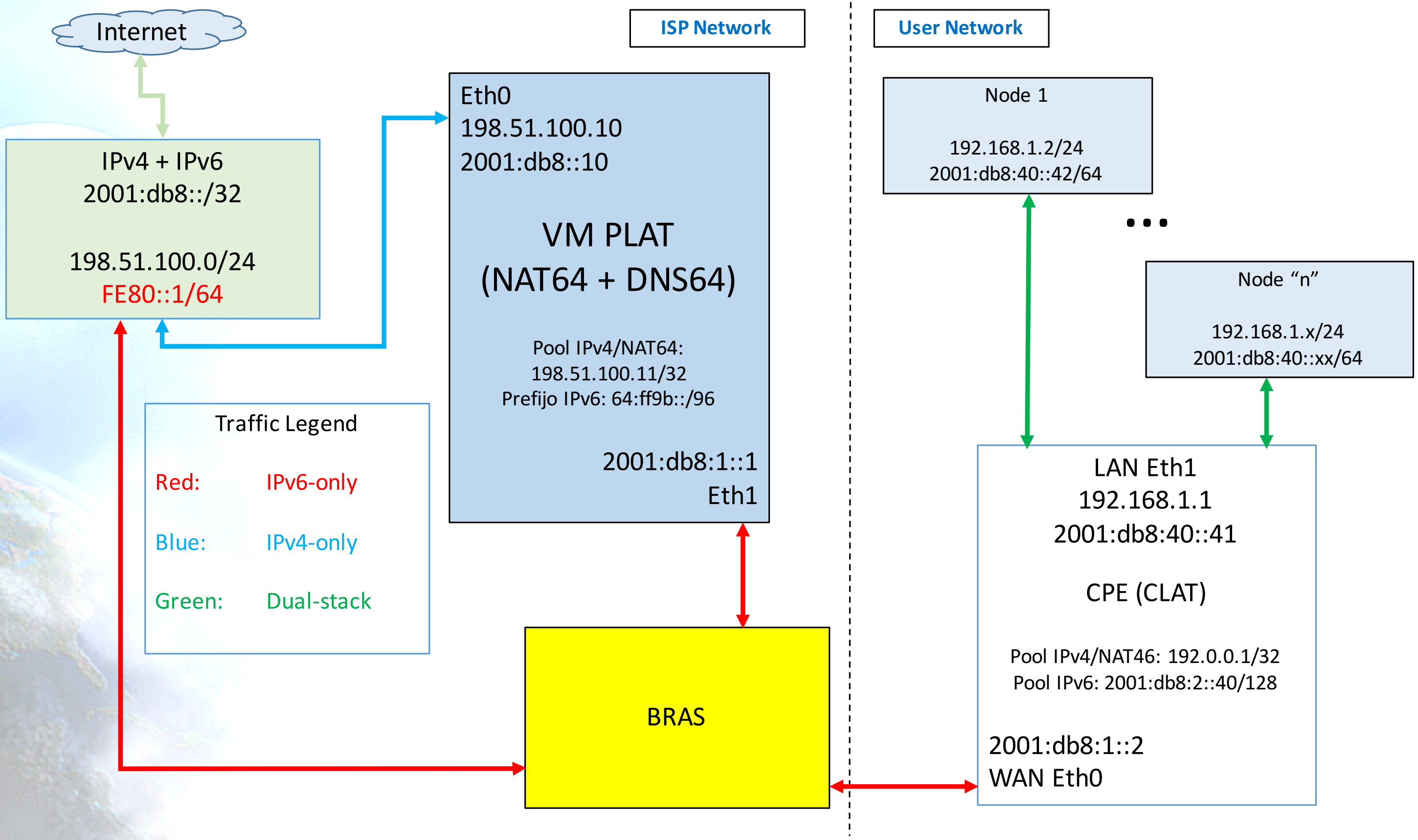
***FaceBook data**

(17/3/2015)

Multiservice Network



Example Residential Customer



ISP Network

User Network

Internet

IPv4 + IPv6
2001:db8::/32

198.51.100.0/24
FE80::1/64

Eth0
198.51.100.10
2001:db8::10

VM PLAT
(NAT64 + DNS64)

Pool IPv4/NAT64:
198.51.100.11/32
Prefijo IPv6: 64:ff9b::/96

2001:db8:1::1
Eth1

Traffic Legend

Red:	IPv6-only
Blue:	IPv4-only
Green:	Dual-stack

BRAS

Node 1

192.168.1.2/24
2001:db8:40::42/64

…

Node "n"

192.168.1.x/24
2001:db8:40::xx/64

LAN Eth1
192.168.1.1
2001:db8:40::41

CPE (CLAT)

Pool IPv4/NAT46: 192.0.0.1/32
Pool IPv6: 2001:db8:2::40/128

2001:db8:1::2
WAN Eth0

NAT64, DNS64
Jool

Jool

- http://jool.mx/
- Open Source SIIT and NAT64 for Linux
- SIIT (RFC7915): Stateless IP/ICMP Translation Algorithm
  - Just “translates 1:1” between IPv4 and IPv6 and back
  - SIIT with EAM (Explicit Address Mapping) allows “rules”
- Stateful NAT64, is a NAT between both
  - So helps with IPv4 address exhaustion

Jool Defined Architectures

- SIIT-DC
- 464XLAT
- SIIT-DC DTM (Dual Translation Mode)

Jool Features

- Runs in Single Interface
  - if needed

- “Node-Based Translation”
  - Using “namespaces” to “wrap” Jool

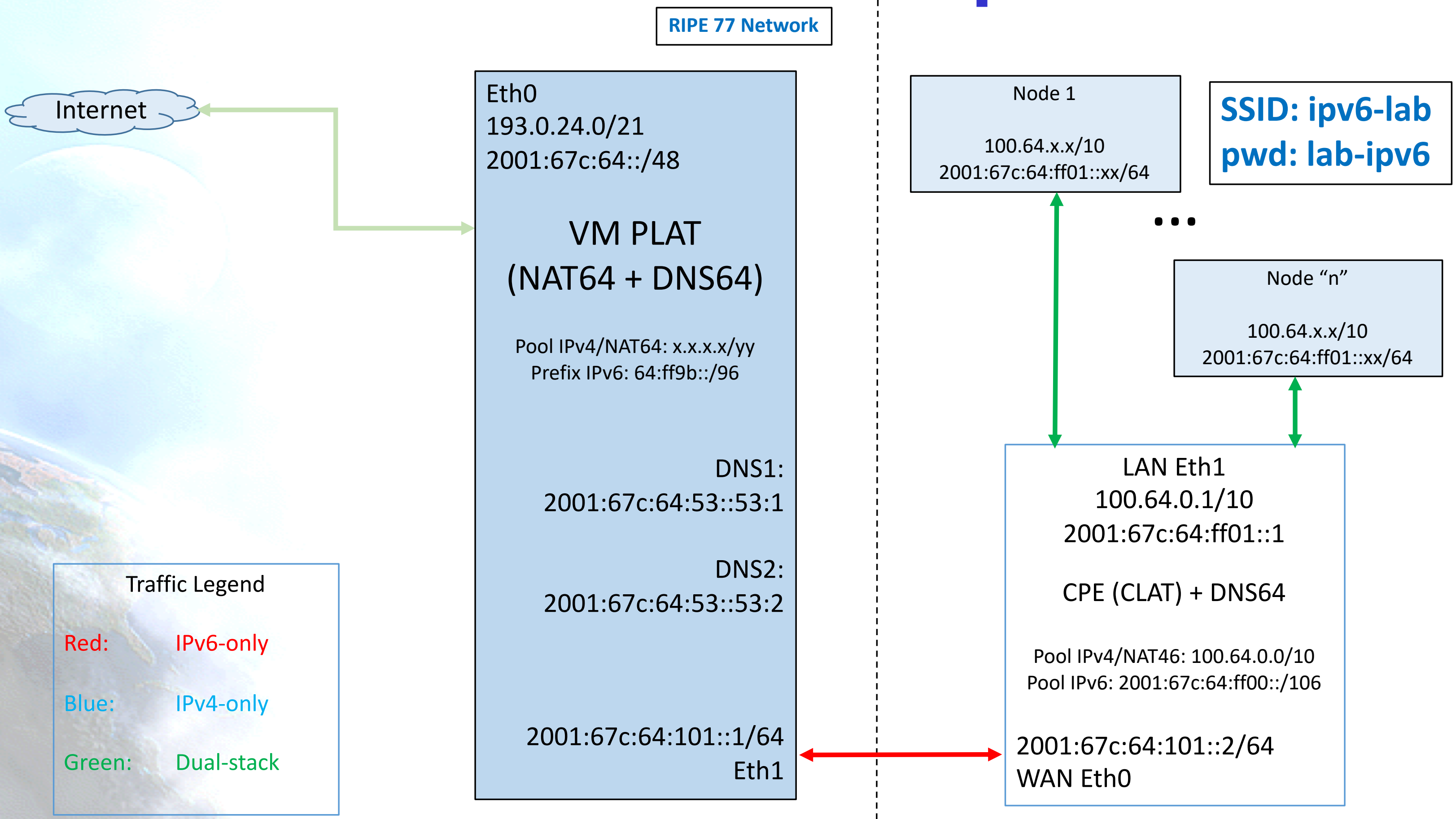
- High-Availability
  - Daemon that allows constant synchronization of sessions across Jool instances

EAMT

- Some examples:

IPv4 Prefix	IPv6 Prefix
192.0.2.1/32	2001:db8:aaaa::5/128
198.51.100.0/24	2001:db8:bbbb::/120
203.0.113.8/29	2001:db8:cccc::/125

Our Demo Setup



RIPE 77 Network

Internet

Eth0
193.0.24.0/21
2001:67c:64::/48

VM PLAT
(NAT64 + DNS64)

Pool IPv4/NAT64: x.x.x.x/yy
Prefix IPv6: 64:ff9b::/96

DNS1:
2001:67c:64:53::53:1

DNS2:
2001:67c:64:53::53:2

2001:67c:64:101::1/64
Eth1

Node 1

100.64.x.x/10
2001:67c:64:ff01::xx/64

SSID: ipv6-lab
pwd: lab-ipv6

...

Node "n"

100.64.x.x/10
2001:67c:64:ff01::xx/64

LAN Eth1
100.64.0.1/10
2001:67c:64:ff01::1

CPE (CLAT) + DNS64

Pool IPv4/NAT46: 100.64.0.0/10
Pool IPv6: 2001:67c:64:ff00::/106

2001:67c:64:101::2/64
WAN Eth0

Traffic Legend

Red: IPv6-only

Blue: IPv4-only

Green: Dual-stack

Demo

```
ping 1.1.1.1
tracert 1.1.1.1
ping www.google.com
ping -4 www.google.com
tracert6 www.google.com
tracert www.google.com
```

Also browse to web site that are/aren’t IPv6 enabled and with literals, for example http://1.1.1.1

In Chrome/Firefox you may want first to install extension “IPvfoo"

NAT64 Setup

```
sudo service network-manager stop
sudo service radvd stop
sudo service isc-dhcp-server stop
sudo service isc-dhcp-server6 stop
sysctl -w net.ipv4.conf.all.forwarding=1
sysctl -w net.ipv6.conf.all.forwarding=1
ethtool --offload br-lan gro off lro off
ethtool --offload eth0 gro off lro off
ethtool --offload eth3 gro off lro off
ip -6 route replace 2001:df9:6::/60 via fe80::a00:27ff:fe6b:e065 dev eth3
modprobe jool pool6=64:ff9b::/96 pool4=220.247.148.2
```

DNS64 Setup

```
/etc/bind/named.conf.options
...
forwarders {
        2001:67c:64:53::53:1;
        2001:67c:64:53::53:2;
        };
dns64 64:ff9b::/96 {
        clients { any; };
        mapped { any; };
        exclude { 0::/3; 4000::/2; 8000::/1; 2001:db8::/32; };
        break-dnssec no;
        };
```

CLAT Setup

```
sudo service isc-dhcp-server start
sudo service radvd start
sudo stop network-manager
sysctl -w net.ipv4.conf.all.forwarding=1
sysctl -w net.ipv6.conf.all.forwarding=1
sysctl -w net.ipv4.ip_forward=1
ethtool --offload eth0 gro off lro off
ethtool --offload eth1 gro off lro off
modprobe jool_siit pool6=64:ff9b::/96
jool_siit --eamt --add 100.64.0.0/10 2001:67c:64:ff00::/106
```

Data-Centers without IPv4!
SIIT-DC

Data-Centers without IPv4!

- Several cases, large content providers with IPv6-only data-centers and more coming …

- Many ways to do that
  - Load Balancing (cost, state, scalability)
  - IPv4 traffic (from Internet) finish in IPv6-only clusters
    - Same RFC1918 space, for IPv4 BGP sessions
    - RFC5549
      - Advertising IPv4 Network Layer Reachability Information with an IPv6 Next Hop
    - IPv4 in IPv6 tunneling, for IPVS (IP Virtual Server)
    - IPv4 link-local (169.254.0.0/16) for Linux and switches

Advantages of IPv6-only

- IPv6 traffic keeps going up
  - Initially more in cellular networks
    - This is a ”more expensive” traffic (radio, energy, bandwidth availability, …)
    - More expensive with IPv4 (“keepalives”) than with IPv6
- If the end-points speak IPv6 there is no NAT
  - Even better, no CGN
- “performance” or ”user-perceived quality of service increases
  - IPv6 40% “faster” than IPv4
    - Response time to complete “HTTP GET”
    - Using HTTP2 and QUIC can increase that performance

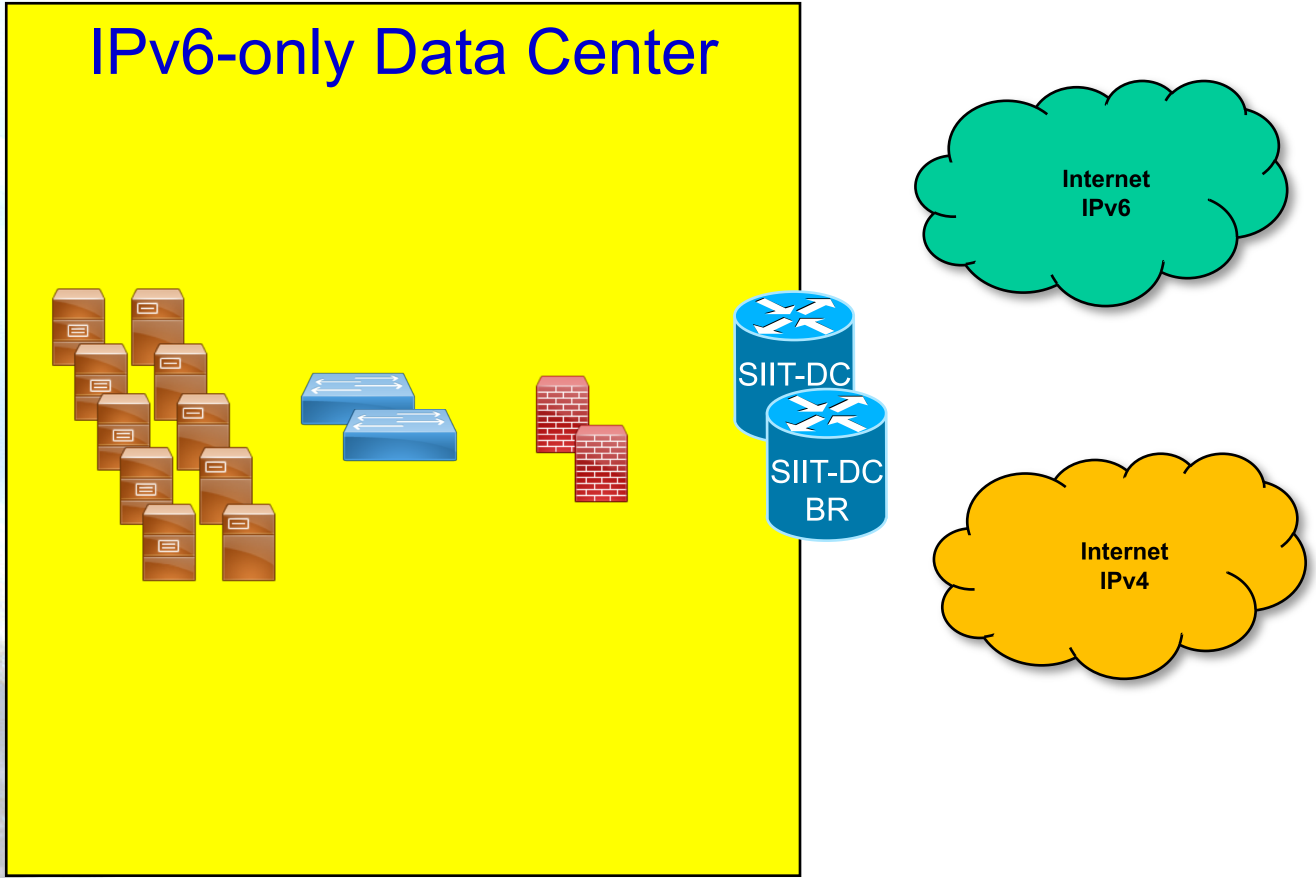
IPv4 or Dual-Stack?

- Against IPv4:
  - Lack of IPv4 addresses
  - Overlapping of private addresses
  - NAT (state)
  - Renumbering (new servers or VMs)
  - Lack of IPv6 support
- Against Dual-Stack:
  - ”Dual” management costs
    - Monitoring, security, human resources, errors, …

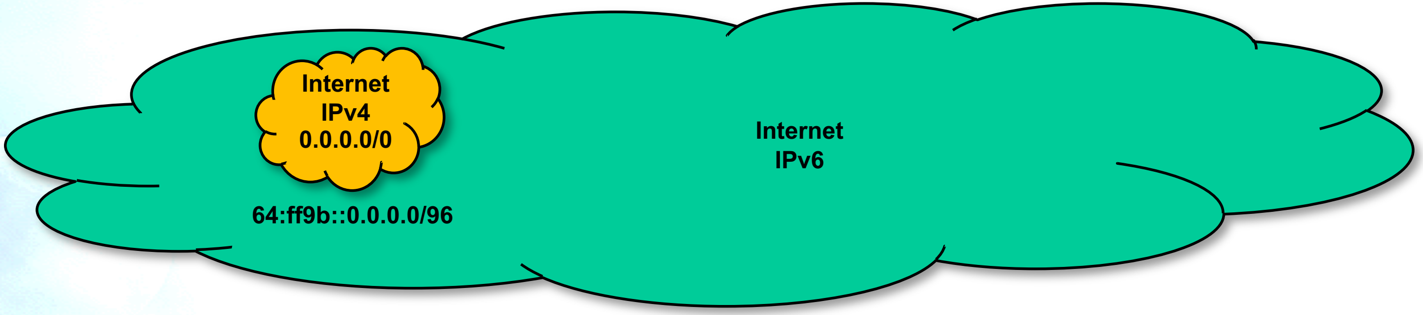
SIIT-DC

- RFC7755 - SIIT-DC: Stateless IP/ICMP Translation for IPv6 Data Center Environments
  - "464XLAT" for the DC
  - No additional software in end-points
- No state!
  - High availability: BGP, ECMP, …
- Keeps source IPv4 address
  - Logging, geolocation, …
- Avoid dual-stack in the DC
  - DC is simplified
- Keeps dual-stack for Internet
  - Service is available for all users
    - IPv4-only, IPv6-only and dual-stack
- Doesn't work with literal addresses neither IPv4-only APIs
  - Not an issue: a DC use DNS!
  - Sorted out as well with RFC7756
    - Stateless IP/ICMP Translation for IPv6 Internet Data Center Environments (SIIT-DC): Dual Translation Mode

Example of DC with SIIT-DC



Mapping all the IPv4 Internet



- An EAM (Explicit Address Mapping) table is configured in the SIIT-DC BR

Translation prefix:	2001:db8:46::/96
IPv4 pool:	192.0.2.0/24

EAM table:

IPv4 Internet address	Address in the DC
192.0.2.1	2001:db8:12:34::1
192.0.2.2	2001:db8:24:68::80
192.0.2.3	2001:db8:24:68::25

Traffic Flow

- Example from IP 203.0.113.50 to 192.0.2.1

IPv4 -> IPv6 translation

	IPv4	IPv6
SRC:	203.0.113.50	2001:db8:46::203.0.113.50
DST:	192.0.2.1	2001:db8:12:34::1

IPv6 -> IPv4 translation

	IPv6	IPv4
SRC:	2001:db8:12:34::1	192.0.2.1
DST:	2001:db8:46::203.0.113.50	203.0.113.50

Support

- Commercial:
  - A10
  - Brocade
  - Cisco
  - F5

- Open Source:
  - Jool
  - Tayga
  - VPP

Thanks!

Contact:

@JordiPalet (The IPv6 Company)
jordi.palet@theipv6company.com