Mecanismos de Transición IPv6

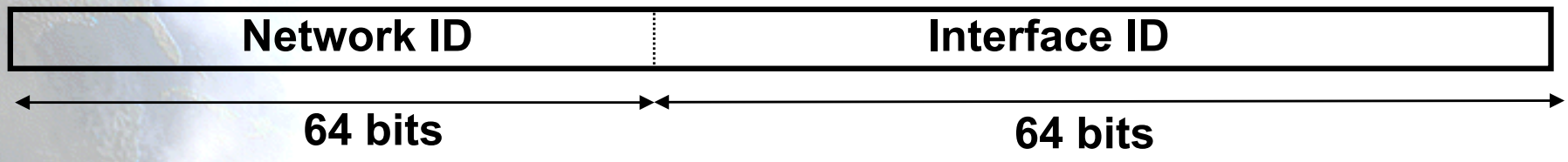
LACNIC
Foz de Iguazú - Brasil
Mayo de 2017

Jordi Palet (jordi.palet@consulintel.es)

Repasos …
ND, SLAAC y DNS

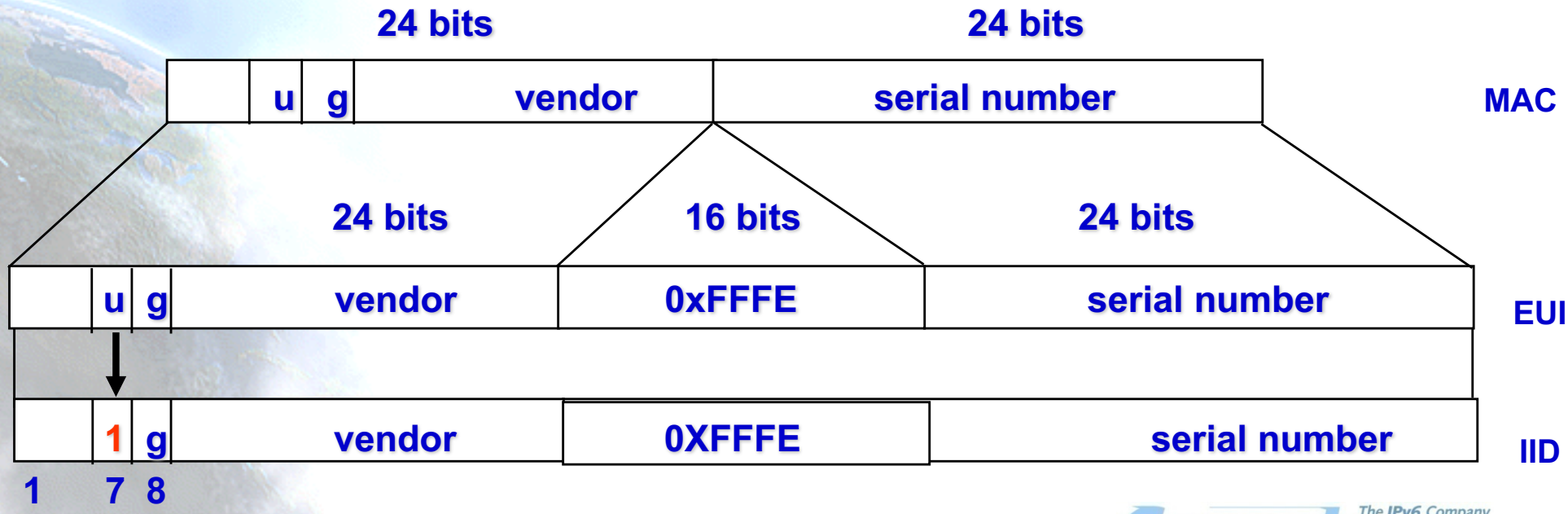
Identificadores de Interfaz

- Los identificadores de interfaz (IID) de una dirección IPv6 Unicast se usan para identificar interfaces en un enlace
- Deben ser únicos en una subred
- Hay IIDs o rangos de IID definidos para usos concretos y no deben ser usados por un nodo IPv6 (ver [RFC5453]



EUI-64

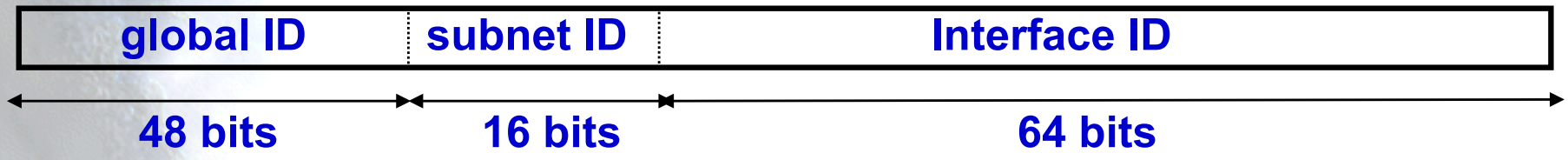
- IEEE define un mecanismo para crear una EUI-64 desde una dirección IEEE 802 MAC (Ethernet, FDDI)
- El IID se obtiene modificando el EUI-64 en el bit u (Universal). Se pone 1 para indicar alcance universal y 0 para indicar alcance local



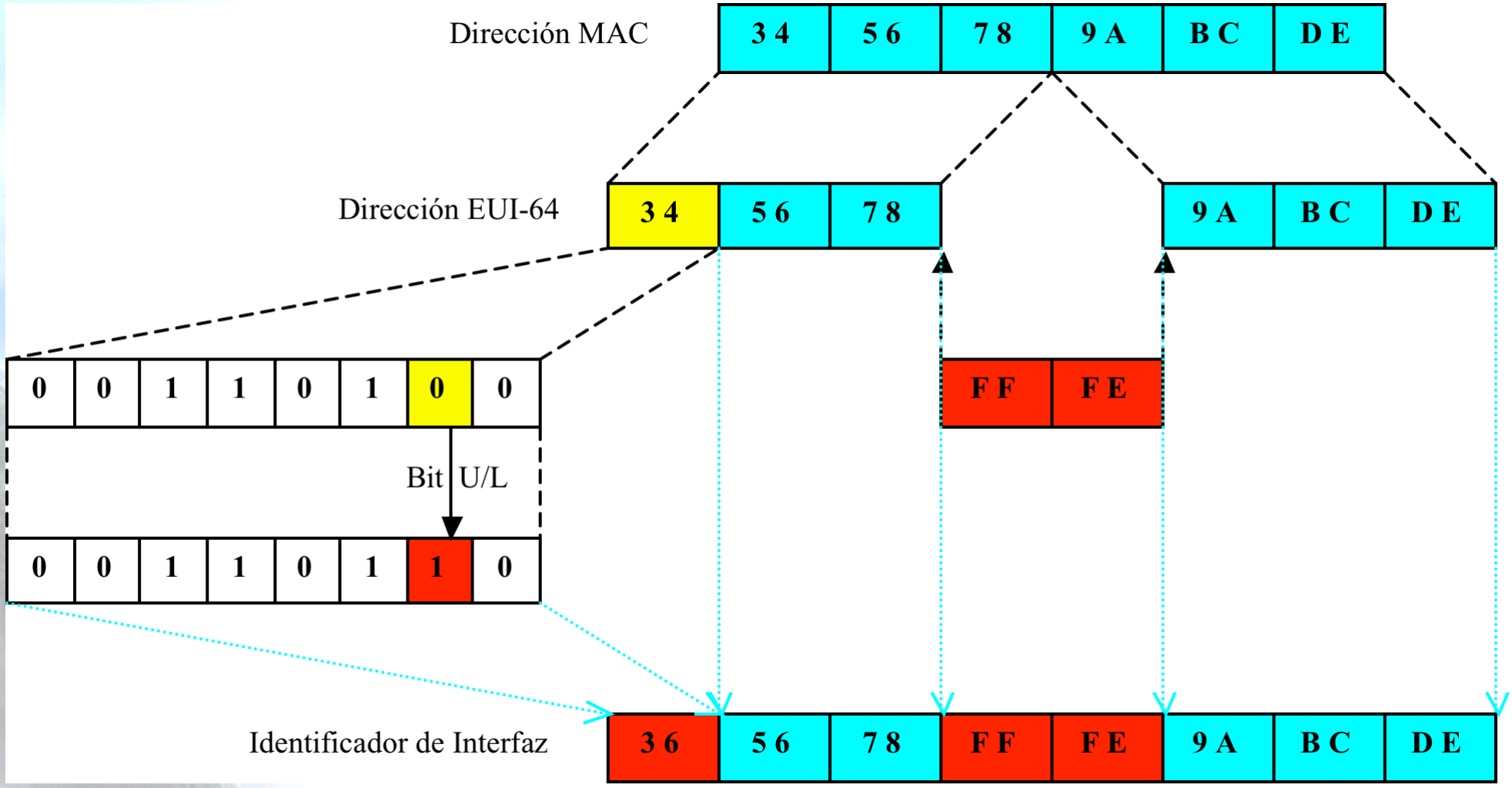
Generación del IID

Los 64-bits de menor peso de las direcciones Unicast pueden ser asignados mediante diversos métodos:

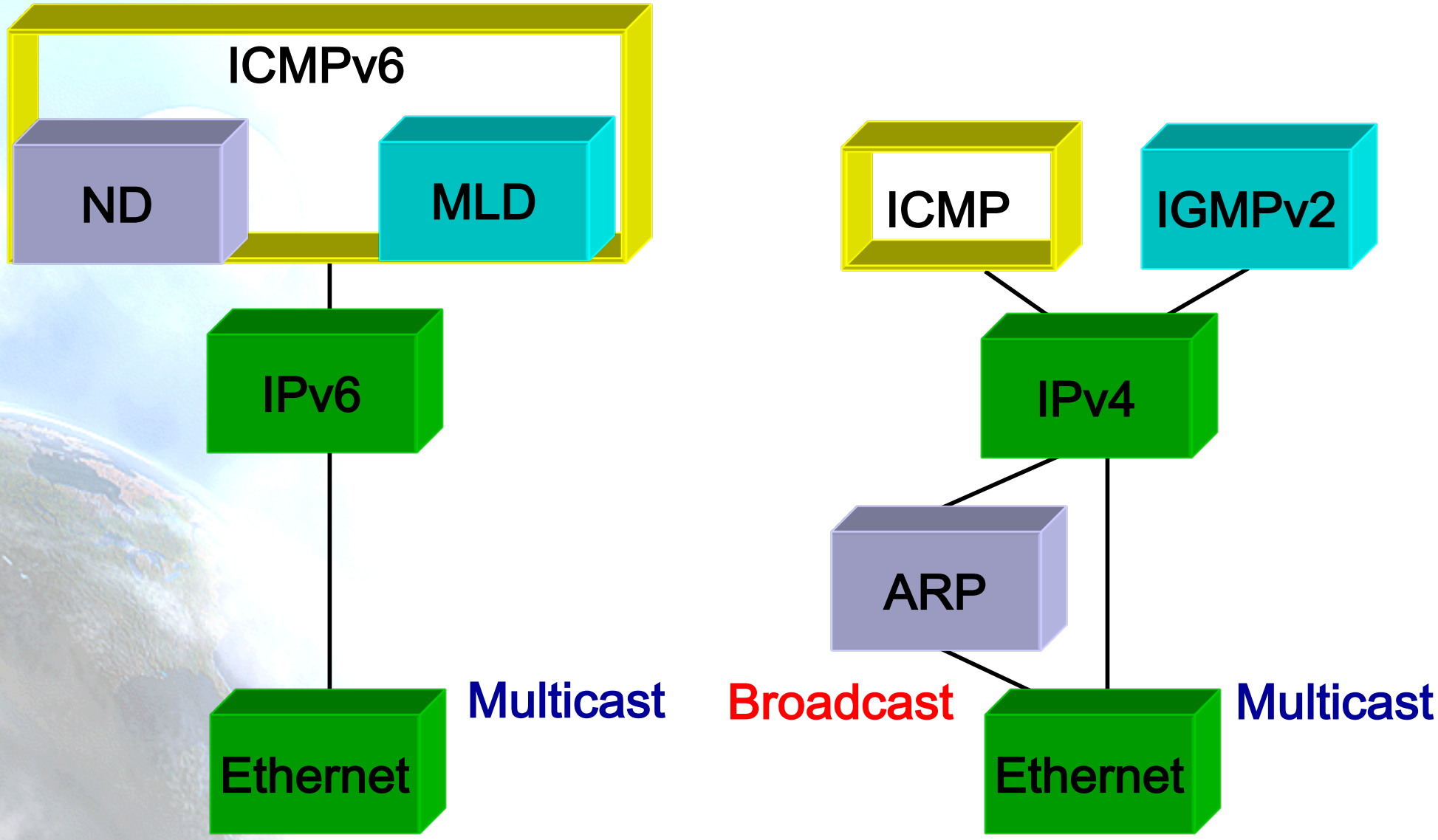
- auto-configuradas a partir de una dirección MAC de 64-bit (FireWire)
- auto-configuradas a partir de una dirección MAC de 48-bit (ejemplo, direcciones Ethernet), y expandida aun EUI-64 de 64-bits
- asignadas mediante DHCP
- configuradas manualmente
- auto-generadas pseudo-aleatoriamente (protección de la privacidad) RFC4941
- “Semantically Opaque Interface Identifiers” RFC7217
- posibilidad de otros métodos en el futuro



EUI-64


Dirección MAC
3 4
5 6
7 8
9 A
B C
D E
Dirección EUI-64
3 4
5 6
7 8
9 A
B C
D E
F F
F E
0 0 1 1 0 1 0 0
Bit U/L
0 0 1 1 0 1 1 0
Identificador de Interfaz
3 6
5 6
7 8
F F
F E
9 A
B C
D E

Plano de Control IPv4 vs. IPv6



Direcciones Multicast Importantes

- FF01::1, FF02::1 — Todos los nodos
- FF01::2, FF02::2, FF05::2 — Todos los encaminadores

- Dirección (SN) multicast a partir de la unicast
  - Si la dirección acaba en "XY:ZTUV"
  - La SN es: FF02::1:FFXY:ZTUV

- Cada nodo IPv6 debe unir la dirección SN a todas sus direcciones unicast y anycast.

ICMPv6 (RFC4443)

- IPv6 emplea el Internet Control Message Protocol (ICMP) como se define en IPv4 (RFC792)
- Aunque se introducen algunos cambios para IPv6: ICMPv6.
- Valor Next Header = 58.
- Se emplea ICMPv6 en los nodos IPv6 para reportar errores encontrados durante el procesamiento de los paquetes y para realizar otras funciones de la capa de Red, tales como diagnósticos (ICMPv6 "ping").
- ICMPv6 es una parte integral de IPv6 y DEBE ser completamente implementado por cada nodo IPv6.

Tipos de mensajes de error ICMPv6

- Destino Inalcanzable (tipo = 1, parámetro = 0)
  - No hay ruta al destino (código = 0)
  - Comunicación con el destino prohibida administrativamente (código = 1)
  - Más allá del ámbito de la dirección origen (código = 2)
  - Dirección Inalcanzable (código = 3)
  - Puerto Inalcanzable (código = 4)
  - Dirección origen falló política ingress/egress (código = 5)
  - Ruta a destino rechazada (código = 6)
- Paquete demasiado grande (tipo = 2, código = 0, parámetro = next hop MTU)
- Tiempo Excedido (tipo = 3, parámetro = 0)
  - Límite de saltos excedidos en tránsito (código = 0)
  - Tiempo de reensamblado de fragmentos excedido (código = 1)
- Problemas de parámetros (tipo = 4, parámetro = offset to error)
  - Campo de cabecera erróneo (código = 0)
  - Tipo no reconocido de “Next Header” (código = 1)
  - Opción IPv6 no reconocida (código = 2)

Mensajes ICMP Informativos

- Echo Request (tipo =128, código = 0)
- Echo Reply (tipo = 129, código = 0)

Type = 128-255	Code	Checksum
Maximum Response Delay		Reserved
Multicast Address		

- Mensajes MLD (Multicast Listener Discovery):
  - Query, report, done (como IGMP para IPv4):

ND (RFC4861)

- Define el protocolo Neighbor Discovery (ND) (Descubrimiento de Vecinos) en IPv6.
- Los nodos usan ND para determinar la dirección de la capa de enlace de los nodos que se sabe que están en el mismo segmento de red y para purgar rápidamente los valores almacenados inválidos.
- Los hosts también usan ND para encontrar encaminadores vecinos que retransmitirán los paquetes que se les envíen.
- Los nodos usan el protocolo para tener conocimiento de los vecinos que son alcanzables y los que no y para detectar cambios de sus direcciones en la capa de enlace.
- ND habilita el mecanismo de autoconfiguración en IPv6.

Interacción Entre Nodos

- Define el mecanismo para solventar:
  - Descubrimiento de encaminadores
  - Descubrimiento de prefijos de red
  - Descubrimiento de parámetros
  - Autoconfiguración de direcciones
  - Resolución de direcciones
  - Determinación del “Next-Hop”
  - Detección de Vecinos Inalcanzables (NUD).
  - Detección de Direcciones Duplicadas (DAD).
  - Redirección del “First-Hop”.

Nuevos Tipos de Paquetes ICMP

- ND define 5 tipos de paquetes:
  - "Router Solicitation" (RS)
  - "Router Advertisement" (RA)
  - "Neighbor Solicitation" (NS)
  - "Neighbor Advertisement" (NA)
  - "Redirect"

Router Advertisements

- En una red (link) con capacidad broadcast, cada encaminador envía periódicamente paquetes multicast RA.
- Un host recibe los RAs de todos los encaminadores, construyendo una lista de encaminadores por defecto.
- El algoritmo de Neighbor Unreachability Detection (NUD) detecta si existen problemas en alcanzar los encaminadores.
- Los RAs contienen una lista de prefijos usados por los hosts para determinar si una dirección destino de un paquete pertenece a dicho link y para la autoconfiguración de direcciones.
- Los RAs y los 'Flags' asociados a cada prefijo permiten a los encaminadores indicar a los hosts como realizar la autoconfiguración (stateless o DHCPv6).

Comparación con IPv4

- IPv6 ND equivaldría a ARP, ICMP Router Discovery e ICMP Redirect en IPv4, con algunas cosas más (NUD).
- ND supone mejoras en muchos aspectos sobre los protocolos usados en IPv4, entre otras:
  - RAs llevan la dirección de la capa de enlace del encaminador, no es necesario resolverla.
  - RAs llevan los prefijos de un enlace, no es necesario un mecanismo para conocer la máscara de red.
  - RAs permiten la Autoconfiguración de direcciones.
  - REDIRECTS llevan la dirección de la capa de enlace del nuevo 'first hop', no es necesario resolverla.
  - El uso de direcciones de enlace local para identificar a los encaminadores, hace que los hosts 'resistan' una renumeración de la red.
  - Usando un 'Hop Limit' de 255 ND es inmune a mensajes ND de fuera del enlace. En IPv4 podían enviar de fuera Redirects y RAs.

Formato Router Advertisement

Bits 8	16			32
Type = 134	**Code = 0**			**Checksum**
Cur Hop Limit	**M**	**O**	**Reserved = 0**	**Router Lifetime**
Reachable Time				
Retrans Timer				
Options …				

- Cur Hop Limit: valor predeterminado que debería ponerse en el campo Hop Count de la cabecera IPv6 de los paquetes que van a ser enviados
- M: 1-bit "Managed address configuration" flag
- O: 1-bit "Other configuration" flag
- Router Lifetime: entero sin signo de16-bits
- Reachable Time: entero sin signo de 32-bits
- Retrans Timer: entero sin signo de 32-bits
- Possible Options: Source LinkLayer Address, MTU, Prefix Information, Flags Expansion (RFC5175)

Formato Router Solicitation

- Cuando arrancan los hosts envían RSs para indicar a los encaminadores que generen un RA inmediatamente.
- Se envía a la dirección multicast que engloba a todos los encaminadores del segmento de red.

Bits 8	16	32
Type = 133	**Code = 0**	**Checksum**
Reserved = 0		
Options …		

- Opciones Posibles: Source Link-Layer Address.

Formato Neighbor Solicitation

- Los nodos envían NSs para obtener la dirección MAC del nodo con el que se pretende comunicar, a la vez que se proporciona la propia dirección MAC del nodo solicitante.
- Los paquetes NSs son multicast cuando el nodo precisa resolver una dirección y unicast cuando el nodo pretende averiguar si un vecino es alcanzable.

Bits 8	16	32
Type = 135	**Code = 0**	**Checksum**
Reserved = 0		
Target Address		
Options …		

- Target Address: La dirección IPv6 objetivo de la solicitud. No debe ser una dirección multicast.
- Opciones Posibles : Source Link-Layer Address.

Formato Neighbor Advertisement

- Un nodo envía NAs como respuesta a un NS y envía NAs no solicitados para propagar nueva información rápidamente.

Bits 8	16	32
Type = 136	Code = 0	Checksum
R S O	Reserved = 0	
Target Address		
Options …		

- Flags:
  - **R**: **Router Flag**=1 indica que el que envía en un encaminador.
  - **S**: **Solicited Flag**=1 indica que se envía como respuesta a un NS.
  - **O**: **Override Flag**=1 indica que deben actualizarse las caches.
- Para NA solicitados, igual al campo “Target Address” del NS. Para un NA no solicitado, la dirección cuya MAC ha cambiado. No puede ser una dirección multicast.
- Posibles Opciones: Target Link-Layer Address (MAC del Tx).

Formato Redirect

- Los encaminadores envían paquetes Redirect para informar a un host que existe otro encaminador mejor en el camino hacia el destino final.
- Los hosts pueden ser redireccionados a otro encaminador mejor pero también pueden ser informados mediante un paquete Redirect que el destino es un vecino.

Bits 8	16	32
Type = 137	**Code = 0**	**Checksum**
Reserved = 0		
Target Address		
Destination Address		
Options …		

- Target Address: La dirección IPv6 del ‘first hop’ que es mejor usar para llegar al ‘Destination Address’ del paquete ICMPv6
- Destination Address: La dirección IPv6 de destino que es redireccionada al ‘target address’ del paquete ICMPv6

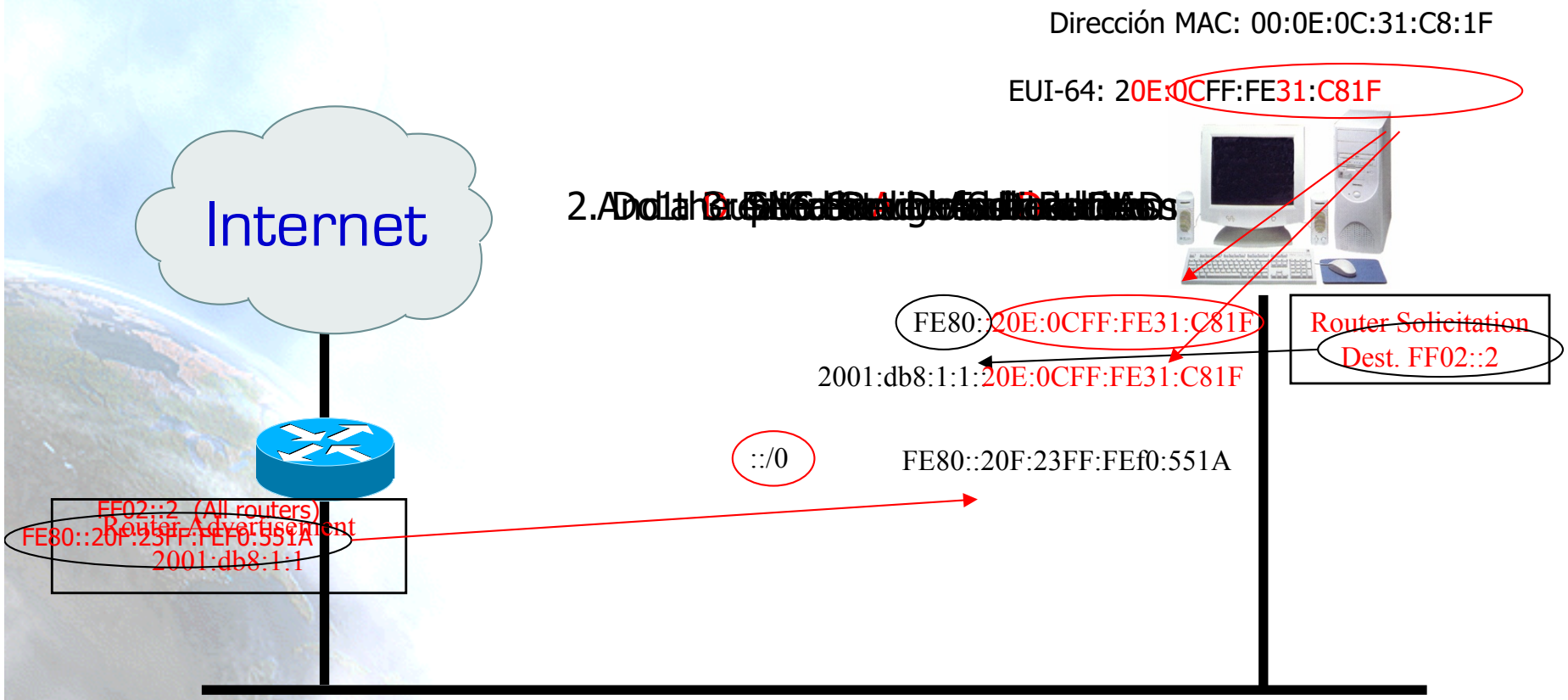
Ejemplo Funcionamiento (1)

- **Neighbor Cache**: Vecinos a los que se les ha enviado tráfico recientemente. Se indexa por la 'on-link unicast IP address'. Cada entrada contiene: dir. capa enlace, si es router/host, información de NUD (reachability state, etc.).
- **Destination Cache**: Mapea IP destino con 'next hop'. Direcciones a las que se ha enviado recientemente.
- **Prefix List**: Contiene los prefijos del enlace. Se basa en los RAs, de donde se saca también el tiempo de validez.
- **Default Router List**: Lista de routers a donde los paquetes 'off-link' deben ser enviados. Cada entrada apunta a una entrada en la Neighbor Cache y tiene un tiempo de validez obtenido del RA (router lifetime).

Autoconfiguración

- El estándar especifica los pasos que un host debe seguir para decidir cómo auto-configurar sus interfaces de red en IPv6
- El proceso de auto-configuración incluye la creación de una dirección IPv6 de ámbito local (link-local) y la verificación de que no está duplicada en el mismo segmento de red, determinando qué información debería ser auto-configurada y en el caso de direcciones, si estas deberían obtenerse mediante “stateful”, “stateless” o ambos
- IPv6 define tanto un mecanismo de auto-configuración de direcciones de tipo “stateful” como “stateless”
- La auto-configuración “stateless” (SLAAC) no precisa de configuración manual en el host, mínima (si acaso alguna) configuración de encaminadores y ningún servidor adicional

Auto-configuración - SLAAC



RA: Flags M y O

- Los flags M y O de los RA indican cómo deben comportarse los hosts con respecto a la autoconfiguración de los parámetros de red
- M indica como configurar la dirección IP
- O indica cómo configurar otros parámetros: DNS, etc.

Dir. / Otros	M	O	Comentario
SLAAC / SLAAC	0	0	Si dual-stack, se puede usar IPv4 para DNS
SLAAC / DHCPv6	0	1	DHCPv6 Stateless
DHCPv6 / SLAAC	1	0	Si dual-stack, se puede usar IPv4 para DNS
DHCPv6 / DHCPv6	1	1	El gateway se aprende del RA

Autoconfiguración Stateless o Serverless (RFC4862)

- El mecanismo “stateless” permite a un host generar su propia dirección usando una combinación de información localmente disponible y de información proporcionada por los encaminadores
- Los **encaminadores anuncian los prefijos de red** que identifican la sub-red asociada a un determinado segmento de red (64 bits)
- Los **hosts generan un identificador de interfaz** que lo identifica de manera única en la subred. Dicho identificador se genera localmente, por ejemplo a partir de la dirección MAC (64 bits)
- Una dirección IPv6 se forma mediante la combinación de ambas informaciones
- En la ausencia de encaminadores, un host puede generar solo las direcciones IPv6 de ámbito local (link-local)
- Las direcciones link-local son suficiente para permitir la comunicación IPv6 entre nodos que están conectados en el mismo segmento de red

Ventajas/Beneficios de la Autoconfiguración Stateless

- La configuración manual de cada máquina antes de conectarla a la red no es necesaria
- Los sitios pequeños compuesto de pocas máquinas conectadas al mismo segmento no necesitarían de un servidor DHCPv6 ni de un encaminador para comunicarse, usarían direcciones link-local
- Un sitio grande con varias subredes no necesitaría de un servidor DHCPv6 para la configuración de direcciones
- Facilita el cambio de prefijo de una sitio mediante el uso de varias direcciones por interfaz y tiempo de vida

Configuración de DNS

- Manualmente
- Con DHCPv4 (clientes dual-stack)
- Por medio de RA (RDNSS)
  - RFC6106
- Dynamic Host Configuration Protocol for IPv6
  - RFC3315 (DHCPv6)
- “Stateless DHCPv6” (RFC3736)
  - Clientes que ya tienen una dirección IPv6

- Dos opciones para configurar la opción RDNSS en los routers:
  - Manualmente
  - Automaticamente, siendo un cliente DHCPv6

Transición y Coexistencia IPv4-IPv6

Técnicas de Transición / Coexistencia

Un amplio abanico de técnicas han sido identificadas e implementadas, básicamente dentro de tres categorías:

(1) doble-pila, para permitir la coexistencia de IPv4 e IPv6 en el mismo dispositivo y redes

(2) técnicas de túneles, para evitar dependencias cuando se actualizan hosts, routers o regiones

(3) técnicas de traducción, para permitir la comunicación entre dispositivos que son sólo IPv6 y aquellos que son sólo IPv4

Todos estos mecanismos suelen ser utilizados, incluso en combinación

Doble-Pila

- Al añadir IPv6 a un sistema, no se elimina la pila IPv4
  - Es la misma aproximación multi-protocolo que ha sido utilizada anteriormente y por tanto es bien conocida (AppleTalk, IPX, etc.)
  - Actualmente, IPv6 está incluido en todos los Sistemas Operativos modernos, lo que evita costes adicionales
- Las aplicaciones (o librerias) escogen la versión de IP a utilizar
  - En función de la respuesta DNS:
    - si el destino tiene un registro AAAA, utilizan IPv6, en caso contrario IPv4
  - La respuesta depende del paquete que inició la transferencia
- Esto permite la coexistencia indefinido de IPv4 e IPv6, y la actualización gradual a IPv6, aplicación por aplicación
- El registro A6 es experimental

Túneles para Atravesar Routers que no Reenvían IPv6

- Encapsulamos paquetes IPv6 en paquetes IPv4 (o en tramas MPLS)
- Muchos métodos para establecer dichos túneles:
  - configuración manual
  - "tunnel brokers" (tipicamente con interfaces web)
  - "6-over-4" (intra-domain, usando IPv4 multicast como LAN virtual)
  - "6-to-4" (inter-domain, usando la dirección IPv4 como el prefijo del sitio IPv6)
- Puede ser visto como:
  - IPv6 utilizando IPv4 como capa de enlace virtual link-layer, o
  - una VPN IPv6 sobre la Internet IPv4

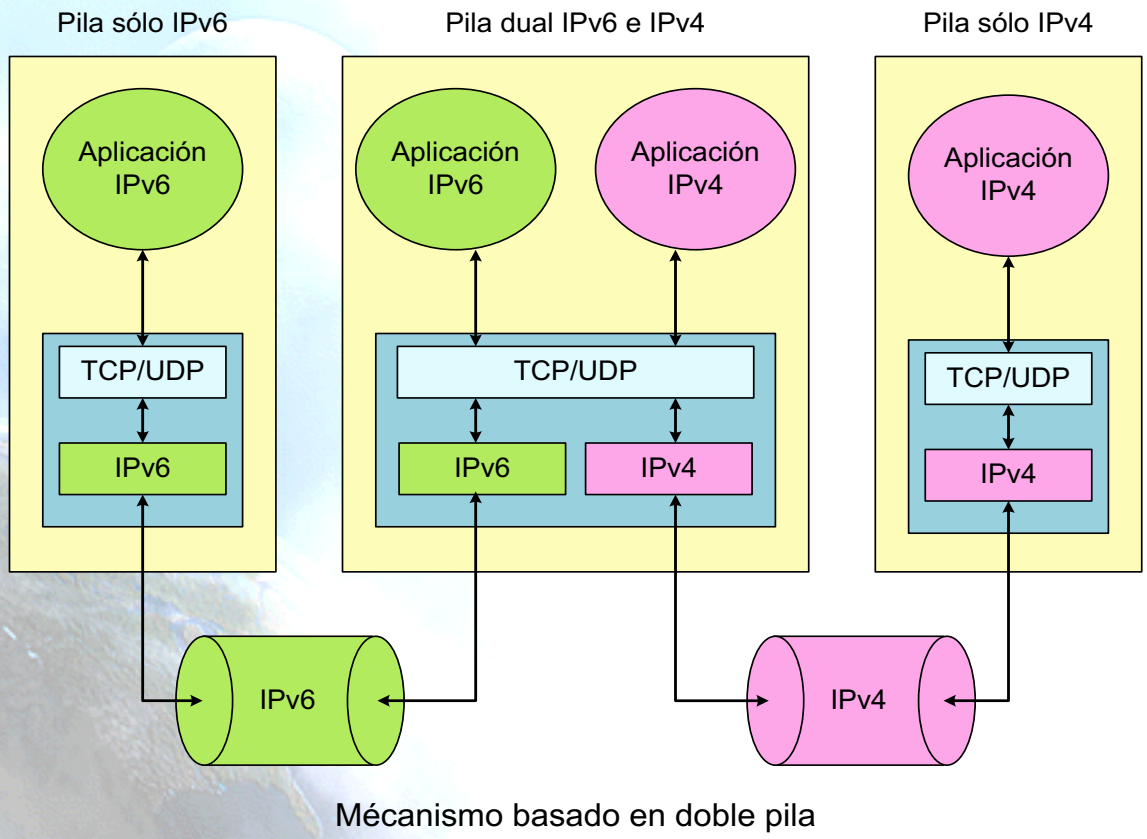
Traducción

- Se puede utilizar traducción de protolocos IPv6-IPv4 para:
  – nuevos tipos de dispositivos Internet (como teléfonos celulares, coches, dispositivos de consumo)
- Es una extensión a las técnicas de NAT, convirtiendo no sólo direcciones sino también la cabecera
  – Los nodos IPv6 detrás de un traductor obtienen la funcionalidad de IPv6 sólo cuando hablan con otro nodo IPv6
  – Obtienen la funcionalidad habitual IPv4 con NAT en el resto de los casos

Mecanismos de transición

- IPv6 ha sido diseñado de tal forma que se facilite la transición y coexistencia con IPv4
- Se han diseñado diferentes estrategias para la coexistencia con redes/nodos IPv4
  - Doble pila, o soporte simultáneo de IPv4 e IPv6
  - Túneles, o encapsulado de IPv6 sobre IPv4 (y viceversa)
    - Son los más utilizados
  - Traducción IPv4/IPv6, como último recurso, dado que no es perfecto

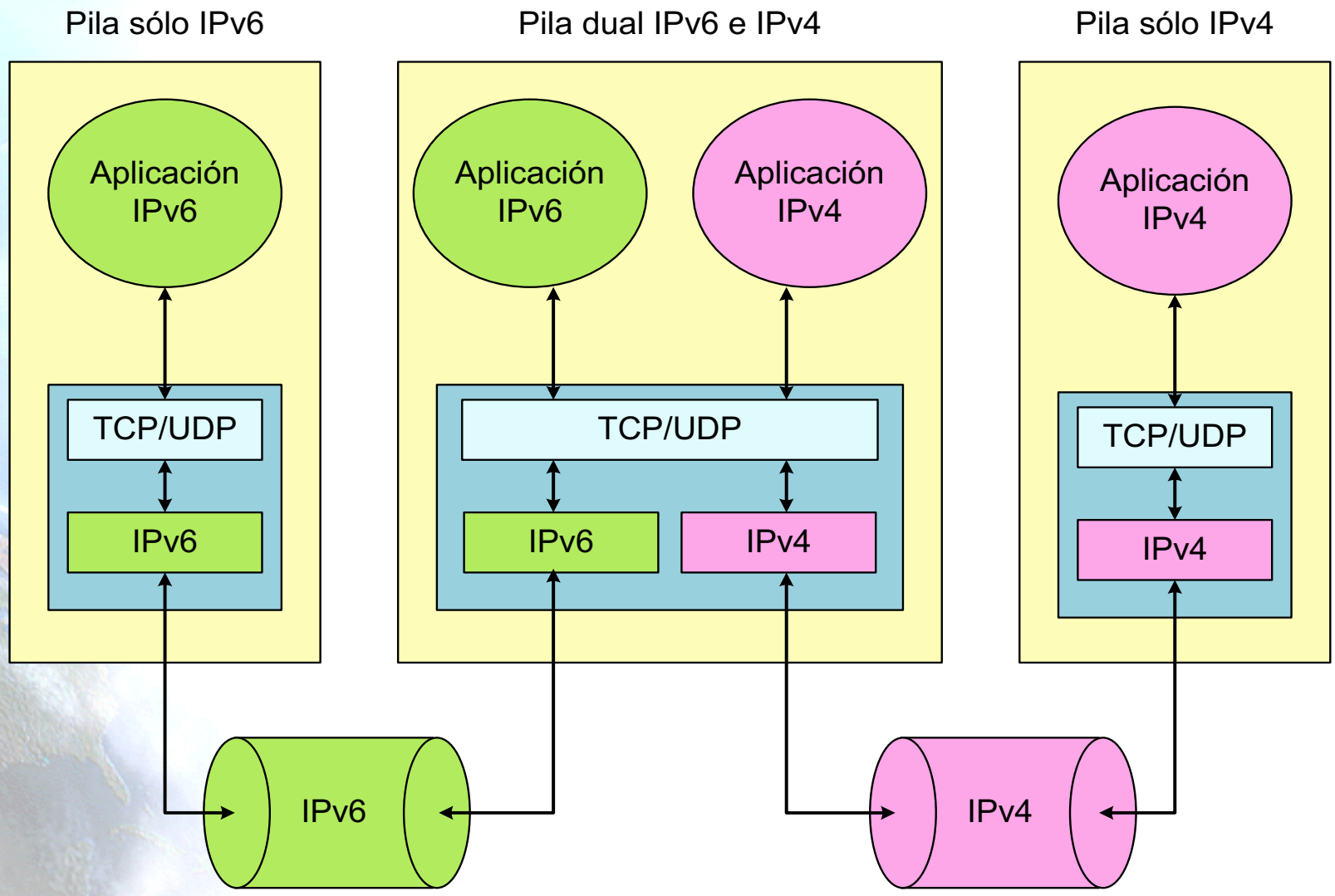
Doble pila (1)



Mécanismo basado en doble pila

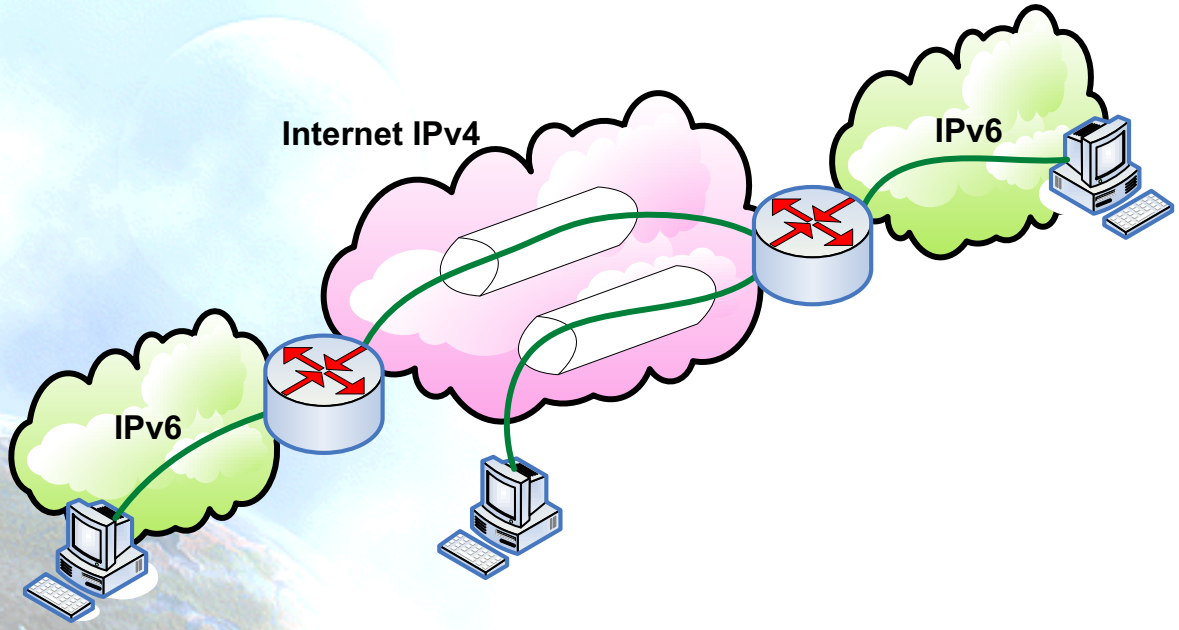
- Los nodos tienen implementadas las pilas IPv4 e IPv6
- Comunicaciones con nodos solo IPv6 ==> Pila IPv6, asumiendo soporte IPv6 en la red
- Comunicaciones con nodos solo IPv4 ==> Pila IPv4

Doble pila (2)

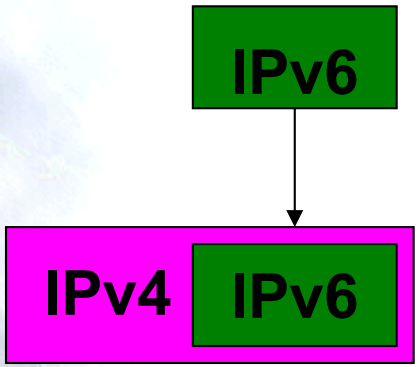


Mécanismo basado en doble pila

Túneles IPv6 en IPv4 (1)

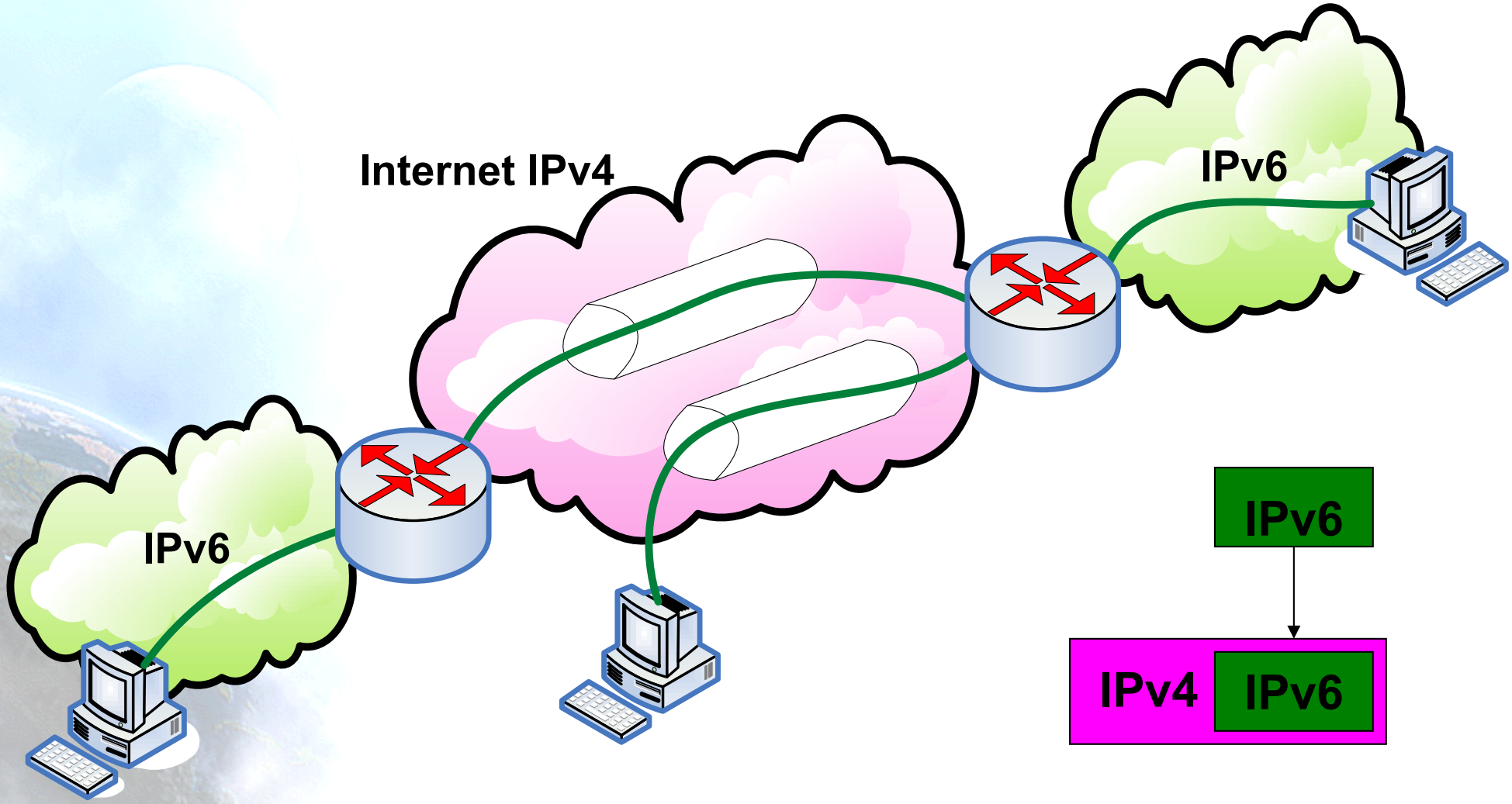


Mécanismo basado en túneles



- Usado para proporcionar conectividad IPv6 en redes que solo tiene soporte IPv4
- Se encapsulan paquetes IPv6 dentro de paquetes IPv4
- Los paquetes resultantes viajan por redes IPv4

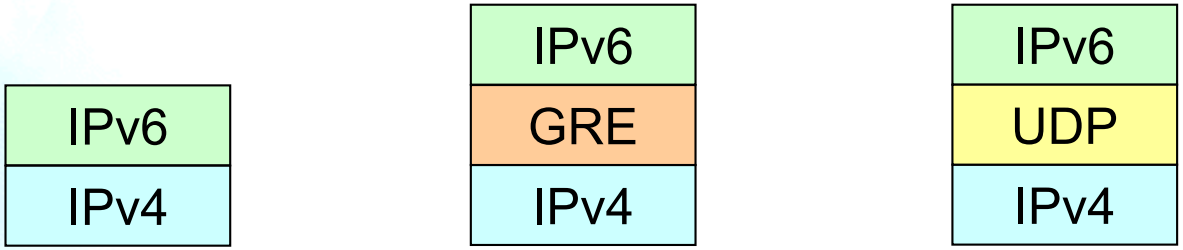
Túneles IPv6 en IPv4 (2)



Mécanismo basado en túneles

Túneles IPv6 en IPv4 (2)

- Existen diversas formas de encapsular los paquetes IPv6

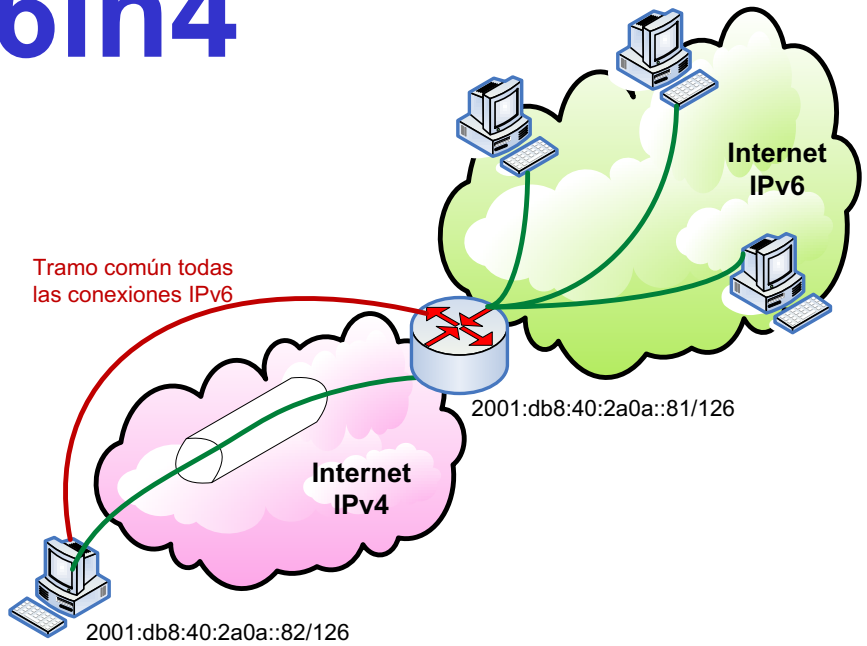


- Existen diversos mecanismos de transición basados en túneles, cada uno con una forma diferente de encapsulación

Túneles IPv6 en IPv4 (3)

- Algunos mecanismos de transición basados en túneles
  - 6in4 [6in4]
  - TB [TB]
  - TSP [TSP]
  - 6to4 [6to4]
  - Teredo [TEREDO], [TEREDOC]
  - Túneles automáticos [TunAut]
  - …
  - ISATAP [ISATAP]
  - 6over4 [6over4]
  - Softwires
  - 6RD
  - NAT64
  - DS-Lite
  - 464XLAT
  - MAP E/T

Túneles 6in4

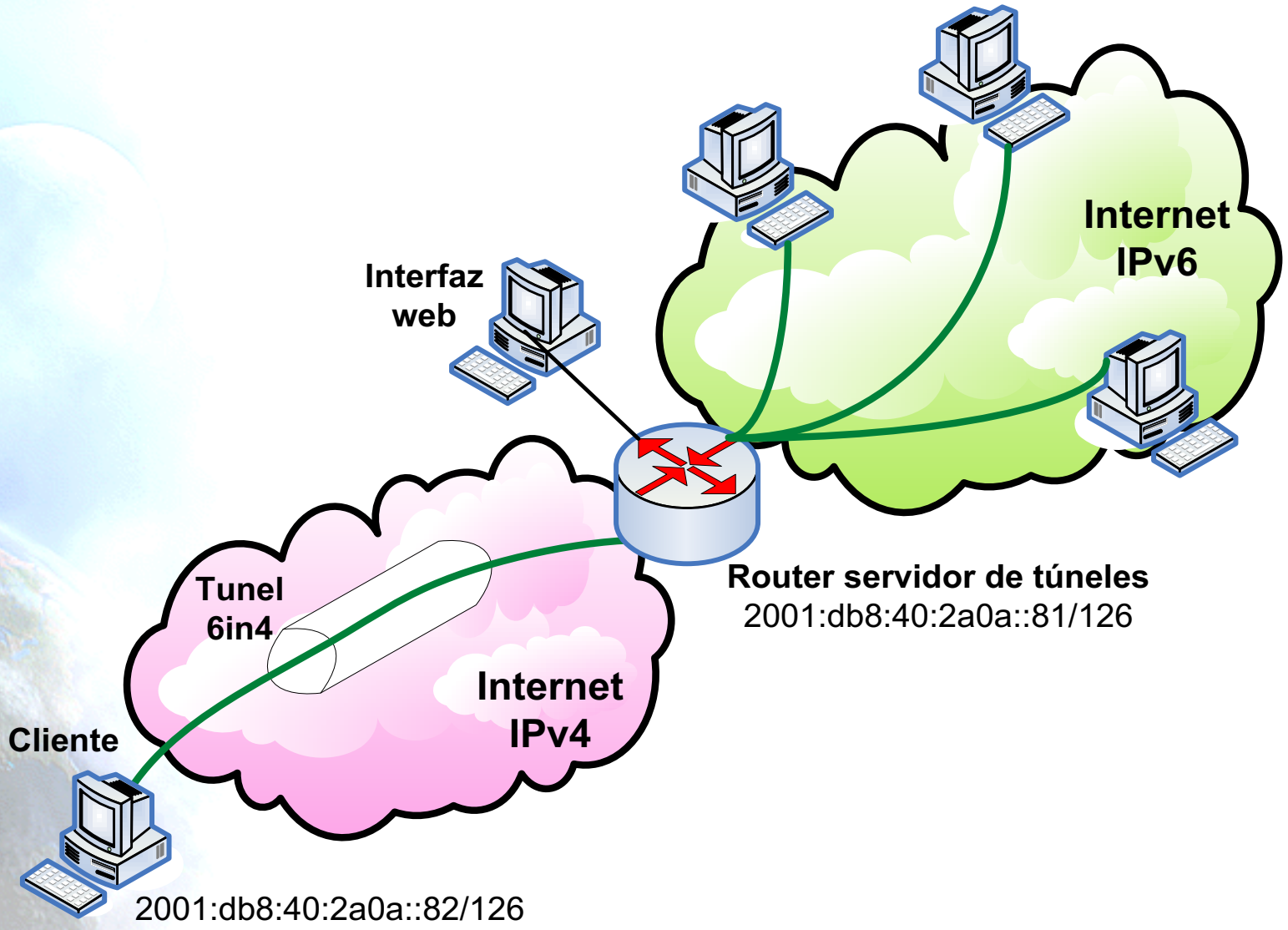


- Encapsula directamente el paquete IPv6 dentro de un paquete IPv4
- Se suele hacer entre
  - nodo final ==> router
  - router ==> router
- Aunque también es posible para
  - nodo final ==> nodo final
- El túnel se considera como un enlace punto-a-punto desde el punto de vista de IPv6
  - Solo un salto IPv6 aunque existan varios IPv4
- Las direcciones IPv6 de ambos extremos del túnel son del mismo prefijo
- Todas las conexiones IPv6 del nodo final siempre pasan por el router que está en el extremo final del túnel
- Los túneles 6in4 pueden construirse desde nodo finales situados detrás de NAT
  - Imprescindible que la implementación de NAT soporte “proto-41 forwarding” [PROTO41]

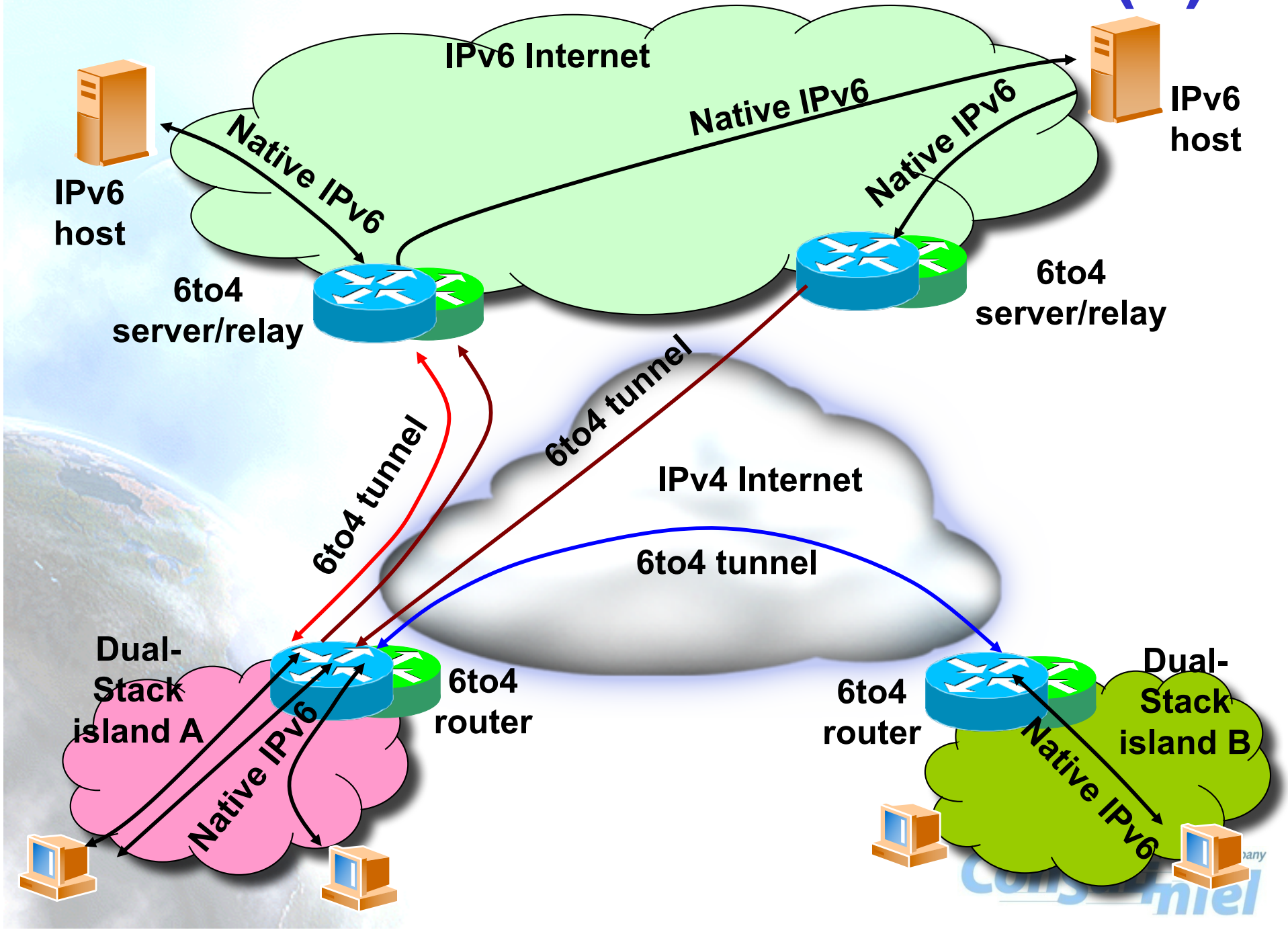
Tunnel Broker: RFC3053 (1)

- Los túneles 6in4 requieren la configuración manual de los equipos involucrados en el túnel
- Para facilitar la asignación de direcciones y creación de túneles IPv6, se ha desarrollado el concepto de Tunnel Broker (TB).
  - Es un intermediario al que el usuario final se conecta, normalmente con un interfaz web
- El usuario solicita al TB la creación de un túnel y este le asigna una dirección IPv6 y le proporciona instrucciones para crear el túnel en el lado del usuario
- El TB también configura el router que representa el extremo final del túnel para el usuario
- En http://www.ipv6tf.org/using/connectivity/test.php existe una lista de TB disponibles
- TSP [TSP] es un caso especial de TB que no esta basado en un interfaz web sino en un aplicación cliente que se instala el cliente y se conecta con un servidor, aunque el concepto es el mismo.

Tunnel Broker: RFC3053 (2)



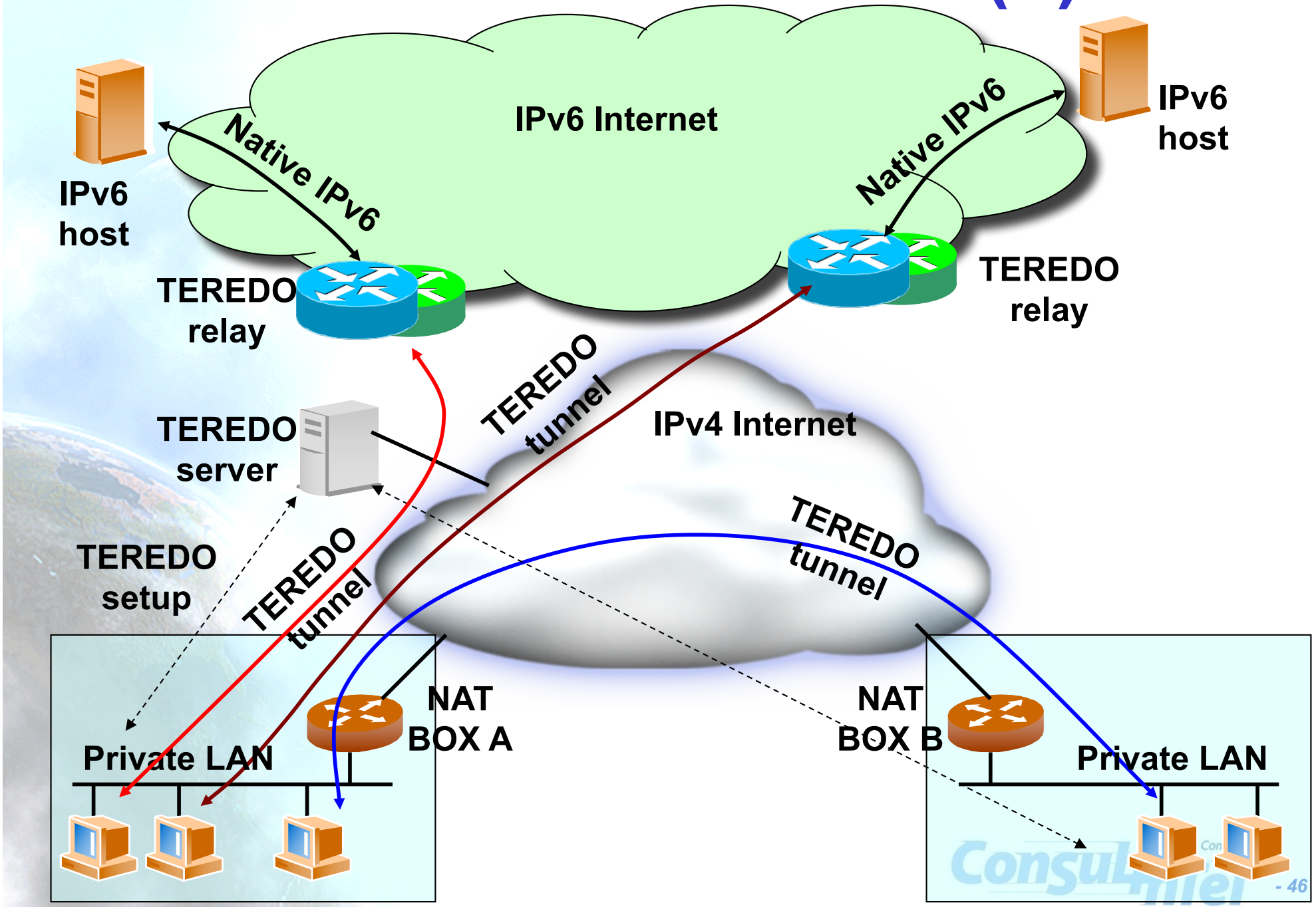
Túneles 6to4: RFC3056 (1)



Túneles 6to4: RFC3056 (2)

- Se trata de un encapsulado de paquetes IPv6 en paquetes IPv4, similar a 6in4
- Diferencias:
  - La dirección IPv6 del cliente no depende del router al que se conecta sino de la dirección IPv4 pública
    - Rango 2002::/16
  - Los paquetes IPv6 de salida del cliente siempre son enviados al mismo “6to4 relay”, sin embargo los paquetes IPv6 de entrada al cliente pueden provenir de otros “6to4 relay” diferentes.
- Prefijo IPv4 anycast (RFC3068):
  - 192.88.99.1/24

Teredo: RFC4380 (1)



Teredo: RFC4380 (2)

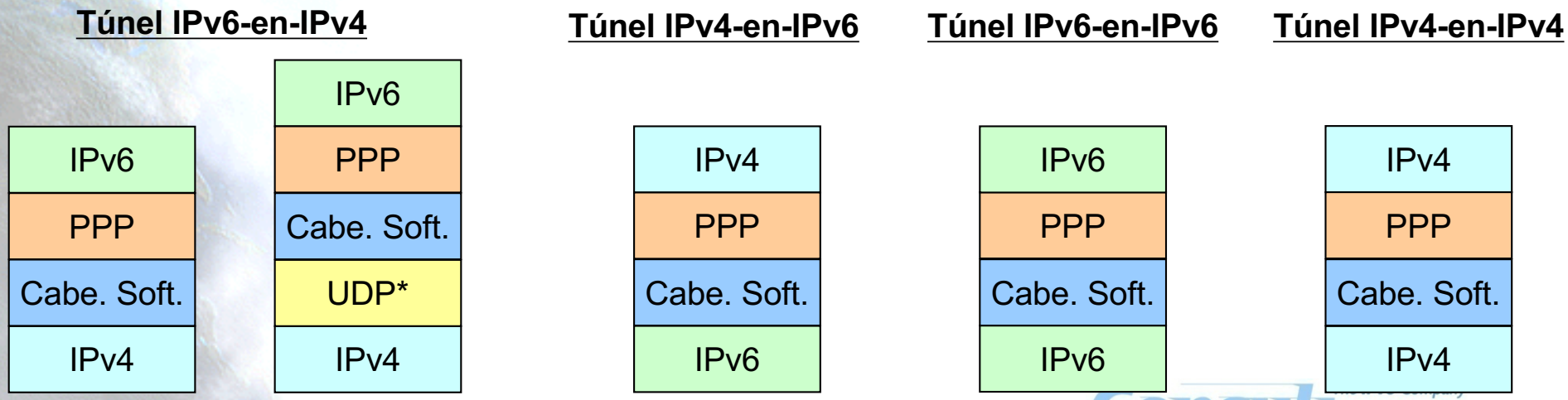
- Teredo [TEREDO] [TEREDOC] está pensado para proporcionar IPv6 a nodos que están ubicados detrás de NAT que no son “proto-41 forwarding”.
  - Encapsulado de paquetes IPv6 en paquetes UDP
- Funciona en NAT de tipo [STUN]
  - Full Cone
  - Restricted Cone
- No funciona en NATs de tipo
  - Symmetric (solventado a partir de Windows Vista)
- Intervienen diversos agentes:
  - Teredo Server
  - Teredo Relay
  - Teredo Client
- El cliente configura un Teredo Server que le proporciona una dirección IPv6 del rango 2001:0000::/32 basada en la dirección IPv4 publica y el puerto usado
  - Si el Teredo Server configurado es además Teredo Relay, el cliente tiene conectividad IPv6 con cualquier nodo IPv6
  - De lo contrario solo tiene conectividad IPv6 con otros clientes de Teredo
- Actualmente Microsoft proporciona Teredo Servers públicos y gratuitos, pero no Teredo Relays

Softwires: RFC4925

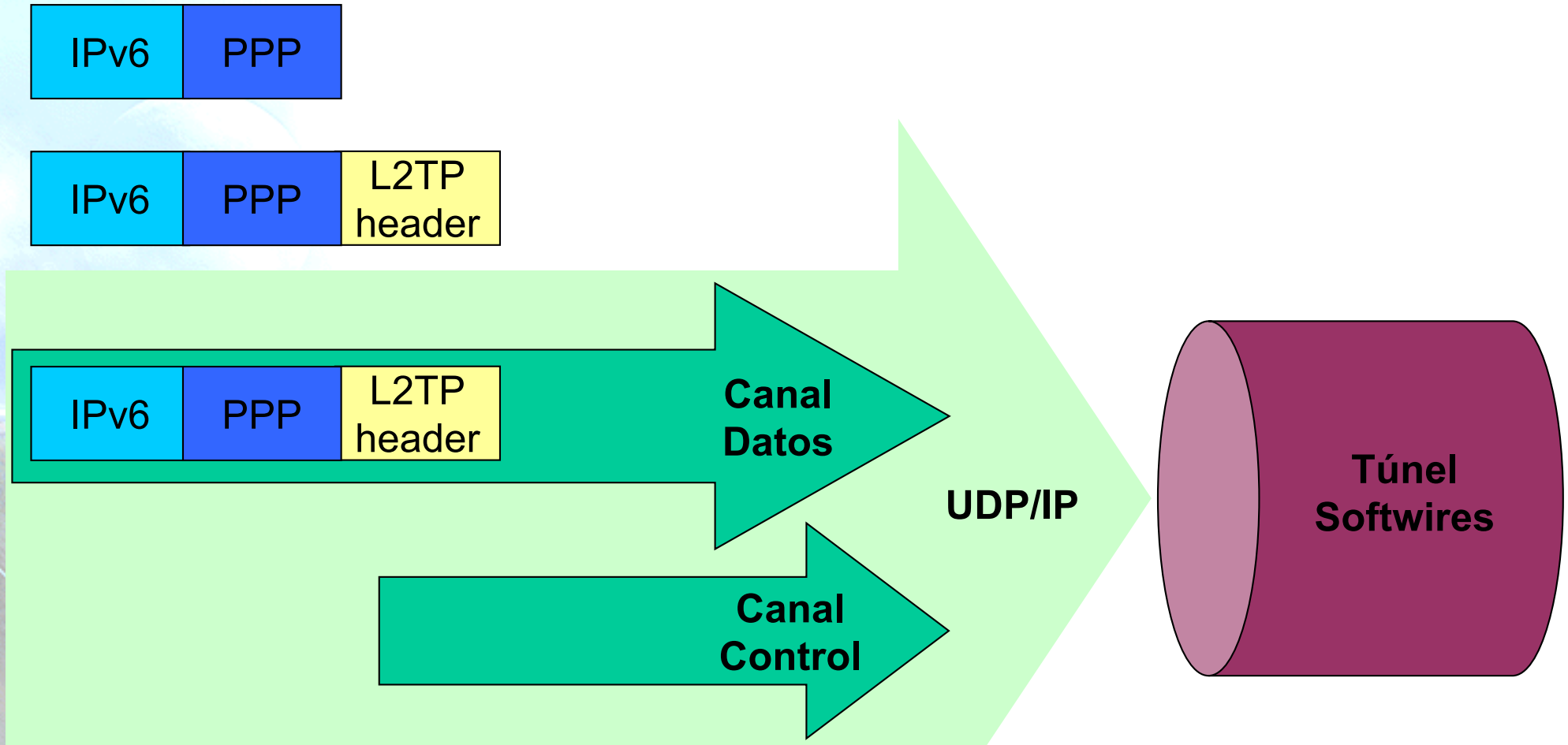
- Protocolo que esta siendo discutido en el grupo de trabajo Softwire del IETF. Presenta las siguientes características:
  - Mecanismo de transición “universal” basado en la creación de túneles
    - IPv6-en-IPv4, IPv6-en-IPv6, IPv4-en-IPv6, IPv4-en-IPv4
    - Permite atravesar NATs en las redes de acceso
    - Proporciona delegación de prefijos IPv6 (/48, /64, etc.)
    - Autenticación de usuario para la creación de túneles mediante la interacción con infraestructura AAA
    - Posibilidad de túneles seguros
    - Baja sobrecarga en el transporte de paquetes IPv6 en los túneles
    - Fácil inclusión en dispositivos portátiles con escasos recursos hardware
  - Softwires posibilitará la provisión de conectividad IPv6 en dispositivos como routers ADSL, teléfonos móviles, PDAs, etc. cuando no exista conectividad IPv6 nativa en el acceso
  - También posibilita la provisión de conectividad IPv4 en dispositivos que solo tienen conectividad IPv6 nativa
- En realidad Softwires no es un nuevo protocolo, sino la definición de cómo usar de una forma diferente protocolos ya existentes con el fin de proporcionar conectividad IPv6 en redes IPv4 y viceversa
- Sofwires se basa en:
  - L2TPv2 (RFC2661)
  - L2TPv3 (RFC3991)

Encapsulamiento de Softwires basado en L2TPv2

- El funcionamiento se especifica en draft-ietf-softwire-hs-framework-l2tpv2
- Existen dos entidades:
  - Softwires Initiator (SI): agente encargado de solicitar el túnel
  - Softwires Concentrator (SC): agente encargado de crear el túnel (tunnel end point)
- Se utiliza PPP para transportar paquetes IPx (x=4, 6) en paquetes IPy (y=4, 6)
  - Opcionalmente se puede encapsular los paquetes PPP en UDP en caso de que haya que atravesar NATs



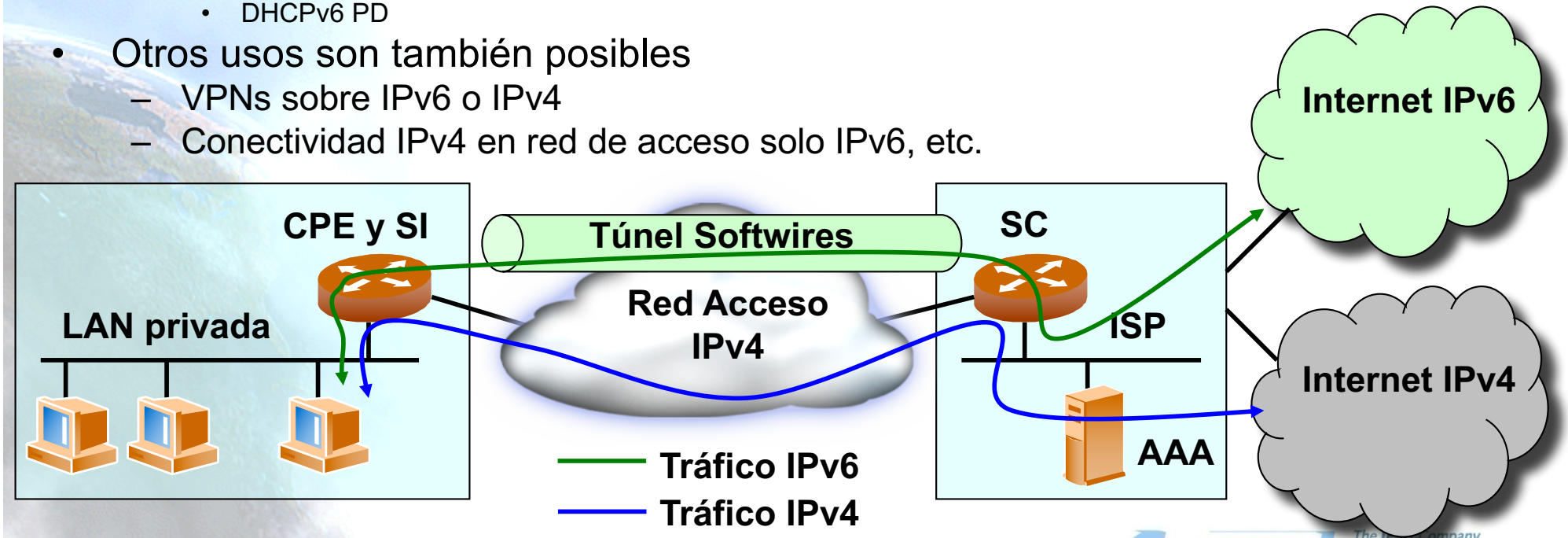
Softwires basado en L2TPv2



- Existe un plano de control y otro de datos
- Se usa PPP como protocolo de encapsulamiento

Ejemplo de uso de Softwires

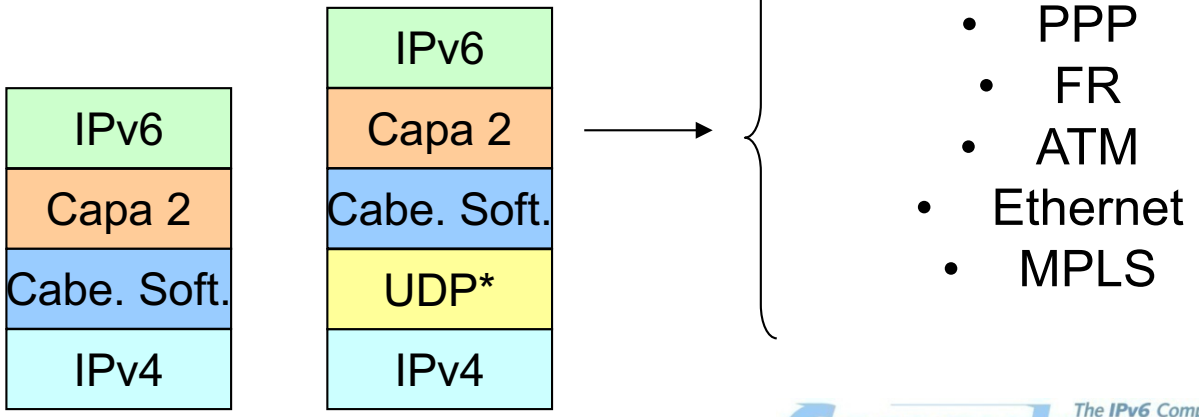
- Un uso típico previsible de Softwires es la provisión de conectividad IPv6 a usuarios domésticos a través de una red de acceso solo-IPv4
  - El SC está instalado en la red del ISP
    - DSLAM, Router de agregación u otro dispositivo
  - El SI está instalado en la red del usuario
    - CPE típicamente. También es posible otro dispositivo diferente en la red del usuario
  - El SC proporciona conectividad IPv6 al SI, y el SI hace de encaminador IPv6 para el resto de la red de usuario
  - Se usa delegación de prefijo IPv6 entre el SC y el SI para proporcionar un prefijo (típicamente /48) a la red del usuario
    - DHCPv6 PD
- Otros usos son también posibles
  - VPNs sobre IPv6 o IPv4
  - Conectividad IPv4 en red de acceso solo IPv6, etc.



Encapsulamiento de Softwires basado en L2TPv3

- Misma filosofía y componentes que con L2TPv2, pero con las particularidades de L2TPv3
  - Transporte sobre IP/UDP de otros protocolos de capa 2 diferentes a PPP
    - HDLC, PPP, FR, ATM, Ethernet, MPLS, IP
  - Formato de cabeceras mejorado para permitir un tratamiento más rápido en los SC
    - Permite velocidades del rango de T1/E1, T3/E3, OC48
  - Mínimo overhead en los paquetes encapsulados (solo de 4 a 12 bytes extra)
  - Otros mecanismos de autenticación diferentes a CHAP y PAP
    - EAP

Túnel IPv6-en-IPv4

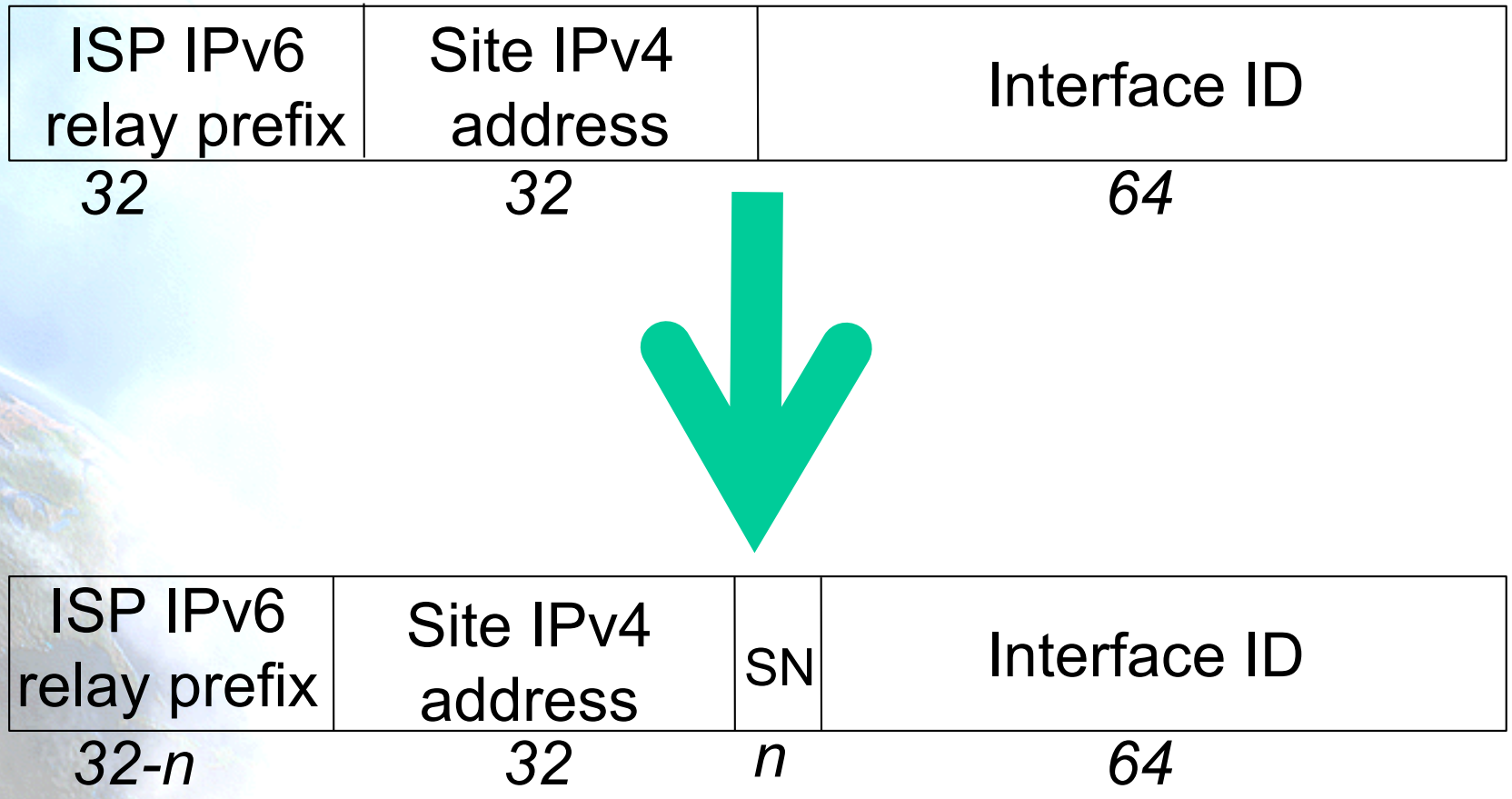


*** Opcional**

6RD: un refinamiento de 6to4 …

- 6RD: IPv6 Rapid Deployment en infraestructuras IPv4
  - 6RD depende de IPv4
- RFC5969
- Implementado por FREE (un ISP Francés)
- Cambios respecto a 6to4:
  - Formato de direcciones
  - Los relés (6rd gateway) está sólo dentro del ISP

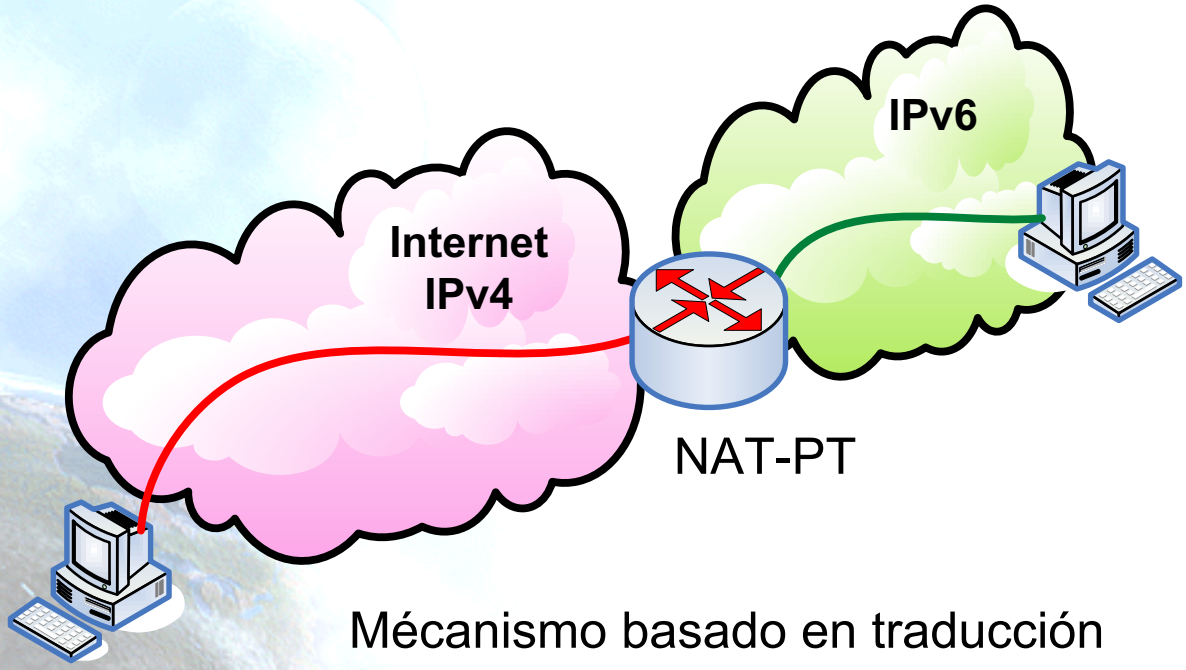
6RD: Formato de Direcciones



6RD: Pros y Contras

- Pros
  - Parece fácil de desplegar SI la red esta bajo control (CPEs, …)
  - Resuelve (?) los problemas de 6to4
    - seguridad, routing asimétrico, …
    - Relés (o gateway's) bajo control del ISP
  - Transparente para el cliente
    - Configuración automática del CPE
- Contras
  - No soportado por las políticas de los RIRs
    - Menos subredes por cliente
  - Cambio o actualización de los CPEs
    - Pocos disponibles
  - Añadir nuevo hardware: 6RD relés/gateway's
    - Pocos disponibles

Traducción IPv4/IPv6



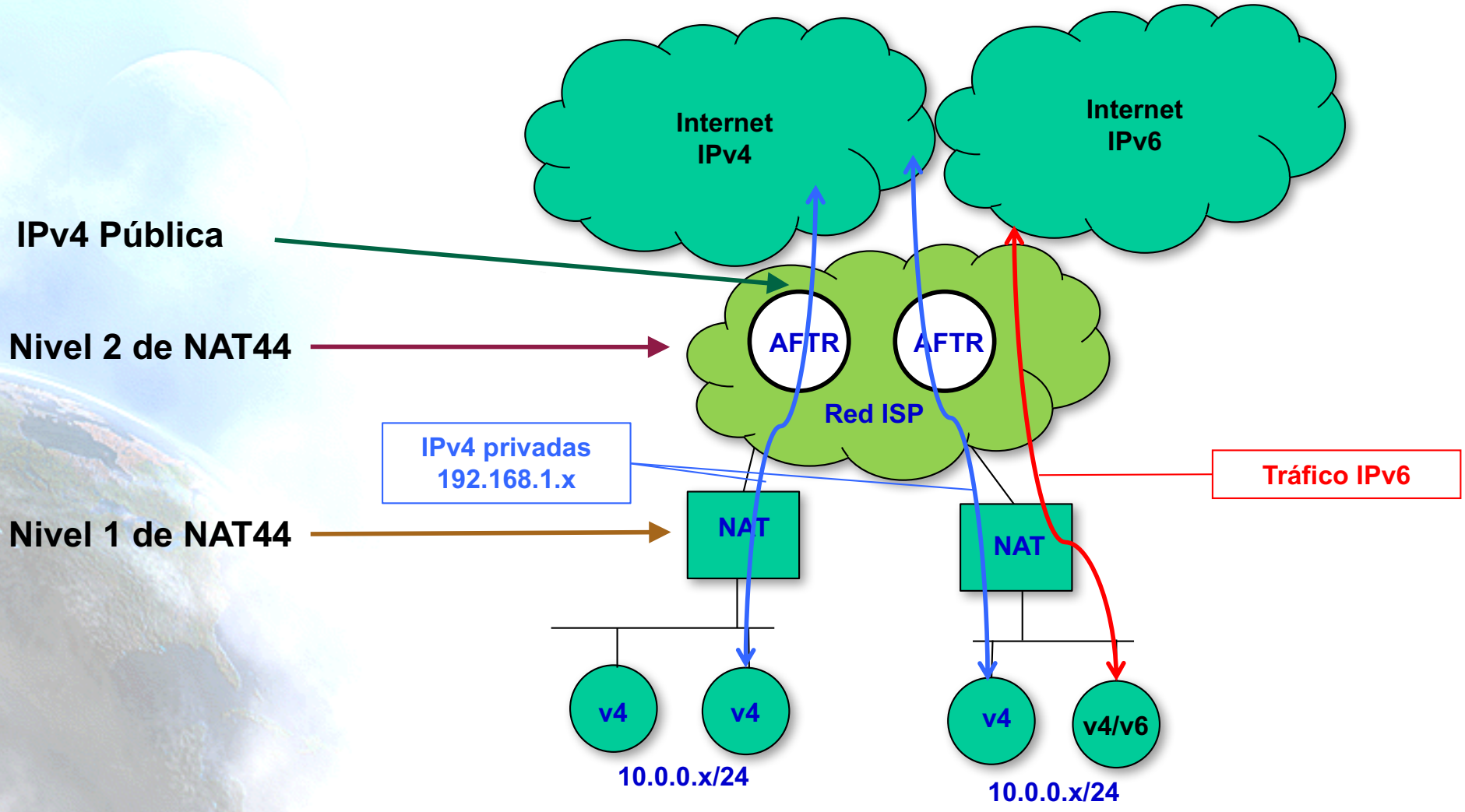
Mécanismo basado en traducción

- Diferentes soluciones, pero tiene en común que tratan de traducir paquetes IPv4 a IPv6 y viceversa
  - [SIT], [BIS], [TRT], [SOCKSv64]
- La más conocida es NAT-PT [NATPT], [NATPTIMPL]
  - Un nodo intermedio (router) modifica las cabeceras IPv4 a cabeceras IPv6
  - El tratamiento de paquetes es complejo
- Es la peor solución puesto que la traducción no es perfecta y requiere soporte de ALGs, como en el caso de los NATs IPv4
  - DNS, FTP, VoIP, etc.

NAT444

- Comúnmente conocido como CGN, CGNAT o LSN
- Sirve para prolongar **artificialmente** la vida de IPv4
- No sirve para desplegar IPv6
  - SIN cambio de CPE
- Compartiendo LAS MISMAS direcciones IPv4 entre varios clientes, combinando:
  - NAT + NAT
- Requiere diversos niveles de NAT
- Se realiza NAT y PAT (Port Address Translation)
- Requiere ALGs (Application Layer Gateways)

Esquema de NAT444



¿Qué rompe CGN?

- UPnP-IGD (Universal Plug & Play - Internet Gateway Device protocol)
- NAT-PMP (NAT Port Mapping Protocol)
- Otros mecanismos de NAT Traversal
- Seguridad
- AJAX (Asyncronous Javascript And XML)
- FTP (ficheros grandes)
- BitTorrent/Limewire (seeding – uploading)
- On-line gaming
- Video streaming (Netflix, Hulu, …)
- Camaras IP
- Tuneles, VPN, IPsec, ...
- VoIP
- Port forwarding
- ...

Ya no tenemos IPv4 …

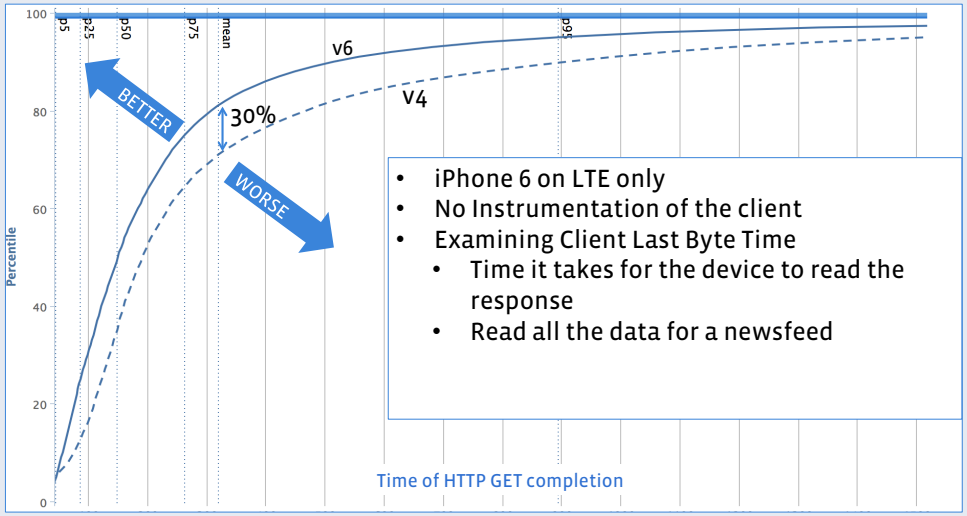
- **El agotamiento de IPv4 imposibilita**
  - **asignar IPv4 a usuarios finales**
  - **asignar IPv4 incluso en redes públicas**
  - **mantener la interoperabilidad de forma escalable con redes solo-IPv4**
- **Como consecuencia, en algunos casos será deseable, desplegar redes "sólo" IPv6**
  - Costes de operación
  - Menos recursos IPv4
  - Prestaciones
  - Eficacia
  - RFCs
  - Otros aspectos …

¿Sólo IPv6?

- **Varios casos, proliferando rápidamente**
- **Ejemplo de FaceBook**
- **Datacenters sólo IPv6**
  - Tráfico interno IPv6, 90% del total (final 2014)
  - +100 Terabits por segundo
  - Previsto: 100% IPv6 en Junio de 2015
  - Permite el uso de FaceBook en redes y clientes sólo-IPv6
  - Tráfico IPv4 (de Internet) terminado en los clusters sólo-IPv6
    - Mismo espacio RFC1918, para sesiones IPv4 BGP
    - Próximamente RFC5549
      - Advertising IPv4 Network Layer Reachability Information with an IPv6 Next Hop
    - IPv4 in IPv6 tunneling, para IPVS (IP Virtual Server)
    - IPv4 link-local (169.254.0.0/16) para Linux y switches

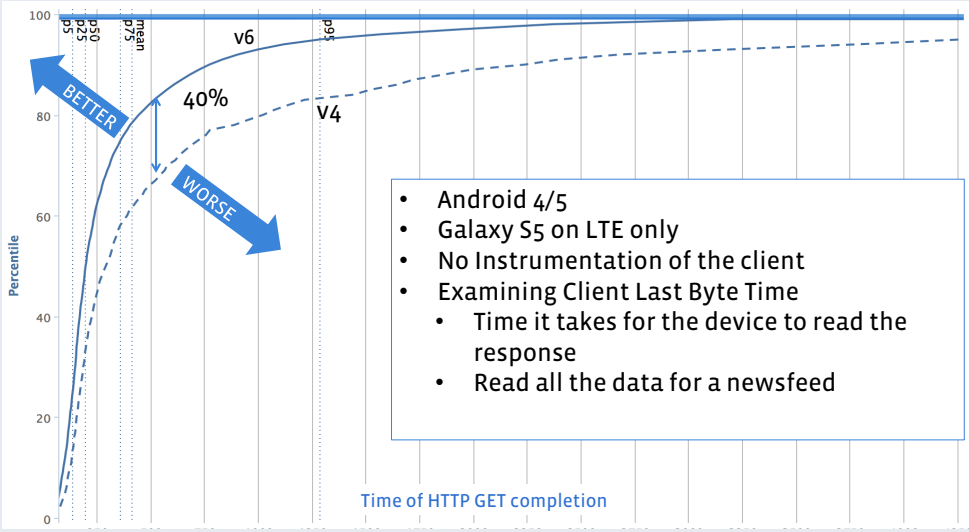
¿Prestaciones?

US Mobile Performance – Dual Stack Provider iOS



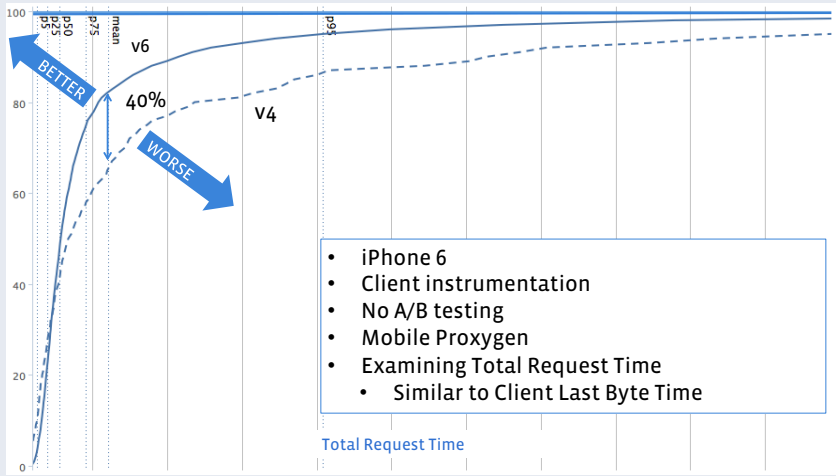
- iPhone 6 on LTE only
- No Instrumentation of the client
- Examining Client Last Byte Time
  - Time it takes for the device to read the response
  - Read all the data for a newsfeed

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

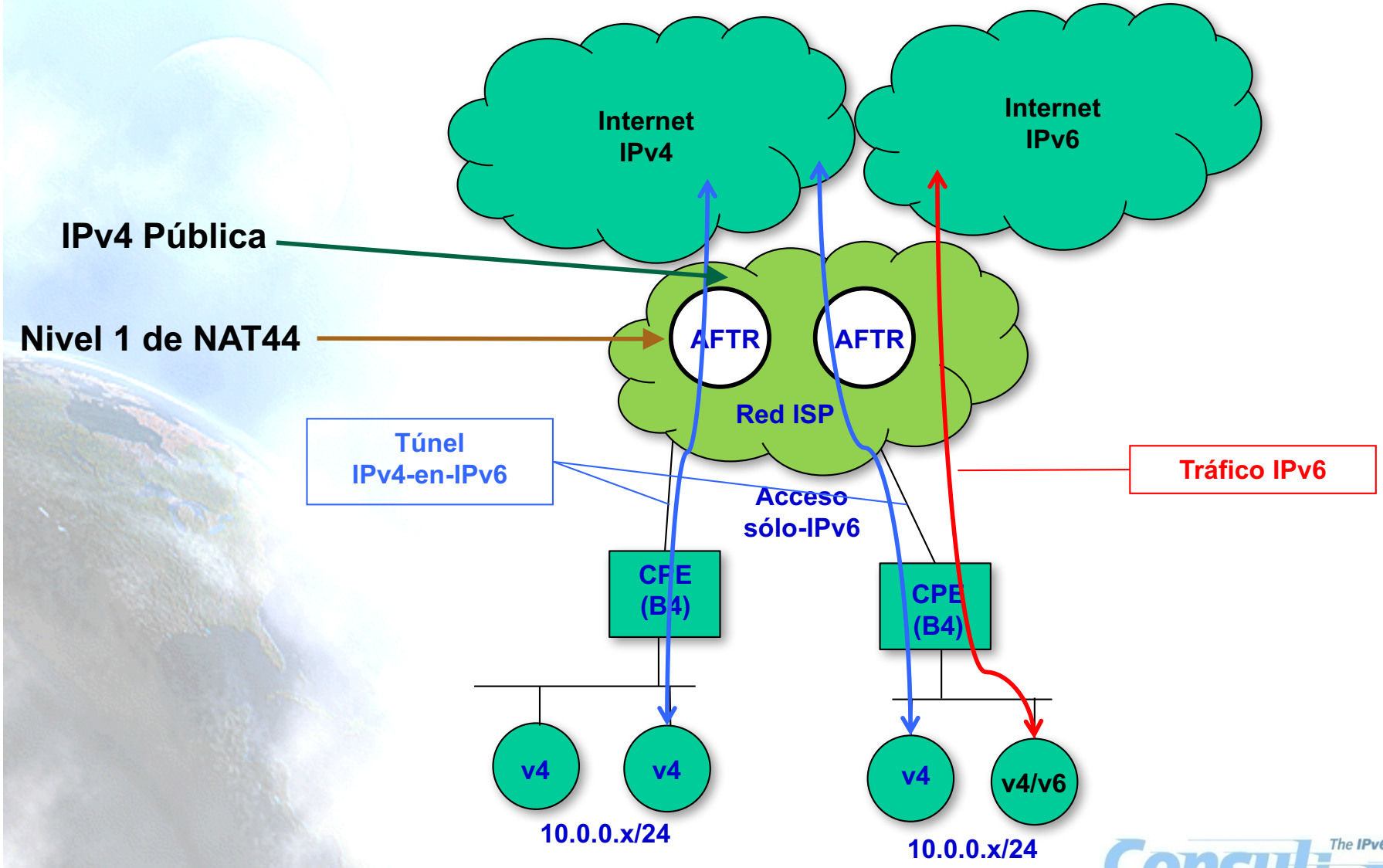
*** Datos presentados públicamente por FaceBook**

(17/3/2015)

Dual Stack Lite (DS-Lite)

- Para resolver el problema del agotamiento de IPv4 y el despliegue de IPv6
- Compartiendo LAS MISMAS direcciones IPv4 entre varios clientes, combinando:
  - Túneles
  - NAT
- No requiere diversos niveles de NAT
- Dos elementos:
  - DS-Lite Basic Bridging BroadBand (B4) - CPE
  - DS-Lite Address Family Transition Router (AFTR)
    - También denominado CGN (Carrier Grade NAT) o LSN (Large Scale NAT)

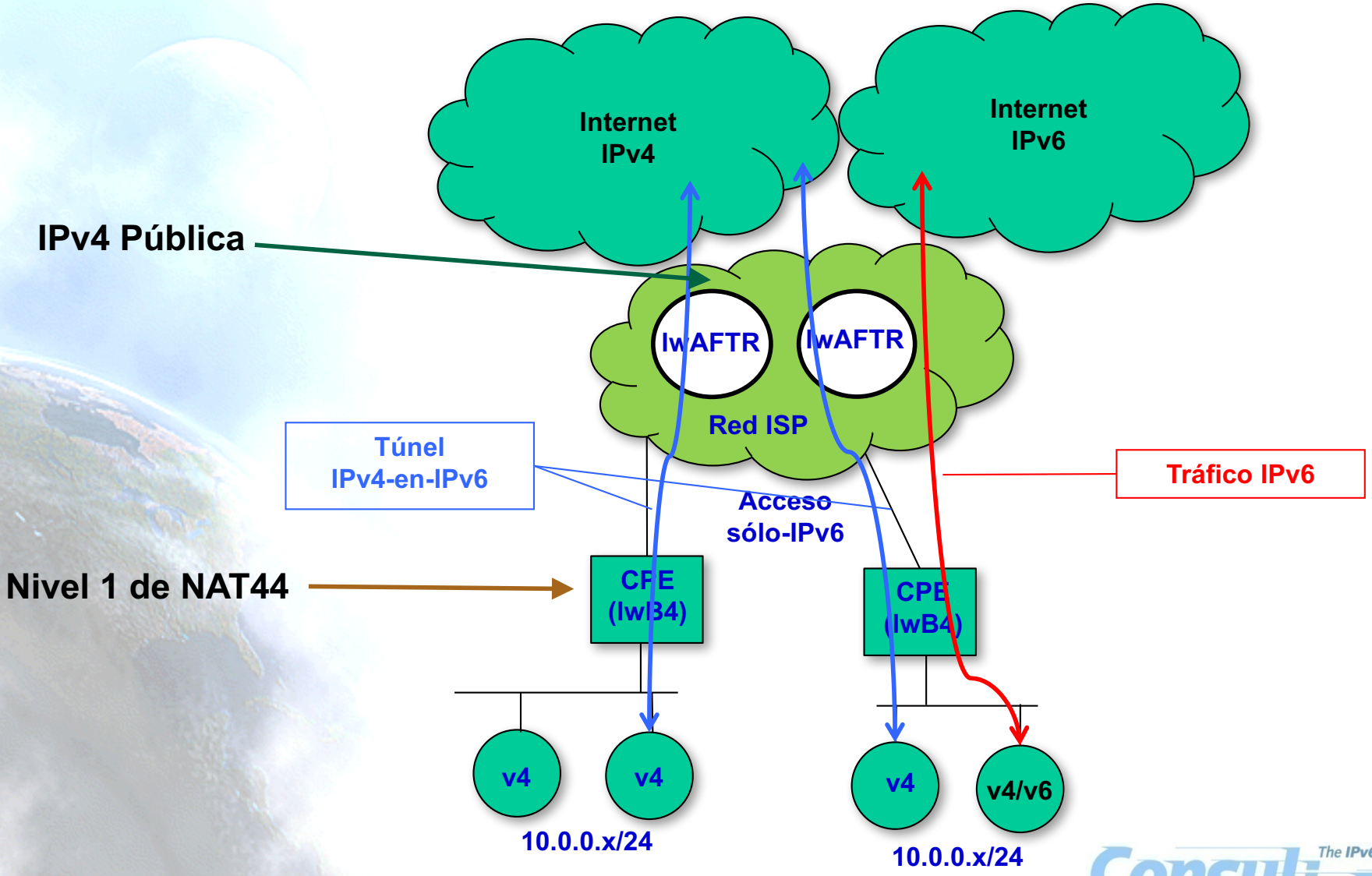
Esquema de DS-Lite



Lightweight 4over6 (lw4o6)

- Similar a DS-Lite -> Cambia posición del NAT
  - Mayor escalabilidad
  - Reduce el logging
- Compartiendo LAS MISMAS direcciones IPv4 entre varios clientes, combinando:
  - Túneles
  - NAT
- No requiere diversos niveles de NAT
- Dos elementos:
  - Lw Basic Bridging BroadBand (lwB4) - CPE
  - Lw Address Family Transition Router (lwAFTR)

Esquema de lw4o6



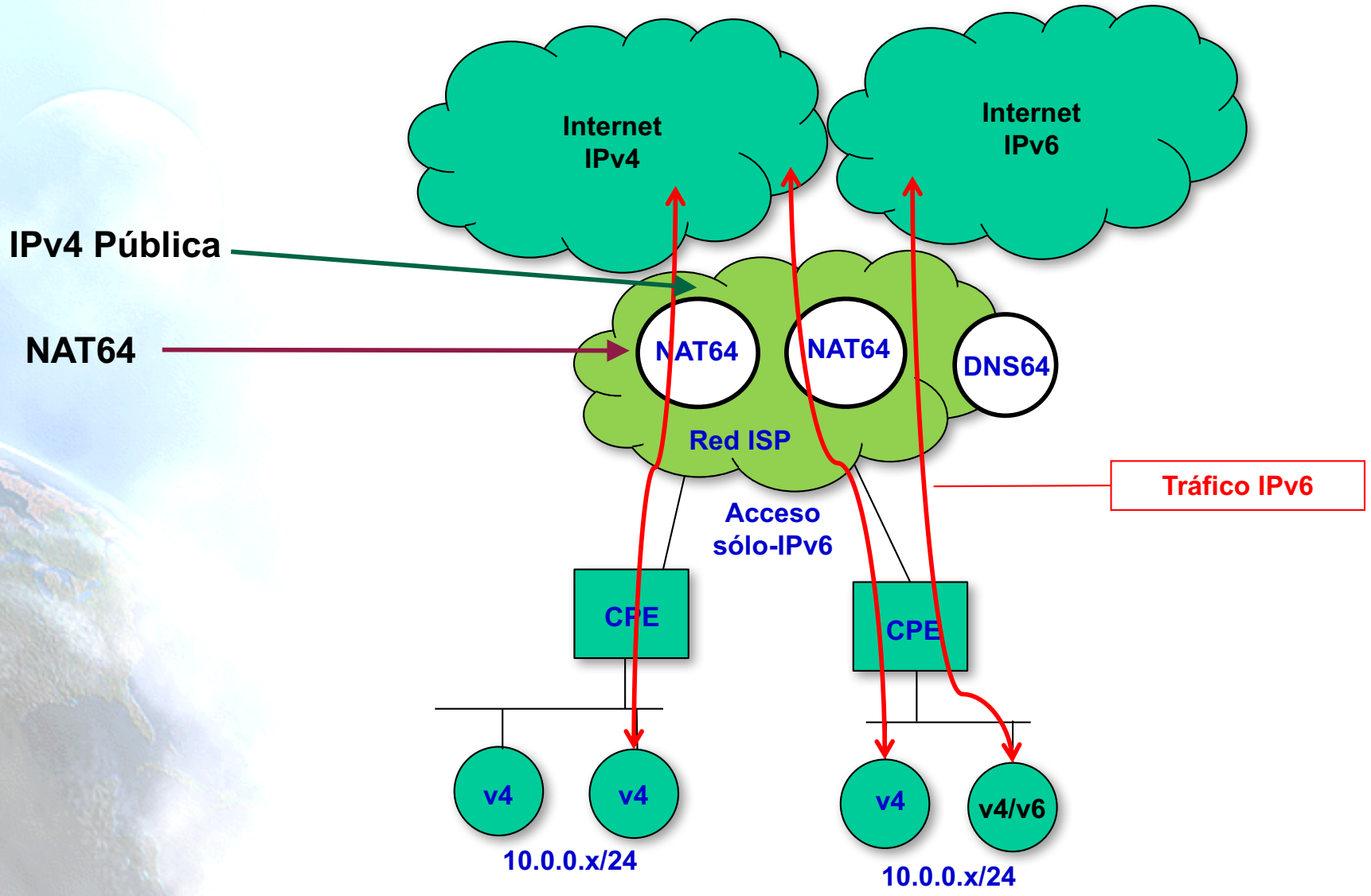
NAT64 (1)

- Cuando los ISPs sólo proporcionan conectividad IPv6, o los dispositivos son sólo-IPv6 (teléfonos móviles).
- Pero aún existen dispositivos sólo-IPv4 en Internet.
- Similar a NAT-PT, pero funcionando “mejor”.
- Elemento opcional, “desacoplado”, DNS64.
- Buena solución si no se requiere IPv4 en el cliente
  - El cliente es IPv6-only
- Algunas aplicaciones no funcionan (Skype …)
  - Peer-to-peer usando referencias IPv4
- Si se usan direcciones literales no funciona
- Si se usan socket APIs no funciona

NAT64 (2)

- NAT64 traduce paquetes IPv6 en IPv4 y viceversa
  - Algoritmo de traducción de cabeceras IP/ICMP
  - Las direcciones IPv4 son algorítmicamente traducidas a/desde direcciones IPv6 usando un algoritmo específico
  - Sólo traduce paquetes unicast con tráfico TCP, UDP e ICMP
  - DNS64 es un mecanismo para sintetizar registros AAAA desde A. La dirección IPv6 contenida en el AAAA sintetizado es generada algorítmicamente desde la dirección IPv4 y el prefijo IPv6 asignado al dispositivo NAT64
- NAT64 permite a múltiples nodos sólo-IPv6 para compartir una dirección IPv4 para acceder a Internet IPv4

Esquema de NAT64



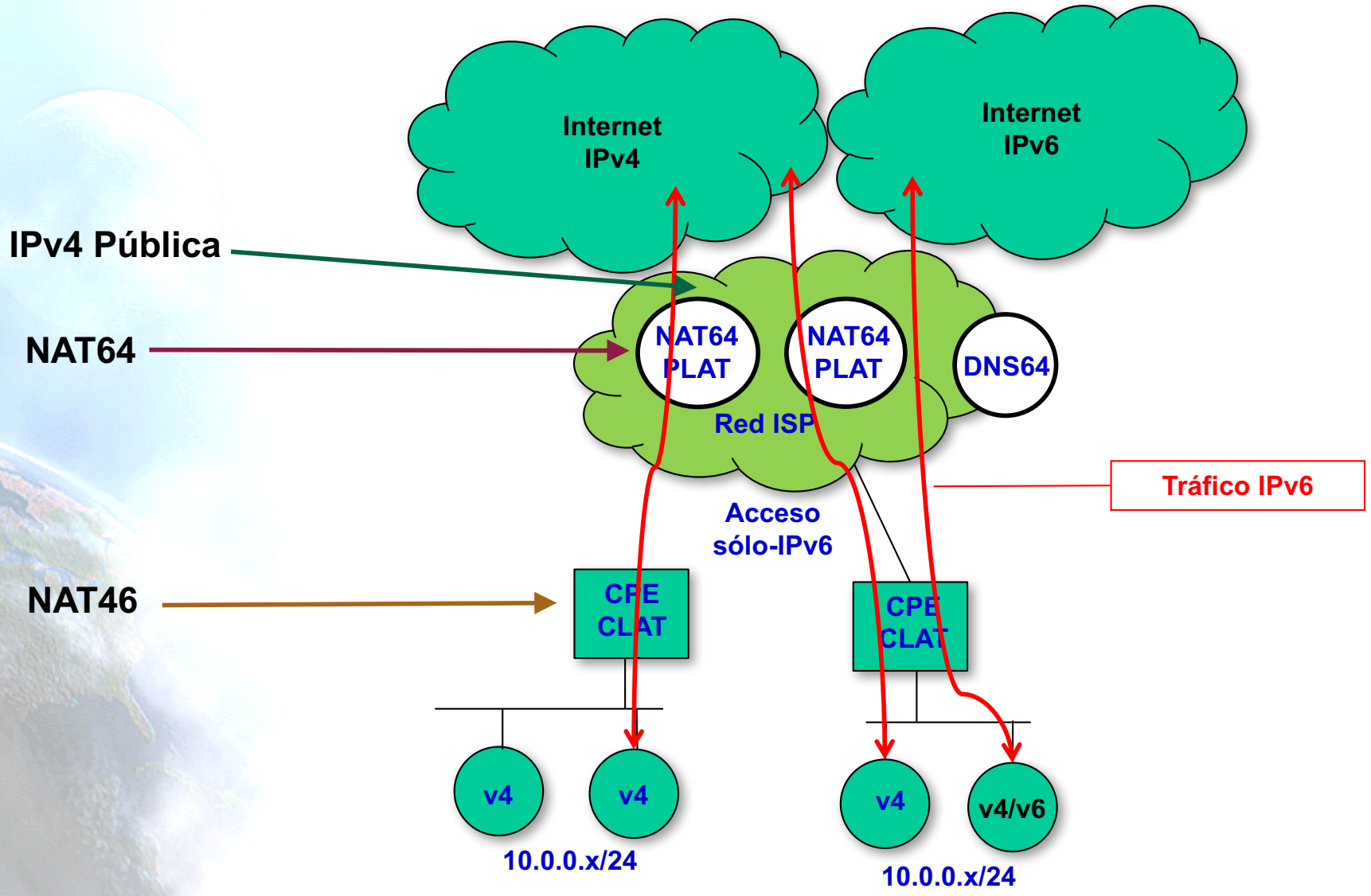
464XLAT

- RFC6877: Combina el uso del RFC6145 y RFC6146
- Uso muy eficiente de recursos IPv4 escasos
  - N*64.000 flujos por cada IPv4
  - Independiza el crecimiento de la red de la disponibilidad de IPv4
- Proporciona servicio IPv4 básico a los clientes sobre una infraestructura sólo-IPv6
  - Funciona con aplicaciones que usan socket APIs y direcciones literales (Skype, etc.)
- Permite ingeniería de tráfico
  - sin “deep packet inspection”
- Despliegue sencillo y disponible
  - Soluciones comerciales y open source

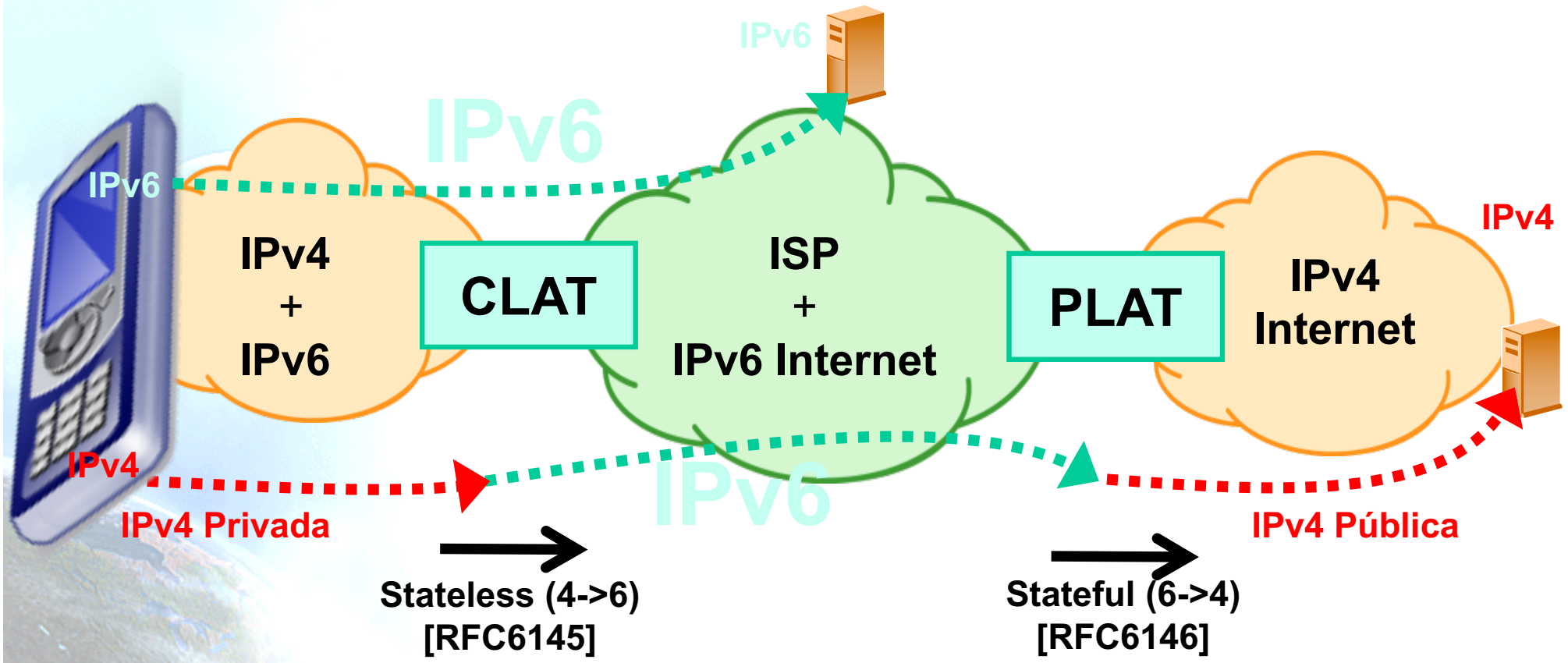
Fallo de apps con NAT64

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
TurboxTax Taxcaster	Broken	NA	
Voxer Walkie Talkie	Broken	NA	YES
Watch ESPN	Broken	1.3.1	
Zynga Poker	Broken	NA	YES

464XLAT


Internet
IPv4
Internet
IPv6
IPv4 Pública
NAT64
NAT64
PLAT
NAT64
PLAT
DNS64
Red ISP
Tráfico IPv6
Acceso
sólo-IPv6
NAT46
CPE
CLAT
CPE
CLAT
v4
v4
v4
v4/v6
10.0.0.x/24
10.0.0.x/24

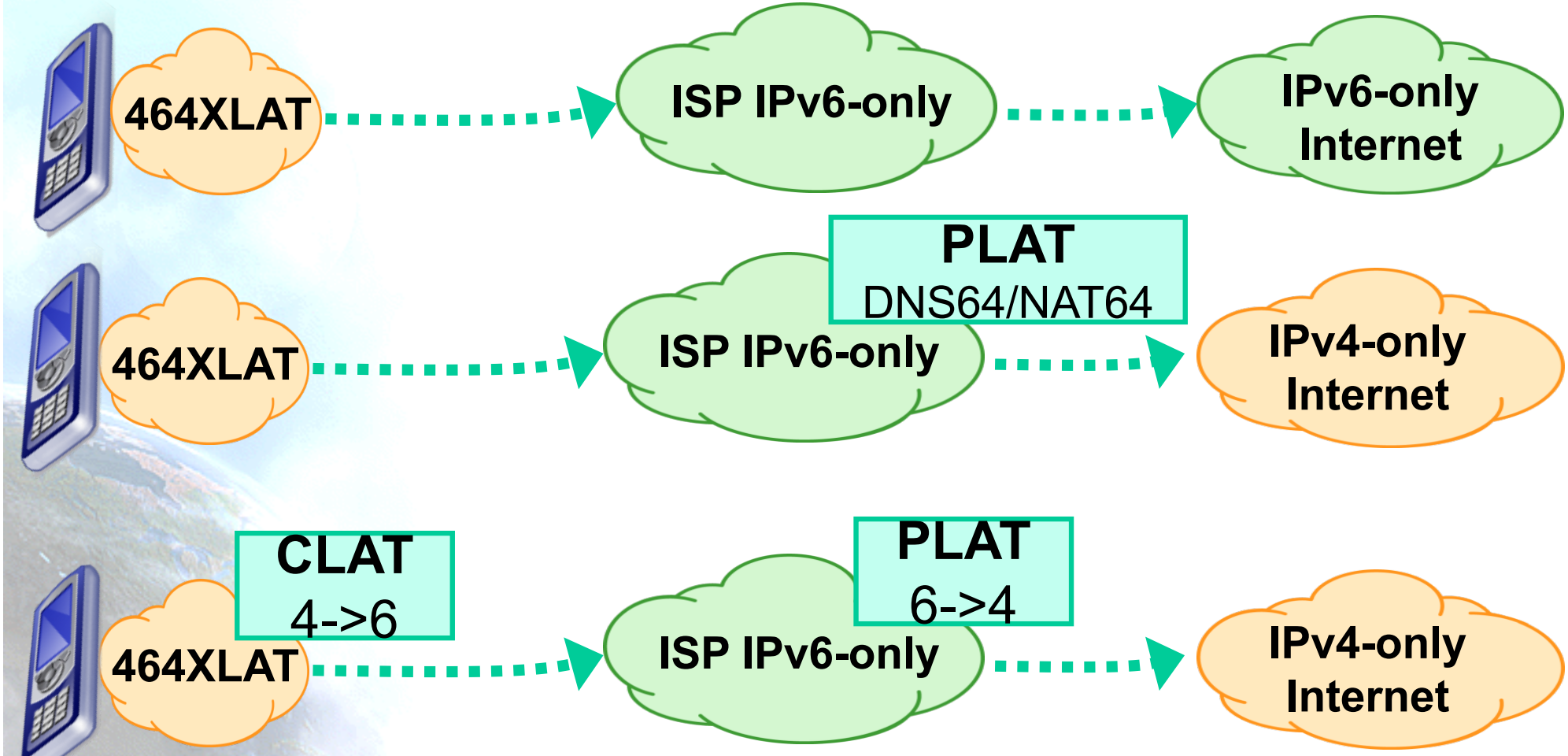
¿Cómo funciona 464XLAT?



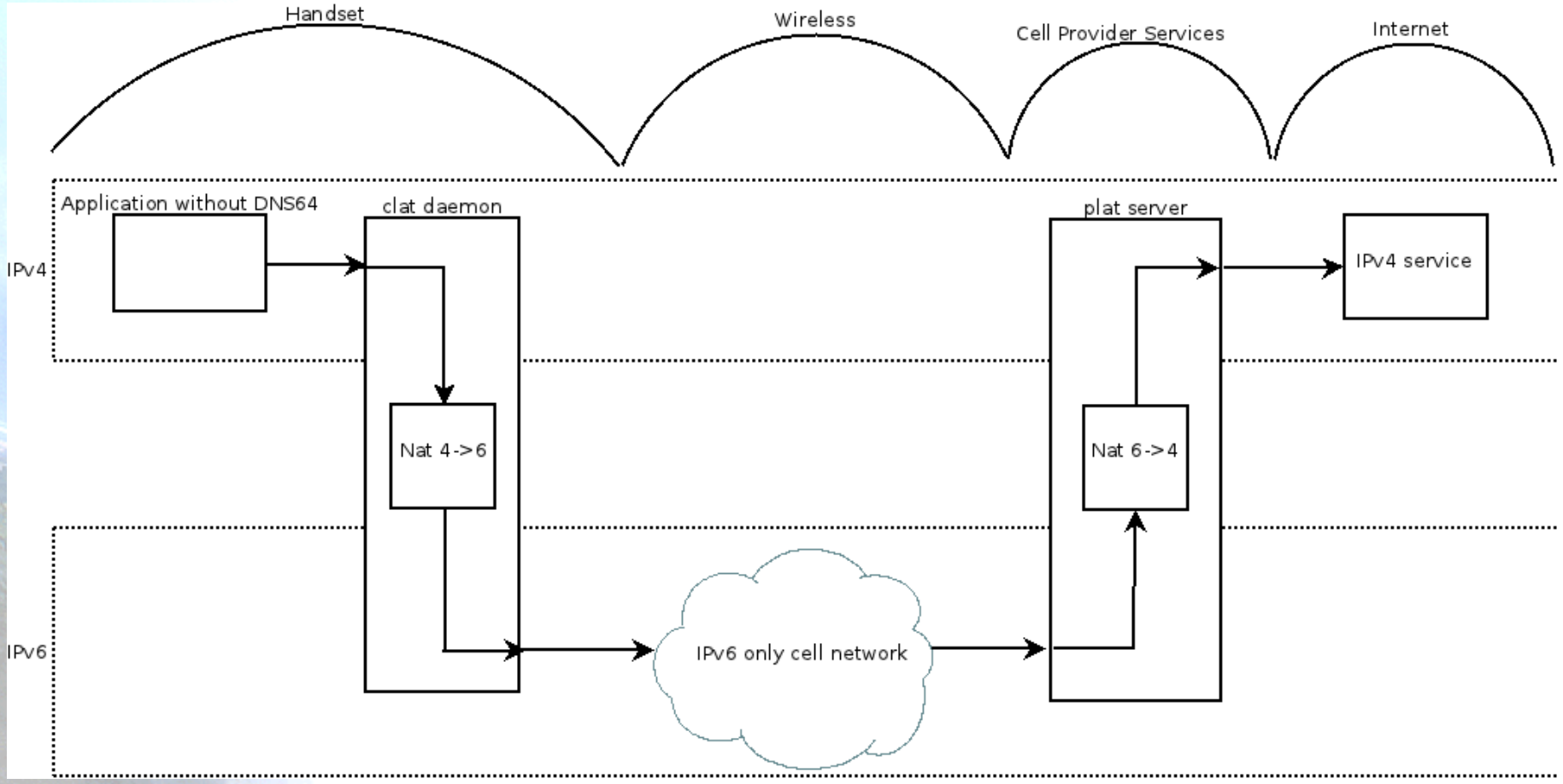
CLAT: Customer side translator (XLAT)

PLAT: Provider side translator (XLAT)

Diversos usos de la Red



Simplicidad



*** Gráfico por Dan Drown**

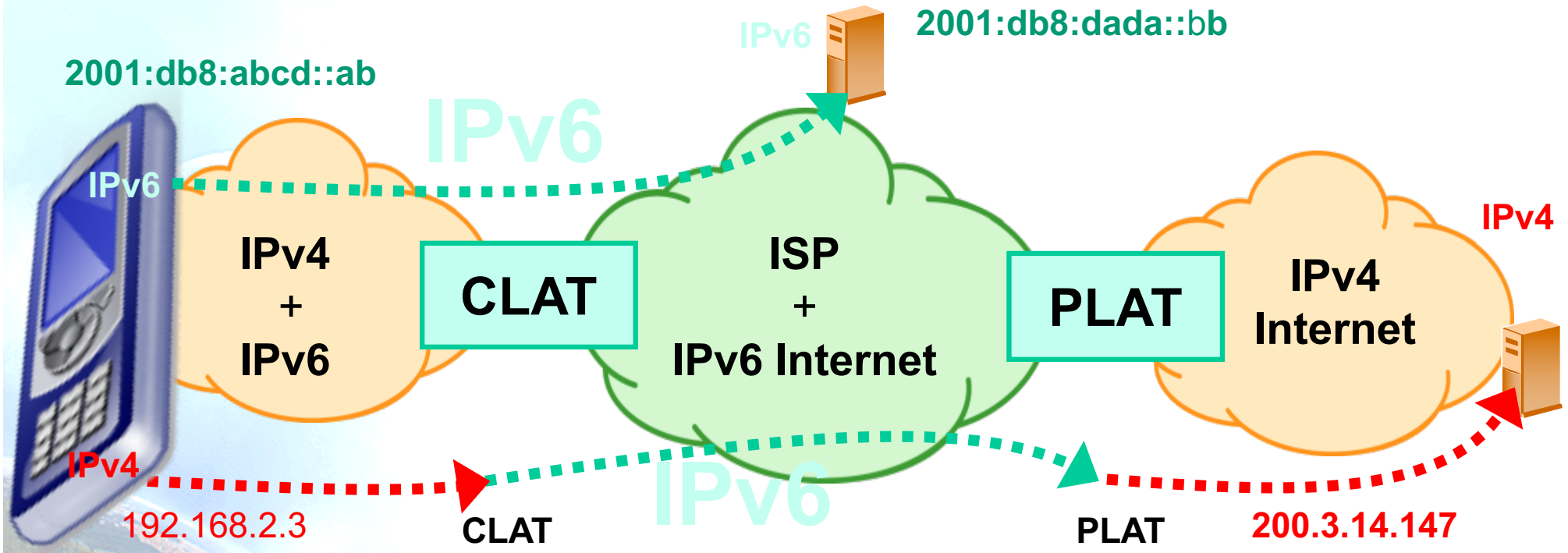
Disponibilidad y Despliegue

- NAT64:
  - A10
  - Cisco
  - F5
  - Juniper
  - NEC
  - Huawei
  - Jool, Tayga, Ecdsys, Linux, OpenBSD, …
- CLAT
  - Android
  - Nokia
  - Windows phone
  - NEC
  - OpenWRT
- Commercial deployments:
  - T-Mobile US: +68 Millions of users
  - Orange
  - Telstra
  - SK Telecom
  - …
  - Big trials in several ISPs (thousands of users)

Coste vs. otras opciones

- Sin CapEx
- Reduce coste de red:
  - No se requiere CGN
  - No hay que “comprar” IPv4
- Lógica de PCFR (Policy and Charging Control Function) para activar IPv6 de forma selectiva en sesiones de clientes en roaming
  - En función del “partner de roaming”
- Sin impacto en billing
  - Truncando direcciones IPv6 en CDRs

Direccionamiento 464XLAT



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

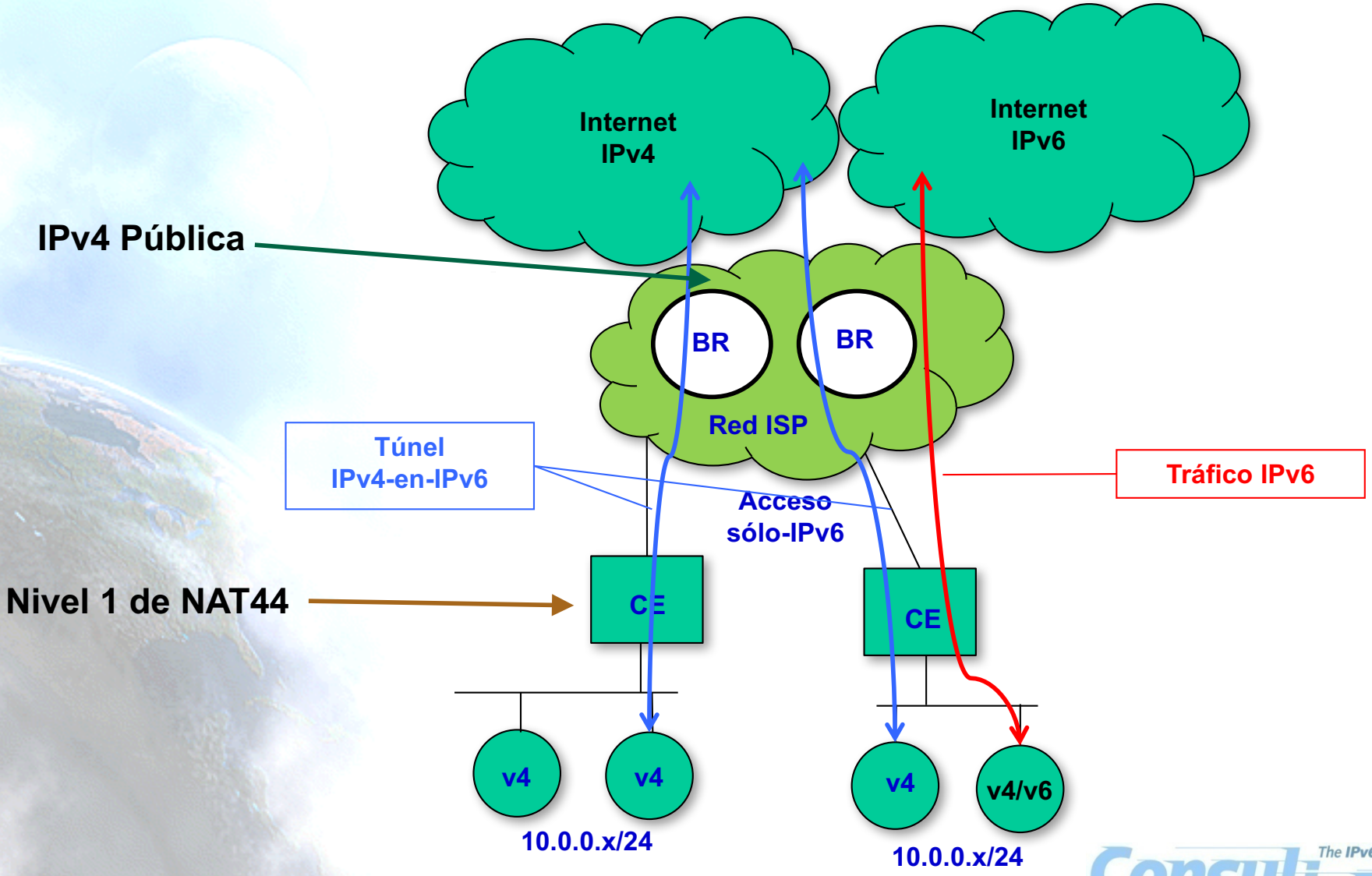
→ **Stateful XLATE [RFC6146]**

IPv4 SRC
192.1.0.1
IPv4 DST
200.3.14.147

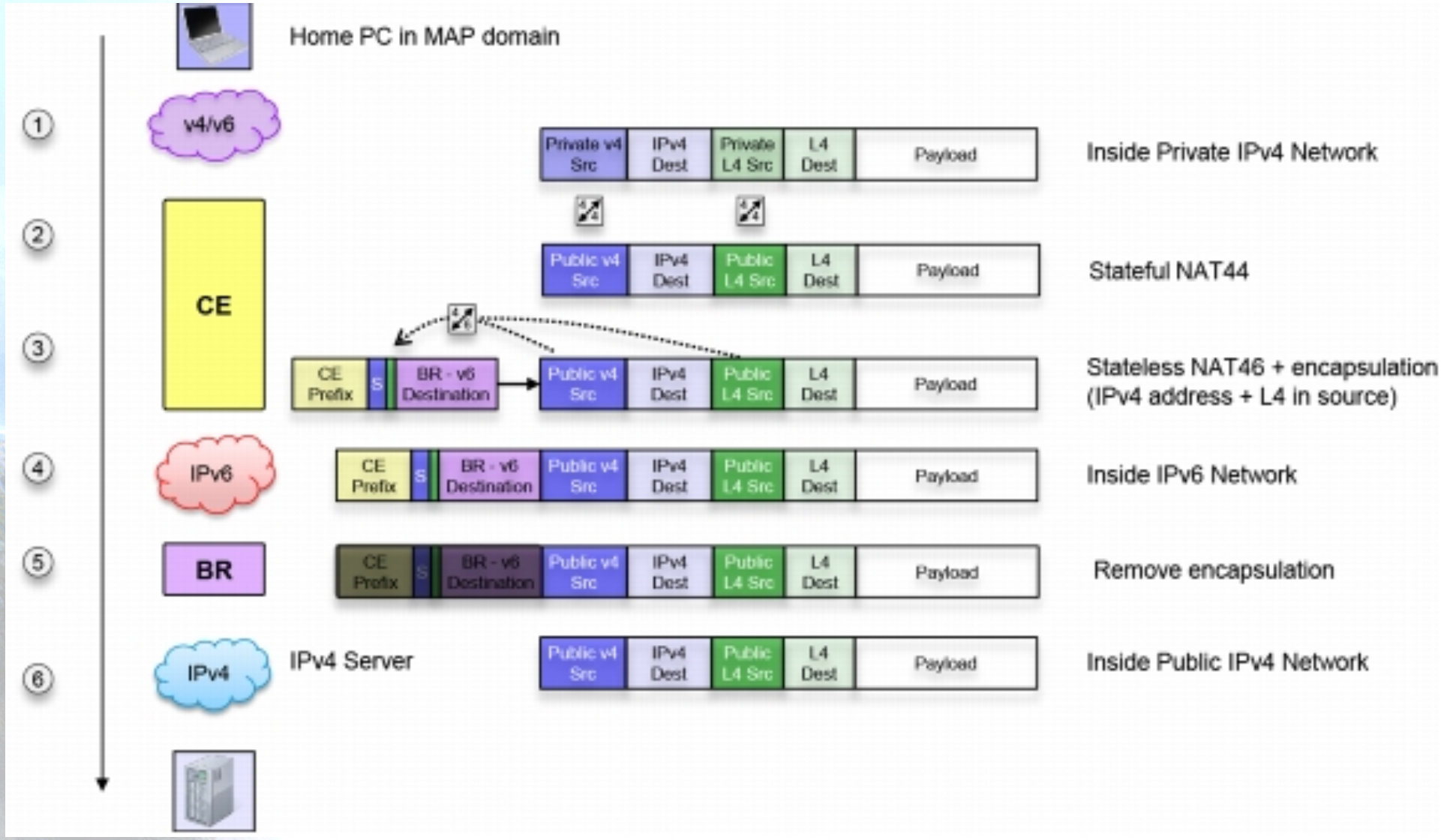
MAP Encapsulation (MAP-E)

- Mapping of Address and Port with Encapsulation
- Es una versión de DS-Lite SIN estado
  - Provisión de un prefijo IPv4, dirección IPv4 o dirección IPv4 “compartida”
  - Mapeado algorítmico entre IPv4 y una dirección IPv6
  - Extiende CIDR a 48 bits (32 IP + 16 puerto)
- El mapeado soporta el encapsulado de paquetes IPv4 en IPv6 tanto en topologías mesh como hubs&spoke, incluyendo el mapeado de direcciones con independencia completa entre direcciones IPv4 e IPv6.
- Dos elementos:
  - MAP Customer Edge (CE)
  - MAP Border Relay (BR)

Esquema de MAP-E



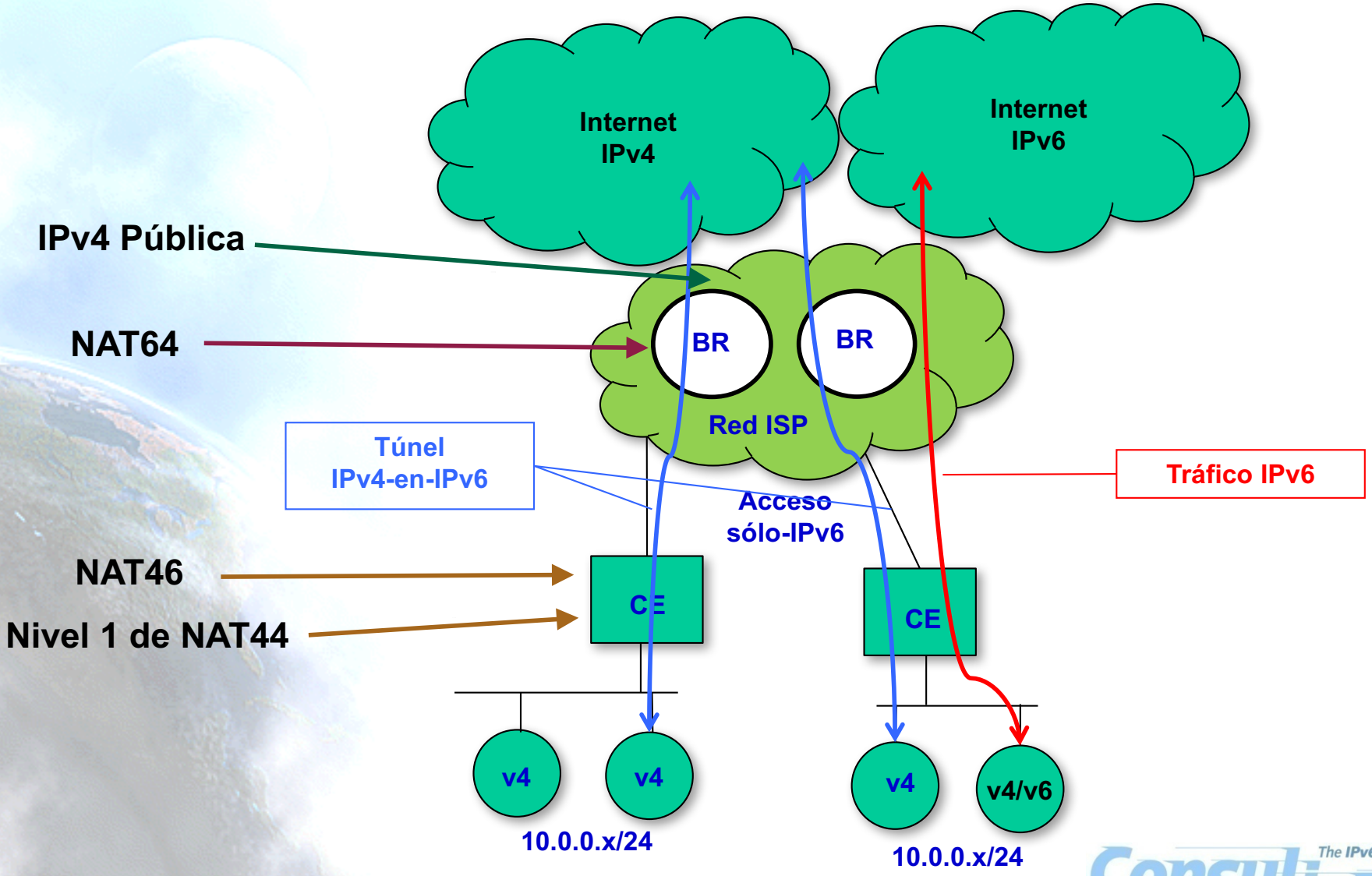
MAP-E Packet Path



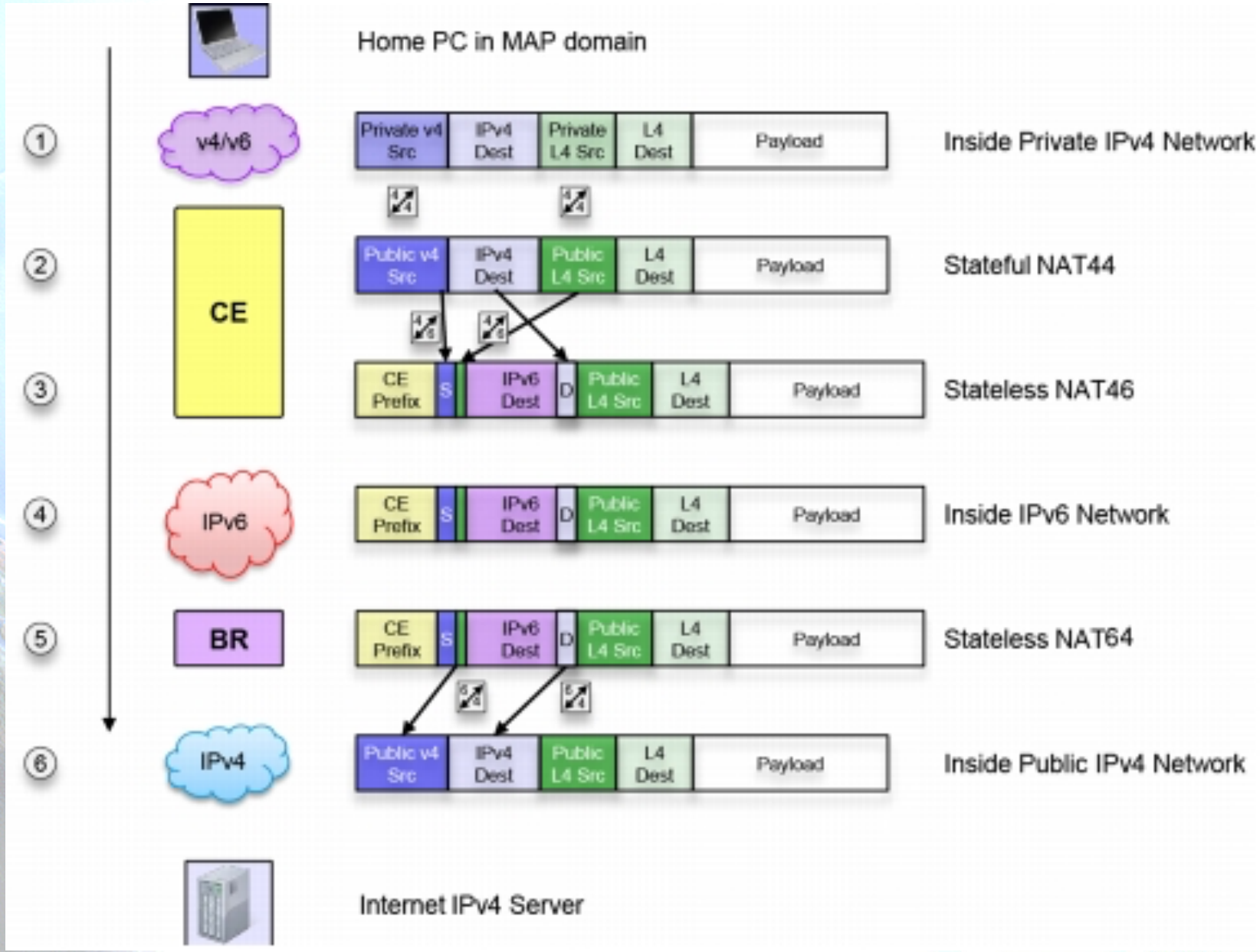
MAP Translation (MAP-T)

- Mapping of Address and Port using Translation
- Similar a MAP-E
- Similar a 464XLAT en el aspecto de la doble traducción NAT46 (CLAT) y NAT64 (PLAT).

Esquema de MAP-T

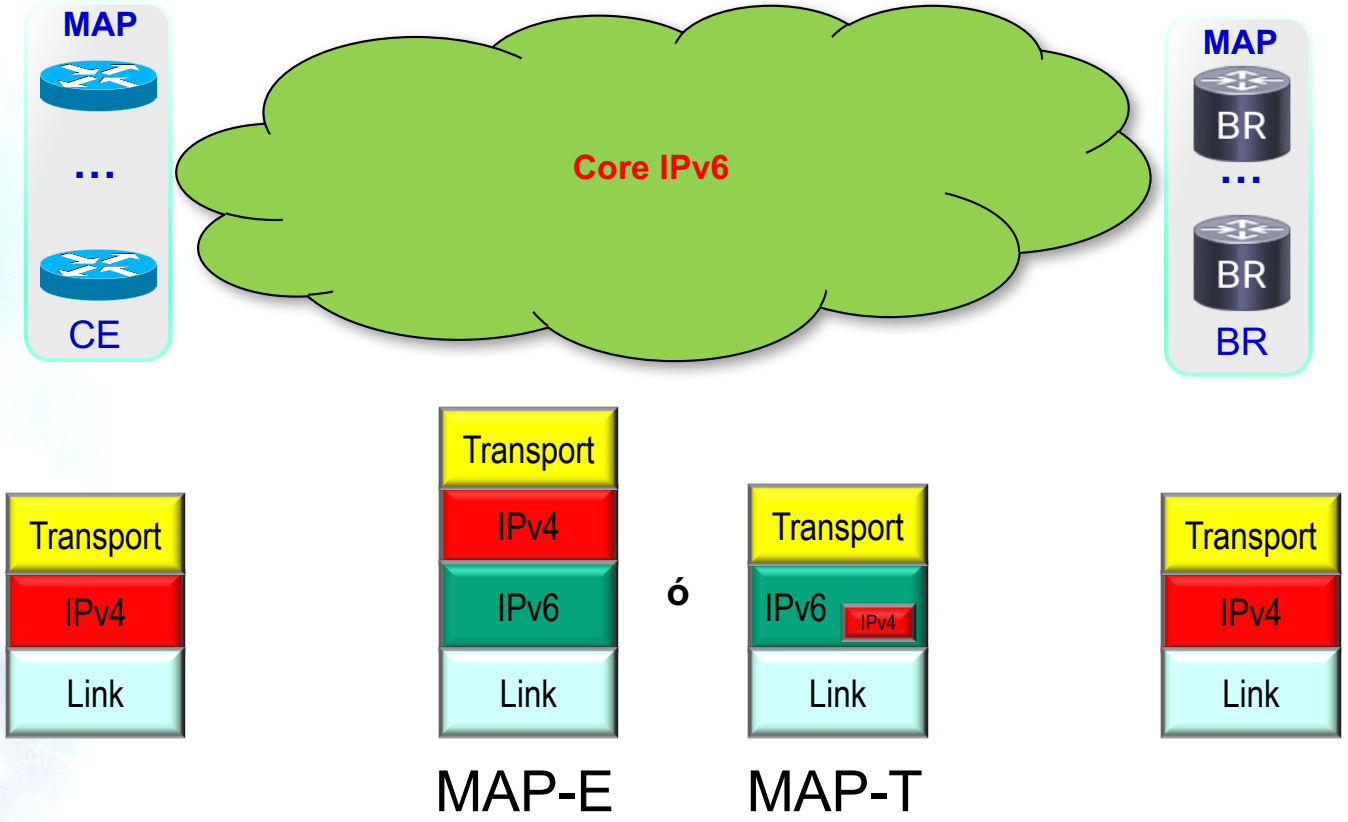


MAP-T Packet Path

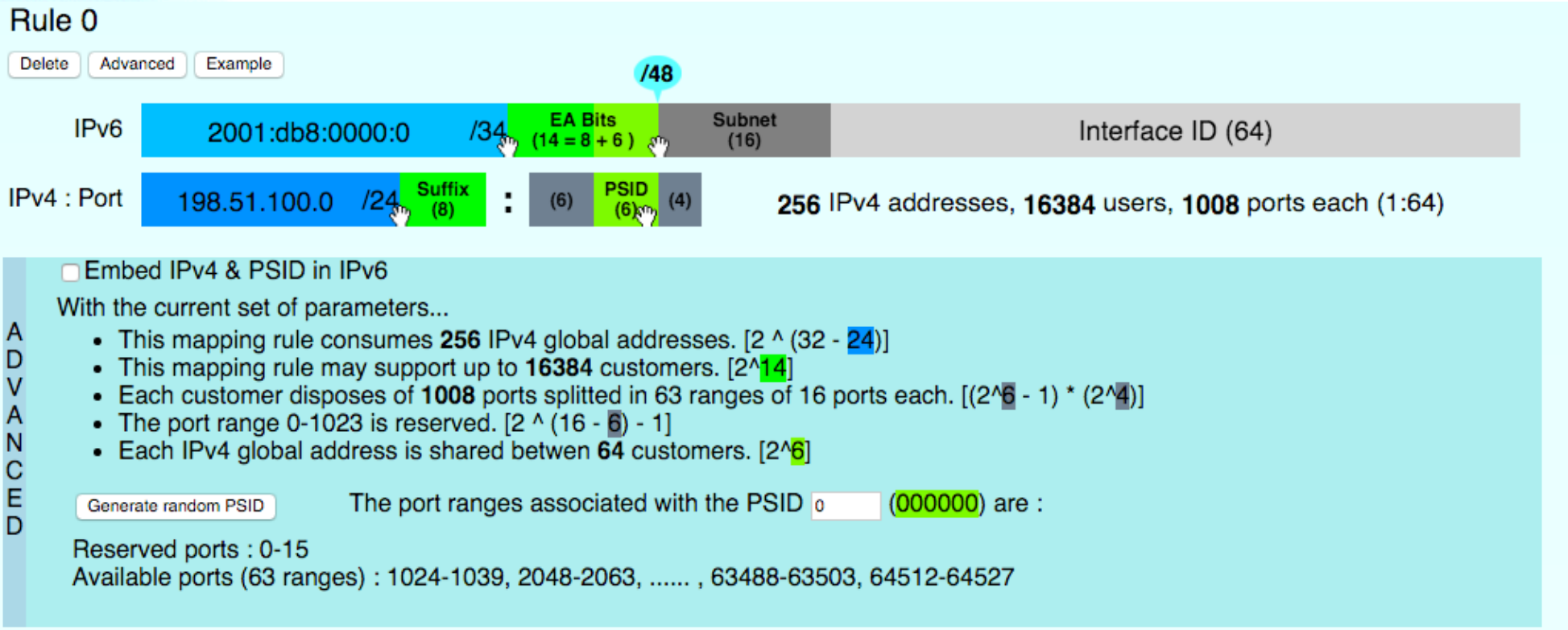


MAP-E vs MAP-T

- MAP-E utiliza 20 bytes adicionales para encapsulado (cabecera del túnel IPv4 en IPv6).

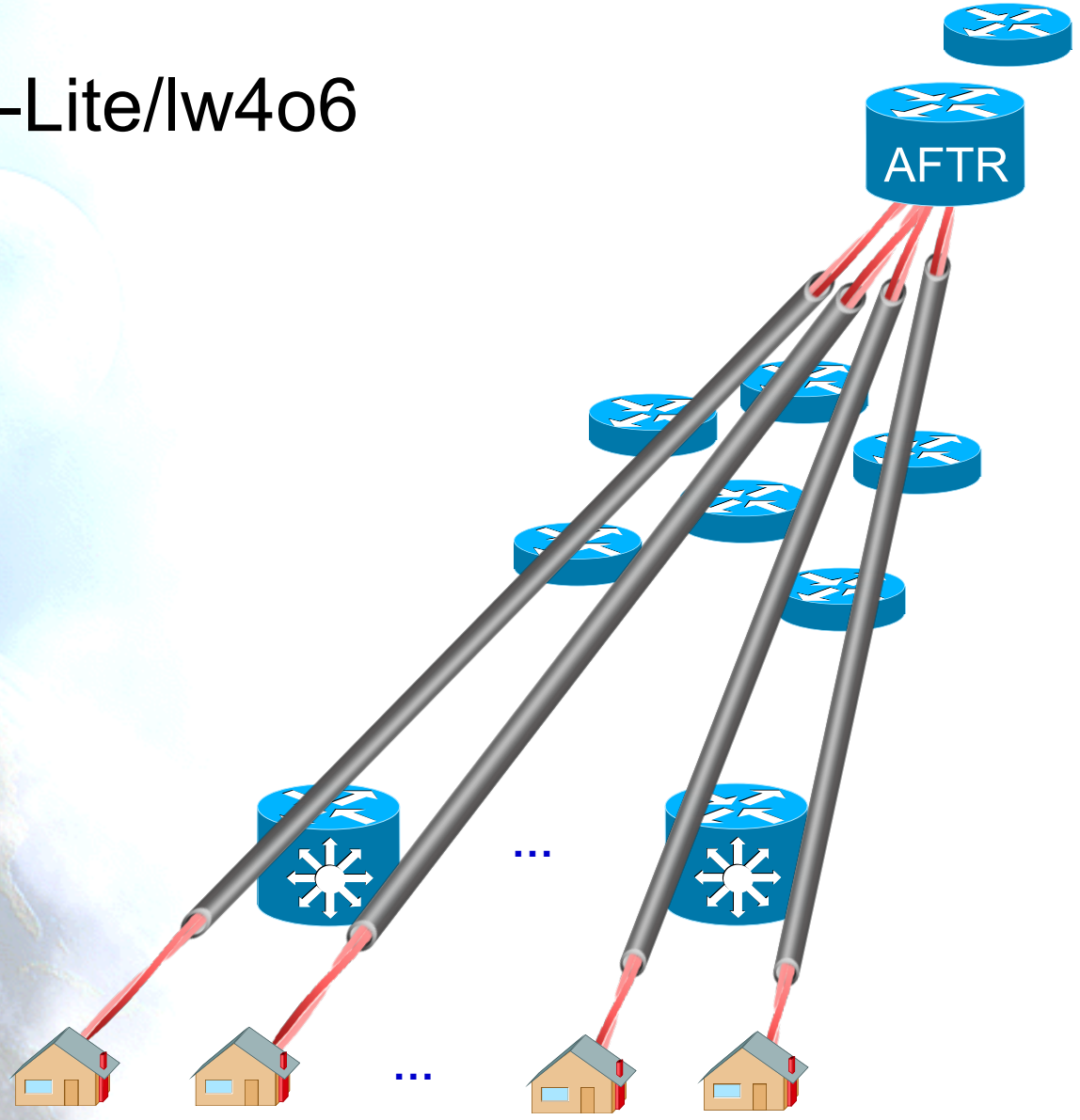


Ejemplo de direccionamiento



Túneles por suscriptores

- DS-Lite/lw4o6



Decenas de prefijos BGP

Millones de túneles

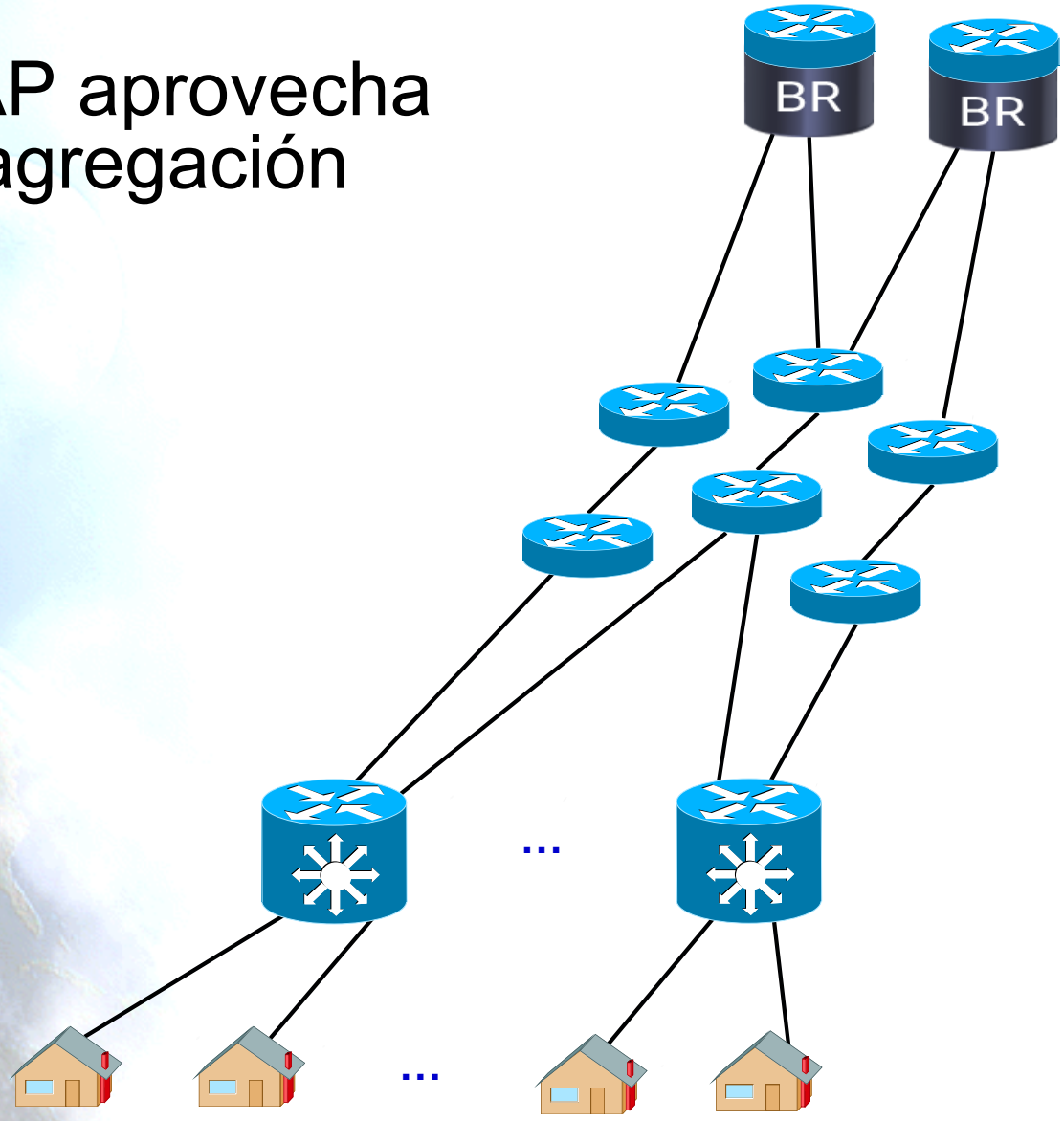
Cientos de prefijos IGP

Miles de rutas por BNG

Millones de suscriptores

Routing IPv6

- MAP aprovecha la agregación



Decenas de prefijos BGP

Decenas de Reglas MAP
NO se requiere CGN

Cientos de prefijos IGP

Miles de rutas por BNG

Millones de suscriptores

Prestaciones DS-Lite vs MAP

Cisco ASR9K

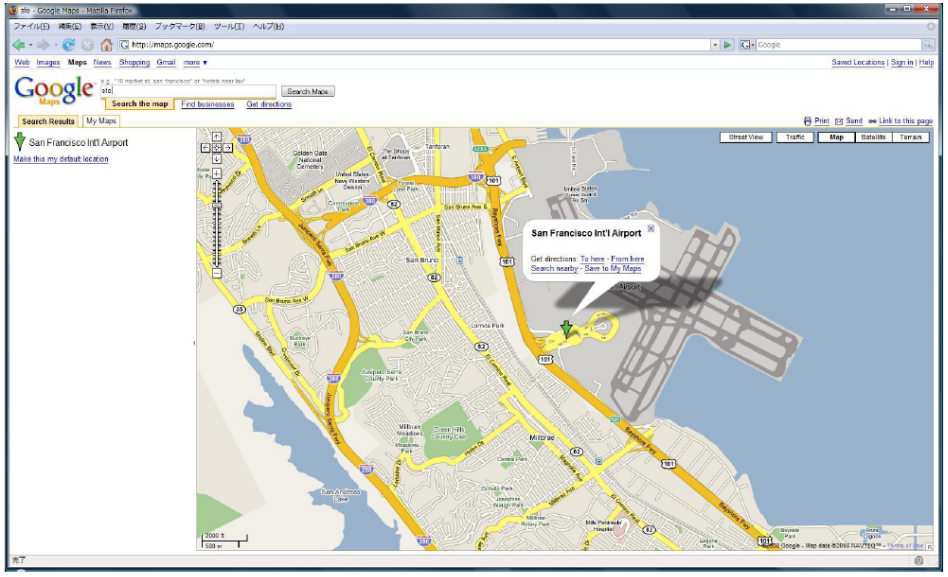
- DS-Lite encamina el tráfico en el ISM Blade
  - 14 Gbps por slot
- MAP NO necesita hacerlo
  - 240 Gbps por slot

Comparativa

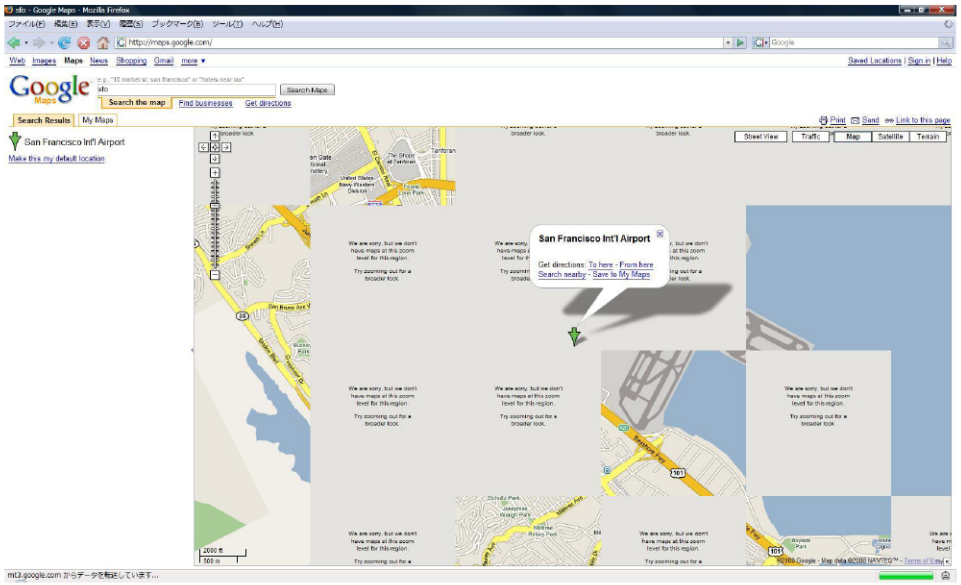
	6RD	Softwires v2	NAT444	DS-Lite	Lw4o6	NAT64	464XLAT	MAP-E	MAP-T
Túnel/Traducción (X)	T 6in4	T 6in4	X	T 4in6	T 4in6	X	X	T 4in6	X
Doble pila LAN	SI	SI	opcional	SI	SI	SI	SI	SI	SI
Multicast IPv4	SI	SI	SI	NO	NO	NO	NO	NO	NO
Red Acceso	IPv4	IPv4	IPv4 /dual	IPv6	IPv6	IPv6	IPv6	IPv6	IPv6
Overhead	20 bytes	40 bytes	-	40 bytes	40 bytes	20 bytes	20 bytes	40 bytes	20 bytes
Impacto plan direccionamiento IPv6	SI	NO	NO	NO	NO	NO	NO	SI	SI
Actualización CPE	SI	SI	opcional	SI	SI	SI	SI	SI	SI
NAT44/NAPT	CPE	CPE	CPE y CGN	CGN	CPE	CPE	CPE	CPE	CPE
Traducción 46/64	-	-	-	-	-	ISP	ISP y/o CPE	-	CPE + ISP
Traducción ISP sin/con estado	-	-	CON	-	-	CON	CON	SIN	SIN
Escalabilidad	Alta	Media	Media	Media	Alta	Alta	Alta	Alta	Alta
Prestaciones	Alta	Baja	Baja	Baja	Alta	Media	Media	Alta	Alta
ALGs	NO	NO	SI	SI	NO	SI	SI	SI	SI
Soporte otros vs sólo-TCP/UDP/ICMP	SI	SI	SI	SI	SI	NO	NO	NO	NO
Comparte "puertos"/IPv4	NO	NO	SI	SI	SI	NO	NO	SI	SI
Agregación IPv6	NO	NO	opcional	SI	SI	SI	SI	SI	SI
Mesh IPv4	SI	SI	SI	NO	NO	NO	NO	SI	SI
Mesh IPv6	SI	NO	opcional	SI	SI	SI	SI	SI	SI
Impacto en logging	NO	NO	SI	SI	NO	SI	SI	NO	NO
Facilidad HA	Alta	Baja	Baja	Baja	Alta	Media	Media	Alta	Alta
Facilidad DPI	Baja	Baja	Alta	Baja	Baja	Alta	Alta	Baja	Alta
Soporte en celular	NO	NO	SI	NO	NO	SI	SI	NO	NO
Soporte en CPEs	SI	SI	SI	SI	SI	SI	SI	SI	SI
	15.5	12.5	10.5	9.5	15	12.5	13	13	13.5

¿Cuantos puertos por usuario?

Max 30 Connections

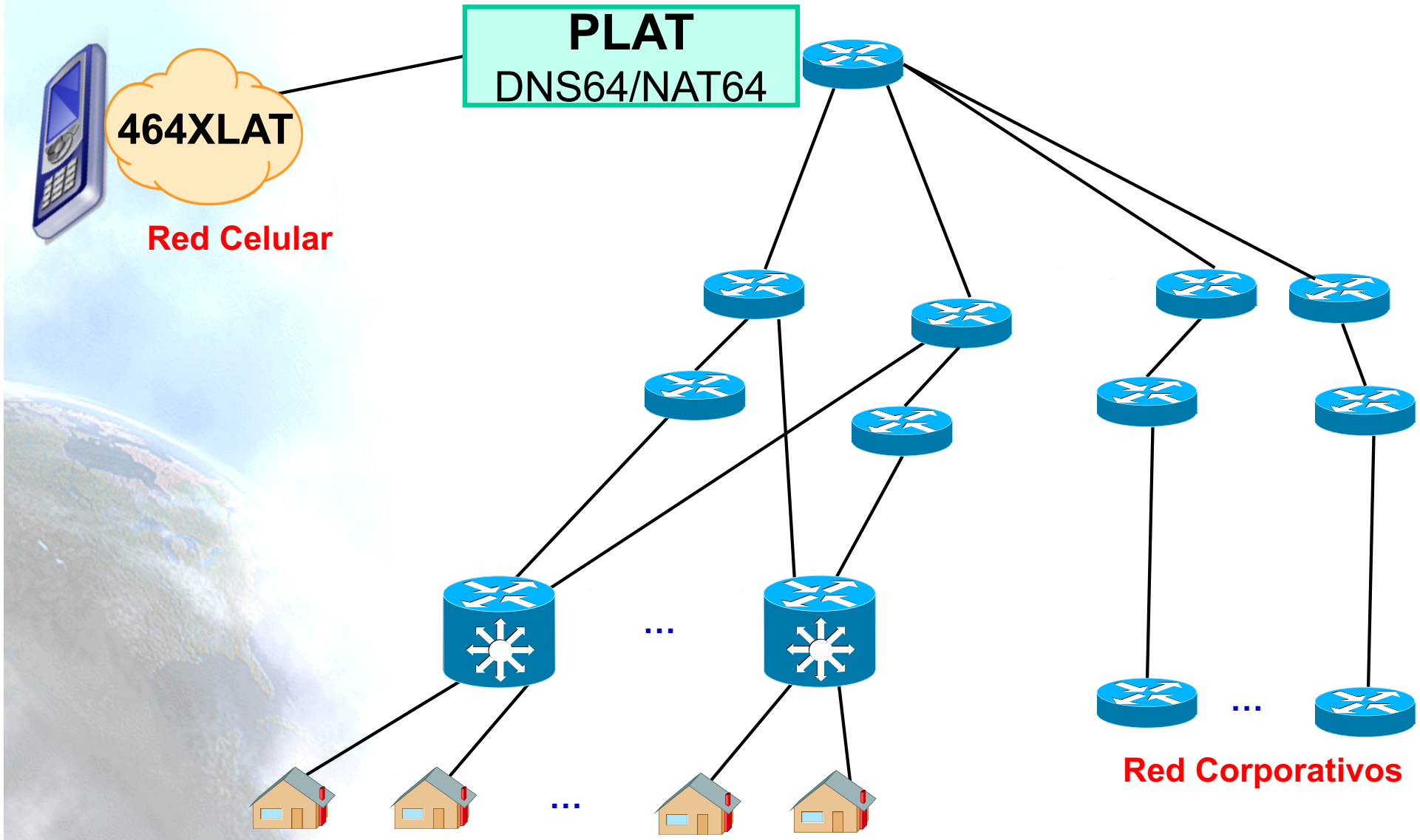


Max 15 Connections

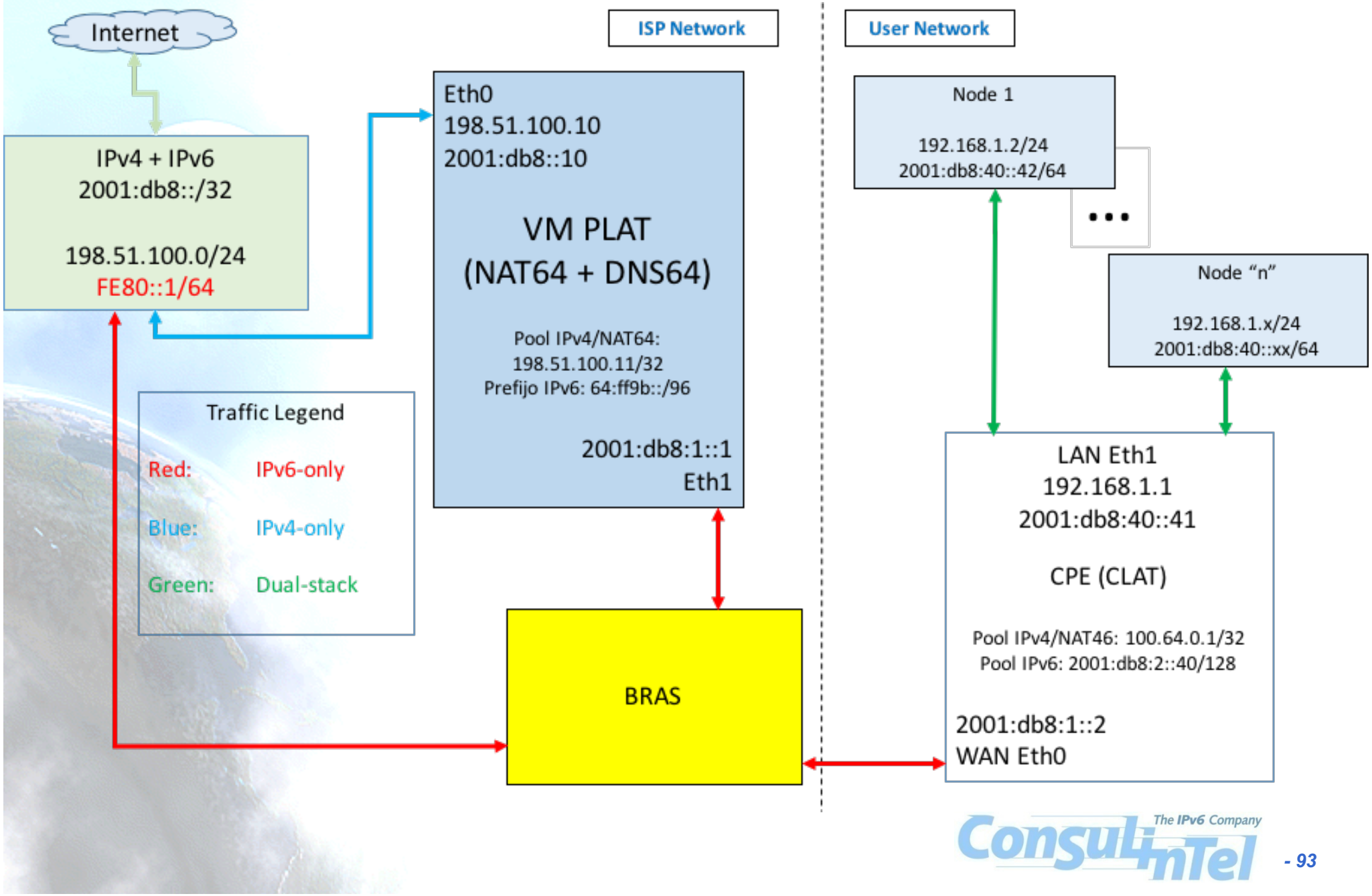


- **Posiblemente un mínimo de 300 por usuario detrás del CPE**
  - **Más si se incrementa el uso de tecnologías como AJAX**
  - **Multiplicar por número medio de usuarios detrás de NAT**
- **Implicaciones de IP/port sharing …**

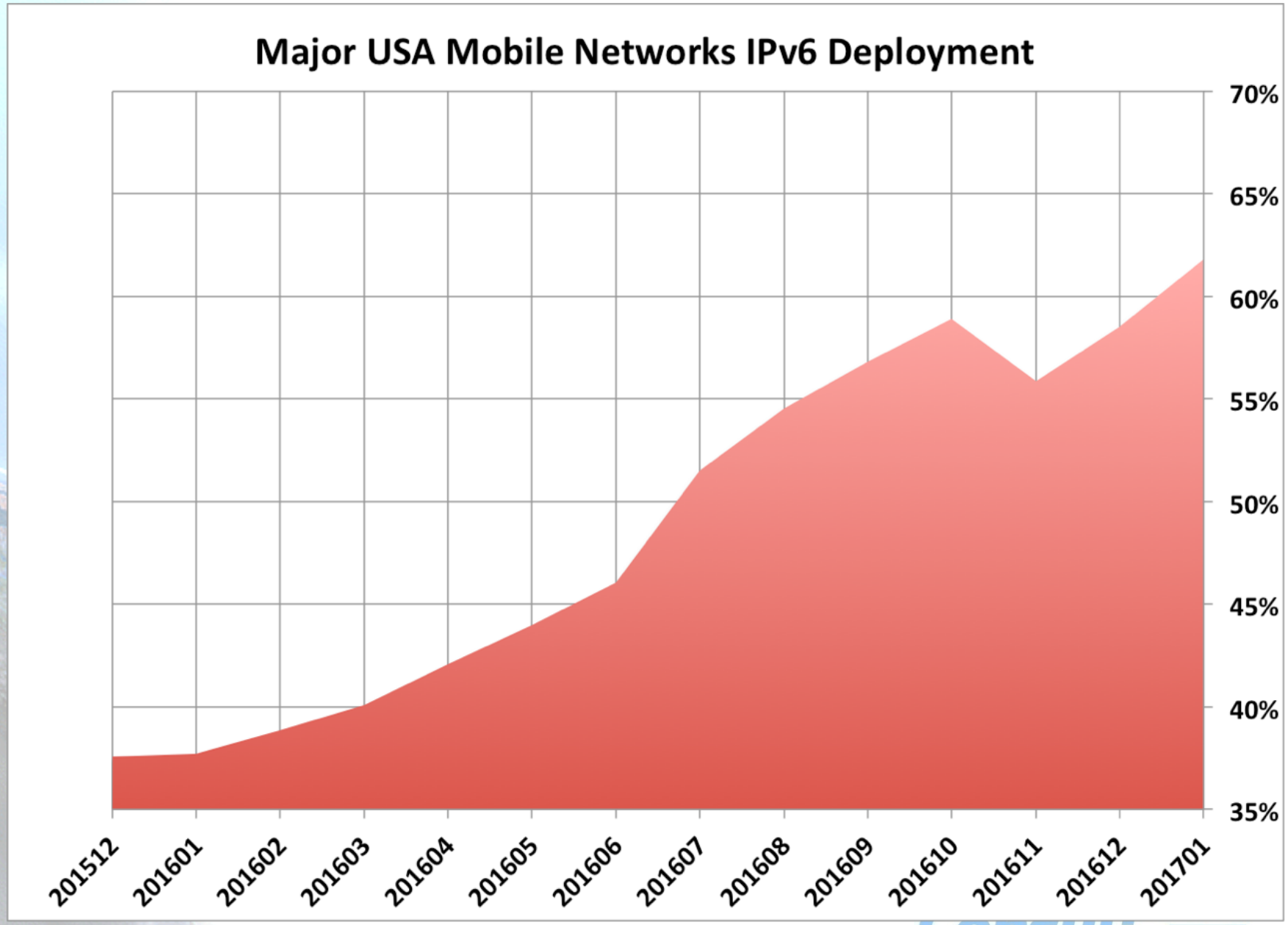
Ejemplo Red Multiservicio



Example Residential Customer



IPv6 in Cellular/US



*ISOC/World IPv6 Launch data

Update of RFC7084

- Basic Requirements for IPv6 Customer Edge Routers
  - Originally include support only for 6RD and DS-LITE
  - Being updated to include support for 464XLAT, MAP T/E, lw4o6, …
- https://tools.ietf.org/html/draft-ietf-v6ops-rfc7084-bis

Gracias !!

Contacto:

- Jordi Palet (Consulintel): jordi.palet@consulintel.es