CPE requirements and IPv6

Ole Trøan,
ot@cisco.com

February 2010

Past and present:

- Worked as an implementer on every aspect of the IOS IPv6 stack. Routing, access, provisioning, ND, DHCP PD, Transition…
- Managed the Japan Development Centre in Tokyo
- Currently working on IPv6 architecture and IPv4 exhaustion
- Editor of the IETF draft on “Basic IPv6 CE requirements”
- Editor of BBF TR-124i2 on IPv6 RG requirements
- Co-author of IETF 6rd mechanism

Agenda

- IPv6 provisioning
- Access networks issues
- CPE requirements
- Unresolved issues
- Real deployment – 6rd

Players in this space

- Cablelabs (DOCSIS)
- BroadBand Forum (TR-124i2)

  Access architectures: TR-101i2, WT-177, WT-187…
- IETF (Basic IPv6 CPE router requirements)

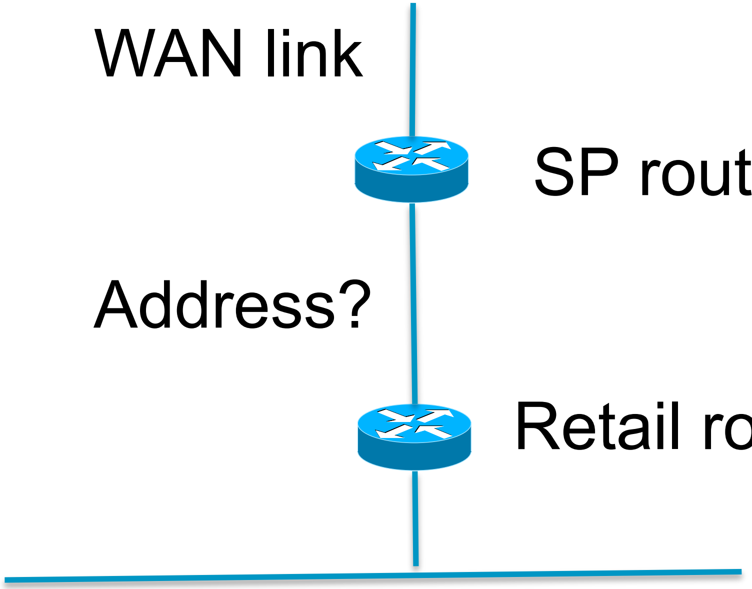
  v6ops

  Homegate BOF / Interim meeting in April
- UPnP forum (IGD)
- HomeGateInitiative
- IPv6 Promotion Council (Japan)
- +++

IPv6 provisioning any different?

- NAT in IPv4:
  - Implicit "Security"
  - Static prefix, even when link is down or before provisioned
  - Multi-homing kind of works
  - "Chained" CPEs kind of works even with plug and play
- Every customer gets a static address block:
  - Route injection / aggregation / fail-overs?
  - Renumbering in the home?
  - Issues before provisioning or when link is down
  - Multi-level NAT44 / Complex topologies?
  - Manually configured prefix/addresses?

Multi-level NAT



WAN link

SP router

Address?

Retail router

LAN link

Address?

Access technologies

- Cable – Cablelabs
- xDSL (PPP, IPoE) – BBF
- GPON - BBF
- 6rd – IETF

Problematic access architectures

- Multi-access link with “pretend” user-isolation

  1:N VLANs are particularly hard.

  Requires some sort of ND proxy, DHCPv6 snooping.

  BBF has been looping on these issues for a long time. Especially support of SLAAC and RGs as bridges.

  DHCP auth

- Access networks are not L3 agnostic. ARP snooping, L2 security features triggered by DHCP... All this must be replicated for IPv6.
- 6rd decouples deploying IPv6 to customers and deploying IPv6 in the access network.

Basic IPv6 CE router requirements

- http://tools.ietf.org/html/draft-ietf-v6ops-ipv6-cpe-router-04
- The absolute minimum “stuff” we could get consensus on. If anyone disagreed then it got taken out.
- Design team with BBF, Cablelabs, IETF
- Passed v6ops WG last call, on its way to IESG
- Planning an advanced version + homegate WG working in this space
- No new inventive work (almost)

CPE requirements (IETF)

WAN side

- Ethernet and PPP (independent IPCP and IPv6CP)
- Addressing: DHCPv6, SLAAC or “unnumbered”
  - Acts as a “host” for the purpose of ND
  - Controlled with M-flag, A-flag
- In “unnumbered” case, create an address from delegated prefix.
- DHCPv6 for DNS, SNTP
- DHCPv6 Prefix Delegation (MUST)
  - CPE will indicate size of prefix required
  - Expectation to give a long-lived /48-/56 to each customer

CPE requirements (IETF)

WAN side 2

- PD route injection and default router selection
  - Unless other information use router discovery (host mode)
  - No dynamic routing protocol by default
- Route injection into ISP IGP
  - Typically done by DHCPv6 snooping
  - State maintenance: Short lease times?
  - Future: BFD echo mode for state maintenance at PE?
- Nothing said about renumbering
  - Instant renumbering in case of failover to another BNG?
  - Short lifetimes are in any case problematic

CPE requirements (IETF)

LAN side

- By default give out ULA addresses and act as a site border router
- Separate /64 from the delegated prefix for each of its LAN side network interfaces
- Must at least support SLAAC, may support DHCPv6 address assignment
- DHCPv6 for DNS options etc.

CPE requirements (IETF)

Security

- Support “Simple security” (NAT equivalent)

  But not statement on default on or not.

- Ingress filtering according to BCP38

- Future:

  Work on “Advanced security”. Basically centrally managed firewall/IPS

  BUT, do you really need filtering on the CPE? What does a firewall give you anyway? IPv6 implementations are “modern” in the sense that they have grown up in the jungle. A firewall doesn’t stop malware anyway.

- “No security” might be the better option

CPE requirements (BBF)

- PD-192 -> WT-192 -> TR-124i2
- In straw ballot now.
- No conflicts with the IETF work. Includes some more features.
- Basically something an SP can use to create an RFP

CPE requirements 2 (BBF)

- Transitioning

  6rd – decouples IPv6 in the access network from delivering IPv6 service to customers.

  Ds-lite – A solution to IPv4 exhaustion in the SP network.

- QoS (marking on tunnels ++)
- Dynamic DNS, DNS proxy
- Detection of existing ULA prefixes on a link
- Rudimentary Hierarchical Prefix Delegation
- DHCPv6 vendor options

CPE requirements 3 (BBF)

- Multicast

  MLDv2
- Security

  Statefull firewall
- Bridging of IPv6 frames with an IPv6 host stack
- Support for RFC4191 (more specific routes)

Unresolved issues

- Multi-homing
- Multi-level NAT

  Chained NATs for IPv4 works, but not for IPv6

  Hierarchical DHCP PD?
- Complex topologies and auto-provisioning?
- Walled garden

  And multi-prefix with non-congruent topologies?

Unresolved issues 2

- IPv6 only nodes

  NAT64, but what about IPv4 literals?

- IPv6 only and DS hosts on the same link

  Do the host choose to use NAT64?

  Or do provisioning pick a DNS64 for IPv6 only hosts and not for DS hosts?

Unresolved issues 3

- DNS server/proxy reverse and forward zones
- Advanced security
- IPv6 transition mechanisms
- Service Discovery / Firewall traversal
- BFD echo mode

Products and deployments

- Various experiments moving towards production in residential space.
- Free (6rd)
- Expect a big 6rd announcement later in the month. As well as Comcast having announced 6rd.
- We’re working with numerous CPE vendors (with IETF or BBF hat on) to get IPv6 and 6rd support.
- IOS has had IPv6 support for a decade
- Linksys coming 2010 or early 2011
- Expect most CPE vendors to have products in 2010/2011.

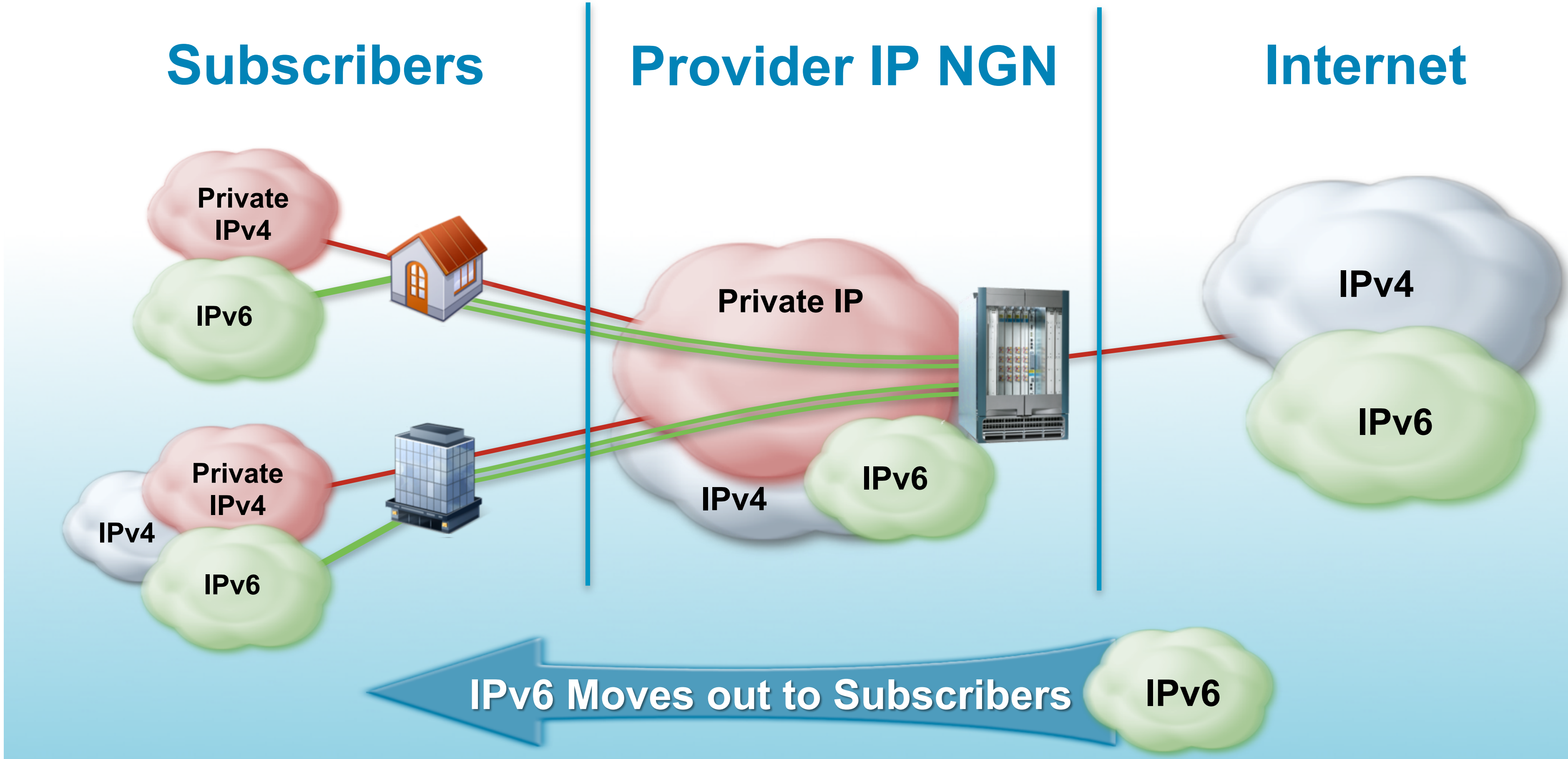
6rd

Ole Trøan,
ot@cisco.com

February 2010

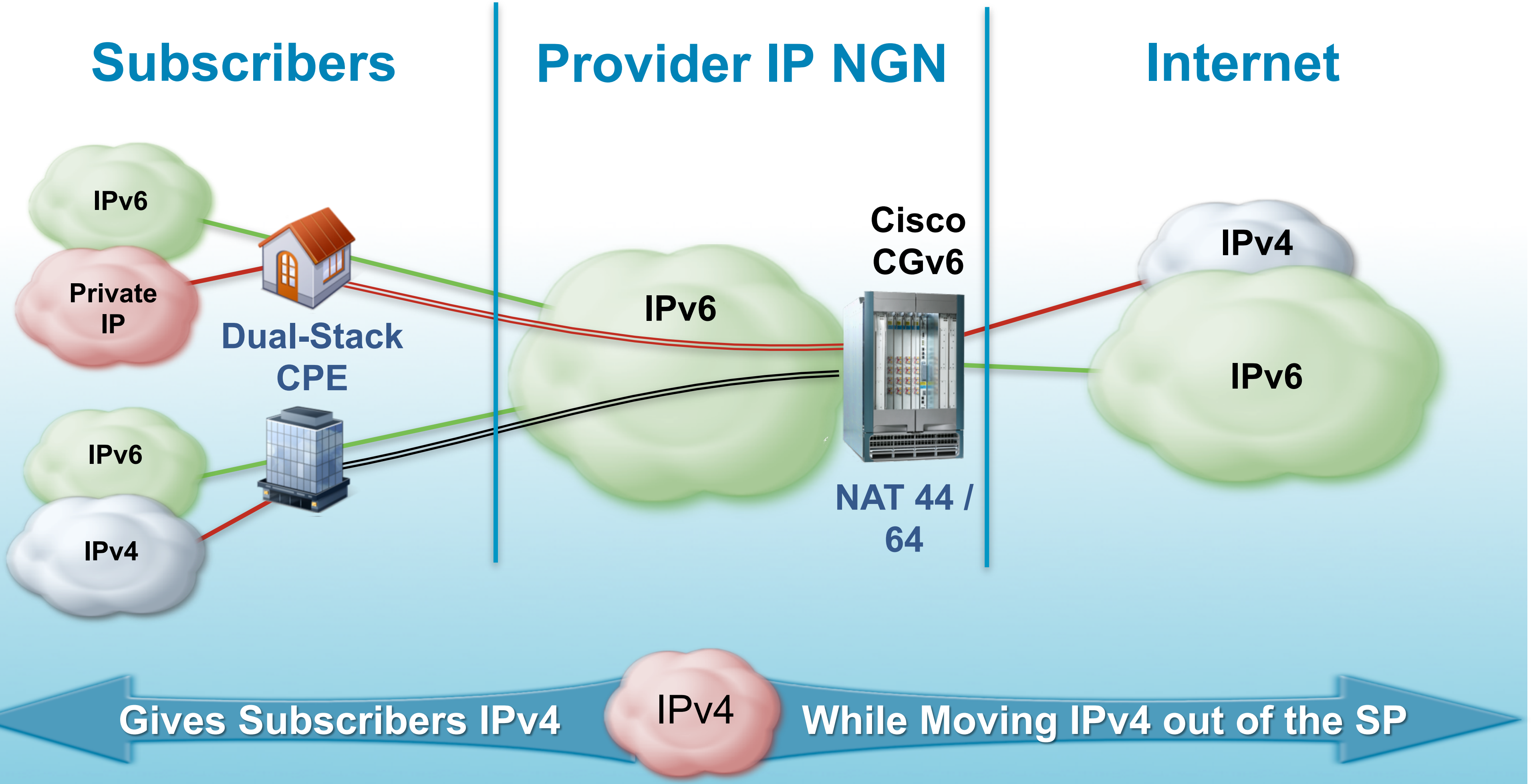
While Connecting IPv6 Islands
IPv6 over IPv4

6PE, 6rd

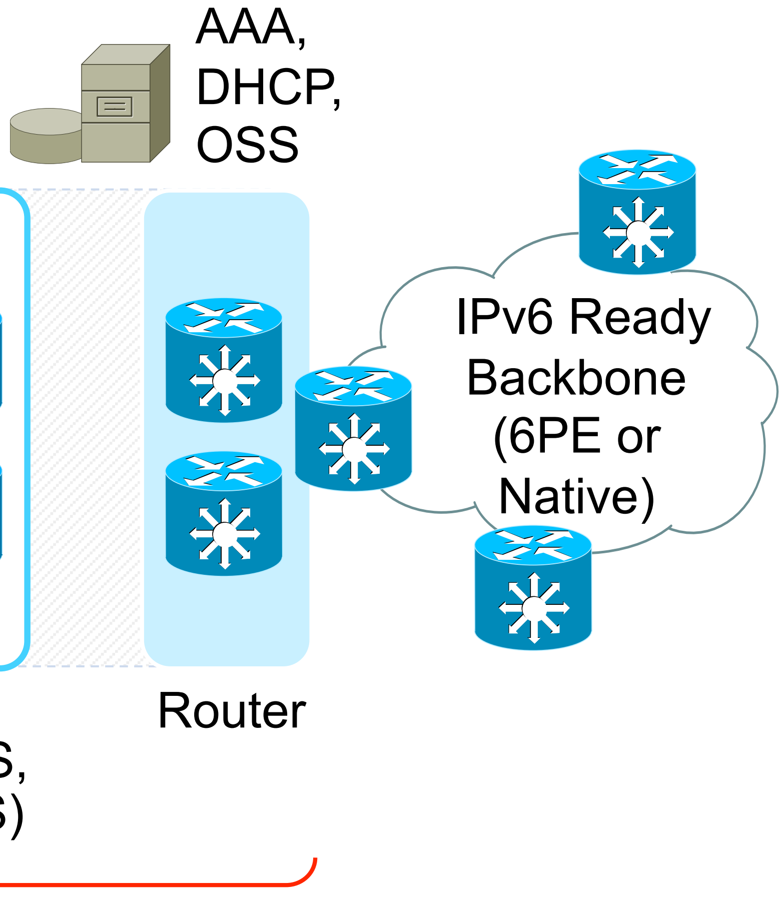


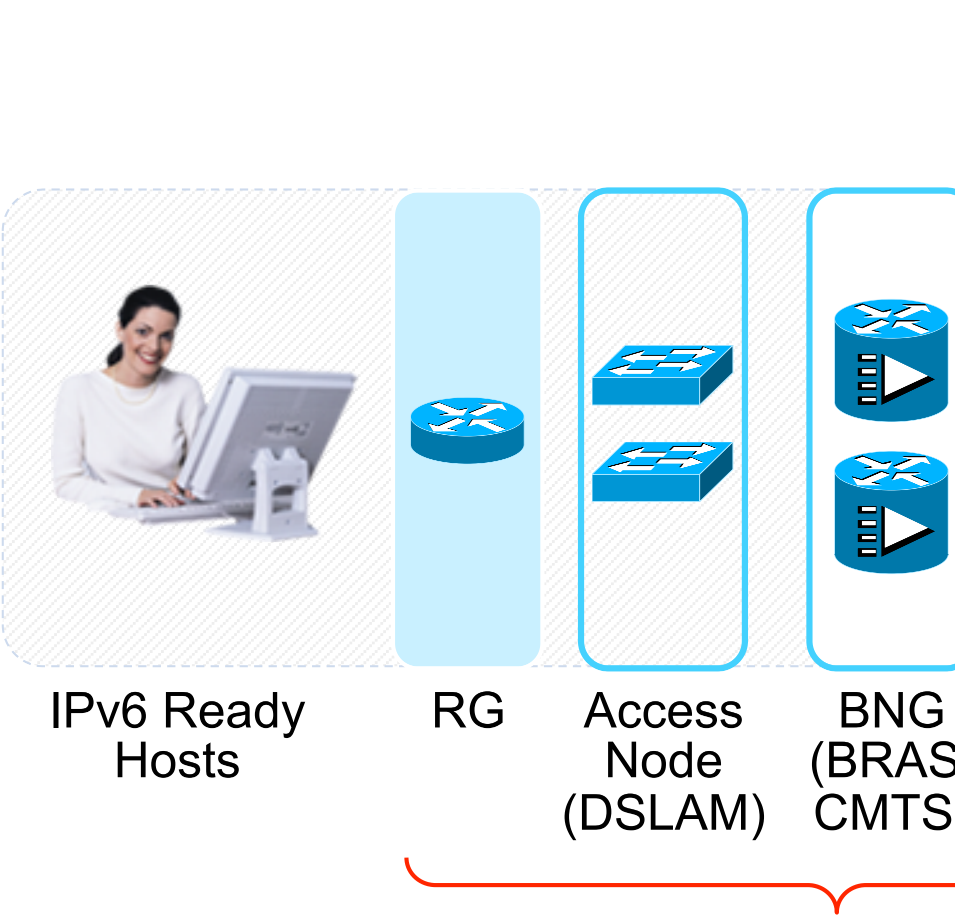
And Move Towards IPv6 Single Stack

"Dual Stack Lite"



Problem: Gap in IPv6 Availabilty



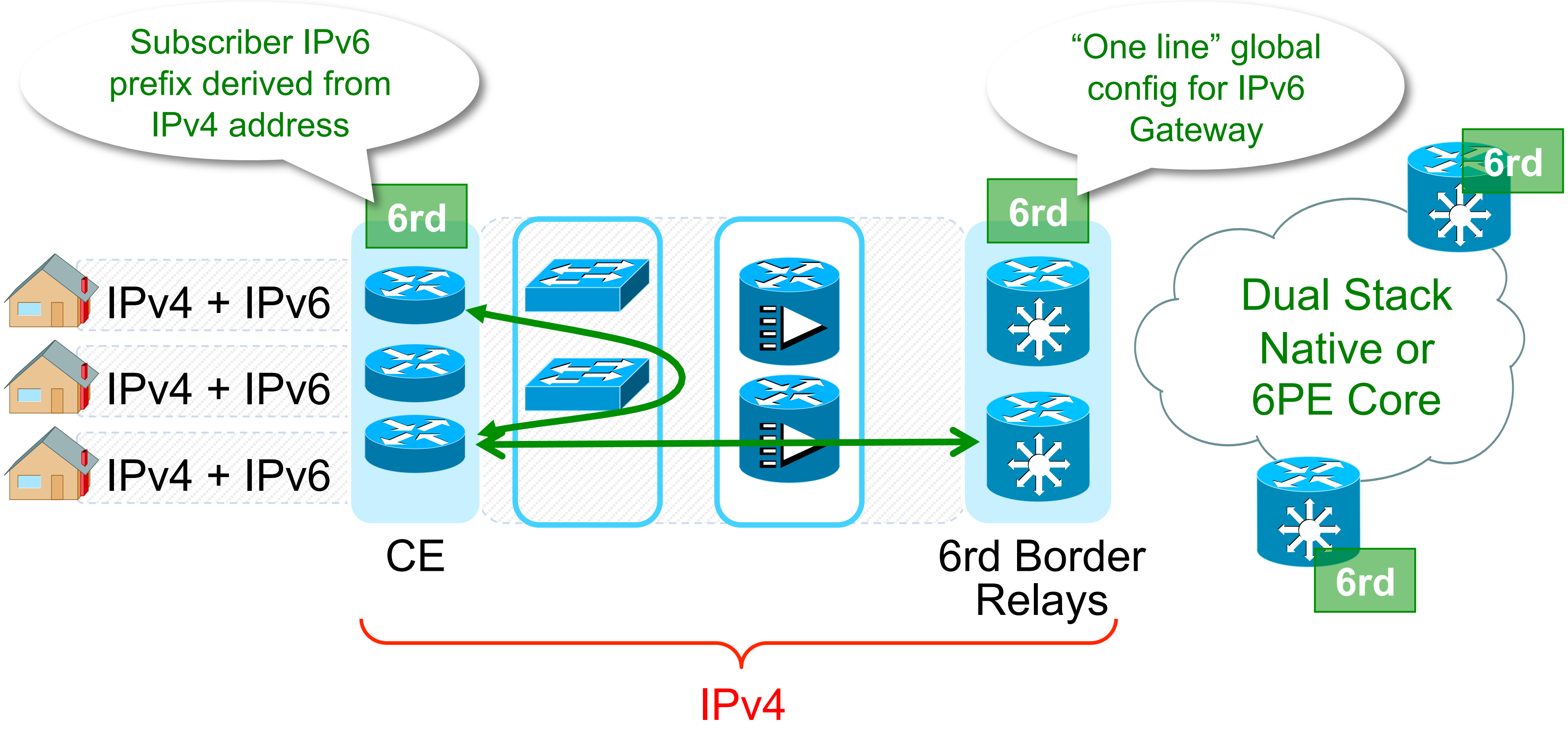


IPv4-Only Access, Aggregation, AAA

6rd: “6PE for the Access Network”

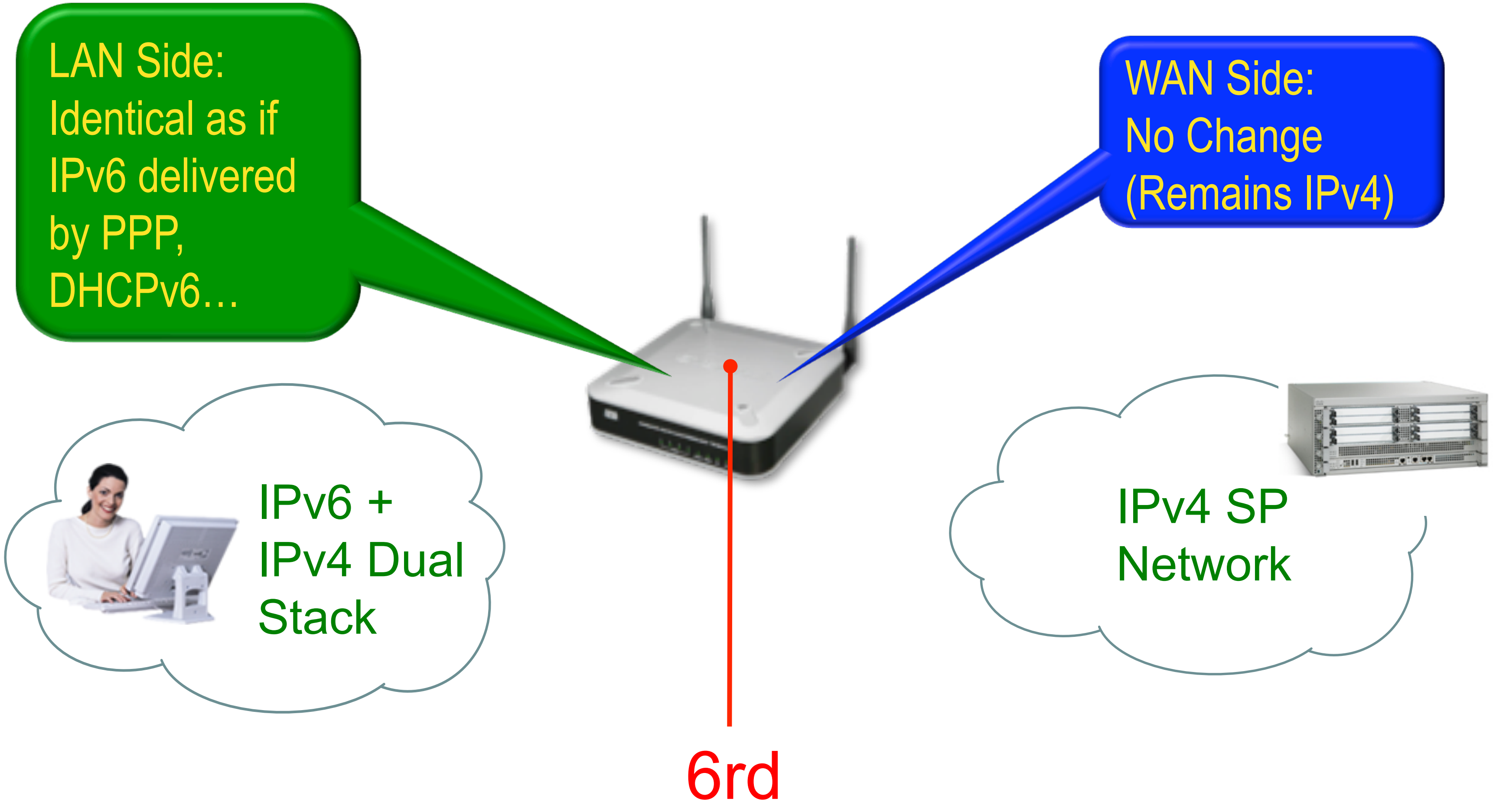
- Native dual-stack service to subscriber sites by leveraging existing access infrastructure, operations…
- Not an IPv6 “trial” service. IPv6 to subscriber is production-quality, native IPv6 + IPv4 dual-stack
- Reuses IPv4 in the SP—No v6 support needed in Access and Aggregation infrastructure, no DHCPv6 servers, no Neighbor Discovery…

6rd in One Slide

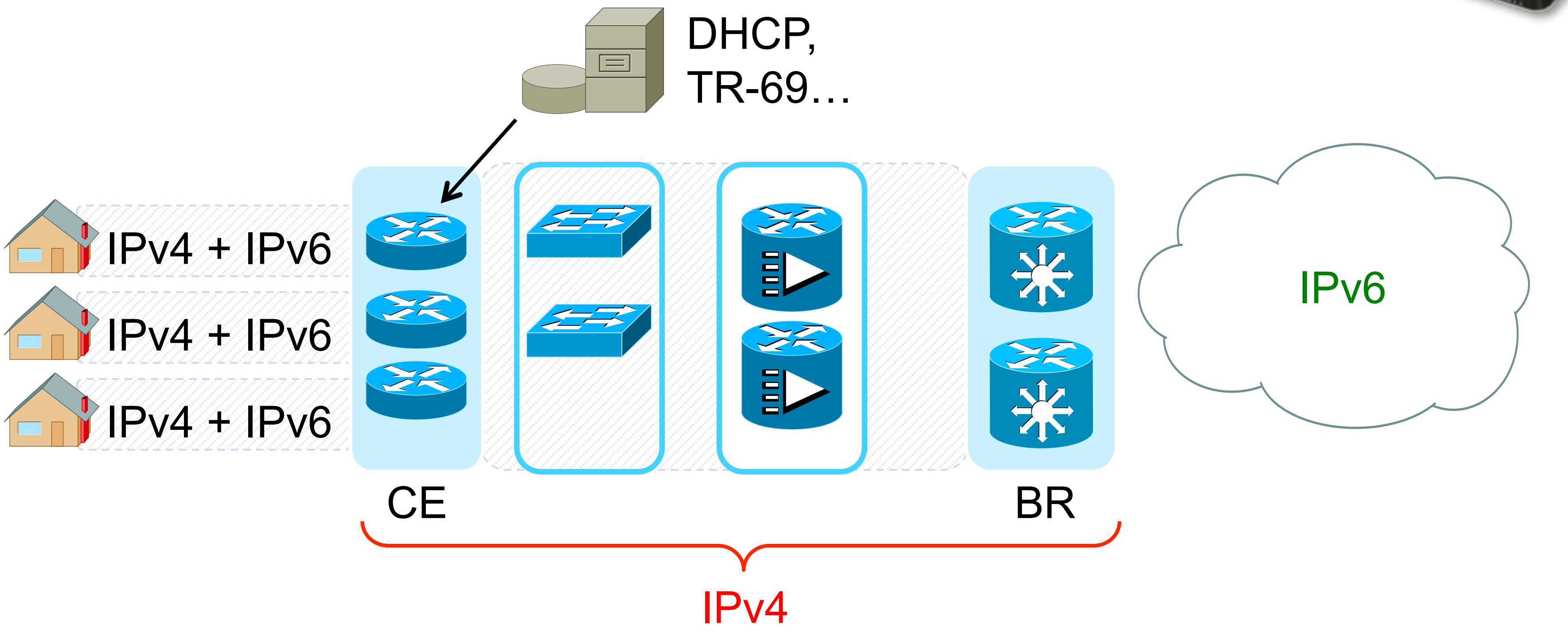

Subscriber IPv6 prefix derived from IPv4 address
"One line" global config for IPv6 Gateway
6rd
6rd
6rd
6rd
IPv4 + IPv6
IPv4 + IPv6
IPv4 + IPv6
Dual Stack Native or 6PE Core
CE
6rd Border Relays
IPv4

- Native dual-stack IP service to the Subscriber
- Simple, stateless, automatic IPv6-in-IPv4 encap and decap functions
- IPv6 traffic automatically follows IPv4 Routing
- BRs placed at IPv6 edge, addressed via anycast for load-balancing and resiliency
- Defined in `draft-ietf-softwire-ipv6-6rd`

Residential Gateway (6rd CE) Implementation

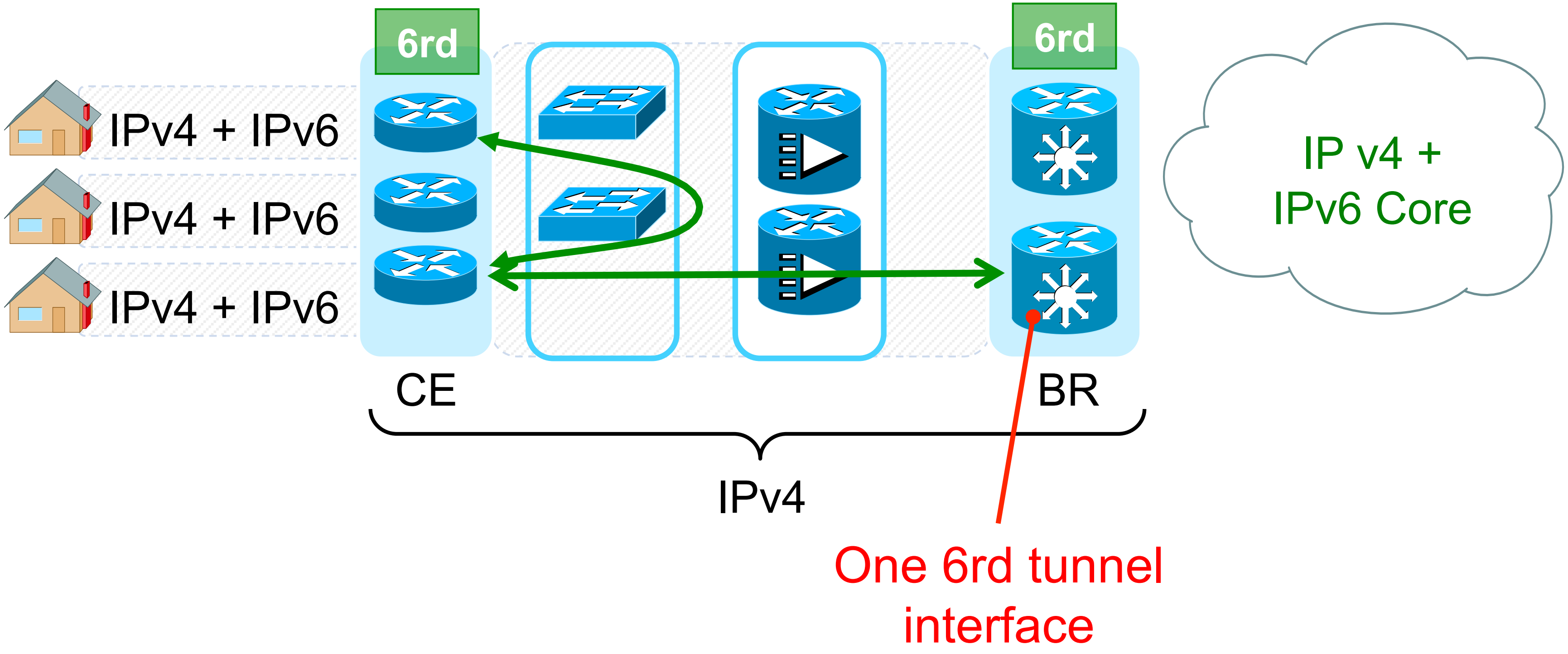


Residential Gateway Configuration



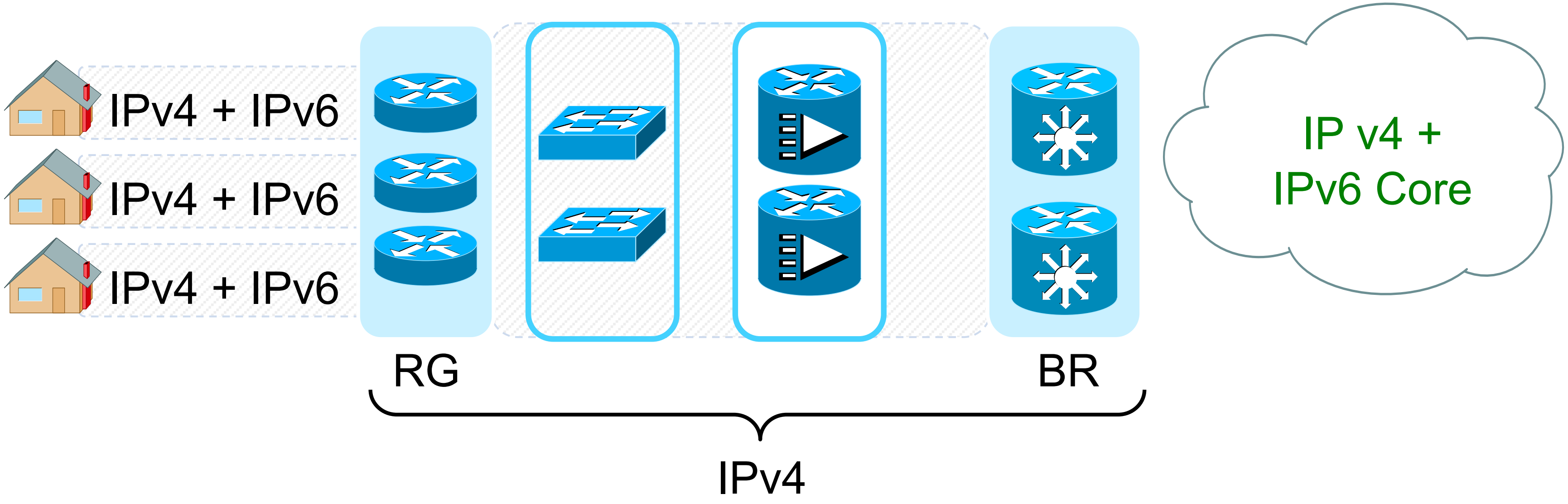
- IPv6 in home configured as for any dual-stack service, 6rd config elements identical for all CEs
  1) ISP 6rd IPv6 Prefix and length (e.g., 2011:100/28)
  2) Common IPv4 bits suffix length (e.g., 0 or 8)
  3) 6rd Relay IPv4 address (e.g., 10.100.100.1 – likely anycast)

Border Relay Implementation



- Single multipoint tunnel interface in Border Relay
- No per-user state, serves ALL users in 6rd Domain

Border Relay Configuration



- BR must have IPv6 reachability (Native, 6PE, GRE Tunnel…)
    1) ISP 6rd IPv6 Prefix and length (e.g., 2011:100/28)
    2) Common IPv4 bits suffix length (e.g., 0 or 8)
    3) 6rd Relay IPv4 address (e.g., 10.100.100.1 – likely anycast)

Protocol Details:
Three essential parts of 6rd:

① IPv6 Prefix Delegation derived from IPv4

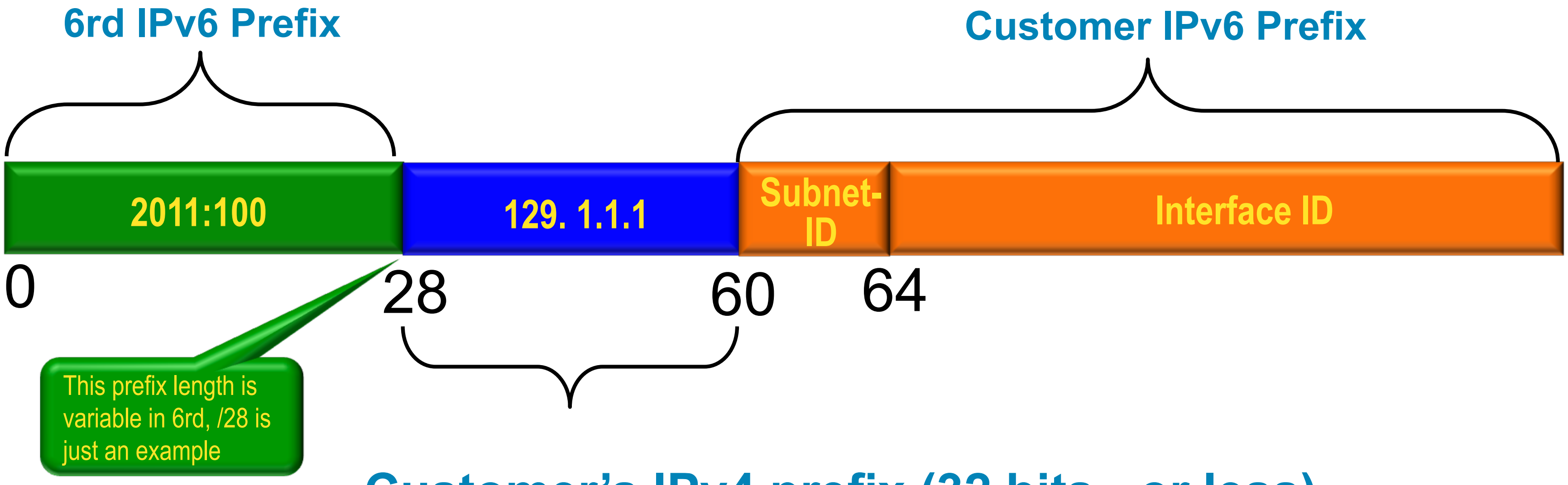
- ➢ Global IPv4 or Natted IPv4 in same deployment

② Stateless mapping and Encapsulation of IPv6 over IPv4 (RFC 4213)

- ➢ IPv4 encapsulation automatically determined from each packet’s IPv6 destination
- ➢ No per-subscriber tunnel state or provisioning

③ IPv4 Anycast to reach Border Routers

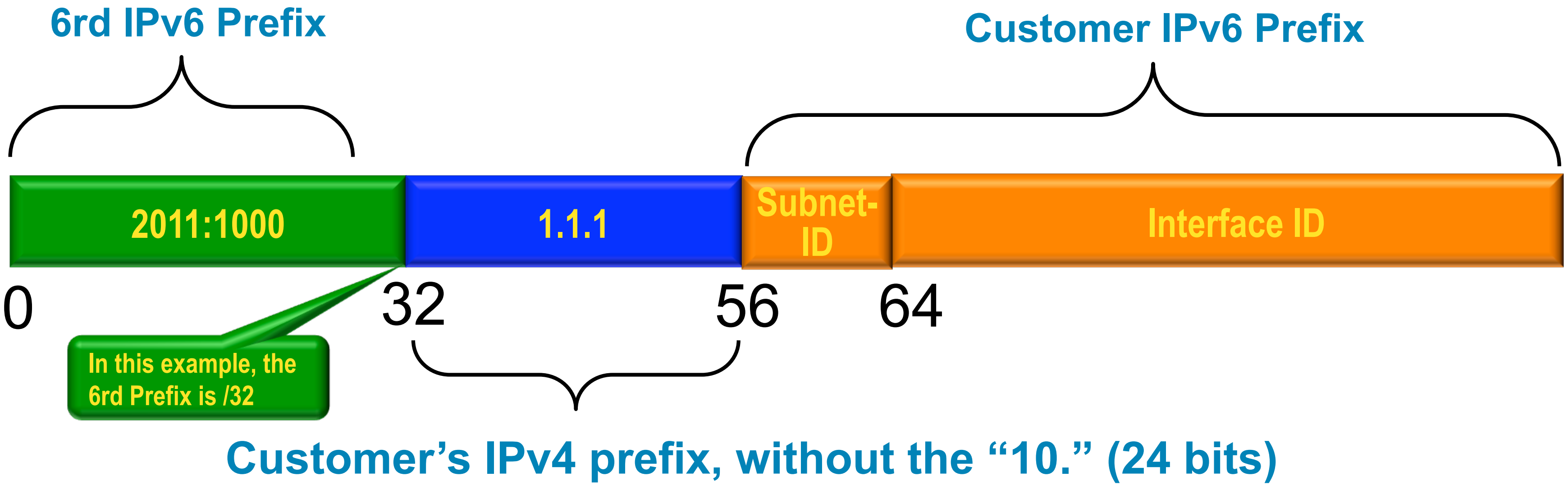
① 6rd Automatic Prefix Delegation (From a Global IPv4 Prefix)



This prefix length is variable in 6rd, /28 is just an example

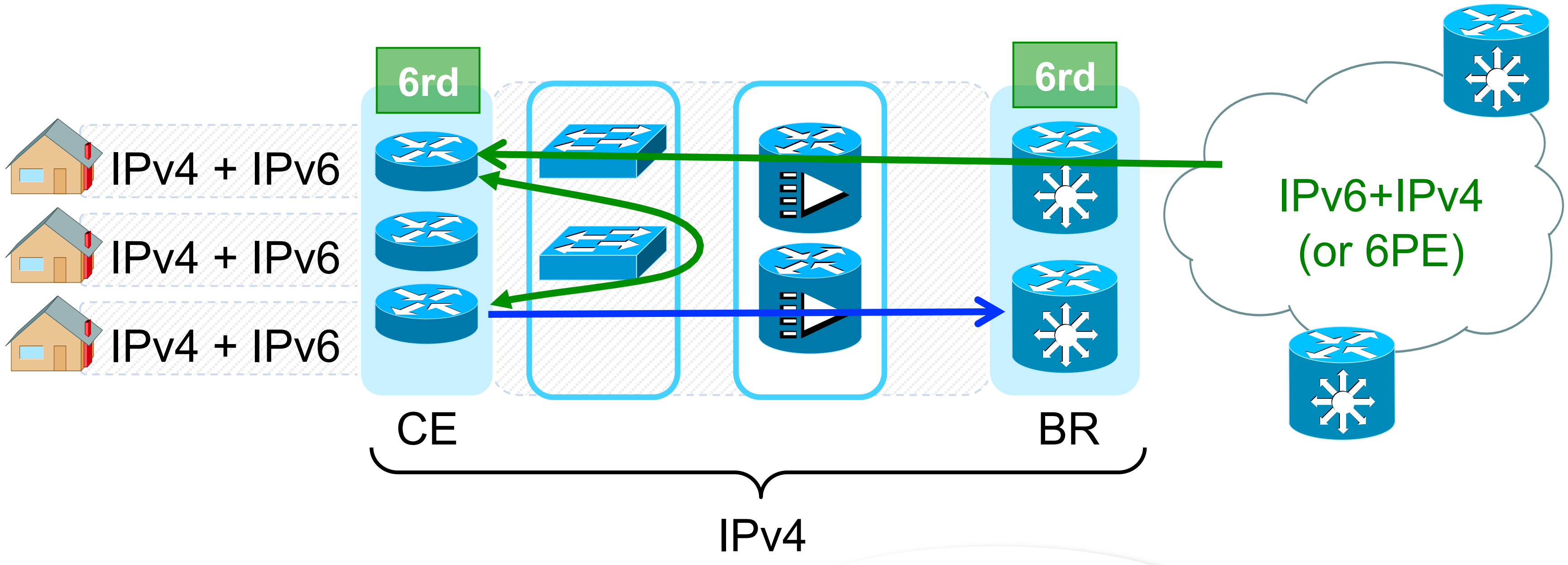
Customer's IPv4 prefix (32 bits - or less)

① 6rd Automatic Prefix Delegation (From a Private IPv4 Prefix)


6rd IPv6 Prefix
Customer IPv6 Prefix
2011:1000
1.1.1
Subnet-ID
Interface ID
0
32
56
64
In this example, the 6rd Prefix is /32
Customer's IPv4 prefix, without the "10." (24 bits)

Any number of bits may be masked off, as long as they are common for the entire domain. This is very convenient when deploying with Private IPv4, but is equally applicable to aggregated global IPv4 space.

② Packet Flow and Encapsulation



IF 6rd IPv6 Prefix

THEN Encap in IPv4 with embedded address

Dest = Inside 6rd Domain

2001:100	8101:0101		Interface ID

ELSE (6rd IPv6 Prefix Negative Match)

ENCAP with BR IPv4 Anycast Address

IPv6 Dest = Outside 6rd Domain

"Not 2001:100…"		Interface ID

③ Border Relay via Anycast

- 6rd is stateless, so no need for packets within a flow to traverse the same Relay
- Allows use of IPv4 routing for load-balancing, resiliency and redundancy
- Border Relays are installed only in strategic locations where native IPv6 is available:

  IPv6 Internet uplinks

  Edge of internal IPv6-enabled network

  BR placement is a function only of IPv6 traffic, not the number of sites

Summary for 6rd

- “6PE for the Access”

  Production-Quality IPv6 by only touching edge points around your network

- Capitalizes on what access networks do well, provisioning and transport of IPv4, adapted for carrying IPv6
- Stateless operation, easy to provision, low overhead
- Proven deployment, standardization well underway
- http://www.cisco.com/go/cgv6

Diving Deeper

- Security

  6rd inherits IPv4 anti-spoofing from the access network.

  Simple rules in the BR and default behavior of the CE ensure amplification attacks cannot occur

- QoS

  By default, CE and BR copy DSCP from IPv6 to IPv4

- Accounting

  Identifying traffic based on protocol 41

- Troubleshooting and Management

  “6rd ping”, BR and CE probing

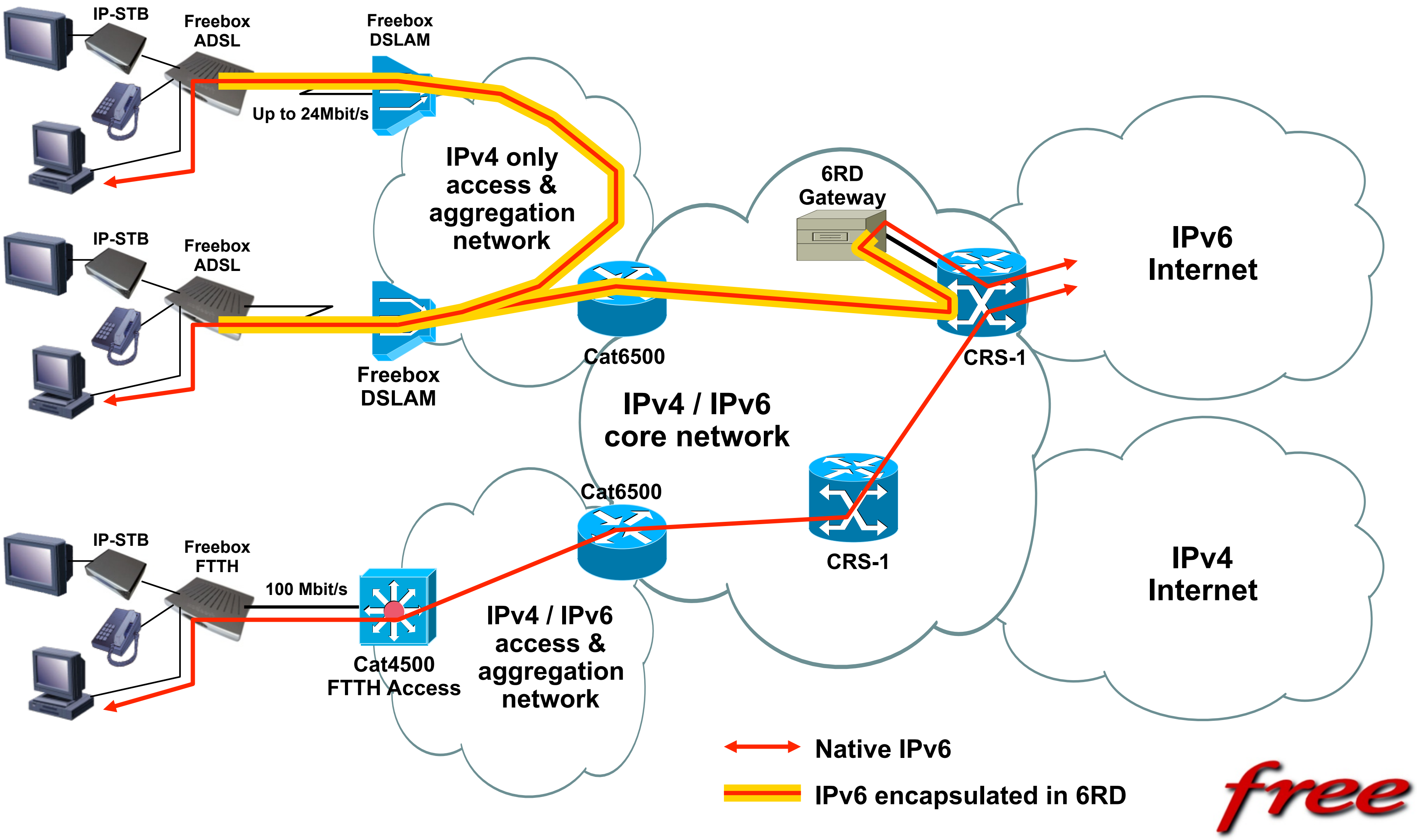
- Transition to Native

  Procedures for transition to native (with or without subscriber renumbering)

Standardization Status

- RFC 5569 describes the Free Telecom deployment based on the original invention by Remi Despres
- draft-ietf-softwire-ipv6-6rd-04.txt Standards Track WG document, entering Last Call now
- On track in the Broadband Forum to be part of their IPv6 Technical Recommendations (WT-192 IPv6 RG Specification)

6rd Deployment



05/05/2009 *IPv6 @ Free* 40

Dual Stack "Lite"

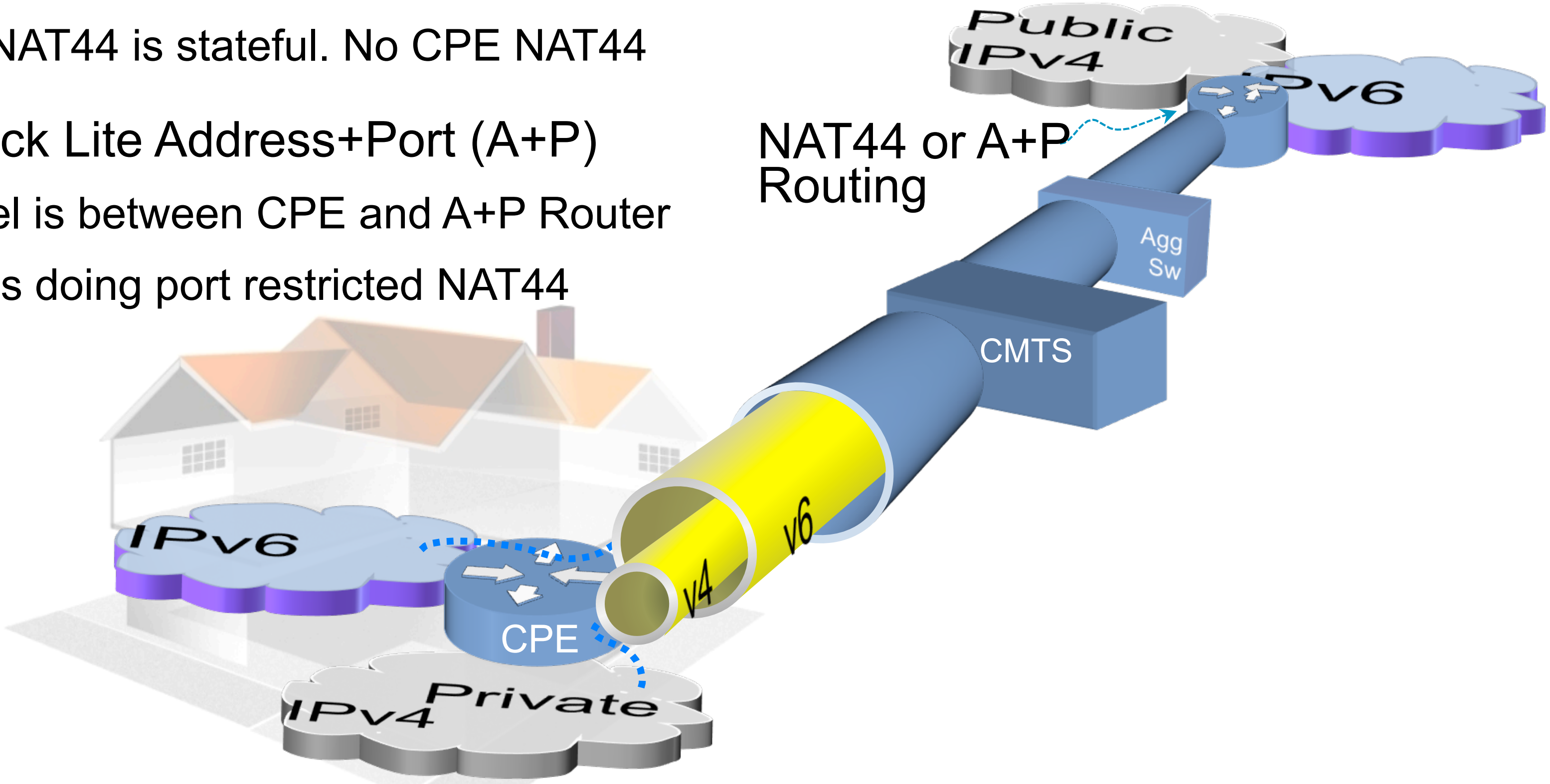
- Tunneling IPv4 using IPv6 transport
- Two common options allowed by: http://tools.ietf.org/html/draft-ietf-softwire-dual-stack-lite-02
- Dual-stack Lite with NAT44

  Tunnel from CPE is to a LSN NAT44 device.

  LSN NAT44 is stateful. No CPE NAT44
- Dual-stack Lite Address+Port (A+P)

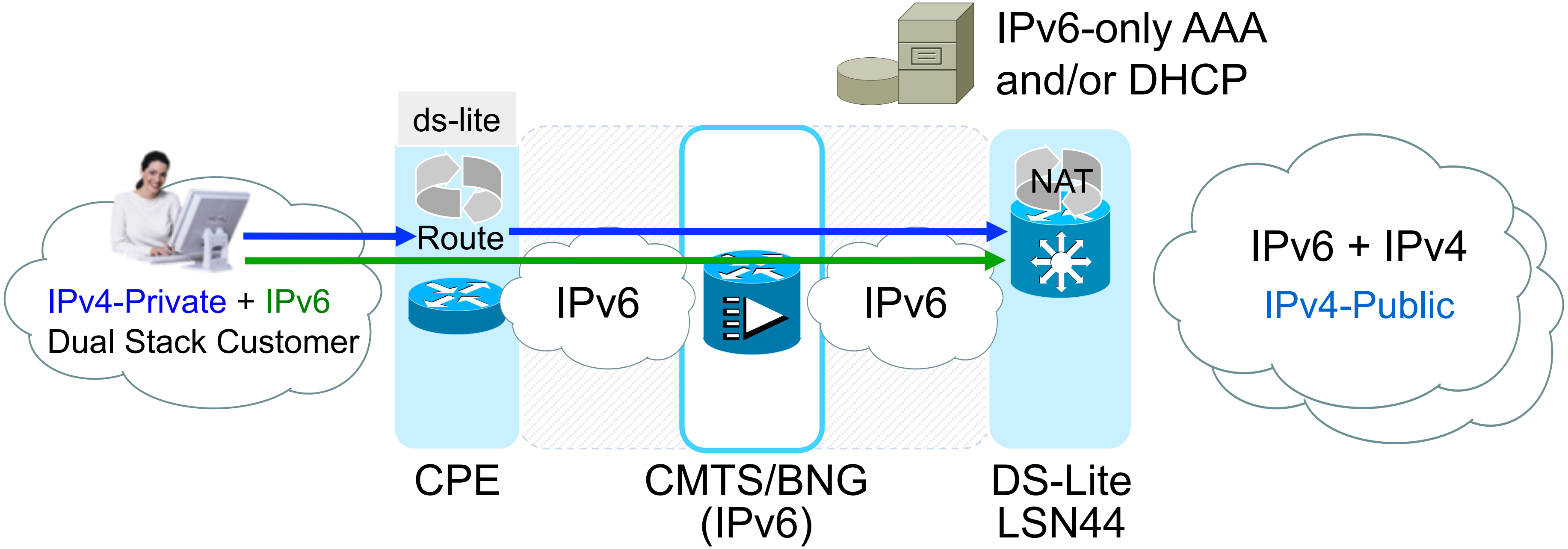
  Tunnel is between CPE and A+P Router

  CPE is doing port restricted NAT44



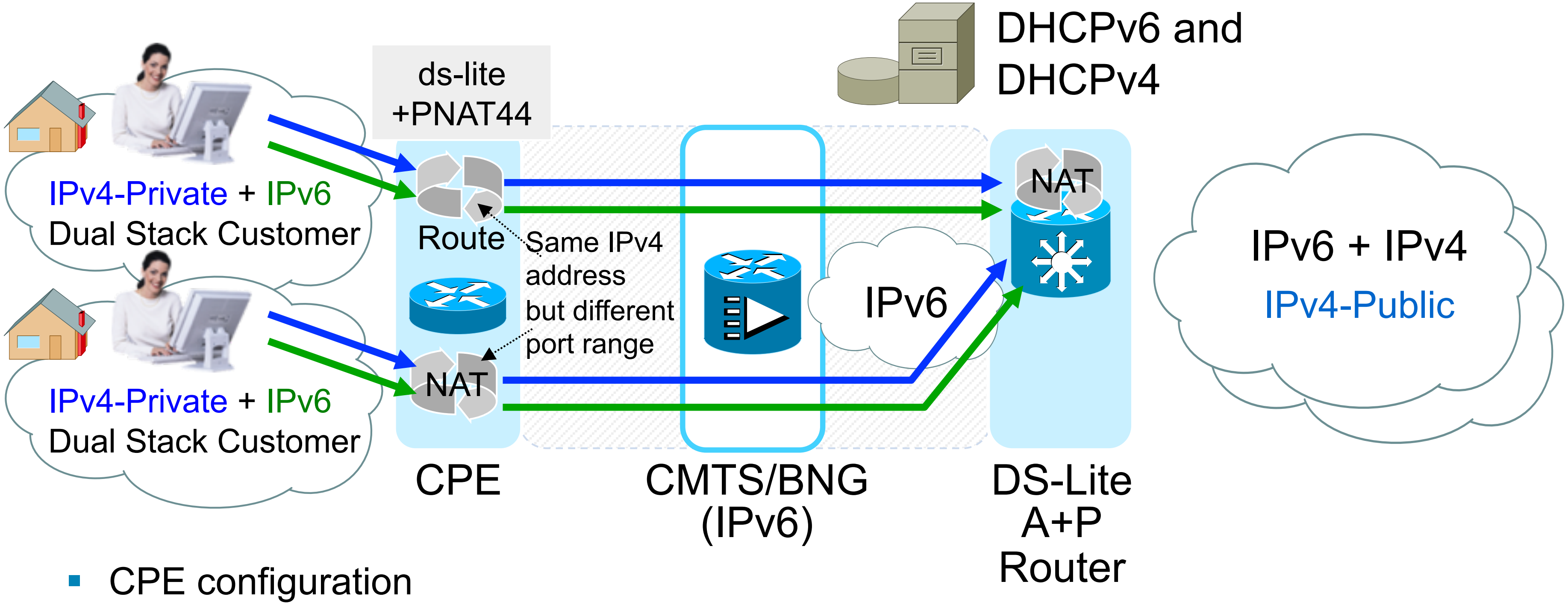
NAT44 or A+P Routing

Dual Stack “Lite” with IPv4 NAT



- CPE configuration
  1) ISP IPv6 Prefix (DHCPv6 or SLAAC assigned)
  2) DS-Lite Tunnel Gateway address (IPv6)
  3) CPE has a *dummy* IPv4 address (eg 0.0.0.1). NAT44 is disabled
- All user sourced IPv4 traffic is routed by the CPE onto point-point ds-lite IPv6 tunnel towards LSN
- LSN44 performs NAT44 function on each subscriber’s IPv6 tunnel.

Dual Stack "Lite" with A+P



- CPE configuration
  1) ISP IPv6 Prefix (DHCPv6 or SLAAC assigned)
  2) DS-Lite Tunnel Gateway address (IPv6)
  3) CPE is dynamically assigned a *public* IPv4 address *and a restricted range* of *IPv4 ports*. Port restricted NAT44 is enabled.
- All user sourced IPv4 traffic is NAT'ed by the CPE into the restricted IPv4 port space and passed onto IPv6 tunnel
- A+P Router performs per user IPv4 port range routing.

DS-Lite Summary

- Dual-stack service to the subscriber
  - Same as 6rd and Dual-Stack
- IPv6-only in Service Provider network
  - “Lite” -> “Less IPv4 To Enable”
  - Frees global IPv4 addresses
  - Reduces need for overlapping private IPv4 space
- Tool for migration to fully native IPv6