WELCOME

IPV6 TRANSITION TECHNOLOGIES

Alastair (AJ) JOHNSON
alastair.johnson@alcatel-lucent.com
February 2014

AGENDA

1. Introduction
2. The transition technologies
   - Native dual stack
   - Dual-Stack Lite (DS-Lite)
   - NAT64 and 464XLAT
   - 6 Rapid Deployment (6rd)
   - Mapping Address and Port (MAP)
   - Other transition technologies
3. Operational deployment considerations
4. Comparison of transition technologies
5. Conclusion

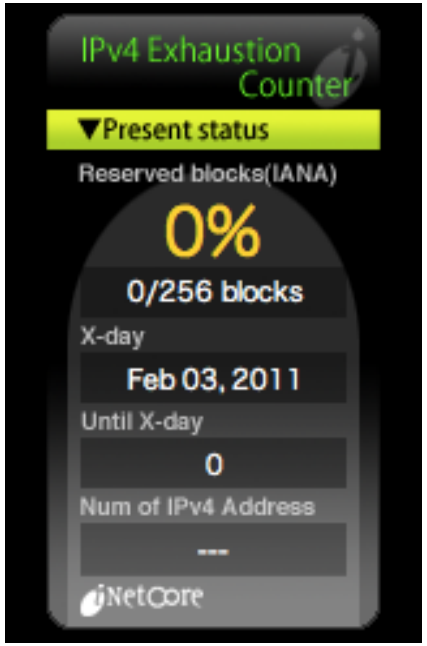
INTRODUCTION

INTRODUCTION
IPV4 EXHAUSTION DRIVES IPV6 ACTIVITY IN THE NETWORK

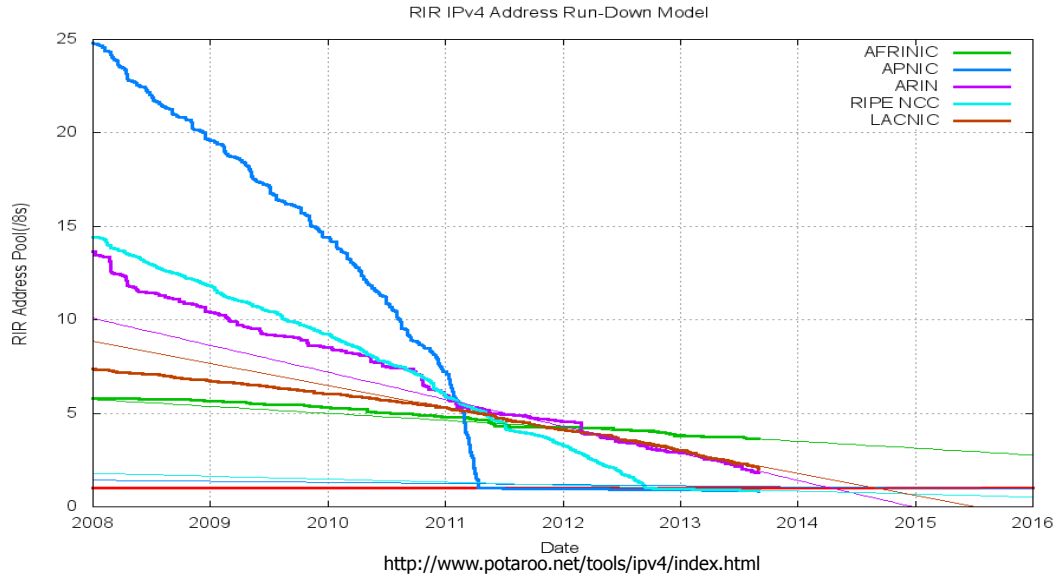
- IANA exhausted the IPv4 available pool on February 3, 2011
  - The final /8 policies went into effect and each RIR received a single /8 prefix
  - Some RIRs will reserve the whole or part of this /8 for IPv6 transition needs

- The RIRs will exhaust their free IPv4 pools over the course of the next few years[1]
  - APNIC exhausted April 2011
  - RIPE NCC exhausted September 14 2012
  - ARIN anticipated to exhaust March 2015
  - LACNIC anticipated to exhaust December 2014
  - AfriNIC anticipated to exhaust July 2021

1 Estimated as of February 2014



http://inetcore.com/project/ipv4ec/index_en.html



http://www.potaroo.net/tools/ipv4/index.html

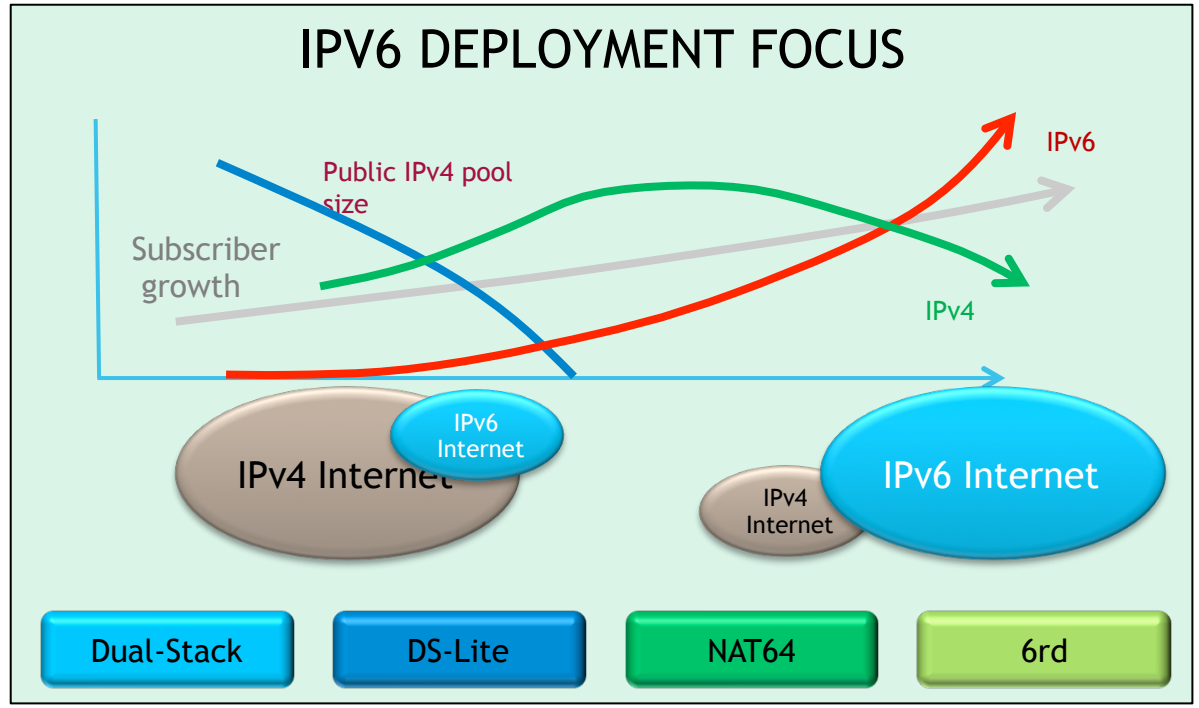
INTRODUCTION
IPV6 DEPLOYMENT – IT IS HAPPENING

- IPv6 deployment ***is happening*** around the world
  - Regional deployments vary – some countries showing significant (>10%) traffic volume over IPv6
  - Some sites exceeding 50% traffic volume over IPv6
- Google provides an interesting barometer of overall IPv6 traffic seen on the Internet at ~3%
  - Some locations will exceed this traffic level
- Significant IPv6 growth in the US market from a few major operators (Verizon Wireless, Comcast, AT&T)
- Some countries in EMEA and CALA also show significant growth, typically from one big operator driving deployment
- Refer to other excellent presentations from the APNIC scientists researching IPv6 connectivity availability

http://www.google.com/ipv6/statistics

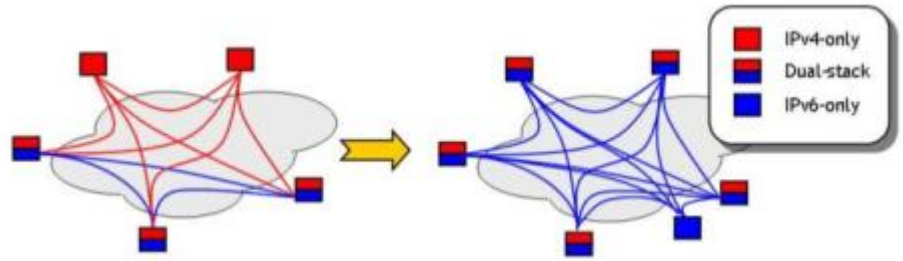
INTRODUCTION
WHAT ARE TRANSITION TECHNOLOGIES

- Access Transition technologies are mechanisms that allow operators to deploy and migrate their subscriber-base to IPv6
- Transition technologies have been developed by the IPv6 community to help accelerate IPv6 deployment, and reduce barriers to IPv6 uptake



- All transition technologies should be evaluated carefully to identify which technology or technologies are the best fit for any given operator to deploy
- Some transition technologies have a ‘long term life’, others are seen as interim solutions to deploy IPv6 quickly while investment or technology catches up
- CPE is one of the most important domains for IPv6 deployment – to support any transition technology, long term strategy, and managing cost
- **Avoiding multiple CPE swaps and migrations should be a key goal for any operator**

INTRODUCTION
WHY USE TRANSITION TECHNOLOGIES

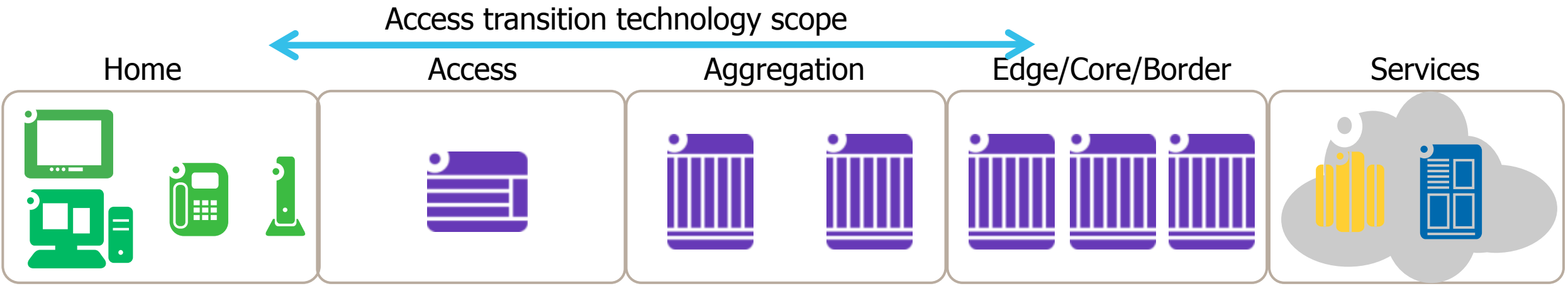


- If IPv6 services are desired, why not deploy native IPv6?
  - This is still the best approach if possible!
- Not always possible:
  - Technology constraints in the network may make native IPv6 deployment difficult or impossible without equipment replacement
  - Wholesale environments might not support IPv6 services currently
  - Desire to roll out IPv6 services as quickly as possible (trial, or overlay services)
  - Previous network architecture decisions may make native IPv6 deployment difficult without network changes (design, test, etc)
  - CPE support and replacement concerns

- **Transition technologies may allow operators to deploy IPv6 services in environments where native deployment is not possible; or to deploy IPv6 services quicker**
  - However, do transition technologies support IPv6 deployment, or IPv4 continuity?

INTRODUCTION
TRANSITION TECHNOLOGIES IN OTHER DOMAINS

- The focus of this presentation is about IPv6 transition technologies and the role they may play in giving subscribers access to IPv6 services
- Other transition technologies may be present in the domains of a typical service provider network, but are not in the scope of this presentation
  - E.g. 6PE/6VPE, service/datacenter ALGs, ABGW-F, etc.



INTRODUCTION
LARGE SCALE NAT

- Large Scale NAT (LSNAT), Carrier Grade NAT (CGNAT), or any other type of service provider IPv4-to-IPv4 based NAT platforms and technologies are **not a transition mechanism to IPv6**

- These technologies are **IPv4 continuity solutions**

- LSNAT is one of several mechanisms that an operator may use to manage IPv4 exhaustion in their network while deploying IPv6 services

- There is much in the way of commonality between IPv6 transition technologies and LSNAT, in that many “transition technologies” are actually providing a way to offer IPv4 services over IPv6 only infrastructure, including IPv4 address sharing
- This presentation will not discuss LSNAT beyond this slide

IPV6 TRANSITION TECHNOLOGIES

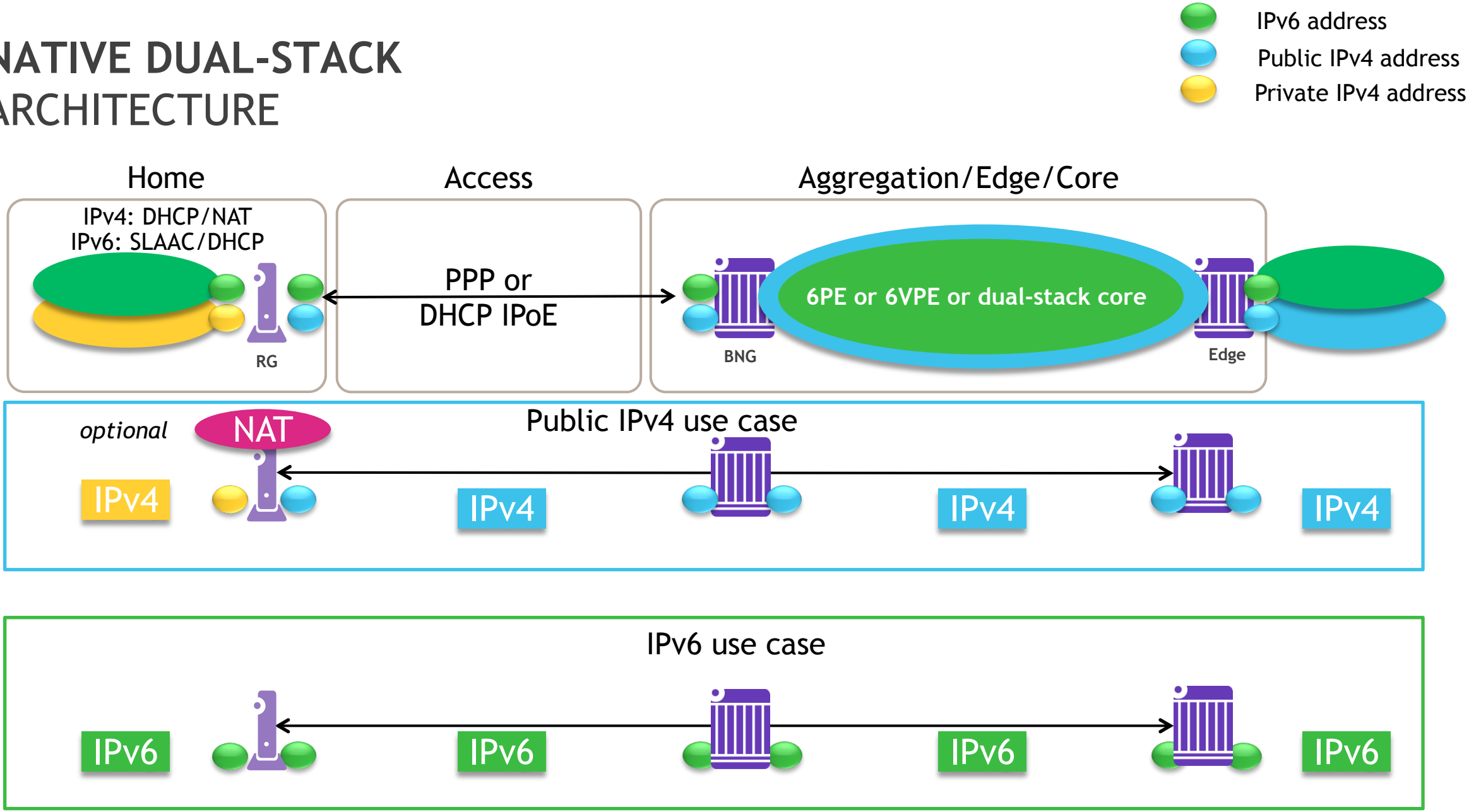
THE IMPORTANCE OF TRANSITION TECHNOLOGY SELECTION

- The Internet community (vendors, operators, IETF, subscribers) have done an excellent job in making sure we have a large number of transition technologies to pick from
- It is important to make sure the technology picked aligns with your engineering, operational and business needs and optimizes investment
- Understand the industry direction and what this means for support of transition technologies
- With so many to choose from, analyzing vendor support in both the CPE and translator domains is critical
  - Avoiding or minimizing vendor lock-in
  - Ensuring long-term support for features in your software deployed in the network

NATIVE DUAL-STACK
INTRODUCTION

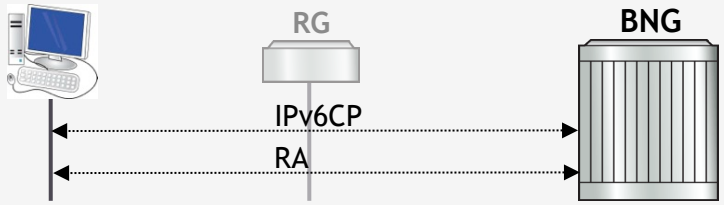
- Deploying IPv6 services as native dual-stack is the best case approach for most operators and subscribers
  - However, it is often the most difficult
- No special encapsulation or tunneling is required
- Native IPv4 and IPv6 services are offered in parallel in the same subscriber session
  - Consistent service edge behavior between IPv4 and IPv6
  - IPv4 addressing is still provided to the subscriber with a potential for very long term sunsetting
- Deployment complexity levels vary in different environments
  - Some networks with minimal or no legacy equipment may find deploying native dual stack services very easy
  - Other networks with older or legacy equipment may find dual stack is not possible due to equipment constraints
  - CPE support is increasing significantly for dual-stack services on PPPoE and IPoE interfaces, including DHCPv6 (with prefix delegation) and SLAAC WAN support
- Ongoing operational considerations
  - What's the impact of running two parallel stacks on the network? Twice the monitoring, reporting, etc…

NATIVE DUAL-STACK ARCHITECTURE

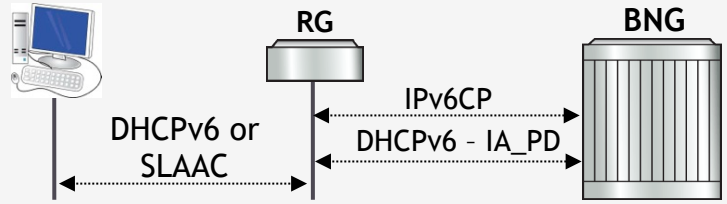


DUAL STACK
PPPOE ACCESS MODELS

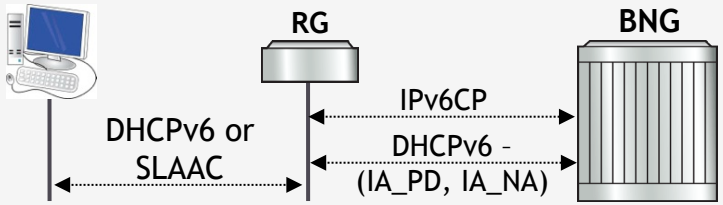
- Bridged residential gateway
  - IPv6CP negotiates interface-id
  - SLAAC for prefix allocation. Linked to PPPoE session



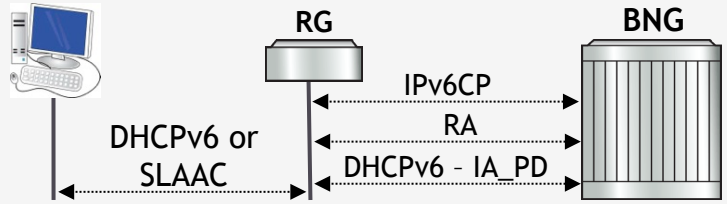
- Routed residential gateway – No WAN IP
  - IPv6CP negotiates interface-id
  - DHCPv6 assigns 'delegated prefix'



- Routed residential gateway – DHCPv6 WAN IP
  - IPv6CP negotiates interface-id
  - DHCPv6 assigns 'delegated prefix' and 'non-temporary address' /128 address



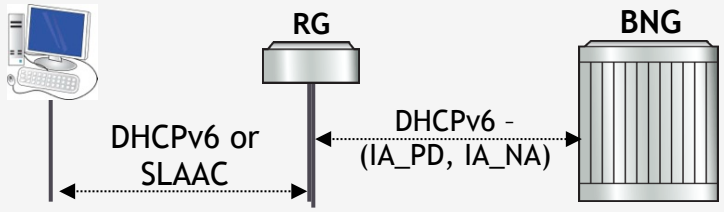
- Routed residential gateway – SLAAC WAN IP
  - IPv6CP negotiates interface-id
  - DHCPv6 assigns 'delegated prefix'
  - SLAAC for WAN IP allocation. Linked to PPPoE session.



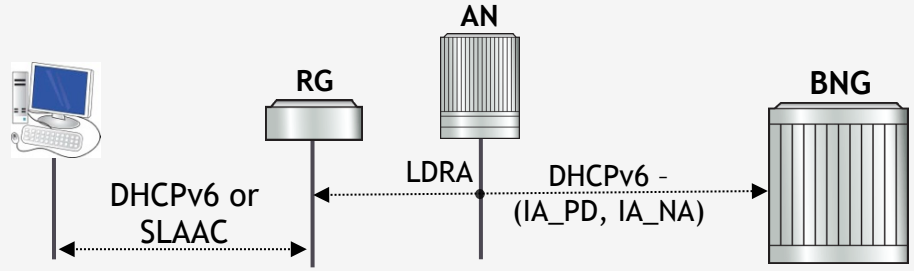
* Only IPv6 operation shown. Regular IPCP for IPv4 allocation

DUAL STACK
DHCP ROUTED ACCESS MODELS

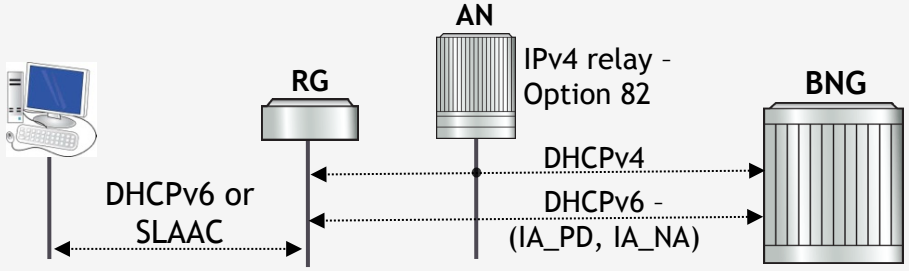
- Routed GW with DHCPv6, 1:1 VLAN
  - DHCPv6 IA_PD, IA_NA (optional)



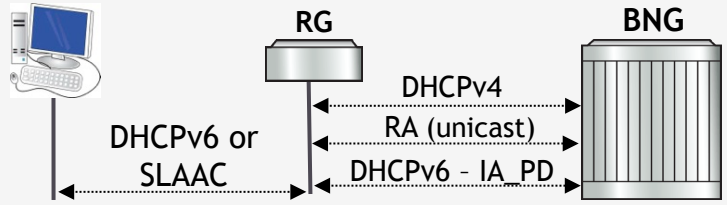
- Routed GW with DHCPv6, N:1 VLAN
  - DHCPv6 IA_PD, IA_NA (optional)
  - LDRA adding DHCPv6 option 18,37 for line-identification



- Routed GW with DHCPv6, N:1, no LDRA
  - DHCPv6 IA_PD, IA_NA (optional)
  - Linked DHCPv4 relay for line-identification



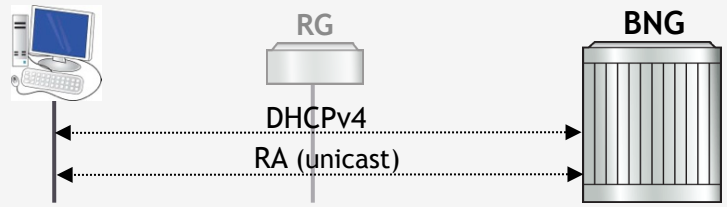
- Routed GW with DHCPv6, SLAAC WAN IP
  - DHCPv6 IA_PD
  - Linked RA for WAN IP allocation



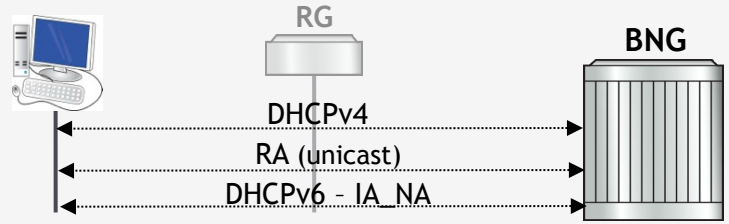
* Only IPv6 operation shown. Regular DHCPv4 for IPv4 allocation

DUAL STACK
DHCP BRIDGED ACCESS MODELS

- Bridged GW with SLAAC, 1:1/N:1 VLAN
  - DHCPv4 and SLAAC RA linking



- Bridged GW with DHCPv6, N:1 VLAN
  - DHCPv4 and SLAAC RA linking (needed to trigger DHCPv6 for some clients)
  - DHCPv6 fo IA_NA allocation



NATIVE DUAL-STACK
DOMAIN IMPACT

ACCESS

- Zero impact in PPPoE environments
  - PPPoE encapsulates traffic; RGs will be enable IPv6 when supported
- Medium to high impact in IPoE environments
  - N:1 VLANs may require network rearchitecture and rely on new features in the access network
- Access node support (DSLAM, OLT, CMTS) becomes very important

SUBSCRIBER EDGE

- High impact – need to support IPv6 services:
  - Subscriber management, queuing, accounting, DHCP-PD, SLAAC(*), etc
- Scaling may be impacted when enabling IPv6 in BRAS/BNG
- Equivalency of features in the subscriber edge node is required – IPv4 & IPv6 should feel the same

* SLAAC for subscriber management is an interesting issue, general industry trend is DHCPv6 based

HOME NETWORK

- Still the most complex domain to manage
- Customer Gateway most likely needs to be replaced
- BBF TR-124i3, RFC7084 specifies the requirements for IPv6 residential gateways
- Vendor support for IPv6 WAN/LAN is increasing significantly, e.g. Technicolor, D-Link, AVM, NEC, ..
- Home network components need to support IPv6
- Internal addressing structure for the home network needs to be considered too

NATIVE DUAL-STACK

- ‘Best-case’ transition design for IPv6 deployment allowing full coexistence of IPv6 and IPv4 services with incremental deployment approach
  - Subscribers take-up IPv6 services as they are enabled or their CPE is replaced
- Subscriber experience is identical regardless of IPv6 or IPv4 service, which are terminated on the same equipment (CPE, BNG) and share queues, SLA, and authorization and accounting policies
- Impact to the customer side of the network is high due to the CPE swap requirement
  - Significant number of CPE today are now IPv6 capable
  - Take advantage of the technology refresh cycle of the last few years
- Broadband Forum TR-177 and TR-187 along with TR-124i3 give excellent references for operators looking to deploy dual-stack
- Depending on topology (IPoE v. PPPoE) the impact in the access/aggregation domains varies:
  - PPPoE is very straightforward to deploy IPv6
  - IPoE does require some changes in the access network
    - If Lightweight DHCPv6 Relay Agent (LDRA) support is required in access nodes
    - N:1 VLAN architecture does place some requirements on CPE behavior and potential requirements around handling Duplicate Address Detection
    - 1:1 VLAN architecture is preferred for IPoE broadband deployment

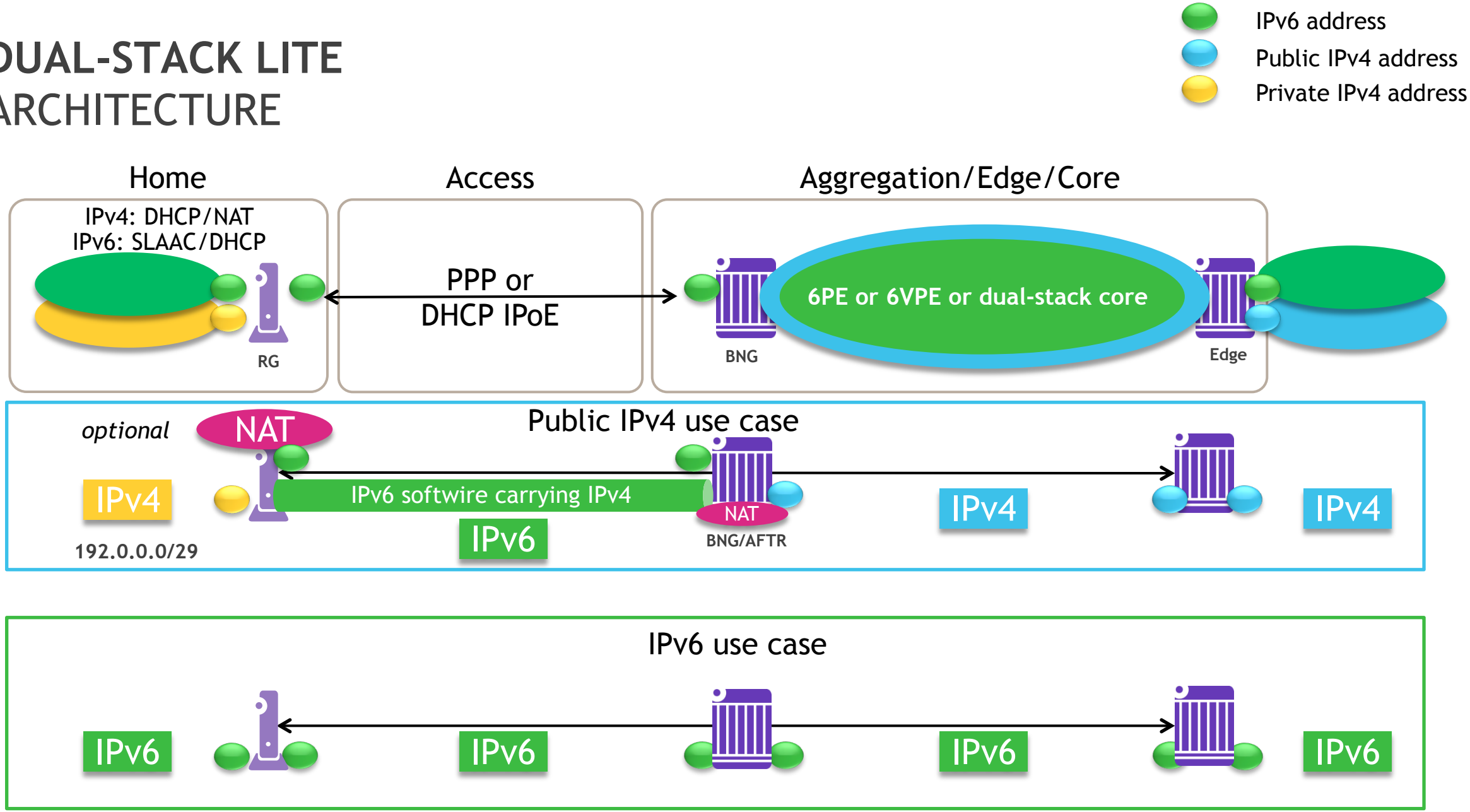
NATIVE DUAL-STACK

- Debate over SLAAC vs. DHCPv6 in the access attachment continues, however general recommendation and approach is DHCPv6 based to align with DHCPv4 model in existing networks
  - May support both to allow for non-RG use cases
- Impact in the subscriber edge (BNG) is variable:
  - Impact to some legacy BNGs may be substantial when dual stack service is enabled impacting scalability, or lack of features for full equivalent IPv4 deployment
  - Operators need to investigate this carefully, however modern BNGs should have no issues when deploying dual-stack services at high subscriber scale
- Dual-stack does have drawbacks in that it may require potential capital investment if equipment forklift upgrades are required, as well as the impact of monitoring two address families in the network (twice the link monitoring, etc)
- Dual-stack does provide an interesting and easy approach to an IPv6-only network by simply turning IPv4 off in the future (and potentially using NAT64, etc)
- Allows status-quo to remain for non-Internet services (e.g. VoIP ATA, CPE/RG management, IPTV services etc) as existing IPv4 path is retained

DS-LITE
INTRODUCTION

- Defined by RFC6333 *Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion*
- Addresses operators that want IPv6-only access networks while providing support for IPv4-only nodes
  - Supporting the view that removing IPv4 from the access network is more efficient than supporting two stacks
- CPE encapsulates IPv4 traffic into IPv4-over-IPv6 tunnel using RFC2473
- Softwire concentrator (AFTR) decapsulates IPv4 packet and performs NAT44 using the unique IPv6 transport address for NAT mapping (LSNAT)
- IPv4 traffic is routed by CPE (Basic Bridging Broad Band element [B4]) to IPv4-over-IPv6 tunnel and is subject to a single NAT operation at the softwire concentrator (or Address Family Transition Router [AFTR])
- IPv6 traffic is routed natively by CPE and BNG
- There is no protocol translation between IPv4 and IPv6

DUAL-STACK LITE ARCHITECTURE



LIGHTWEIGHT DUAL-STACK LITE (LWDS-LITE/LW4OVER6)
…IF YOU THOUGHT IT WASN'T LIGHT ENOUGH, IT GETS LIGHTER

- Built upon *draft-penno-softwire-sdnat* and *draft-ietf-softwire-lw4over6*
- The principle is to combine DS-Lite with an A+P-esque (RFC6346) approach
  - Distribute the NAT function between the B4 element and the AFTR element
  - Remove translation state from the service provider network
- B4 element learns its outside IPv4 address and port range during network attachment
  - B4 performs IPv4 NAT into the outside IPv4 address and port range, and encapsulates it into the softwire to the AFTR
  - AFTR performs anti-spoofing checks and forwards the packet if successful
  - In the return path, the AFTR checks the IPv4 address and destination port, and forwards over the appropriate softwire to the correct B4
- Distributes state between the CPE (B4) and the service provider (AFTR)
  - From per-flow to per-subscriber at the AFTR

DS-LITE
DOMAIN IMPACT

ACCESS	• Medium to high impact depending on topology and technology
	• Access network becomes **single stack IPv6 only**
	  • All upgrades that a native dual-stack scenario requires are also required for DS-Lite
	• Removing IPv4 from the access network becomes interesting
	  • All CPE attaching to the network must support DS-Lite and IPv6 attachment

SUBSCRIBER EDGE	• AFTR node(s) are needed in the network
	  • May be colocated in the BNG or a dedicated element
	  • LSNAT and support infrastructure is required
	• BNG must support all requisites for implementing IPv6 subscriber management
	• Older equipment that does not support IPv6 will need to be replaced
	• Considerations for Lawful Intercept and DPI are needed

HOME NETWORK	• Still the most complex domain to manage • Customer Gateway (DSL modem/router, cable modem, etc) most likely needs to be replaced – must support IPv6-only WAN and DS-Lite – vendor support increasing • Home network components need to support IPv6 • Internal addressing structure for the home network needs to be considered too • IPv4 NAT at the customer gateway is removed

DUAL-STACK LITE

- DS-Lite targets the case where operators wish to remove IPv4 from the access-aggregation and subscriber management edge
  - Single-stack IPv6 between CPE and BNG
  - Continue to support IPv4 connectivity through a classic NAT44 function
- Moving to all-IPv6 in the access network has significant impact in the CPE domain
  - CPE must be upgraded to support IPv6 WAN and all associated connectivity (management, VoIP, IPTV etc)
  - NAT function is removed from CPE which potentially reduces cost (CPU/memory) in maintaining NAT state
  - CPE are commercially available today that support DS-Lite and vendor support is continuing to increase
- Access network and subscriber management edge must support IPv6 in the same manner of dual-stack deployment
  - Some tricks used in dual-stack networks for binding IPv6 subscribers to IPv4 identity are not possible in IPv6-only networks
  - DS-Lite typically assumes an IPoE deployment but could be used in the PPP case as well

DUAL-STACK LITE

- AFTR nodes need to be located near subscriber traffic – typically in or adjacent to the BNG
  - Avoid hauling traffic to centralized locations in the network which may impact TE or interface scaling in the network core
  - A potential drawback to non-BNG located AFTR is that any DPI or other IPv4 classification may be forced to occur at AFTR or elsewhere in the network, potentially stranding existing investment
- DS-Lite moves the NAT44 function out of the RG and into the SP environment, the SP must support NAT translation logging as the subscribers share a common LAN IPv4 prefix (192.0.0.0/29)
- DS-Lite forces re-architecture of existing service offerings such as VoIP and IPTV
  - May need to be moved to native IPv6 services to avoid transiting AFTR nodes in the network
  - Services transiting AFTR may present a significant bandwidth bottleneck (e.g. multicast traffic)
- DS-Lite generally implies a migration where entire Access Nodes (or regions) are migrated at once, rather than per-subscriber migration
- DS-Lite provides an interesting and easy approach to an IPv6-only network by simply turning IPv4 off in the future when it is no longer required

NAT64
INTRODUCTION

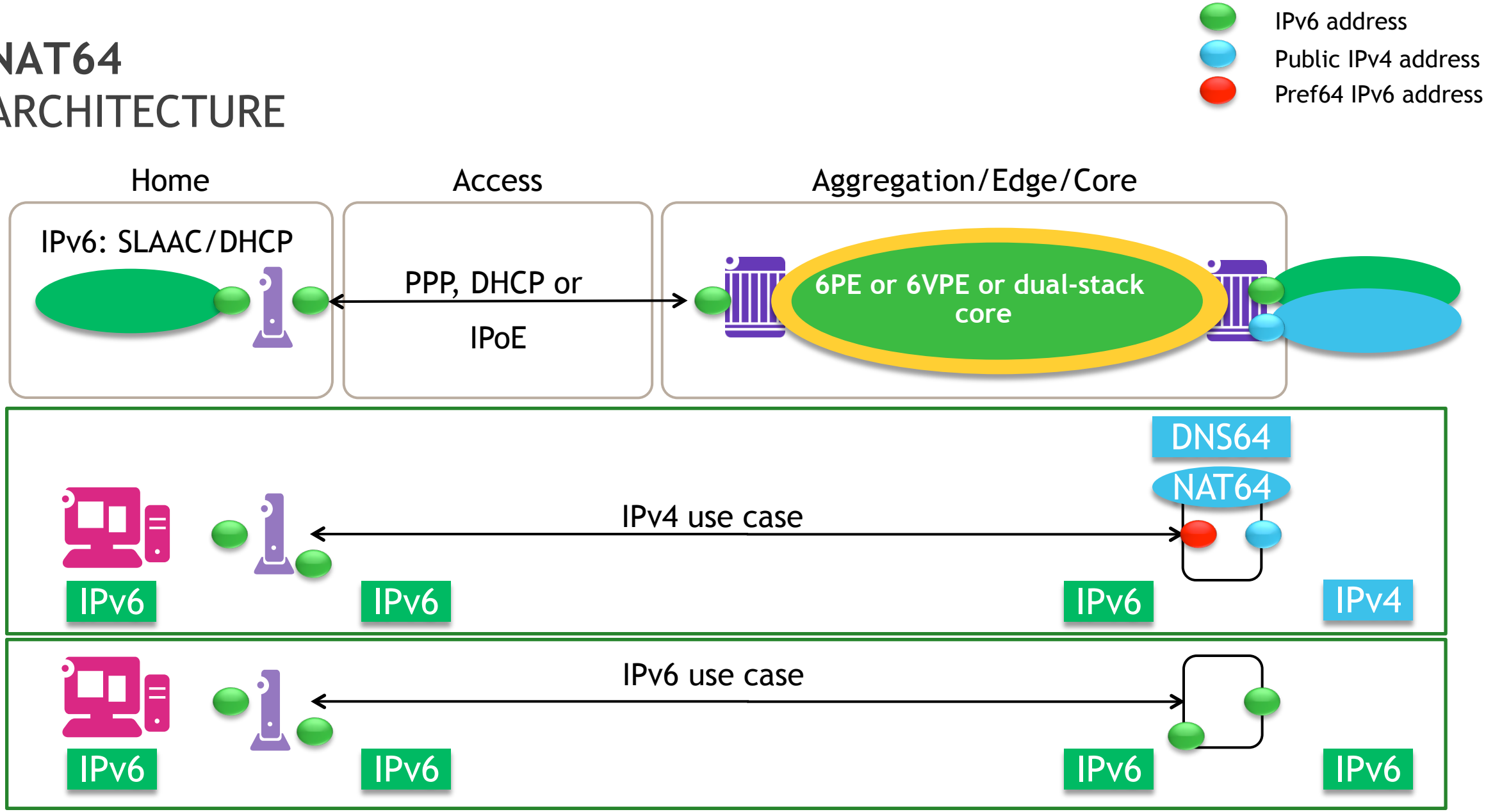
- Addresses operators who want IPv6-only access networks, but providing support for IPv4-only servers or content
  - Implies a well behaved, well understood CPE/UE – and ideally a minimal set of applications
  - Does **not** support IPv4-only hosts attaching to the network

- CPE/UE connects to IPv4 hosts through a synthesized IPv6 address, provided by a DNS64 engine
  - Well known prefix 64:ff9b::/96 is used to map IPv4 server addresses
  - Any client that cannot use a DNS64 server or provide local DNS64 resolution will not be able to connect to the IPv4 server, e.g. no more connecting by IP address

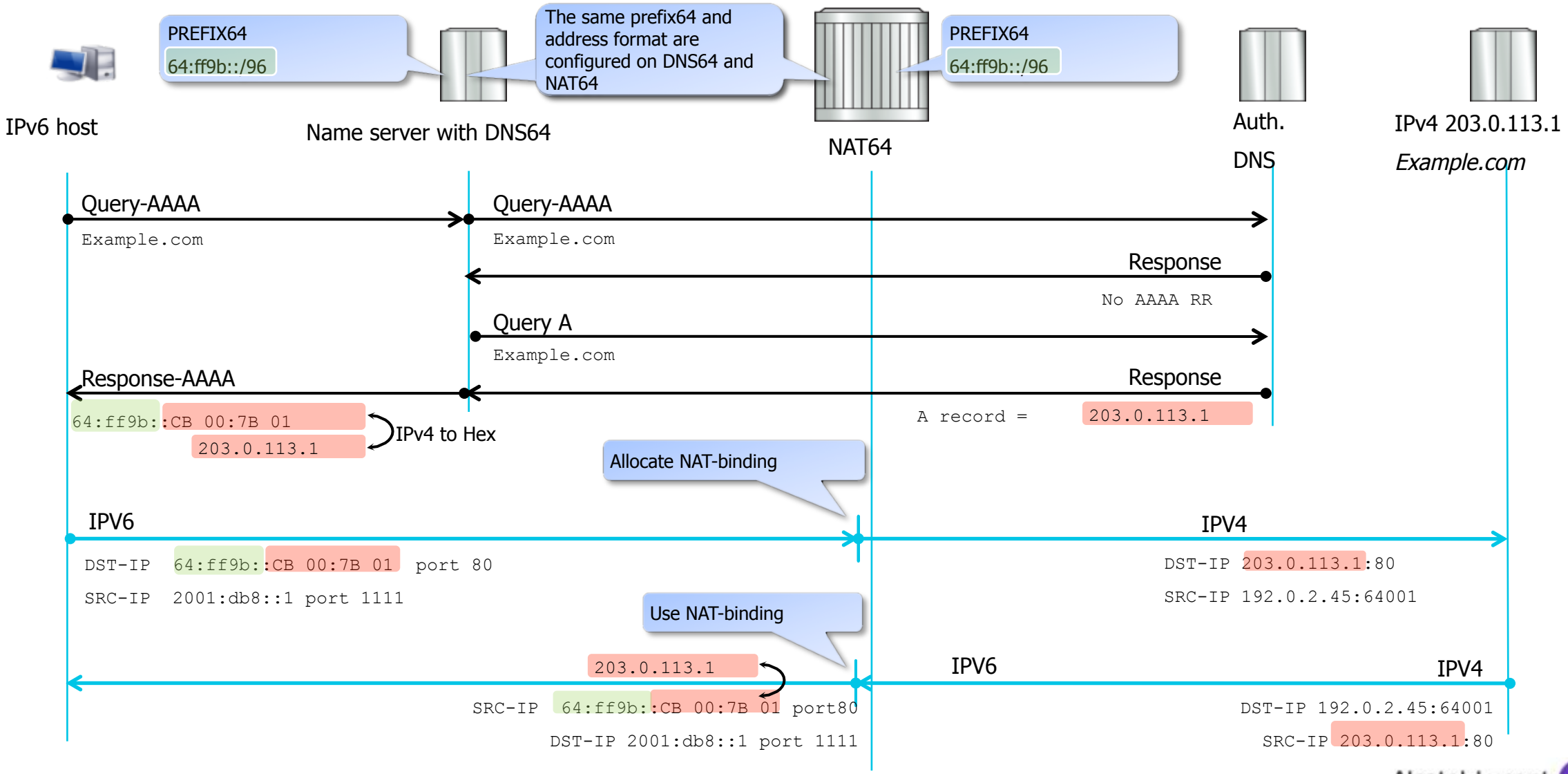
- IPv6 traffic is routed natively by CPE and BNG

- NAT64 can be used as a PLAT in the 464XLAT architecture

NAT64
ARCHITECTURE



NAT64
FLOWCHART



NAT64
DOMAIN IMPACT

ACCESS	• Medium to high impact depending on topology and technology • Access network becomes **single stack IPv6 only** so all upgrades that a native dual-stack scenario requires are also required for NAT64 • **All devices attaching to the network must support IPv6, including in-home**

SUBSCRIBER EDGE	• NAT64 is needed in the network
	  • May be colocated in the BNG or a dedicated element
	  • If a separate element, load balancing of NAT64 traffic must be considered
	• DNS64 node must also be deployed
	• BNG must support all requisites for implementing IPv6 subscriber management
	• Lawful Intercept and DPI must be considered

HOME NETWORK	• Customer Gateway (DSL modem/router, cellphone, cable modem, etc) most likely needs to be replaced – must support IPv6-only WAN • **Home network components must support IPv6** • Internal addressing structure for the home network needs to be considered too • IPv4 NAT at the customer gateway is removed – and direct IPv4 support may be removed • DNSSEC support will break with a DNS64 in the middle of the DNS chain • Typically only useful or talked about for wireless environments at the moment

NAT64

- NAT64 targets the case where operators wish to remove IPv4 from the access-aggregation and subscriber management edge
  - Single-stack IPv6 between CPE and BNG
  - Continue to support IPv4 connectivity through NAT64 function
  - NAT64 is considered impractical for wireline environments, but shows promise in wireless environments
- Moving to all-IPv6 in the access network has significant impact in the CPE domain
  - CPE must be upgraded to support IPv6 WAN and all associated connectivity (management, VoIP, IPTV etc)
  - NAT function is removed from CPE which potentially reduces cost (CPU/memory) in maintaining NAT state
- Access network and subscriber management edge must support IPv6 in the same manner of dual-stack deployment
  - Some tricks used in dual-stack networks for binding IPv6 subscribers to IPv4 identity are not possible in IPv6-only networks
- NAT64 nodes must be located near subscriber traffic – typically in or adjacent to the BNG
  - Avoid hauling traffic to centralized locations in the network which may impact TE or interface scaling in the network core
  - All classification/DPI on the IPv4 side of traffic should be preserved through the NAT64 to enforce end-to-end behavior in the SP network

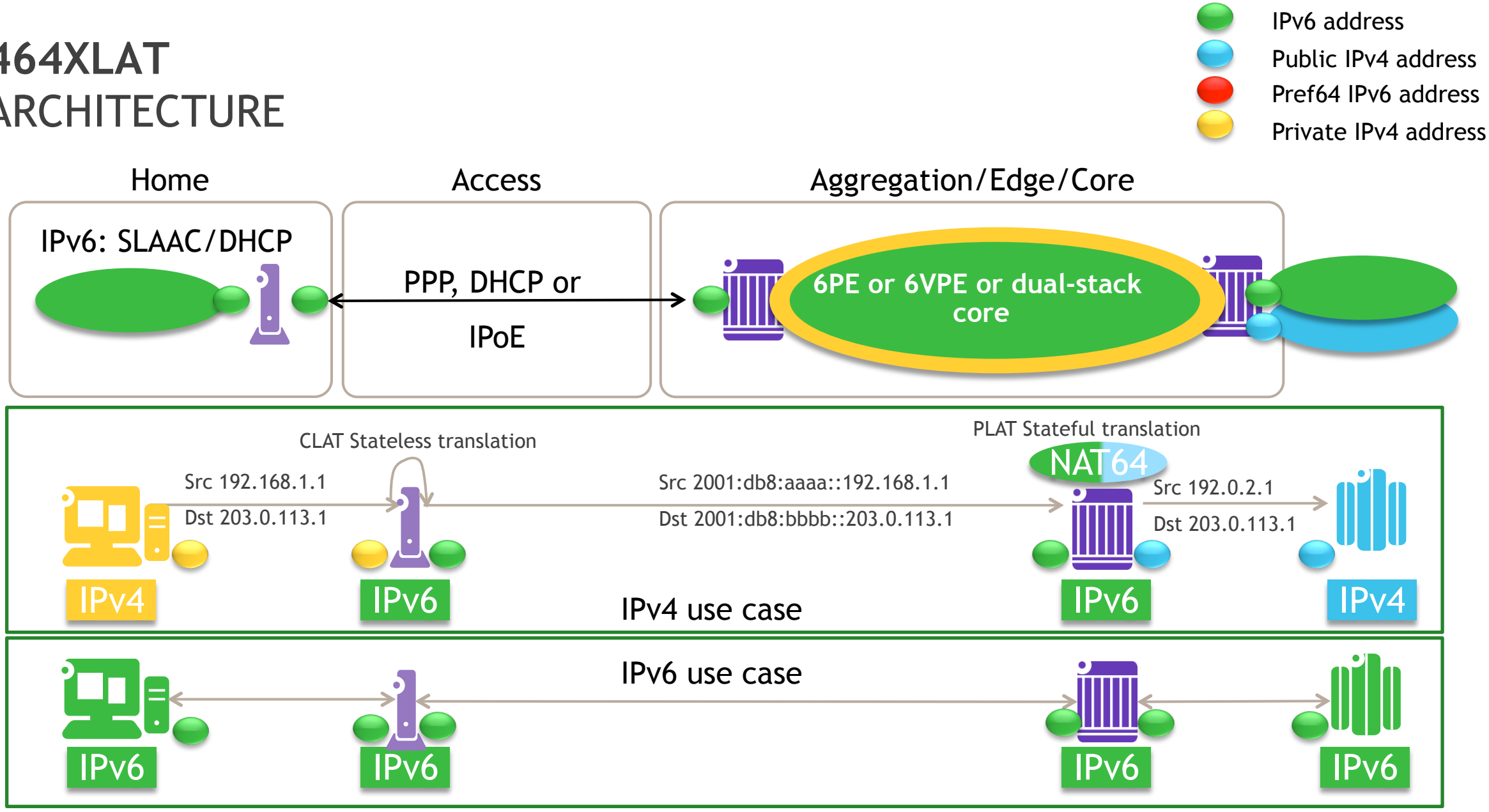
NAT64

- A DNS64 node must also be deployed to provide the DNS synthesis
  - Translate DNS responses with only A-records into AAAA-records with the well-known Pref64 prefix
  - Major DNS vendors support DNS64 today
- NAT64 breaks some applications
  - Those that rely on IPv4 literals (e.g. attempting to establish a socket directly to 192.0.2.1)
  - Some experiments have been conducted with IPv6-only networks and NAT64 environments to document behavior
- NAT64 forces the re-architecture of existing service offerings such as VoIP and IPTV
  - May need to be moved to native IPv6 services to avoid transiting AFTR nodes in the network
  - Services transiting AFTR may present a significant bandwidth bottleneck (e.g. multicast traffic)
- NAT64 generally implies a migration where entire Access Nodes (or regions) are migrated at once, rather than per-subscriber migration
- NAT64 provides an interesting and easy approach to an IPv6-only network by simply turning IPv4 off in the future when it is no longer required

464XLAT
EXTENDING NAT64

- 464XLAT is documented in RFC6877 and defines an architectural approach to implementing NAT64 with IPv4 continuity in an IPv6-only access network
- A combination of stateless (RFC6145) and stateful (RFC6146) translation technologies, split across the Customer Located Address Translator (CLAT) and Provider Located Address Translator (PLAT) functions
- Bridges the gaps that are in NAT64:
  - Removes the dependency on DNS64 in the service provider network
  - Allows support for IPv4-only applications on the UE or home network
  - Allows support for IPv4-only devices (e.g. Windows 98, un-touched Windows XP) in the home network
- Requires the residential gateway or UE to support the CLAT functionality
  - Can be retrofitted into some existing devices with a software upgrade or package
- Has been deployed in production environments

464XLAT ARCHITECTURE



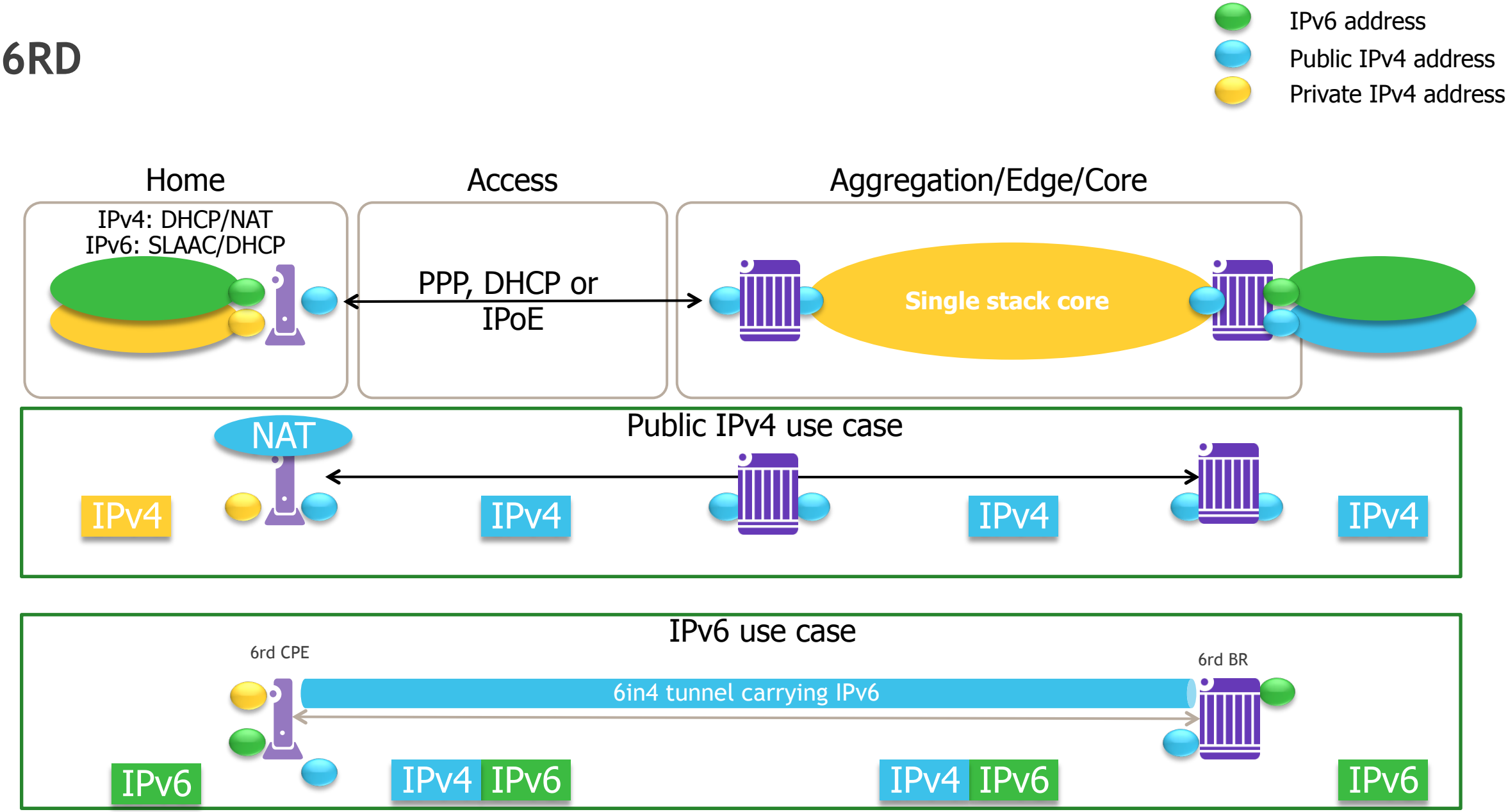
464XLAT

- 464XLAT makes NAT64 much more deployable for many operators
  - NAT64 was previously seen as only useful in tightly constrained networks where IPv6-only hosts existed, and protocols that needed to be translated were simple (HTTP, SMTP, etc)
  - 464XLAT allows IPv4-only hosts in the home network to continue to function – important for embedded devices that do not support IPv6 at all
    - Gaming consoles
    - Thermostats
    - Etc
- IPv6 services stay as native IPv6 as per the dual-stack, DS-Lite, and NAT64 deployment use-cases
- IPv4 services can be retained, and offer the benefit of address sharing/overloading at the CLAT function
  - All associated drawbacks and operational considerations associated with this in other deployment approaches apply!
- No use of tunneling makes encapsulation and MTU management much more straight-forward
- Traffic engineering of CLATs can be handled in the same way as NAT64
  - Anycast
  - Use of different prefixes (distributed to the CPE via configuration methods)
  - Etc
- 464XLAT has been deployed at T-Mobile USA – refer to Cameron Byrne's presentation

6RD
INTRODUCTION

- 6 Rapid Deployment – RFC5969 *IPv6 Rapid Deployment on IPv4 Infrastructures*
- A tunneling technology based loosely on 6to4
- 6rd allows IPv6 to be deployed over existing IPv4-only access networks, without any forklift upgrades to the access, aggregation, or subscriber management networks
- All addresses are automatically discovered by the CPE, while the BR address may be statically configured or discovered via a variety of mechanisms (e.g. dhcp option)
- Fits well for wireline network environments where a CPE swap or upgrade is easy, but access networks are complex or expensive to modify (or are third party)
- Device-to-device traffic may be routed directly, and not through the BR when staying within a 6rd domain
- 6rd has plans under discussion for eventual sunsetting in favor of native IPv6 (dual or single stack)

6RD



6RD
DOMAIN IMPACT

ACCESS	• No impact for 6rd – access network remains **exactly the same** for the initial deployment • Subsequent migrations may still be required to get to ultimate end-state • E.g. native dual-stack, DS-Lite, or similar

SUBSCRIBER EDGE	• Border Relay is needed in the network
	  • May be colocated in the BNG or a dedicated element
	  • Load balancing of elements should be considered, as well as traffic engineering
	• No change to subscriber management at the BNG
	• Potential loss of visibility of tunneled traffic at BNG
	• Lawful Intercept and DPI need to be considered

HOME NETWORK	• Customer Gateway (DSL modem/router, cable modem, etc) most likely needs to be replaced or upgraded – must support 6rd. • Many RGs are shipping 6rd support today • Home network components need to support IPv6 for native services • IPv4 NAT at the customer gateway is still present • Potential MTU impact for tunnels – potentially higher WAN MTU or frag-support required • Useful for environments where the access network can't be touched (wholesale..)

6RD

- 6rd targets the case where operators wish to immediately deploy IPv6 to their subscriber base, but cannot enable it in the native access
- As 6rd encapsulates IPv6 in IPv4, it can be deployed across any existing IPv4 network
  - Some constraints faced by operators that drive 6rd deployment include legacy Access Nodes (e.g. DSLAMs) that cannot support forwarding IPv6 packets, or older access technologies (e.g. DOCSIS 1.1) that cannot support IPv6
  - L3 wholesale access environments that cannot support IPv6 are another common barrier to deployment
- Significant impact in the CPE domain as the CPE must be upgraded to support 6rd
  - CPE are commercially available today that support 6rd and vendor support is continuing to increase
- Access network and subscriber management edge face no changes
- 6rd Border Relays must be deployed in the network
  - BRs should be located near subscriber traffic (e.g. in or adjacent to the BNG)
  - Avoid hauling traffic to centralized locations in the network which may impact TE or interface scaling in the network core
  - A potential drawback to non-BNG located 6rd BR is that any DPI or other IPv4 classification may be forced to occur at 6rd BR or elsewhere in the network, potentially stranding existing investment or impacting service provider operations

6RD

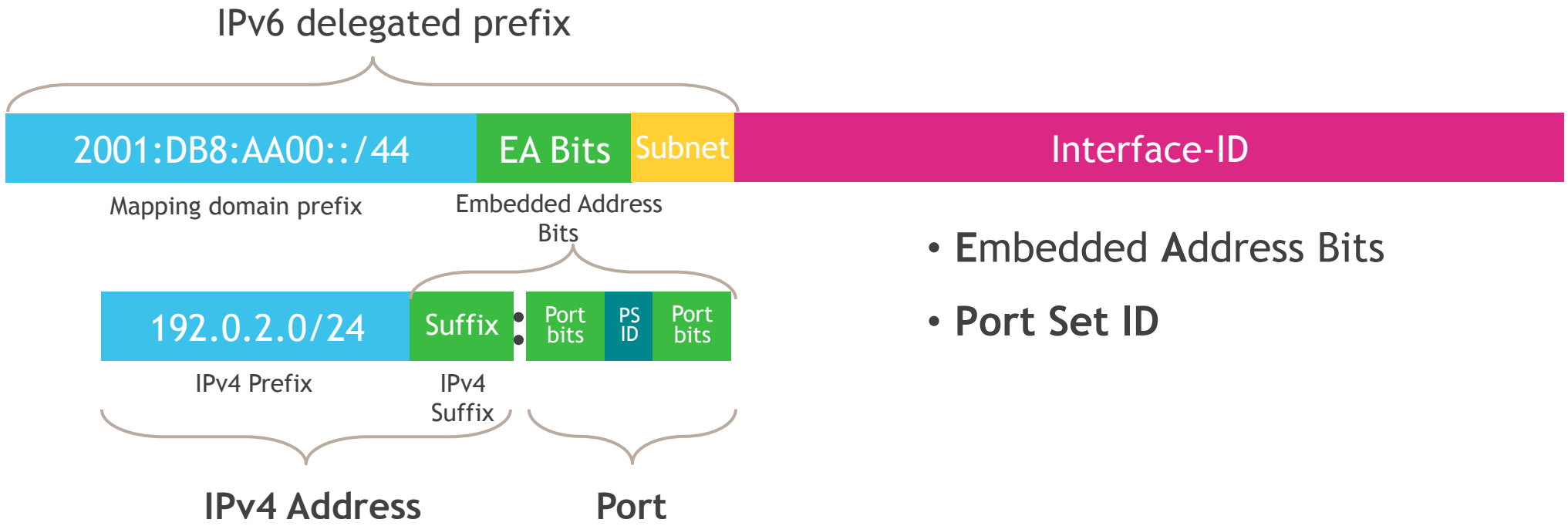
- 6rd may automatically derive the subscriber prefix with variable length subnetting (e.g. 48-56-64) based on the IPv4 address
  - Operators must consider exactly how many IPv4 bits they wish to stuff into the IPv6 prefix
  - Does this approach impacts any RIR allocated IPv6 prefixes?
- 6rd does not force re-architecture of existing service offerings such as VoIP and IPTV which may remain on the existing IPv4 service
- 6rd can be deployed incrementally with no impact to the subscriber base as and when CPE are upgraded to support 6rd
- 6rd does not solve the long term problem of removing IPv4 from the access network or moving to native IPv6 services, however some discussion for this is being undertaken in the IETF
- Potential MTU issues may occur with the tunnel, but may be mitigated by increasing WAN MTU or implementing fragmentation in the 6rd BR and CPE

MAPPING ADDRESS AND PORT (MAP)
INTRODUCTION

- Mapping Address and Port (MAP) refers to two similar technologies:
  - MAP-Encapsulation (MAP-E), defined in *draft-ietf-softwire-map*
    - An approach that uses IPv4-in-IPv6 **encapsulation** to transport IPv4 packets over IPv6 and a mechanism for mapping between IPv6 address and IPv4 addresses and transport layer ports
    - Standards track document
  - MAP-Translation (MAP-T), defined in *draft-ietf-softwire-map-t*
    - An approach that uses **translation** between IPv4 and IPv6 address families to support IPv4 over an IPv6 network and a mechanism for mapping between IPv6 address and IPv4 addresses and transport layer ports
    - Experimental track document
- MAP is an approach that uses stateless address sharing at the service provider Border Router (BR) and stateful address sharing at the CPE; while transporting packets over IPv6
  - Leveraging distributed statefulness
  - Leveraging IPv6 route aggregation
  - Provides a mapping between IPv6 addresses and IPv4 addresses

MAP
ADDRESS MAPPING

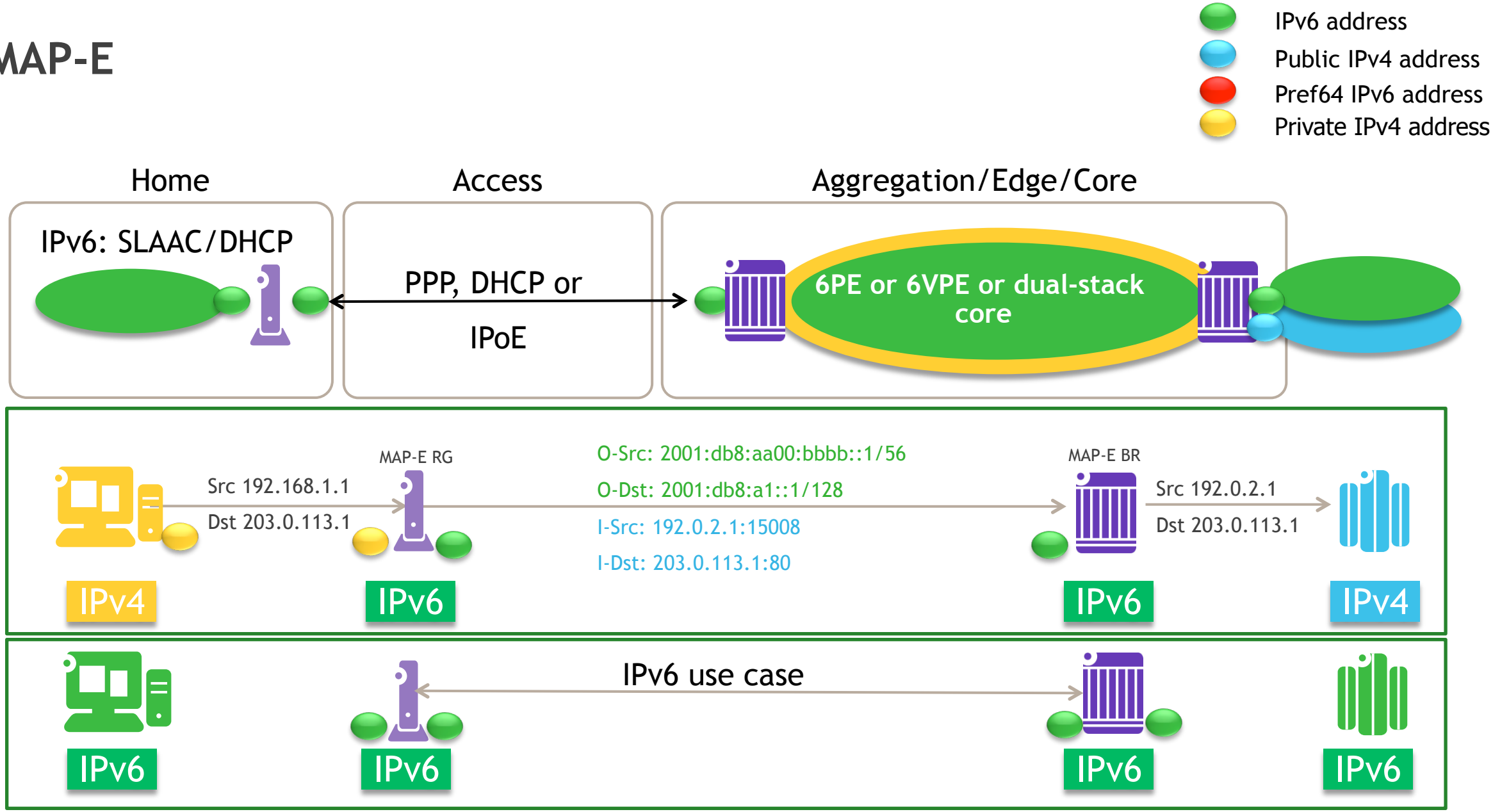
http://6lab.cisco.com/map/MAP.php



- **Embedded Address Bits**
- **Port Set ID**

- Example above shows a /44 prefix allowing for 4096 /56s (e.g. GPON OLT), and a /24 for address overloading
  - 12 EA bits identify the CPE
  - 8 of those bits are used in the host portion, 4 of them in the port range
  - Allows for 4096 subscribers to share 256 IPv4 addresses, with 4032 ports each (1:16)
- CPE must source traffic from within the correct port range (corresponds to PSID)

MAP-E



MAP-E SUMMARY

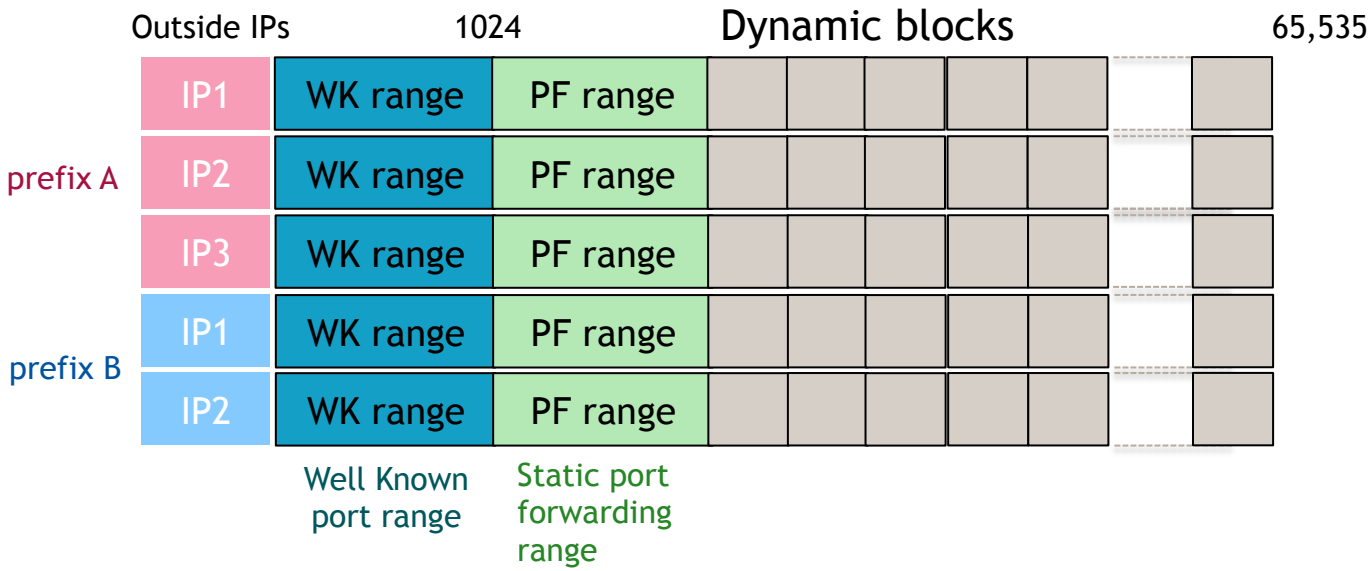
- MAP-E provides a mechanism for operators to provide IPv4 services in an IPv6-only network, without requiring translation state to be kept in the service provider environment
- By distributing the NAT functionality to the CPE, the processing requirement is distributed across many devices
  - CPU and memory distribution efficiency
  - No state to be managed in the SP core which makes failover between BRs much easier
- While address mapping architecture makes the IPv4-IPv6 identity easy, the Service Providers can only identify shared IPv4 to subscriber mapping if they are provided with the source port used by the subscriber
  - If the source port is not provided the mapping cannot be made as with other technologies
  - This may require some operators to deploy detailed flow logging
- CPE support is currently limited
  - Roadmap item for CPE
- Dynamic port block extension is currently not possible
  - Port exhaustion may become a challenge
- Fragmentation and reassembly presents performance and challenges if the subscriber link cannot support the IPv6 overhead with full IPv4 packet sizes

OTHER TRANSITION TECHNOLOGIES

- Not talked about today, but are other technologies that are used:
  - 6in4
    - Very simple, lightweight tunneling of IPv6 over IPv4 (proto-41), defined in RFC4213. Forms the basis of encapsulation for many other transition technologies
  - 6to4
    - Leverages simple 6in4 tunneling combined with a reserved prefix (2002::/16) to allow for IPv4-derived IPv6 prefixes, which allows tunneling over the IPv4 network between endpoints using the IPv4 address that is embedded in the prefix. This can be combined with anycast based relays to connect 6to4 networks to the native IPv6 network
  - 4in6
    - Encapsulates IPv4 into IPv6, defined in RFC2473. Forms the basis of encapsulation for other technologies
  - SIIT
    - Stateless IP/ICMP translation that allows an IPv6-only host to talk to an IPv4-only host, via header translation. Leveraged by other transition technologies
  - (d)IVI
    - Stateless IPv4-IPv6 transition technology that allows for 1:1 mapping of IPv6-to-IPv4 and for N:1 mapping of IPv6-to-IPv4

OPERATIONAL CONSIDERATIONS

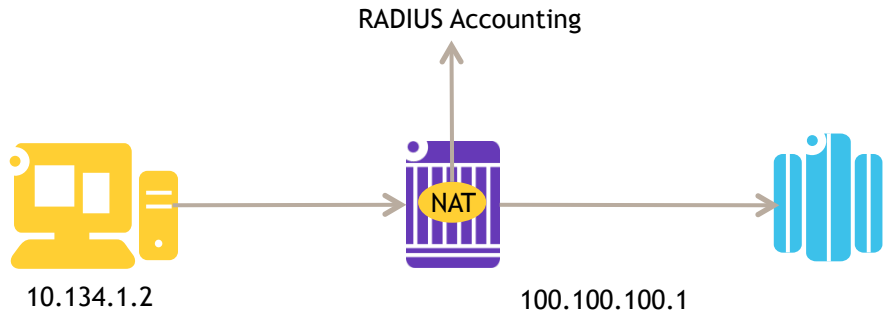
LOGGING OF DYNAMIC MAPPINGS



- Logging of dynamic mappings is required for subscriber traceability
  - Controlling logging transaction volume is essential
- Configurable port-block size can vastly reduce the amount of logging information as events are only generated on block allocation or deletion
- Retention timers allow outside IP to subscriber mapping to persist beyond the last block deletion
- Typical logging options:
  - Simple logging (e.g Syslog)
  - RADIUS accounting logging
  - Full flow based logging

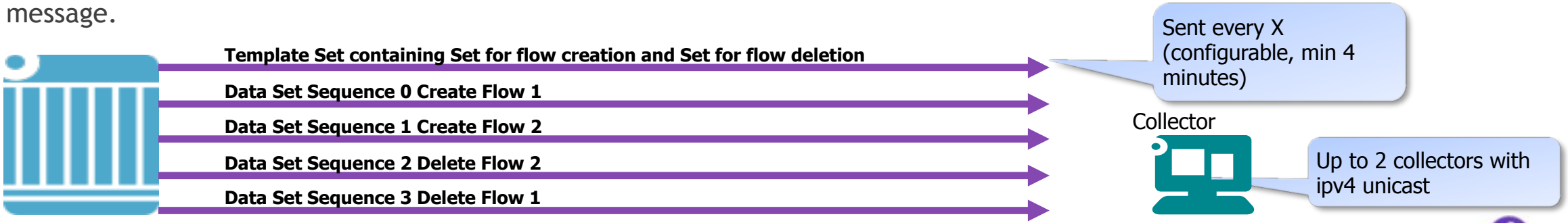
RADIUS ACCOUNTING EXAMPLE

```
2  {Item Worker 8} <radiusItem.request-decode> Request decode:
        NAS-IP-Address = 10.0.222.132
        NAS-Identifier = "7750-132"
        NAS-Port = 69271552
        Acct-Status-Type = Start
        Acct-Multi-Session-Id = "4fabf3a71bcfede039600727739600727"
        Acct-Session-Id = "4fabf3a71bcfede08fabf4a739600727"
        Called-Station-Id = "00-00-00-00-01-01"
        User-Name = "LSN44@10.134.1.2"
        Alc-Subsc-ID-Str = "LSN44@10.134.1.2"
        Alc-Serv-Id = 99
        Framed-IP-Address = 10.134.1.2
        Anonymous = v6527-a140-00000064
        Anonymous = v6527-a141-64646401
        Alc-Nat-Port-Range = "100.100.100.1 8080-8583 router 100"
        Acct-Input-Packets = 0
        Acct-Output-Packets = 0
        Acct-Input-Octets = 0
        Acct-Output-Octets = 0
        Acct-Input-Gigawords = 0
        Acct-Output-Gigawords = 0
        Acct-Session-Time = 0
        Event-Timestamp = "2012/05/10 18:58:15"
```



FLOW LOGGING EXAMPLE

- Why is it required?
  - Port range logging is the best scalable logging solution, however could be desirable to log per flow
- How is it done?
  - NAT flow logging allows LSN to export the creation and deletion of NAT flows to an external server
  - Uses (e.g.) IETF IPFIX NetFlowv10 RFC5101 format. Data structures are contained in RFC5102
    - UDP streams are stateless due to the significant volume of transactions, however they do contain sequence numbers such that packet loss can be identified
  - IPFIX defines two different type of messages that will be sent from the IPFIX exporter:
    - Template Set – an IPFIX message that defines fields for subsequent IPFIX messages but contains no actual data of its own
    - Data Sets – here the data is passed using the previous Template Set message to define the fields. This means an IPFIX message is NOT passed as sets of TLV, but instead data is encoded with a scheme defined through the Template Set message.



FLOW LOGGING EXAMPLE
DRAFT-IETF-BEHAVE-IPFIX-NAT-LOGGING

Timestamp	64 bit
natInstanceId	32
vlanID	16
ingressVRFID	32
sourceIPv4Address	32
postNATSourceIPv4Address	32
protocolIdentifier	8
sourceTransportPort	16
postNAPTsourceTransportPort	16
destinationIPv4Address	32
postNAPTdestinationTransportPort	16
sourceIPv6Address	128
destinationIPv6Address	128
postNATSourceIPv6Address	128
postNATDestinationIPv6Address	128

internalAddressRealm	8
externalAddressRealm	8
natEvent	8
portRangeStart	16
portRangeEnd	16
natPoolId	32
natLimitEvent	32

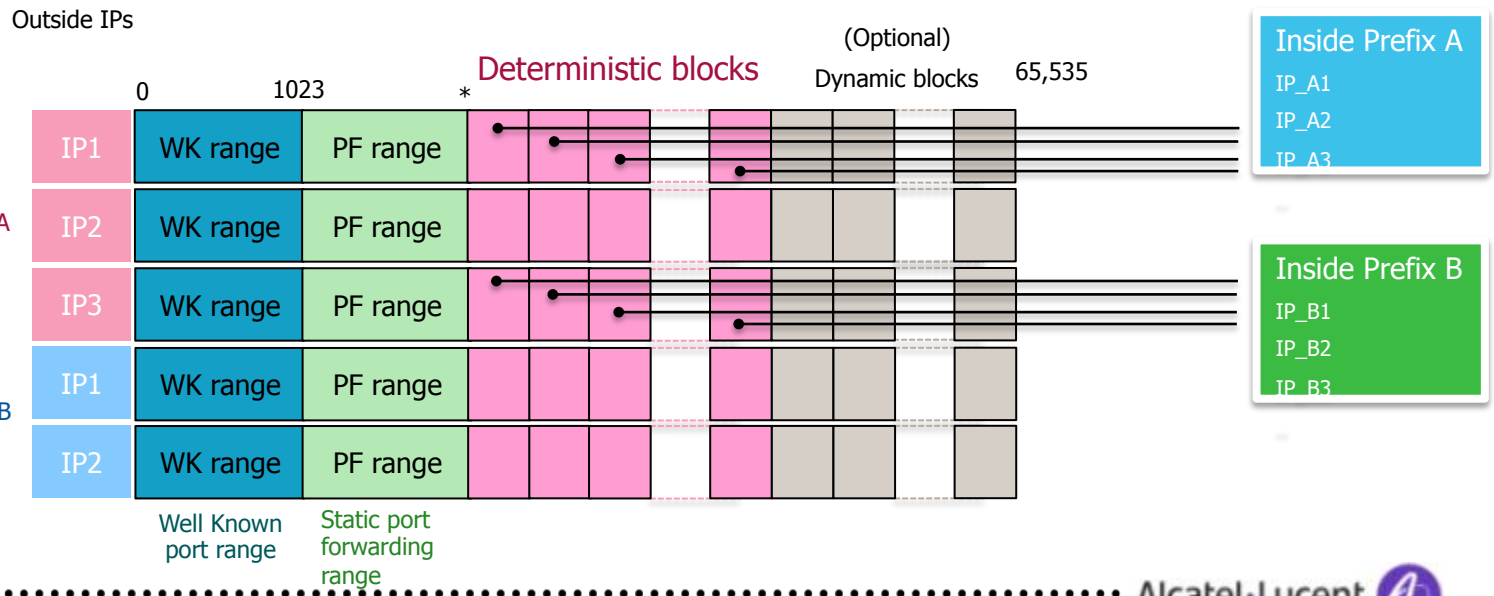
Such as:
- NAT44 Create
- NAT44 Delete
- NAT addresses exhausted
- NAT64 Create
- ...

FLOW LOGGING
NAT64 CREATE/DELETE EXAMPLE

Timestamp	1393261126
natInstanceId	1234
vlanID/ingressVRFID	3003
sourceIPv6Address	2001:db8:abba:baba:1234::1
postNATSourceIPv4Address	192.0.2.27
protocolIdentifier	6 [TCP]
sourceTransportPort	3333
postNAPTsourceTransportPort	4963
destinationIPv6Address	64:ff9b::CB00:711A
postNATDestinationIPv4Address	203.0.113.26
destinationTransportPort	80
postNAPTdestionationTransportPort	80
internalAddressRealm	3003
externalAddressRealm	1234

DETERMINISTIC NAT

- Deterministic NAT allows operators to have a consistent, reversible mechanism for subscriber mapping to port blocks
  - Each subscriber is permanently mapped to an outside IP and a dedicated (deterministic) port block based on algorithm defined in *draft-donley-behave-deterministic-cgn*
- This permanent mapping is referred to as “deterministic port block”
- No support for overbooking; all inside IP addresses are permanently mapped to an outside IP + port range
- No need for logging as the reverse mapping can be obtained using a known formula
- Subscriber ports can be expanded by allocating an extra dynamic port block (logging required)



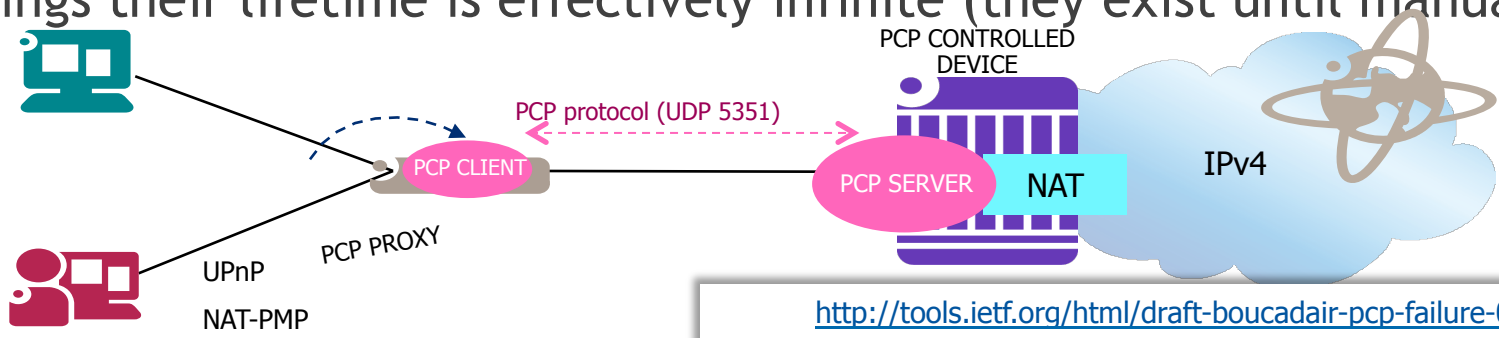
PCP (PORT CONTROL PROTOCOL)
DRAFT-IETF-PCP-BASE-XX

- PCP allows applications to create mappings from an external IP address and port to an internal IP address and port. These mappings are required for successful inbound communications destined to machines located behind a NAT or a firewall.
- After creating a mapping for incoming connections, it is necessary to inform remote computers about the IP address and port for the incoming connection
  - This is usually done in an application-specific manner
  - For example, a computer game might use a rendezvous server specific to that game
- Reduces NAT-friendly keepalives:
  - PCP learns (and influence) the NAT mapping lifetime. This helps reduce bandwidth on the subscriber's access network, traffic to the server, and battery consumption on mobile devices.
- PCP outdates ALGs:
  - ALGs create mappings for applications that establish additional streams or accept incoming connections
  - ALGs incorporated into NATs may also modify the application payload.
  - Industry experience has shown that these ALGs are detrimental to protocol evolution.
- PCP subject to transport layer protocols with ports and without ports

PCP ALLOWS AN APPLICATION TO CREATE ITS OWN MAPPINGS IN NATS AND FIREWALLS
DEPLOYABLE FOR NAT444/NAT64/DS-LITE OR FIREWALL

PCP (PORT CONTROL PROTOCOL)
DRAFT-IETF-PCP-BASE-XX

- Mappings:
- Implicit dynamic mappings: traffic such as an outgoing TCP SYN or outgoing UDP packet. Not originally designed for creating NAT (or firewall) state (pass through a NAT device)
- Explicit dynamic mappings are created as a result of explicit PCP MAP and PEER requests.
- Explicit static mappings are created by manual configuration
- Implicit and explicit dynamic mappings are dynamic and have a lifetime
- Explicit static mappings their lifetime is effectively infinite (they exist until manually removed)

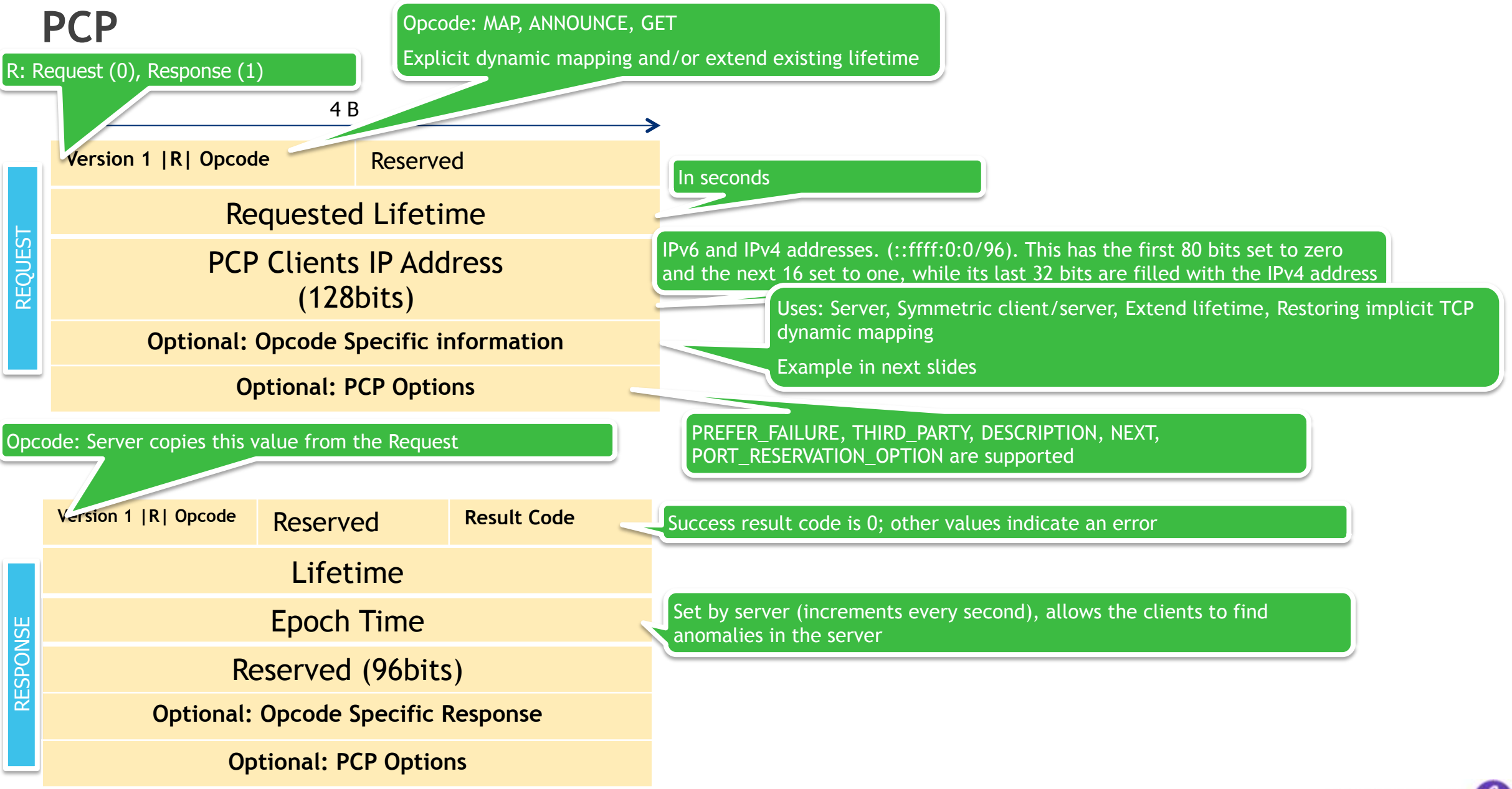


http://tools.ietf.org/html/draft-boucadair-pcp-failure-02

http://tools.ietf.org/html/draft-boucadair-pcp-rtp-rtcp-02

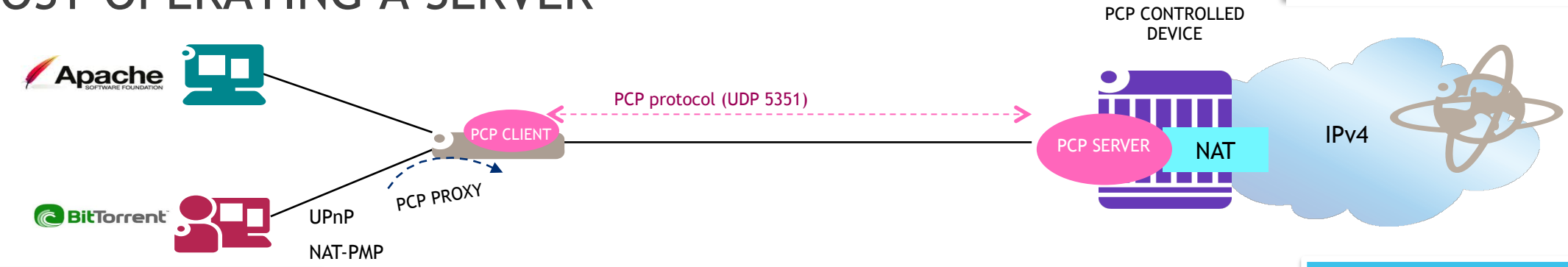
http://tools.ietf.org/html/draft-boucadair-pcp-extensions-01

PCP



PCP EXAMPLE
HOST OPERATING A SERVER

*Header not shown:
Includes Lifetime
PCP Option: NO PREFER_FAILURE and THIRD_PARTY (when DSLite, UPnP and NAT-PMP)



OPCODE* (MAP) REQ

4 B

Protocol (6)	Reserved (24 bits, all 0s)
Internal Port (1500)	Suggested Port (1500)
Suggested External IP Address (128 bits) ::ffff:135.203.19.215	

OPCODE (MAP) RESP

4 B

Protocol (6)	Reserved (24 bits, all 0s)
Internal Port (1500)	Assigned Port (1500)
Suggested External IP Address (128 bits) ::ffff:10.203.1.20	

MAP REQ (Renew Lifetime)

MAP RESP (Renew Lifetime)

Server listens for traffic on a port, never initiates:
- Needs to create mapping in the NAT
- Announce the port mapping in a Rendevouz (DNS, SIP message, etc)

ADDRESS MANAGEMENT
HOW MUCH OVERLOAD IS TOO MUCH?

- Defining subscriber port blocks for IPv4 is a common operational challenge
- The answer is usually “it depends”
  - Subscriber type/network type
  - How much logging do you need to do
  - What algorithm are you using for address overloading (deterministic, MAP, LSN-alike, etc)
- Ideally, you do not want subscribers to exhaust port blocks as subscriber experience is impaired
  - The infamous “Google Maps will not load tiles” example
  - How to handle this in distributed technologies?
- Comfortable amounts to start with for most operators are 2:1, 5:1
  - 10:1 and even 20:1 **may** be sustainable in some environments
  - 100:1 or above may introduce operational challenges
- Analysis of subscriber types and traffic flows is required and cautious implementation is always recommended

LAWFUL INTERCEPT

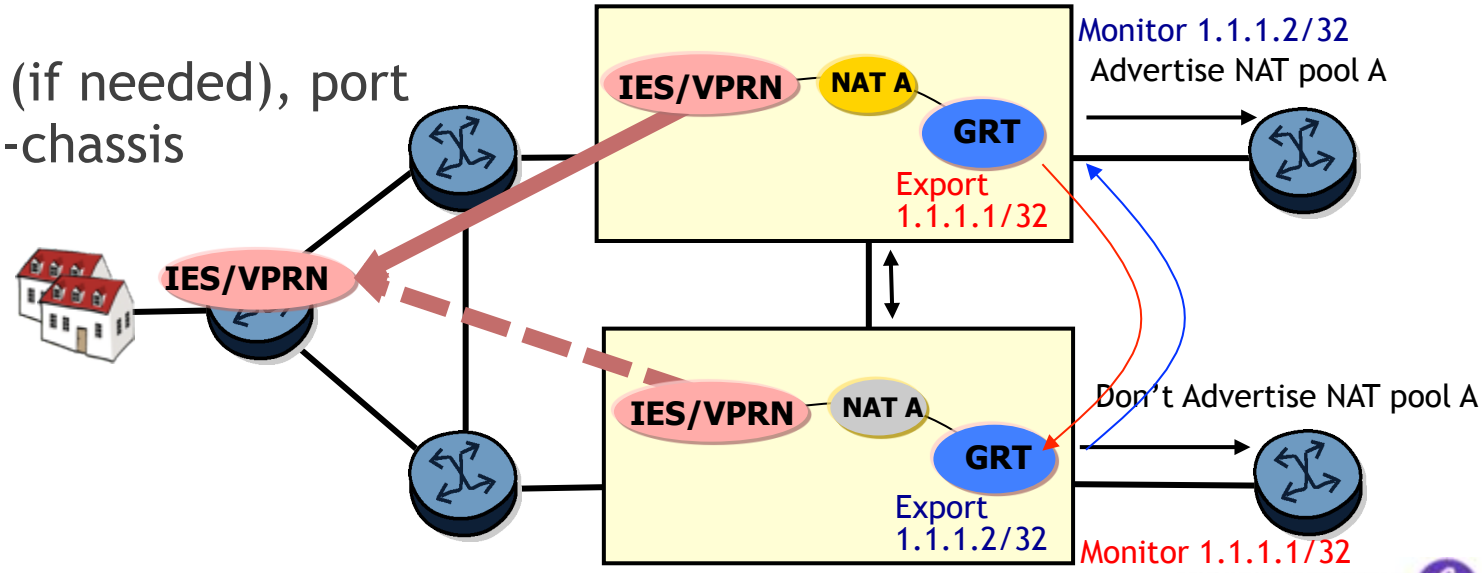
- Lawful Intercept with transition technologies can become an interesting challenge
- LEAs tend to want intercepts to be un-tampered-with (i.e. subscriber traffic must be original)
  - Does this mean we want inside packets or outside packets?
  - In the case of technologies which involve CPE, can we be assured of receiving all traffic?
  - Can the LEA digest encapsulated traffic?
- How to correlate packets with a subscriber?
  - Using outside IP information is most useful – with assured correlation with the subscriber
- This is a new area that many government authorities are investigating at the moment

CPE

- CPE support is the biggest requirement for deploying a transition technology
- Many CPE vendors are now supporting the most popular transition technologies, and in many cases, multiple transition technologies
- Time to market for new technologies remains a concern
- CPE and AFTR/LSN host functionality interop becomes important, particularly for encapsulated protocols where fragmentation behavior is important
- Some known CPE that support a variety of transition mechanisms:
  - D-Link
  - Technicolor
  - NEC
  - Zyxel
  - Linksys

NETWORK RESILIENCY

- NAT or transition technology resiliency becomes a key concern
- Many approaches that may be taken
  - Within the NAT node, state may be preserved across multiple service engines, with service engines providing re-hashing in the event of failure
  - Stateless technologies obviously have the advantage here (state sync is very challenging)
  - For some technologies, anycast based network load balancing is very feasible to provide resiliency and load balancing
- Synchronization of subscriber state (if needed), port forwarding etc is required for multi-chassis



COMPARISON OF TRANSITION TECHNOLOGIES

METHODS OF TRANSITION

Home device	Access network	Destination	Solutions
IPv4	IPv4	IPv4 Internet	Dual-Stack
IPv6	IPv6	IPv6 Internet	
IPv4	IPv6	IPv4 Internet	DS-Lite
IPv6	IPv6	IPv4 Internet	NAT64 Stateful
IPv6	IPv4	IPv6 Internet	6RD

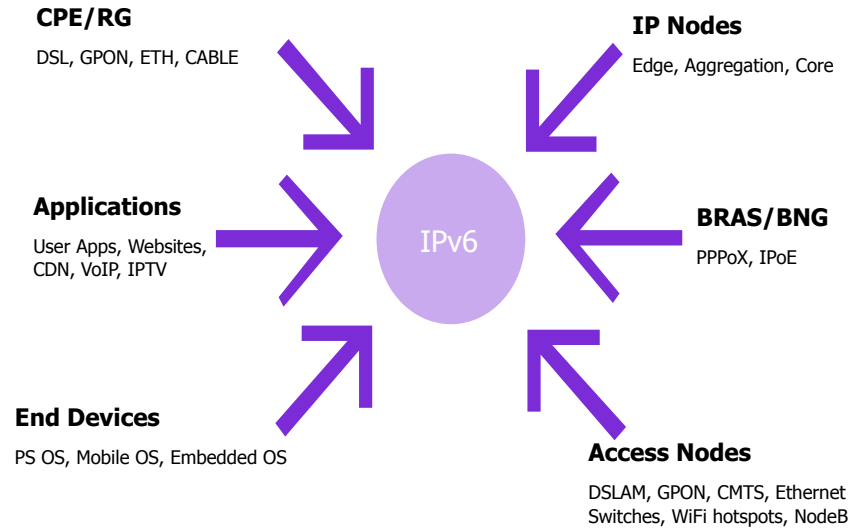
SUMMARIES AND COMPARISON

	Native Dual Stack	DS-Lite	NAT64	6rd
CPE	Almost always CPE change	CPE change and support for DS-Lite	CPE change (IPv6 only)	CPE change
End user impact	OK – not much changes	OK – not much changes	NOK – any IPv4-only devices (or partial-IPv6) are impacted. No non-DNS64 support	OK – not much changes
Pro	'Simple' technology with no transition or tunneling involved	Single address family in the access network	Single address family in the access network	Single address family in the access network Quick to deploy
Con	Cost of supporting dual-stack networks Device support Deployment time	All the effort of deploying dual-stack + extra Extra DS-Lite AFTR needed Traffic obfuscation in the network Device support	Application brokenness with IPv4-literals NAT logging required Will only work for IPv6-supporting hosts	Traffic obfuscation in the network Device support Not necessarily a 'long term' solution
Most suitable for…	Deployment everywhere! Best long term option that gives the widest support for both address families ➔Wireline, Wireless	New build environments where both removing IPv4 from and deploying IPv6-only access is feasible. ➔Wireline	New build environments where IPv6-only access is acceptable and the majority of content will work through NAT64/DNS64 ➔ Wireless environments	Legacy environments that cannot support native IPv6 access, and are willing to trade-off multi-stage migrations over the long term ➔ Wireline environments

Every transition technology employs translation – applications **will** be affected

CONCLUSION

CONCLUSION



- IPv6 deployment and transition technologies are a multi-dimensional problem

- There are a lot of transition technologies available with varying levels of support
- Operators should carefully evaluate which technology is most appropriate to meet their needs
  - The field is still changing!
- **The transition technology should align with the long term vision of the operator – generally this should look towards native IPv6 support**
- It might take multiple iterations to get to a long term view of native IPv6 (with transitional support for IPv4) – but it is important to try minimizing this from an investment and complication perspective

REFERENCES

Document	Location
Broadband Forum TR-177 IPv6 in Context of TR-101	http://www.broadband-forum.org/technical/download/TR-177.pdf
Broadband Forum TR-187 IPv6 for PPP Broadband Access	http://www.broadband-forum.org/technical/download/TR-187.pdf
Broadband Forum TR-124i3 Functional Requirements for Residential Gateway Devices	http://www.broadband-forum.org/technical/download/TR-124_Issue-3.pdf
RFC6333 Dual-Stack Lite Broadband Deployments	http://tools.ietf.org/html/rfc6333
RFC6052 IPv6 Addressing of IPv4/IPv6 Translators	http://tools.ietf.org/html/rfc6052
RFC6146 Stateful NAT64	http://tools.ietf.org/html/rfc6146
RFC6147 DNS64 DNS Extensions	http://tools.ietf.org/html/rfc6147
RFC5969 IPv6 Rapid Deployment on IPv4 Infrastructures	http://tools.ietf.org/html/rfc5969
draft-despres-softwire-4rd-u IPv4 Residual Deployment	http://tools.ietf.org/html/draft-despres-softwire-4rd-u-06
RFC7084 Basic Requirements for IPv6 Customer Edge Routers	http://tools.ietf.org/html/rfc7084
RFC6586 Experiences from an IPv6 Only Network	http://tools.ietf.org/html/rfc6586
464XLAT Experiences from T-Mobile USA	https://sites.google.com/site/tmoipv6/464xlat
RFC6219 CERNET IVI Translation Design and Deployment	http://tools.ietf.org/html/rfc6219
RFC6145 IP/ICMP Translation Algorithm	http://tools.ietf.org/html/rfc6145
RFC6144 Framework for IPv4/IPv6 Translation	http://tools.ietf.org/html/rfc6144
draft-townsley-v6ops-6rd-sunsetting Sunsetting for 6rd	http://tools.ietf.org/html/draft-townsley-v6ops-6rd-sunsetting-00
IPv6 CPE at the ARIN GetIPv6 Wiki	http://getipv6.info/index.php/Broadband_CPE
RIPE IPv6 CPE Survey	https://www.ripe.net/data-tools/ripe-labs/ipv6-cpe-survey

www.alcatel-lucent.com