IPv6 (Internet Protocol version 6) (parte II)

2020

Facultad de Informática
UNIVERSIDAD NACIONAL DE LA PLATA

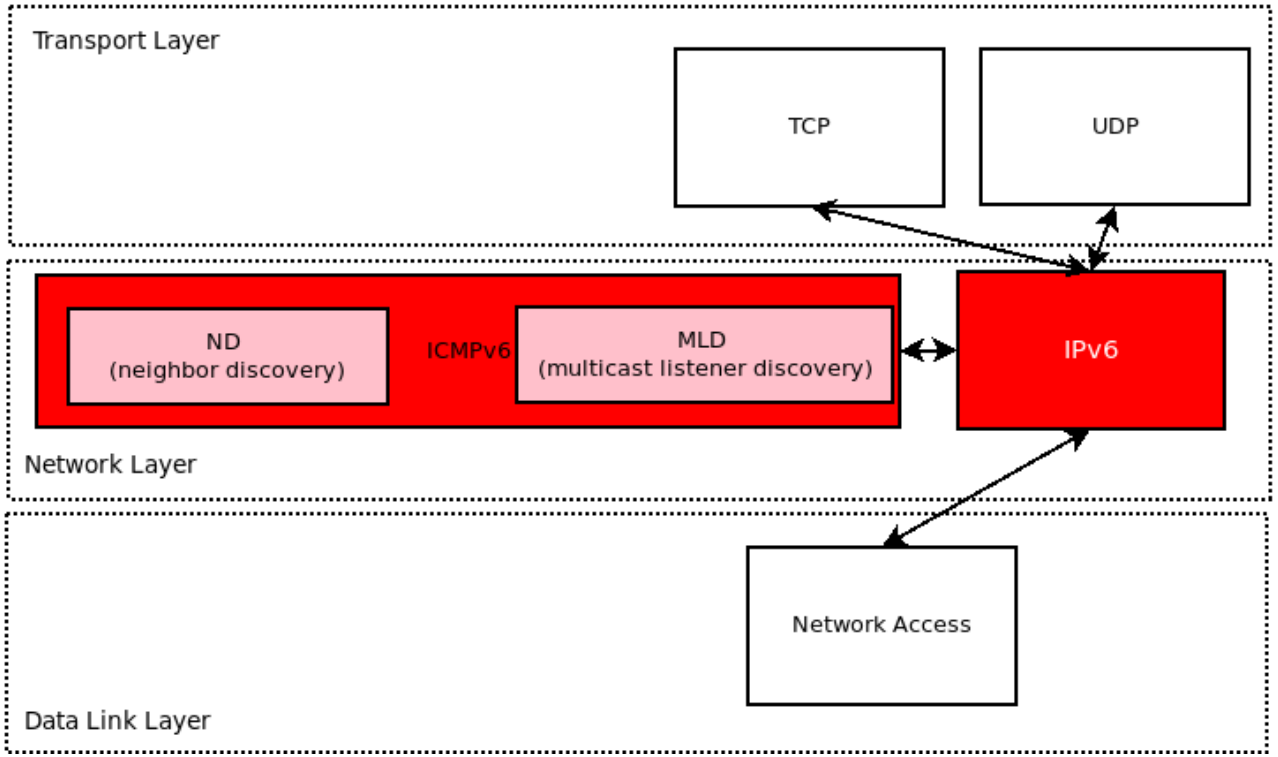
Contenidos

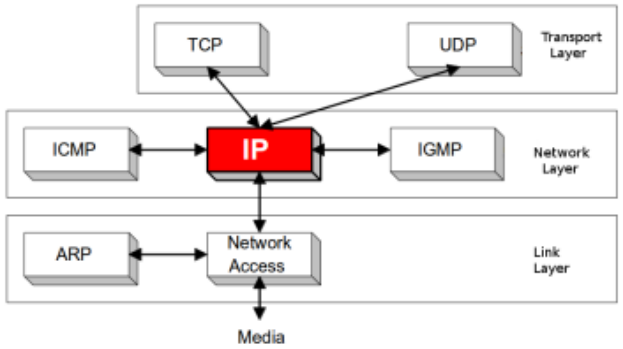
Contenidos

Contenidos

IPv6 Stack

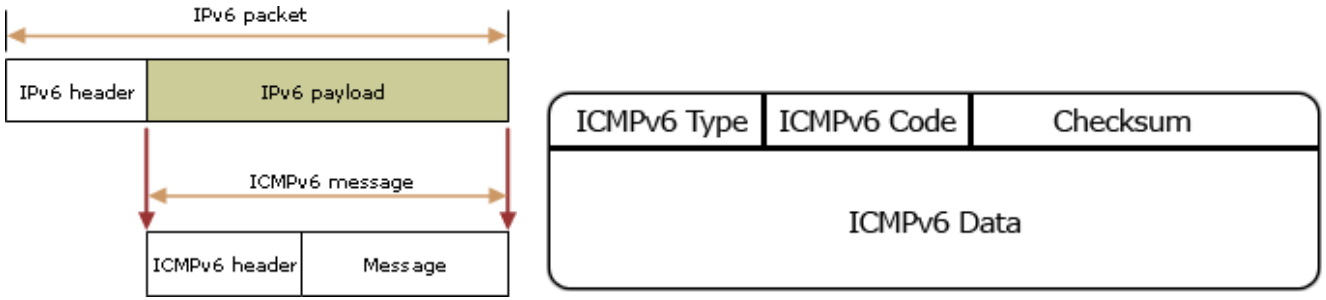
- Cambia el plano de control según IPv4





Internet Control Message Protocol v6 (ICMPv6)

- ICMPv6 definido en RFC-4443.
- Parte fundamental del stack IPv6.
- Resuelve:
  - Multicast Listener Discovery (MLD), reemplazo de IGMP.
  - Neighbor Discovery Protocol (NDP), reemplazo de ARP y mensajes Router Discovery(Auto-configuración), Redirect.
  - Mensajes de control de ICMP: informativos (ping), errores.



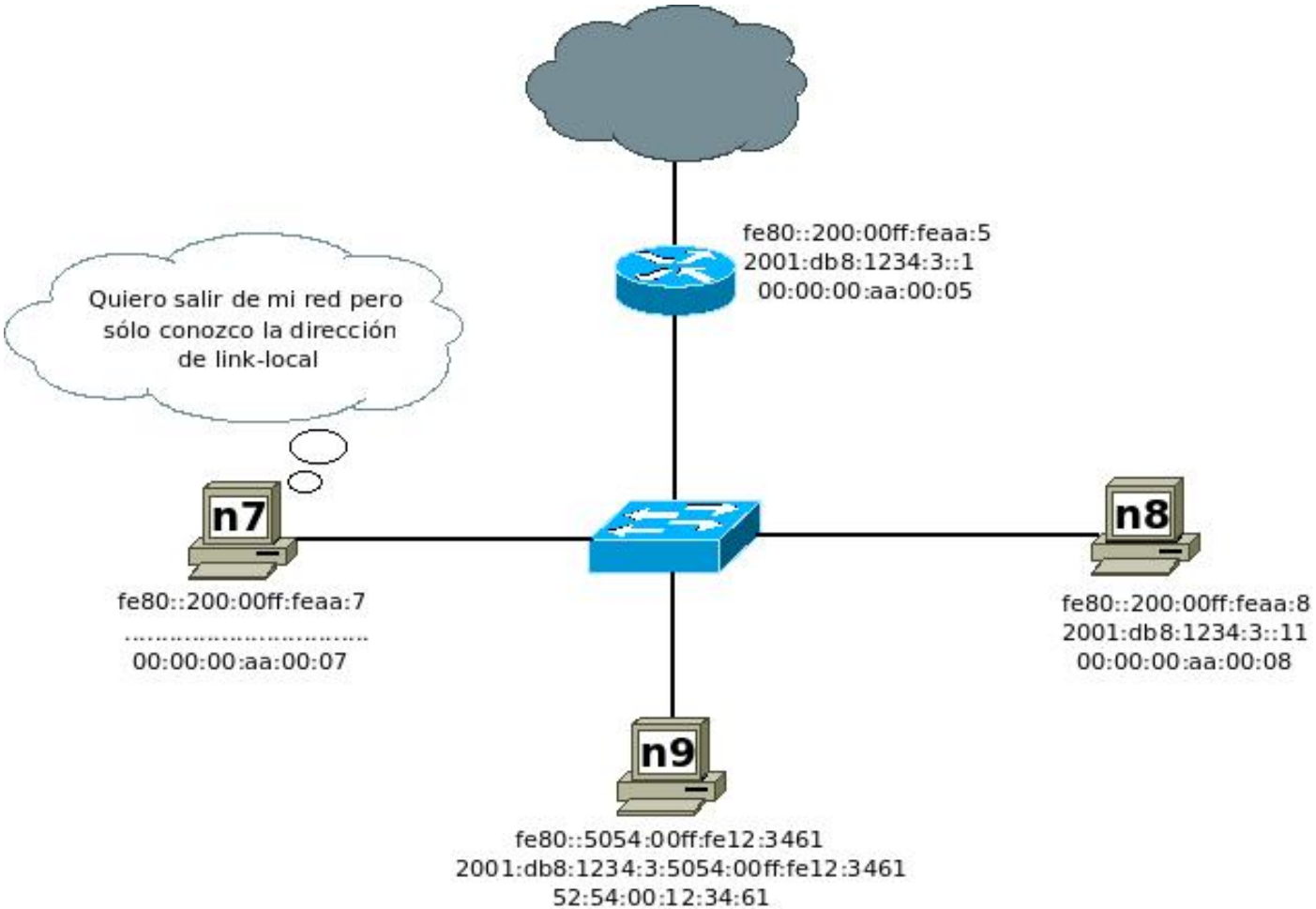
fuente: http://www.microsoft.com, http://www.ipv6.com

IPv6 Stateless Autoconfiguration

- Parte del NDP.
- Reemplaza la configuración manual de direcciones IPv6 en IPv4.
- Alternativa básica a DHCPv6, pero sin estados, **SLAAC**.
- El router anuncia uno ó más prefijos de red mediante mensajes Router Advertisement RA (134).
- Se pueden solicitar bajo demanda Router Solicitation RS (133).
- Los hosts auto-configuran su dirección de link-local y solicitan el prefijo a algún router de la red.
- Una vez obtenido se auto-configuran generando su propias direcciones, previo realizar DAD (Duplicate Address Detection).
- Determinan y configuran el default gateway a partir de los Router Advertisement recibidos.
- Router Advertisement puede llevar opciones de configuración del DNS, RFC-6106.

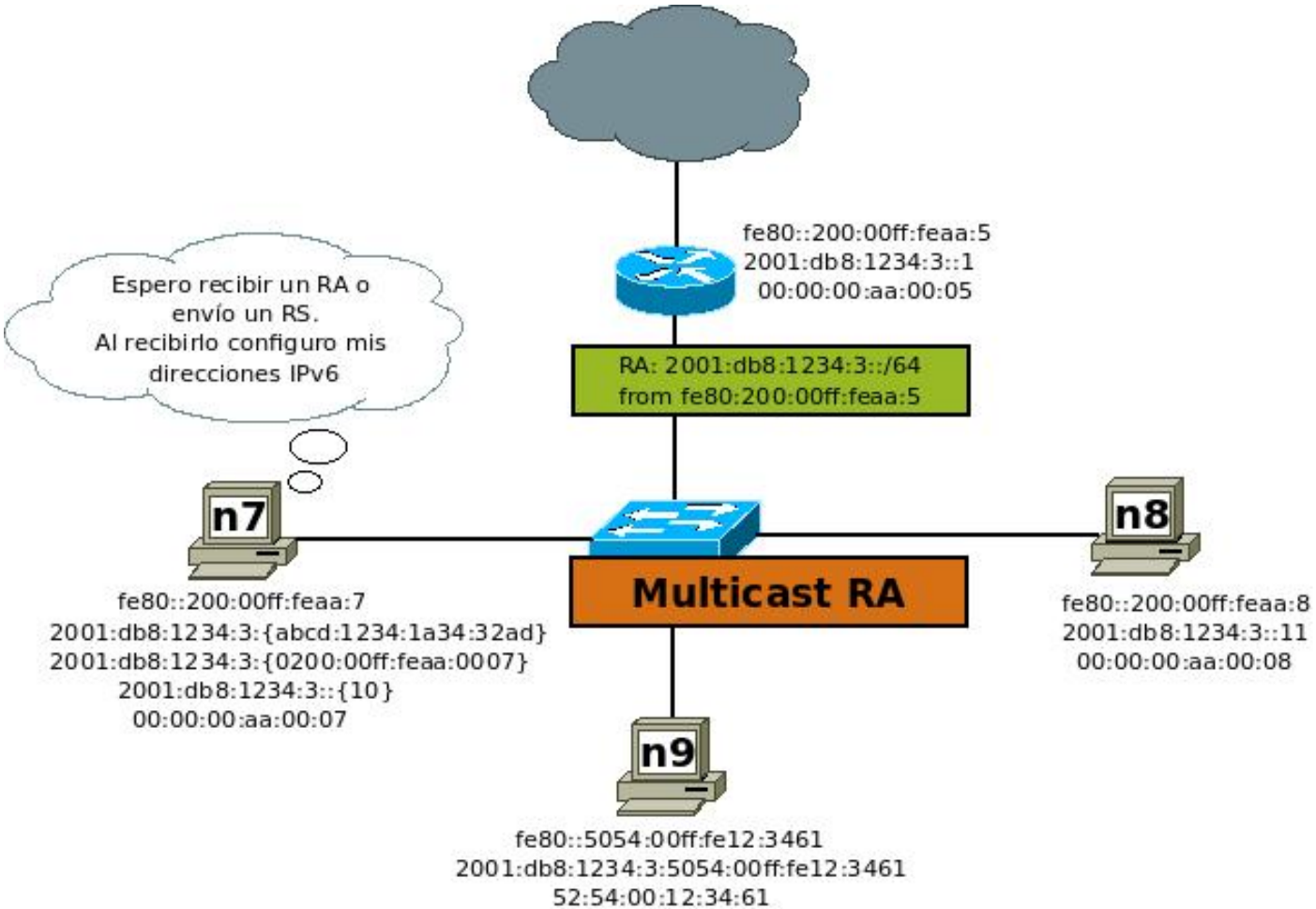
Router Discovery

- Aprendizaje de su propia configuración



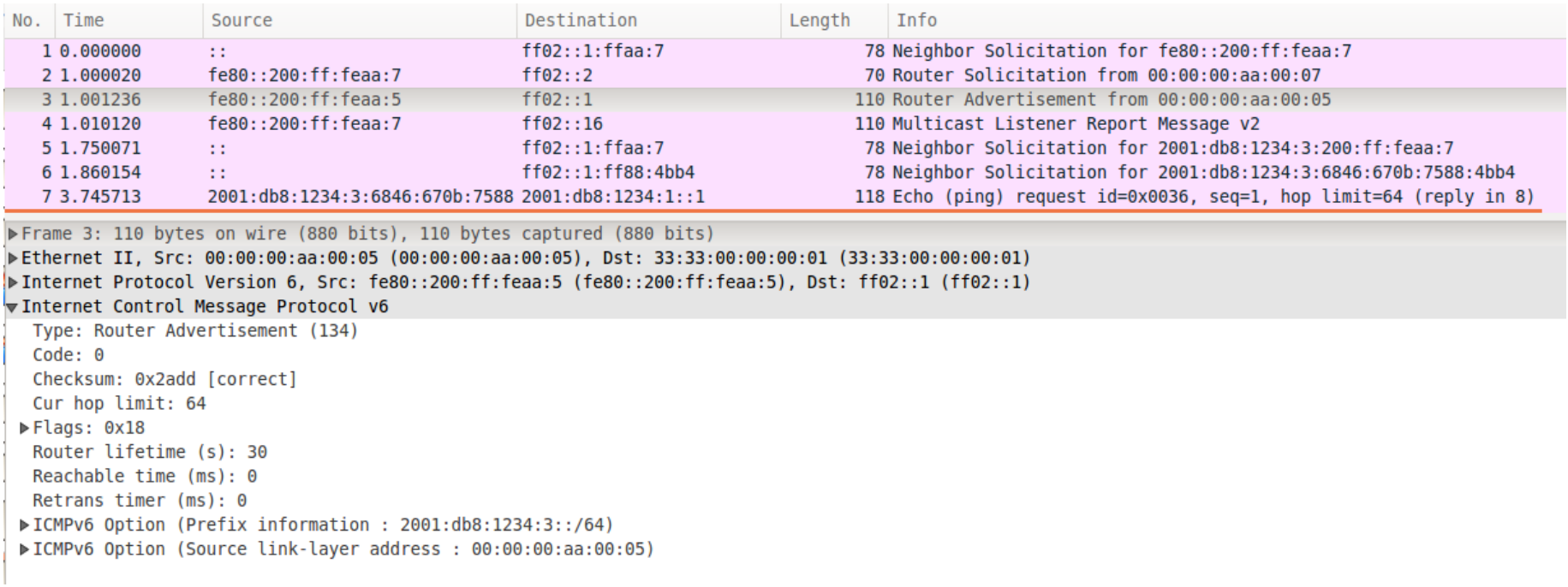
Router Discovery

- Aprendizaje de su propia configuración



Mensajes RS y RA

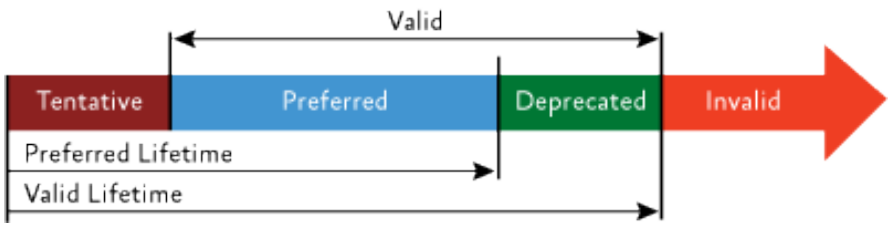
```
root@n7:/# ifconfig eth0 down;sleep 1;ifconfig eth0 up
...
root@n7:/# sleep 4;ping6 2001:db8:1234:1::1
PING 2001:db8:1234:1::1(2001:db8:1234:1::1) 56 data bytes
64 bytes from 2001:db8:1234:1::1: icmp_seq=1 ttl=63 time=0.251 ms
64 bytes from 2001:db8:1234:1::1: icmp_seq=2 ttl=63 time=0.265 ms
...
```



captures/ipv6-rs-ra-dad-icmp–i.pcap, rows: 3 captures/ipv6-rs-ra-dns.pcap, row: 3

Detalles mensaje RA - RFC 4861

- Flags: L(on-link), A(Autonomous), R(Router).
  - L, indica que esta asignado a una interfaz (si no esta no se asume off-link).
  - A, sirve para autoconf. global.
  - R, indica que es router, sirve para NUD (Neighbor Unreach. Detection).
- Otros parámetros: Tiempo de vida válido y preferido (valid and preferred lifetime).



fuente: http://www.microsoft.com

IPv6 Autoconfiguration DHCP6

- Configuración manual es posible (routers, servers).
- Configuración automática, hay variantes: SLAAC, DHCP6.
- Combinaciones:
  - SLAAC solo, hoy puede obtener conf. básica para Internet, Prefijo, Router y DNS (otros MTU).
  - SLAAC solo, algunos equipos no soportan la opción de DNS, requieren DHCP6.
  - SLAAC + DHCP6, conf. básica más parámetros extras por DHCP6.
  - DHCP6 solo, requiere RA para Router/Gateway de la red.

IPv6 Autoconfiguration DHCP6 (cont.)

- Flags en RA:
  - O bit - Other Configuration Flag, RFC4861, indica que puede usar DHCP6 para obtener otros parámetros, por ejemplo DNS info.
  - M bit - Managed Address Configuration Flag, RFC4861, indica que usa DHCP6.
- Combinaciones:
  - $O = 0, M = 0$, conf. vía SLAAC, stateless; si hay DHCP6 no lo usaría.
  - $O = 1, M = 0$, conf. vía SLAAC, stateless; por DHCP6 obtiene parámetros adicionales. *AdvOtherConfigFlag*.
  - $O = *, M = 1$, conf. vía DHCP6 stateful; salvo router. *AdvManagedFlag*.

Datagrama DHCP6 (Wireshark)

No.	Time	Source	Destination	Length	Info
24	47.468477	fe80::200:ff:feaa:5	ff02::1	110	Router Advertisement from 00:00:00:aa:00:05
25	52.675902	fe80::200:ff:feaa:5	ff02::5	90	Hello Packet
26	52.720484	fe80::200:ff:feaa:7	ff02::1:2	118	Solicit XID: 0x28f60b CID: 000100011e44f6e6000000aa0007
27	52.722313	fe80::200:ff:feaa:5	ff02::1:ffaa:7	86	Neighbor Solicitation for fe80::200:ff:feaa:7 from 00:00
28	52.722390	fe80::200:ff:feaa:7	fe80::200:ff:feaa:5	86	Neighbor Advertisement fe80::200:ff:feaa:7 (sol, ovr) is
29	52.722461	fe80::200:ff:feaa:5	fe80::200:ff:feaa:7	204	Advertise XID: 0x28f60b IAA: 2001:db8:1234:3::1ed6 CID:
30	52.724592	fe80::200:ff:feaa:7	ff02::1:2	164	Request XID: 0x1afc22 CID: 000100011e44f6e6000000aa0007
31	52.725194	fe80::200:ff:feaa:5	fe80::200:ff:feaa:7	204	Reply XID: 0x1afc22 IAA: 2001:db8:1234:3::1ed6 CID: 0001

```
▶ Frame 24: 110 bytes on wire (880 bits), 110 bytes captured (880 bits)
▶ Ethernet II, Src: 00:00:00:aa:00:05 (00:00:00:aa:00:05), Dst: 33:33:00:00:00:01 (33:33:00:00:00:01)
▶ Internet Protocol Version 6, Src: fe80::200:ff:feaa:5 (fe80::200:ff:feaa:5), Dst: ff02::1 (ff02::1)
▼ Internet Control Message Protocol v6
    Type: Router Advertisement (134)
    Code: 0
    Checksum: 0x2a9d [correct]
    Cur hop limit: 64
  ▼ Flags: 0x58
      0... .... = Managed address configuration: Not set
      .1.. .... = Other configuration: Set
      ..0. .... = Home Agent: Not set
      ...1 1... = Prf (Default Router Preference): Low (3)
      .... .0.. = Proxy: Not set
      .... ..0. = Reserved: 0
    Router lifetime (s): 30
    Reachable time (ms): 0
```

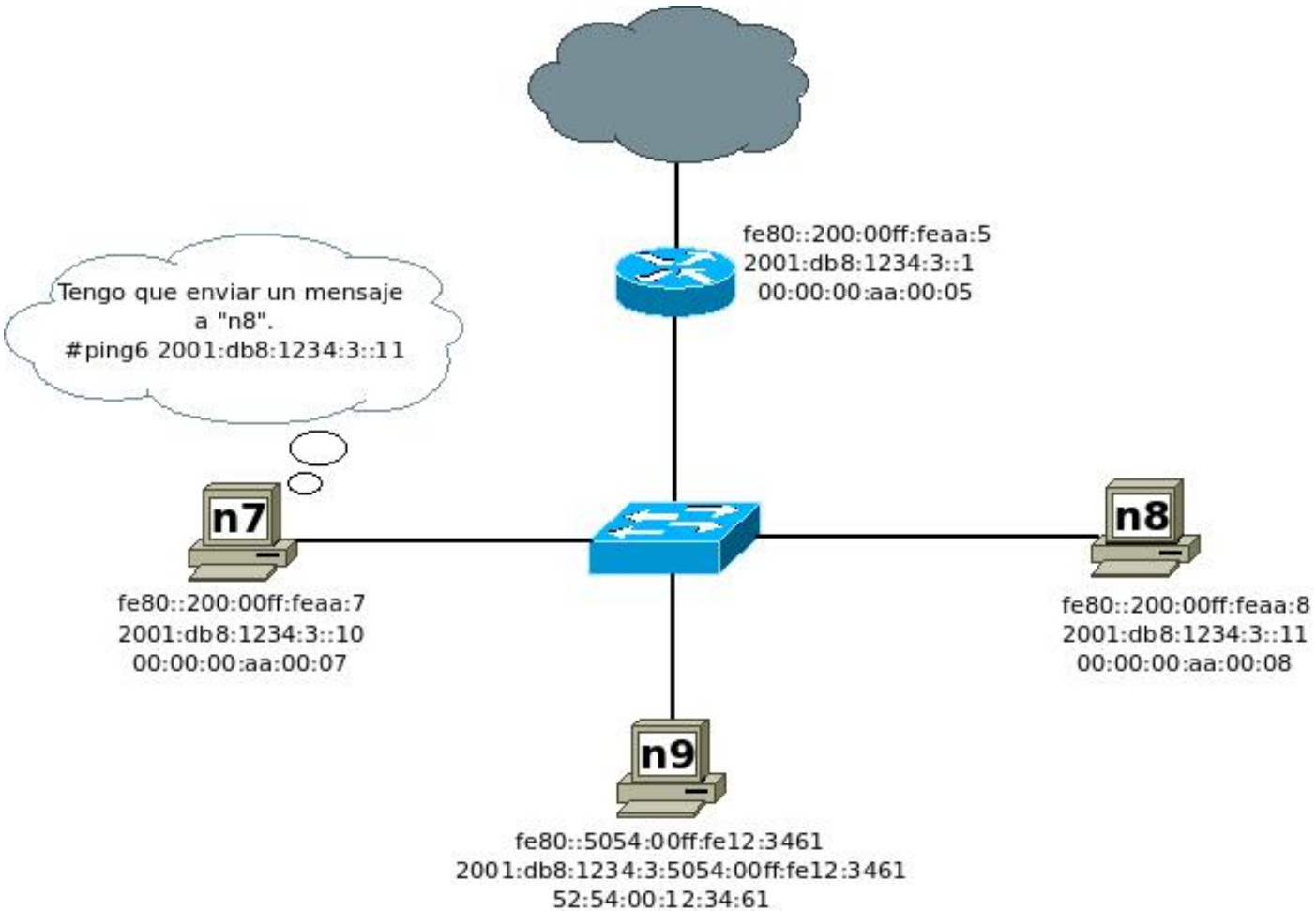
captures/ipv6-radvd-dhcp6.pcap, rows: 24,26,29,...

Neighbor Discovery

- Reemplaza básicamente al protocolo ARP de IPv4.
- Mapea direcciones lógicas (IPv6) a direcciones de Hardware (MAC, EUI-48, EUI-64).
- Trabaja conjuntamente con Ethernet (u otros protocolos de L2 multiacceso con broadcast: Bluetooth, 802.11, (Token-Ring, FDDI).
- Trabaja de forma dinámica, auto-aprendizaje, sin configuración.
- Puede configurarse de forma estática.
- Definido en RFC-4861.
- 2 tipos de mensajes: Neighbor Solicitation NS(135) y Neighbor Adv. NA(136).

Neighbor Discovery (cont.)

- Aprendizaje de direcciones:



Neighbor Discovery - NS y NA

- “n7” debe recurrir a un Neighbor Solicitation (NS).
- Como no sabe la MAC debe enviar un multicast L2 y L3.

```
root@n7:/# ip -6 addr show dev eth0
27: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qlen 1000
    ...
    inet6 2001:db8:1234:3::10/64 scope global
       valid_lft forever preferred_lft forever
    inet6 fe80::200:ff:feaa:7/64 scope link
       valid_lft forever preferred_lft forever

root@n7:/# ip -6 neigh show
fe80::200:ff:feaa:5 dev eth0 lladdr 00:00:00:aa:00:05 router STALE

root@n7:/# ping6 -c 5 2001:db8:1234:3::11 -I 2001:db8:1234:3::10
```

Neighbor Discovery NS

- Mensajes NS (Neighbor Solicitation):

No.	Time	Source	Destination	Length	Info
1	0.000000	2001:db8:1234:3::10	ff02::1:ff00:11	86	Neighbor Solicitation for 2001:db8:1234:3::11 from 00:00:00:aa:00:07
2	0.000151	2001:db8:1234:3::11	2001:db8:1234:3::10	86	Neighbor Advertisement 2001:db8:1234:3::11 (sol, ovr) is at 00:00:00:a
3	0.000161	2001:db8:1234:3::10	2001:db8:1234:3::11	118	Echo (ping) request id=0x004e, seq=1, hop limit=64 (reply in 4)
4	0.000186	2001:db8:1234:3::11	2001:db8:1234:3::10	118	Echo (ping) reply id=0x004e, seq=1, hop limit=64 (request in 3)

```
▶ Frame 1: 86 bytes on wire (688 bits), 86 bytes captured (688 bits)
▶ Ethernet II, Src: 00:00:00:aa:00:07 (00:00:00:aa:00:07), Dst: 33:33:ff:00:00:11 (33:33:ff:00:00:11)
▼ Internet Protocol Version 6, Src: 2001:db8:1234:3::10 (2001:db8:1234:3::10), Dst: ff02::1:ff00:11 (ff02::1:ff00:11)
  ▶ 0110 .... = Version: 6
  ▶ .... 0000 0000 .... .... .... .... .... = Traffic class: 0x00000000
    .... .... .... 0000 0000 0000 0000 0000 = Flowlabel: 0x00000000
    Payload length: 32
    Next header: ICMPv6 (58)
    Hop limit: 255
    Source: 2001:db8:1234:3::10 (2001:db8:1234:3::10)
    Destination: ff02::1:ff00:11 (ff02::1:ff00:11)
▼ Internet Control Message Protocol v6
    Type: Neighbor Solicitation (135)
    Code: 0
    Checksum: 0xf8db [correct]
    Reserved: 00000000
    Target Address: 2001:db8:1234:3::11 (2001:db8:1234:3::11)
  ▼ ICMPv6 Option (Source link-layer address : 00:00:00:aa:00:07)
      Type: Source link-layer address (1)
      Length: 1 (8 bytes)
      Link-layer address: 00:00:00:aa:00:07 (00:00:00:aa:00:07)
```

captures/ipv6-ns-na-icmp.pcap, rows: 1

Neighbor Discovery NS, NA Addresses

```
** Request: Neighbor Solicitation (NS)

From: MAC:<my-MAC-address>         , IPv6:<my-Link-Local-Address> / IPv6: <my-global-Address>
To:   MAC:<MAC-multicast-address> , IPv6:<solicited-node_multicast_address>

<solicited-node-multicast-address>(SNMA): ff02:0000:0000:0000:0000:0001:ff00:0000/104
                                          + last 24bits IPv6
<MAC-multicast-address>  (MMA)          :  33:33: + last 32bits IPv6 mcast address

n7: mac address:                 00:00:00:aa:00:07
    ipv6 address:                2001:db8:1234:3::10/64
    ipv6 link-local address:     fe80::0200:00ff:fe00:0007
"n7" Try to discover:             2001:db8:1234:3::11/64 MAC address ("n8")

SNMA: ff02:0000:0000:0000:0000:0001:ff00:0000/104 + 00:00:11
      = ff02:0000:0000:0000:0000:0001:ff00:0011
        https://tools.ietf.org/html/rfc4291
MMA:  33:33:                  + ff:00:00:01 =  33:33:ff:00:00:11
                                https://tools.ietf.org/html/rfc2464#section-7
                                https://tools.ietf.org/html/rfc6085
                                https://tools.ietf.org/html/rfc7042
From: MAC: 00:00:00:aa:00:07 , IPv6: fe80::0200:00ff:fe00:0007 / 2001:db8:1234:3::10
To:   MAC: 33:33:ff:00:00:11   IPv6: ff02:0000:0000:0000:0000:0001:ff00:0011

** Reply: Neighbor Advertisement (NA)

From: MAC:<my-MAC-address>             , IPv6:<my-IPv6-Requested-Address>
To:   MAC:<MAC-who-request-address>    , IPv6:<IPv6-who-request-address>

From: MAC: 00:00:00:aa:00:08 , IPv6: fe80::0200:00ff:fe00:0008 / 2001:db8:1234:3::11
To:   MAC: 00:00:00:aa:00:07   IPv6: fe80::0200:00ff:fe00:0007 / 2001:db8:1234:3::10
```

Neighbor Discovery NA

- “n8” procesa el requerimiento y responde con un Neighbor Advertisement (NA):

No.	Time	Source	Destination	Length	Info
1	0.000000	2001:db8:1234:3::10	ff02::1:ff00:11	86	Neighbor Solicitation for 2001:db8:1234:3::11 from 00:00:00:aa:00:07
2	0.000151	2001:db8:1234:3::11	2001:db8:1234:3::10	86	Neighbor Advertisement 2001:db8:1234:3::11 (sol, ovr) is at 00:00:00:a
3	0.000161	2001:db8:1234:3::10	2001:db8:1234:3::11	118	Echo (ping) request id=0x004e, seq=1, hop limit=64 (reply in 4)
4	0.000186	2001:db8:1234:3::11	2001:db8:1234:3::10	118	Echo (ping) reply id=0x004e, seq=1, hop limit=64 (request in 3)

```
▶ Frame 2: 86 bytes on wire (688 bits), 86 bytes captured (688 bits)
▶ Ethernet II, Src: 00:00:00:aa:00:08 (00:00:00:aa:00:08), Dst: 00:00:00:aa:00:07 (00:00:00:aa:00:07)
▼ Internet Protocol Version 6, Src: 2001:db8:1234:3::11 (2001:db8:1234:3::11), Dst: 2001:db8:1234:3::10 (2001:db8:1234:3::10)
  ▶ 0110 .... = Version: 6
  ▶ .... 0000 0000 .... .... .... .... .... = Traffic class: 0x00000000
    .... .... .... 0000 0000 0000 0000 0000 = Flowlabel: 0x00000000
    Payload length: 32
    Next header: ICMPv6 (58)
    Hop limit: 255
    Source: 2001:db8:1234:3::11 (2001:db8:1234:3::11)
    Destination: 2001:db8:1234:3::10 (2001:db8:1234:3::10)
▼ Internet Control Message Protocol v6
    Type: Neighbor Advertisement (136)
    Code: 0
    Checksum: 0x54ef [correct]
  ▶ Flags: 0x60000000
    Target Address: 2001:db8:1234:3::11 (2001:db8:1234:3::11)
  ▼ ICMPv6 Option (Target link-layer address : 00:00:00:aa:00:08)
      Type: Target link-layer address (2)
      Length: 1 (8 bytes)
      Link-layer address: 00:00:00:aa:00:08 (00:00:00:aa:00:08)
```

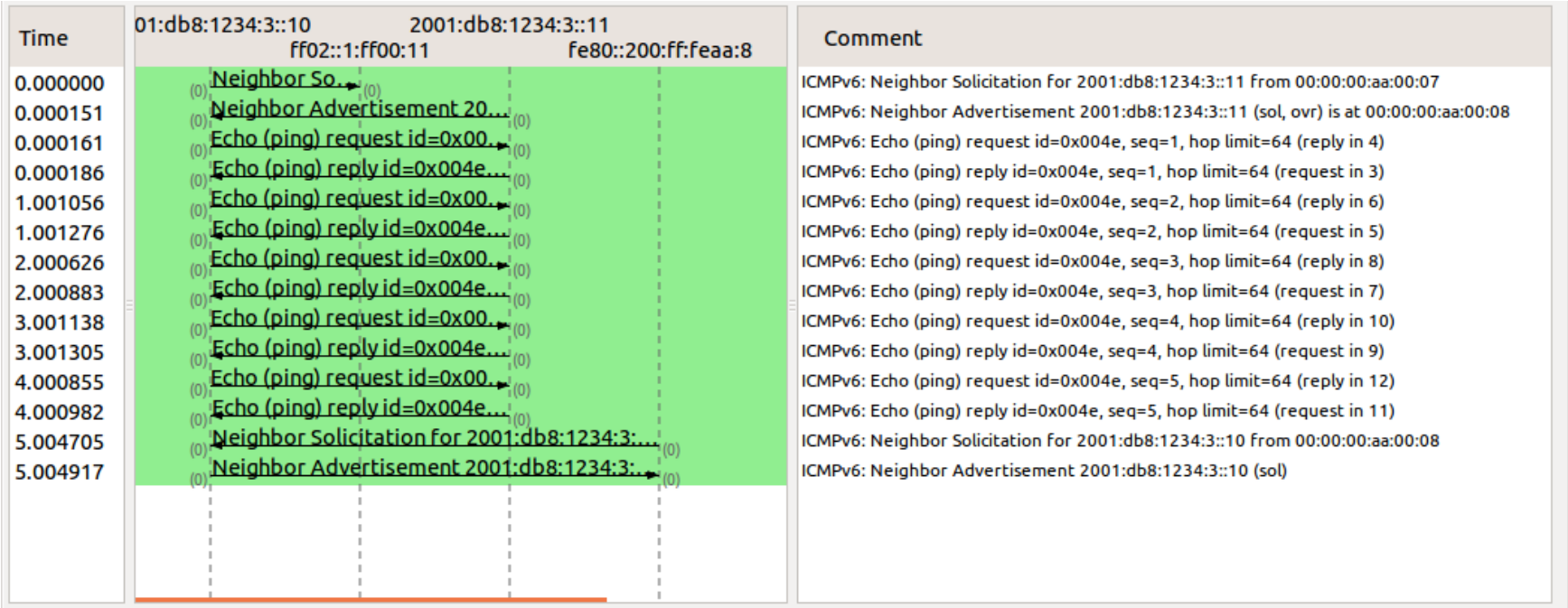
captures/ipv6-ns-na-icmp.pcap, rows: 2

Neighbor Discovery NS, NA y PING-PONG

- “n7” debe recurrir a un Neighbor Solicitation:

```
root@n7:/# ping6 -c 5 2001:db8:1234:3::11 -I 2001:db8:1234:3::10
PING 2001:db8:1234:3::11(2001:db8:1234:3::11) from 2001:db8:1234:3::10 : 56 data bytes
64 bytes from 2001:db8:1234:3::11: icmp_seq=1 ttl=64 time=0.222 ms
64 bytes from 2001:db8:1234:3::11: icmp_seq=2 ttl=64 time=0.273 ms
64 bytes from 2001:db8:1234:3::11: icmp_seq=3 ttl=64 time=0.344 ms
64 bytes from 2001:db8:1234:3::11: icmp_seq=4 ttl=64 time=0.219 ms
64 bytes from 2001:db8:1234:3::11: icmp_seq=5 ttl=64 time=0.181 ms
--- 2001:db8:1234:3::11 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 4000ms
rtt min/avg/max/mdev = 0.181/0.247/0.344/0.059 ms

root@n7:/# ip -6 neigh show
2001:db8:1234:3::11 dev eth0 lladdr 00:00:00:aa:00:08 REACHABLE
fe80::200:ff:feaa:8 dev eth0 lladdr 00:00:00:aa:00:08 REACHABLE
fe80::200:ff:feaa:5 dev eth0 lladdr 00:00:00:aa:00:05 router STALE
```

Neighbor Discovery Flow



captures/ipv6-ns-na-icmp.pcap

Neighbor Discovery NA (cont.)

- Mensajes NA (Neighbor Advertisement):
- Solicitado o no solicitado.
- Solicitado, respuesta a RS Flag: (S)Solicited=1.
- NO-solicitado, para actualizar caches Flag: (O)Override=1.
- Flag (R)Router=1, lo envía el router.

Neighbor Discovery NS (Router)

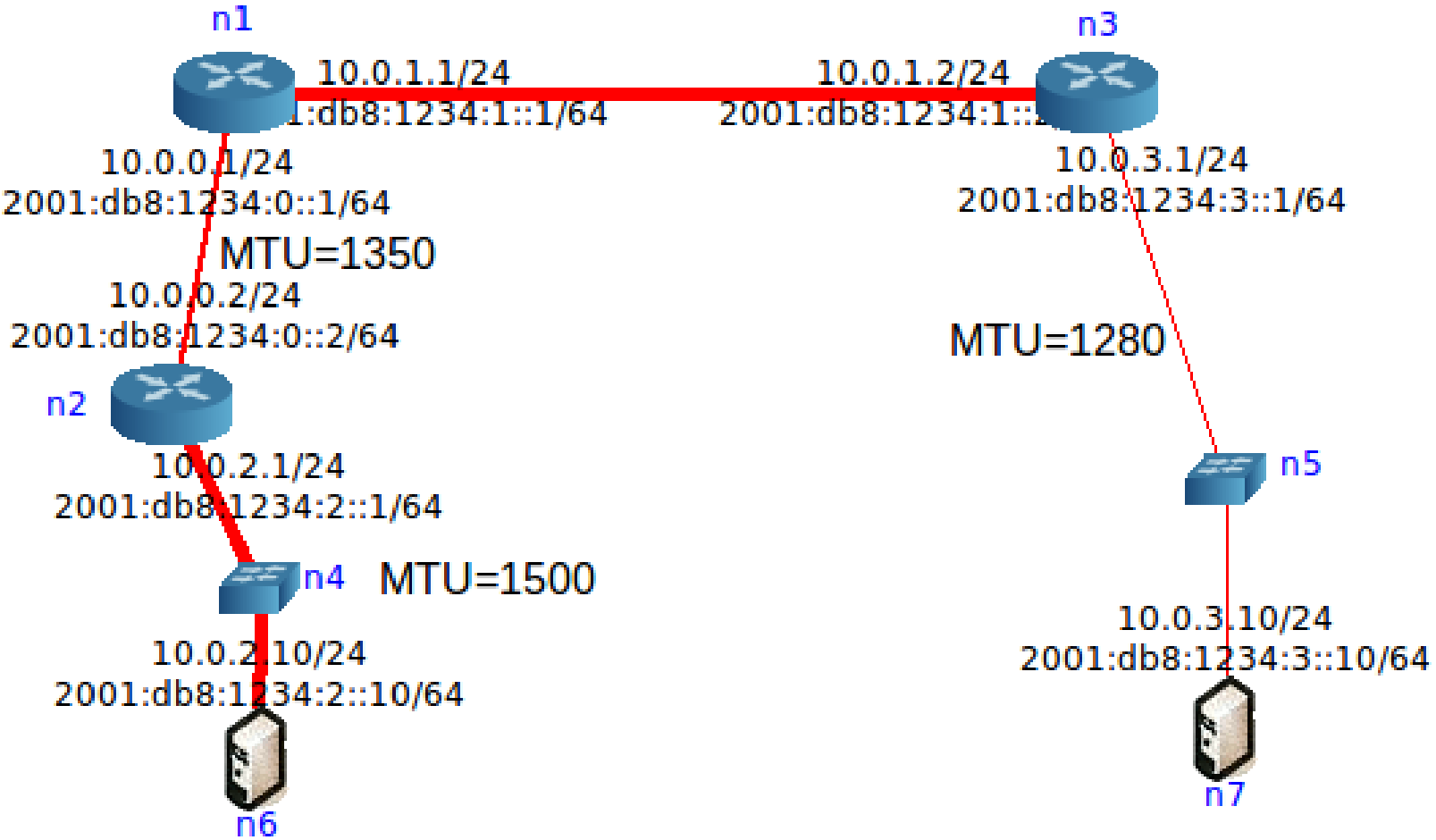
- “n7” debe recurrir a un Neighbor Solicitation, si no tiene router MAC/stale:

```
root@n7:/# ip -6 neigh show
fe80::200:ff:feaa:5 dev eth0 lladdr 00:00:00:aa:00:05 router STALE

root@n7:/# ping6 -c 5 2001:db8:1234:3::1 -I 2001:db8:1234:3::10
PING 2001:db8:1234:3::1(2001:db8:1234:3::1) from 2001:db8:1234:3::10 : 56 data bytes
64 bytes from 2001:db8:1234:3::1: icmp_seq=1 ttl=64 time=0.906 ms
64 bytes from 2001:db8:1234:3::1: icmp_seq=2 ttl=64 time=0.303 ms

root@n7:/# ip -6 neigh show
2001:db8:1234:3::1 dev eth0 lladdr 00:00:00:aa:00:05 router REACHABLE
fe80::200:ff:feaa:5 dev eth0 lladdr 00:00:00:aa:00:05 router STALE

captures/ipv6-ns-na-router-icmp.pcap, row:4
```

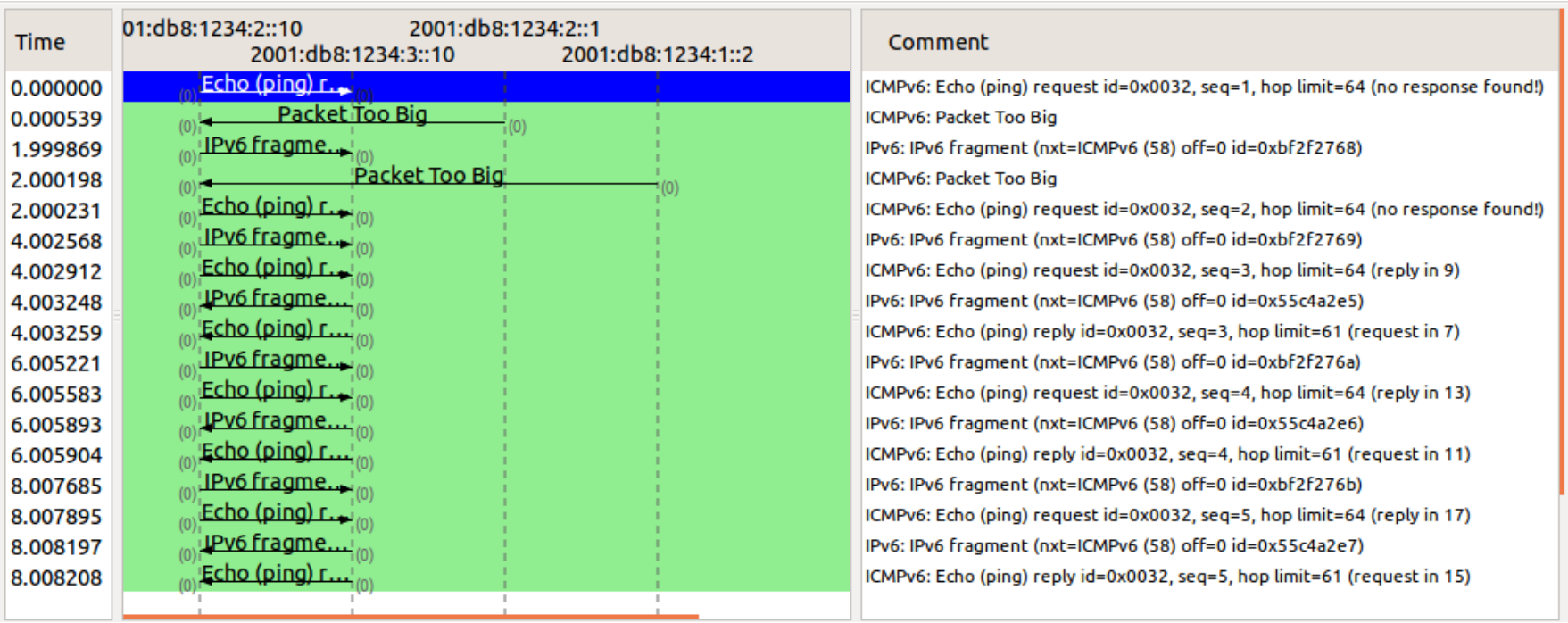
PMTU Discovery

- MTU (Maximum Transmission Unit) depende de L2.
- Cada Link L2 puede tener su MTU Link-MTU.
- A lo largo de un camino se puede establecer Path MTU (PMTU).
- Para IPv6 min MTU=1280B (RFC-2460, RFC-8200), Para IPv4 68B(RFC-791).
- Recomendado, (estandarizado) 1500B.
- Las implementaciones hacen el PMTU discovery (RFC-1981).
- Se utiliza ICMPv6 Error: “Message Too Big”.
- Una vez determinado el PMTU se fragementa de extremo a extremo.
- Se puede saltear y usar directamente 1280, no es óptimo.
- TCP trata de usar MSS (Maximum Segment Size).

PMTU Discovery (cont.)



PMTU Discovery (cont.)



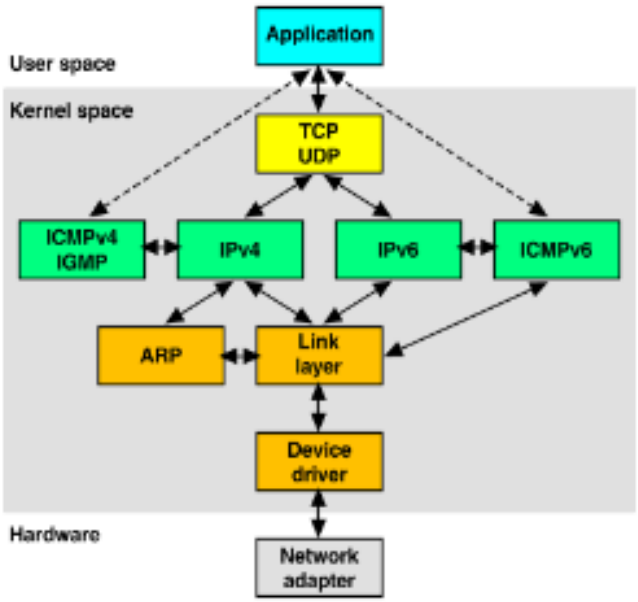
captures/ipv6-frag-doble-nocont–i.pcap

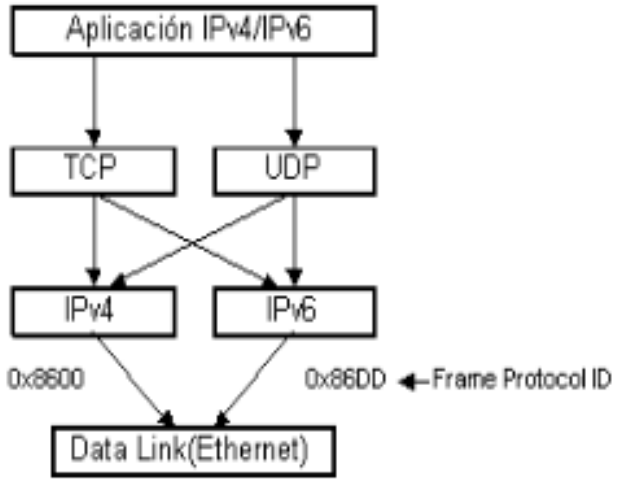
Transición/Coexistencia de IPv4 a IPv6

- No va a existir un día “D” para cambiar de IPv4 a IPv6.
- Amplio abanico de técnicas disponibles:
  - Doble pila (Dual-Stack).
  - Técnicas de tunneling.
  - Técnicas de traducción IPv6 $\leftarrow\rightarrow$ IPv4.
  - Técnicas combinadas.

Dual Stack IPv4/IPv6

- No se elimina la pila IPv4.
- Técnica multi-protocolo como con: NetBIOS, Appletalk, IPX.
- Actualmente, la mayoría de los OSs soportan IPv6.
- La aplicación, biblioteca de código elige cual usar (registros de DNS AAAA y A).
- RFC-6555, “Happy Eyeballs: Success with Dual-Stack Hosts”.
- Van a co-existir por “mucho” tiempo.



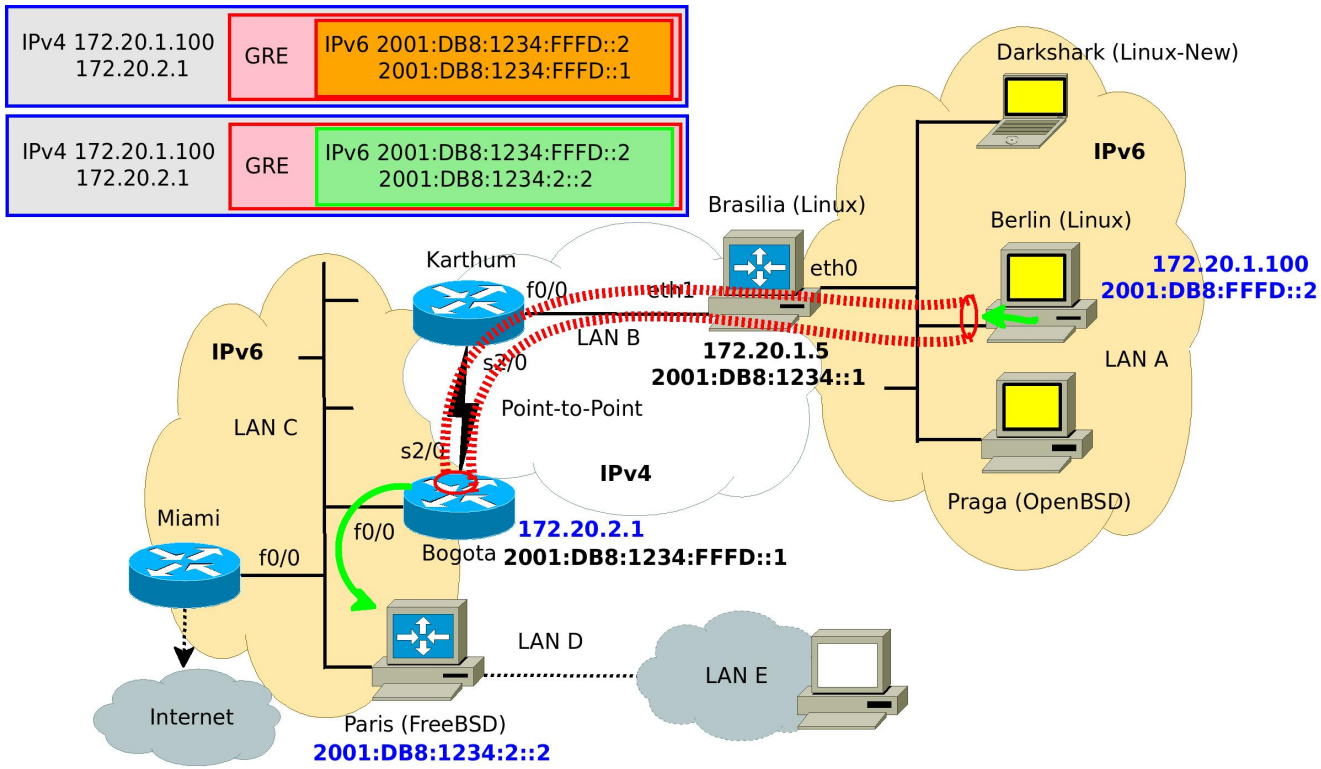


Túneles IPv4/IPv6

- Encapsular IPv6 en IPv4 donde no hay cobertura
- Se requiere doble pila IPv4/IPv6 en los routers
- Pueden ser:
- **Manuales:**
  - GRE (point-to-point).
  - SIT (Simple Internet Tunnel) IP-IP - IPv6-IPv4 (point-to-point).
  - 6in4.
- **Tunnels Brokers, interfaces web de ayuda:**
  - 6in4+TB (https://tunnelbroker.net/)
- **Automáticos:**
  - 6to4 (No funciona sobre NAT)
  - ISATAP (no funciona sobre NAT)
  - Teredo (sobre NAT)

Túneles IPv4/IPv6

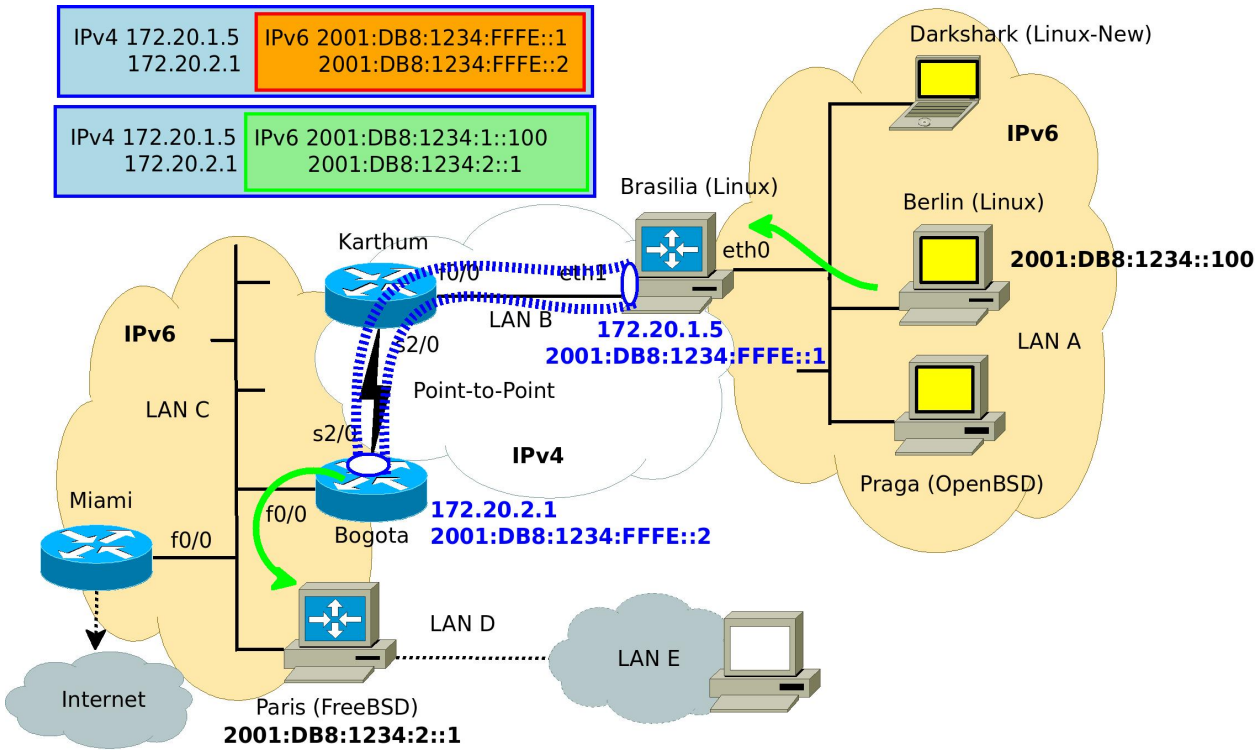
- Manuales (en caso intranet con GRE):
  - Encapsula en GRE (47) General Routing Encapsulation.
  - Punto a punto entre host y router o entre routers.



captures/113-ipv6-tun-gre-Berlin-e0.pcap. row:1

Túneles IPv4/IPv6 (cont.)

- Manuales (en caso intranet con SIT):
  - Encapsula directamente en IPv4, proto 41(IPv6).
  - Punto a punto entre host y router o entre routers.



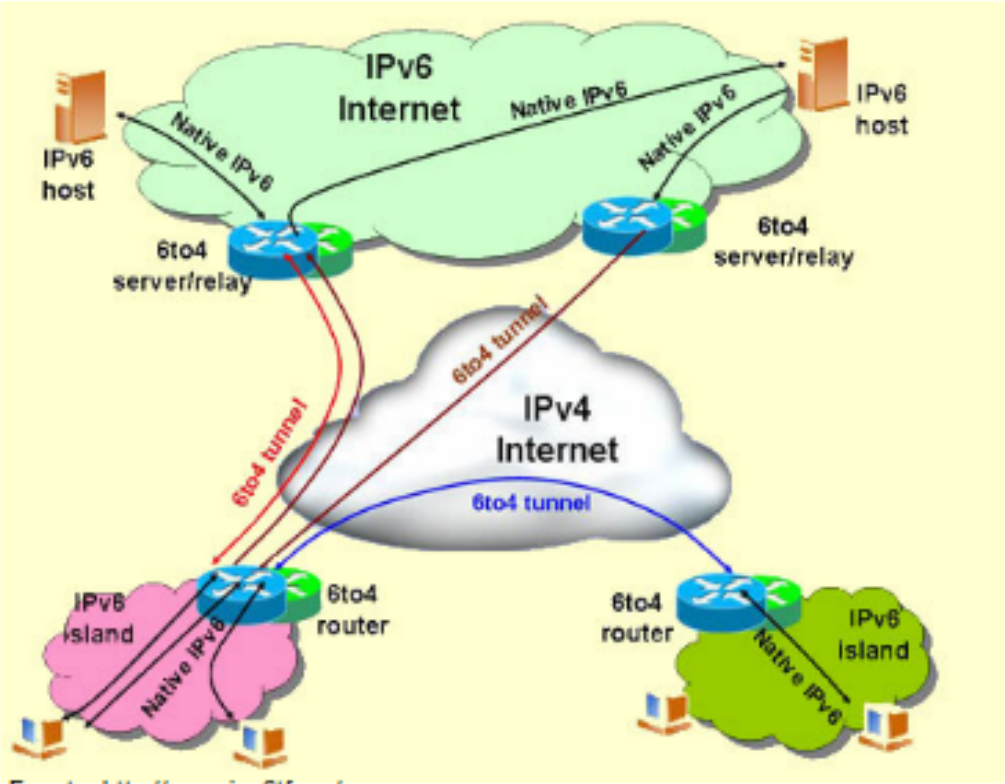
captures/110-ipv6-tun-sit.pcap, row:4 captures/111-ipv6-tun-sit-Bogota-S00.pcap, row:8

Túneles IPv4/IPv6 (cont.)

- sit/6in4 con TB (Tunnel Broker):
  - Encapsula en IPv4 como SIT.
  - Se consideran Punto a Punto.
  - Direcciones de ambos extremos en el mismo bloque, por ejemplo /126.
  - Para funcionar con NAT requiere que no se filtre proto 41 (IPv6 en IPv4).
  - Agregan un site web que determina la conf. (TB).
  - El usuario entra al site y obtiene la conf.

Túneles IPv4/IPv6 (cont.)

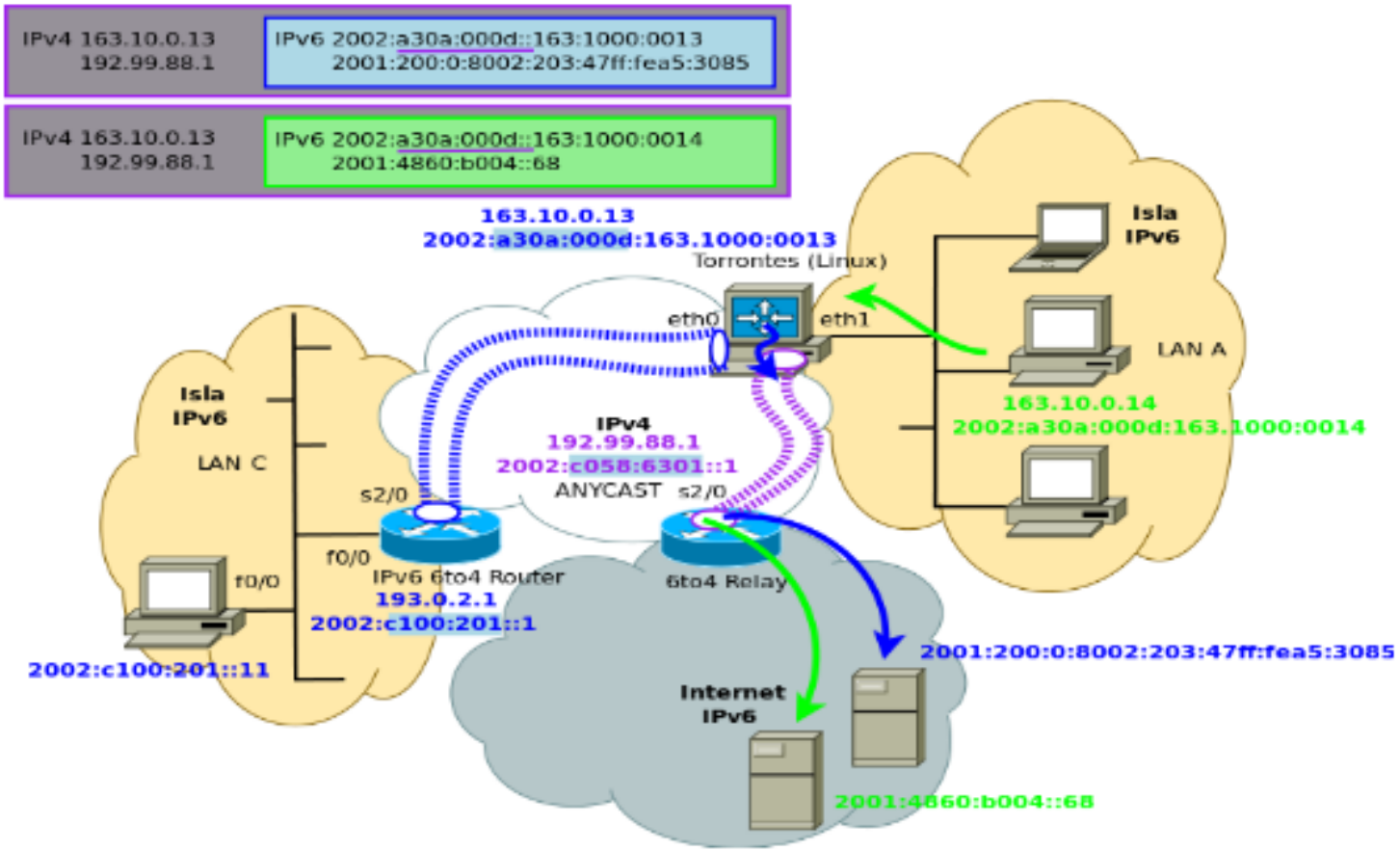
- 6to4: sobre direcciones IPv4 globales:
  - Se usa el bloque: 2002::/16
  - Paquetes de salida siempre al 6to4 relay.
  - Paquetes de vuelta, pueden usar otro origen.
  - Prefijo 6to4 relays anycast: 192.88.99.1/24



Fuente: http://www.ipv6tf.org/

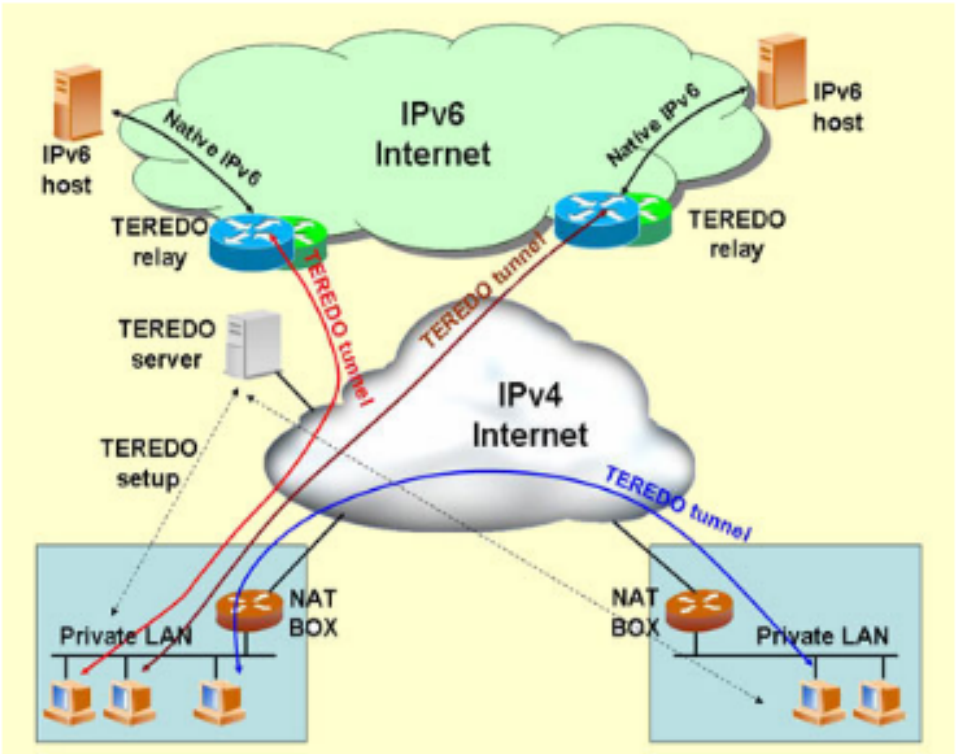
Túneles IPv4/IPv6 (cont.)

- 6to4: en Intranet



Túneles IPv4/IPv6 (cont.)

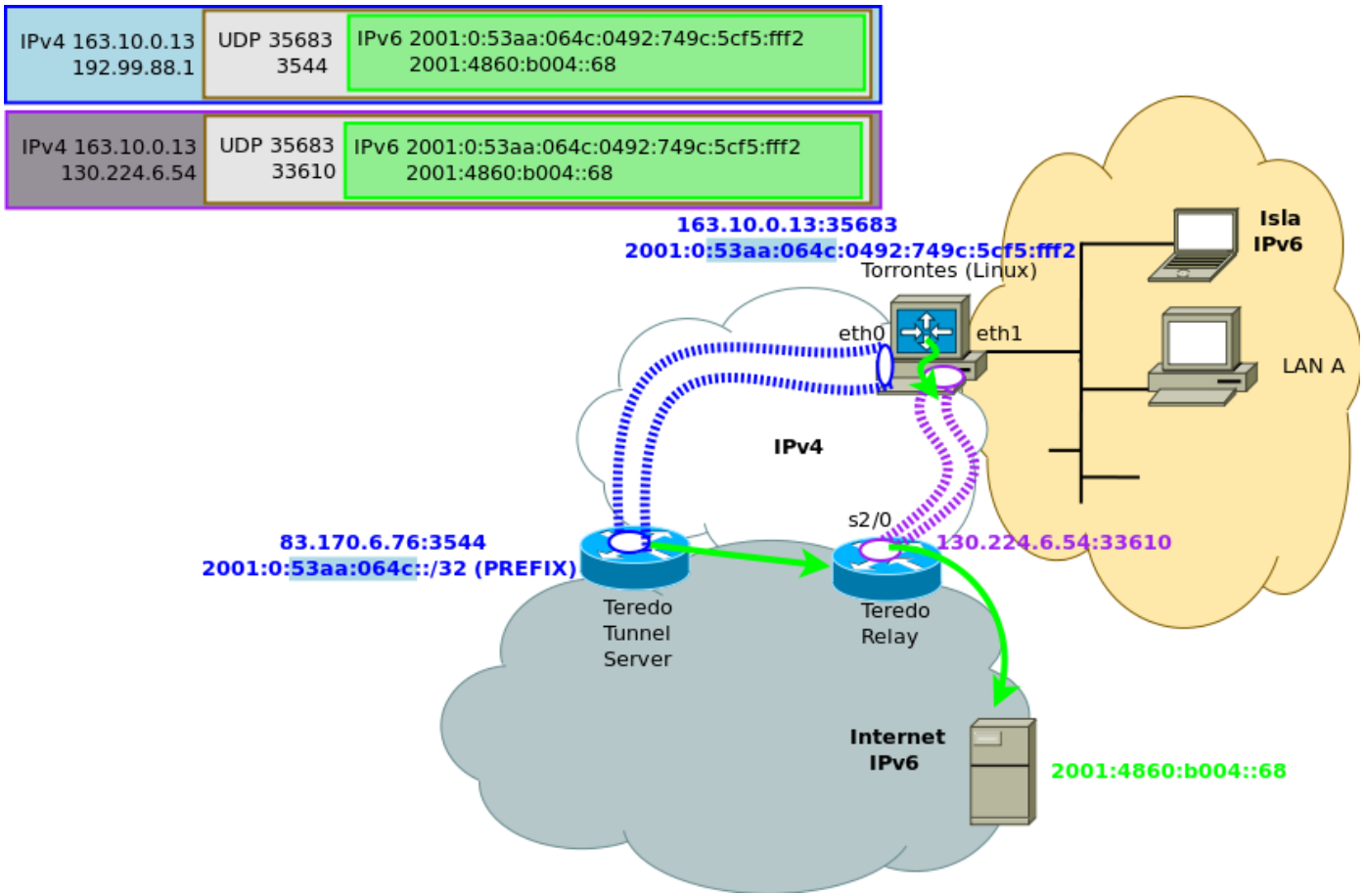
- Teredo:
  - funciona sobre direcciones IPv4 privadas, con NAT y sin Proto 41.
  - Encapsula el datagramas UDP.
  - Se configura el cliente contra un Server Teredo.
  - El Server Teredo proporciona acceso a los Teredo Relay.



Fuente: http://www.ipv6tf.org/

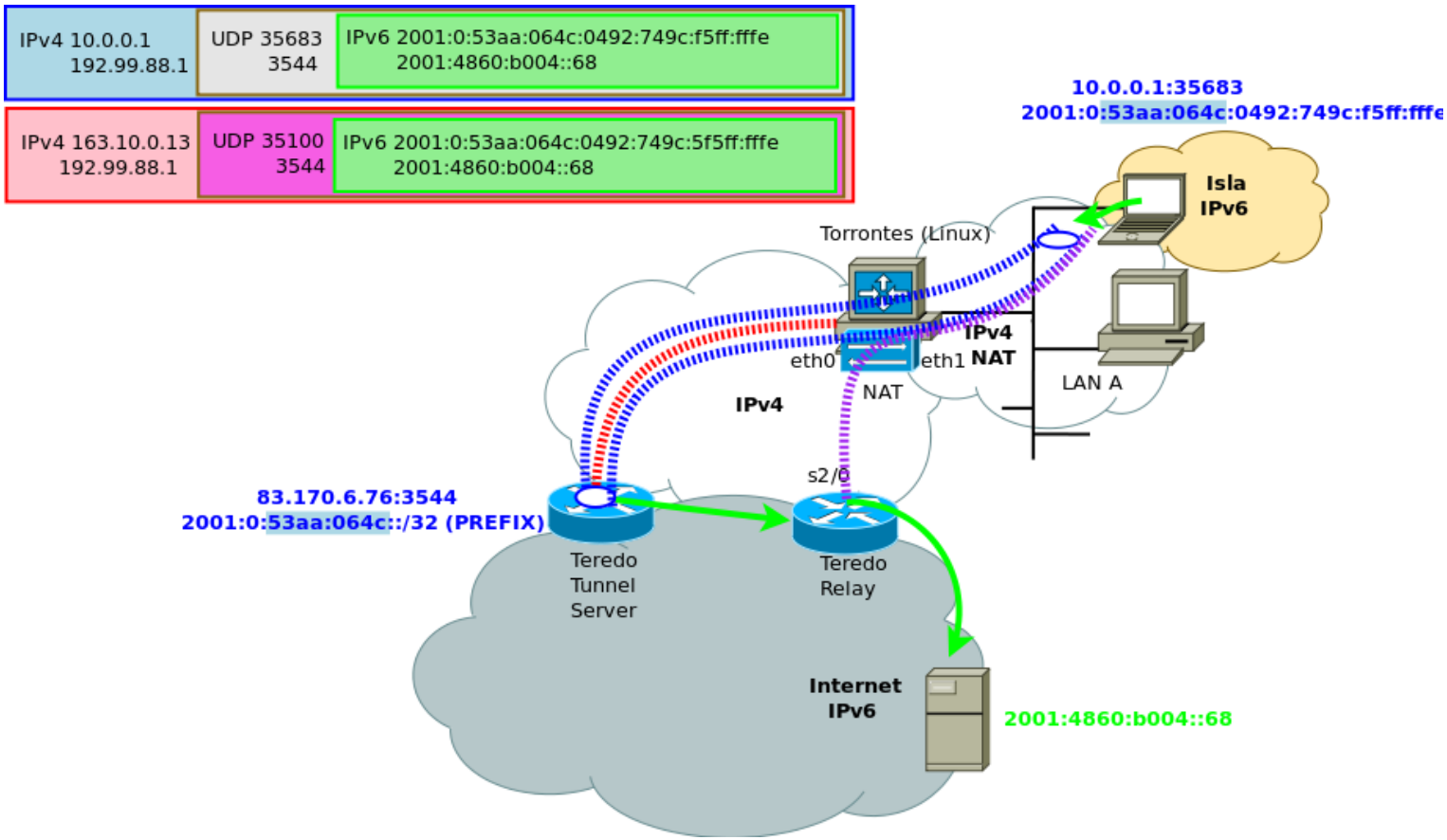
Túneles IPv4/IPv6 (cont.)

- Teredo: sobre direcciones IPv4 públicas.



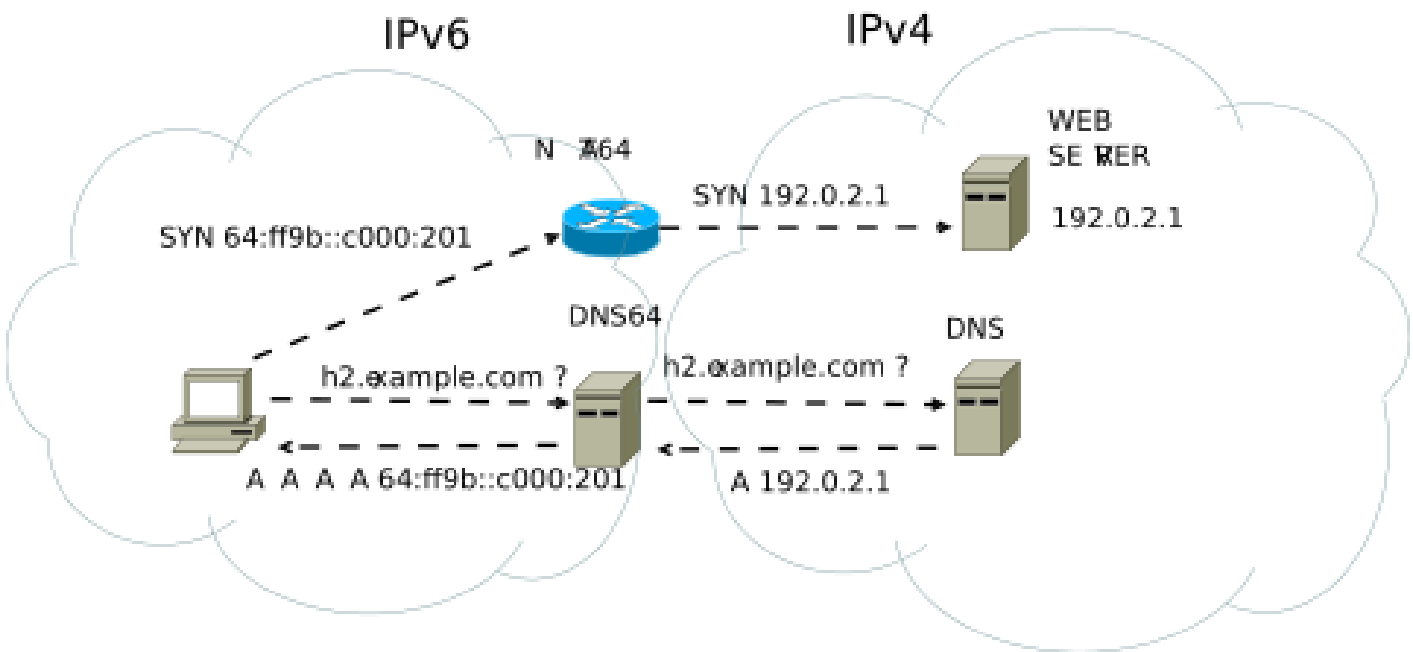
Túneles IPv4/IPv6 (cont.)

- Teredo: sobre direcciones IPv4 privadas.



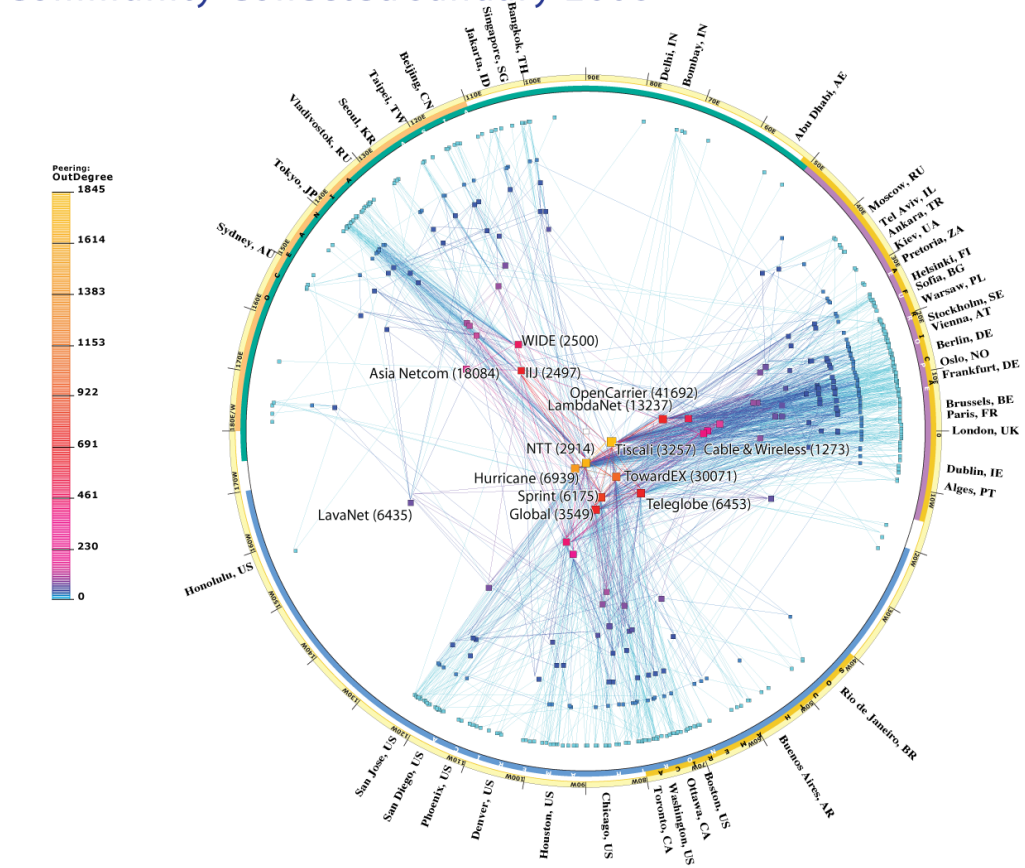
Técnicas de Traducción

- NAT64(NATPT) + DNS64.
- No es una técnica recomendada, rompe principio end-2-end.

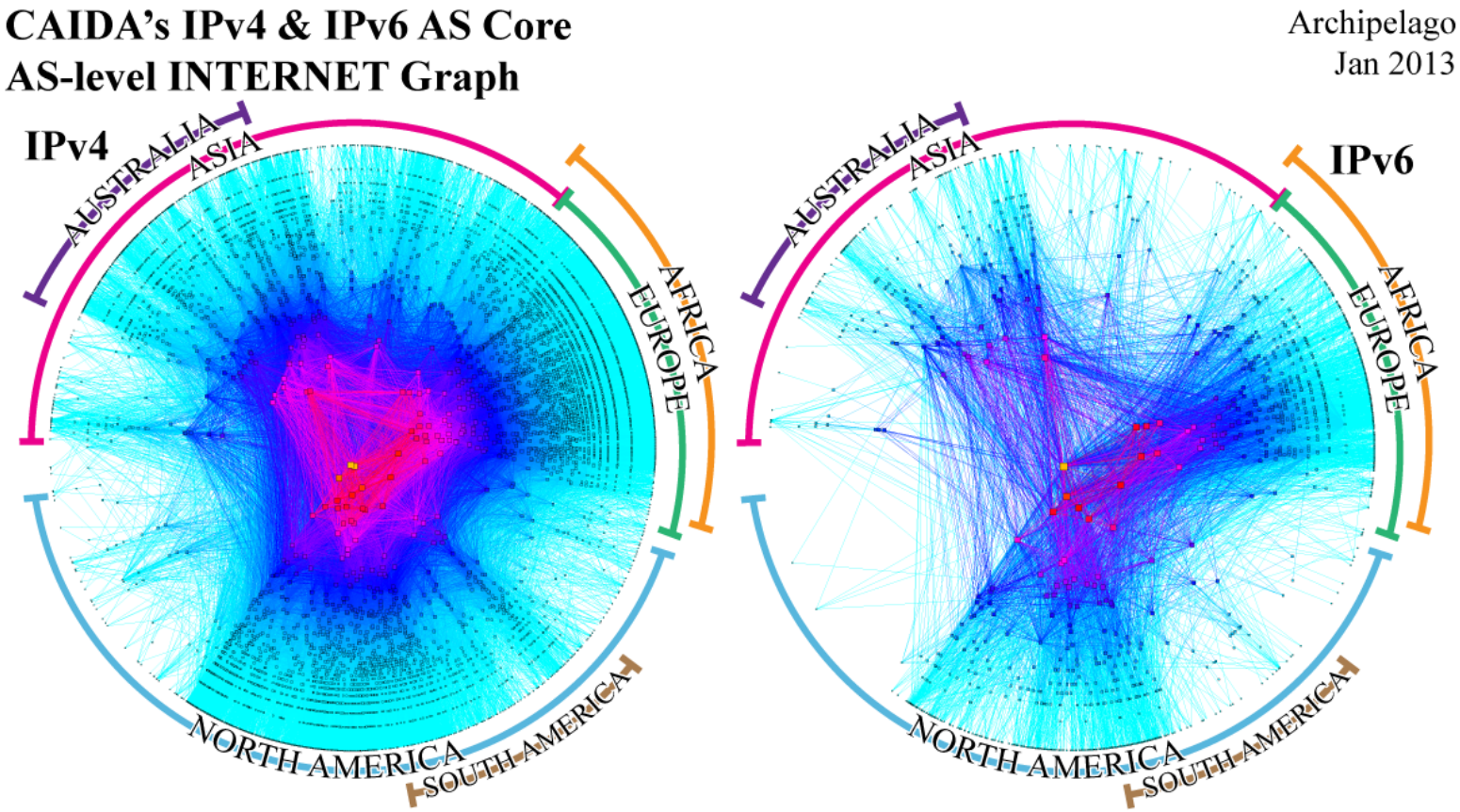


fuente: http://www.wikipedia.com

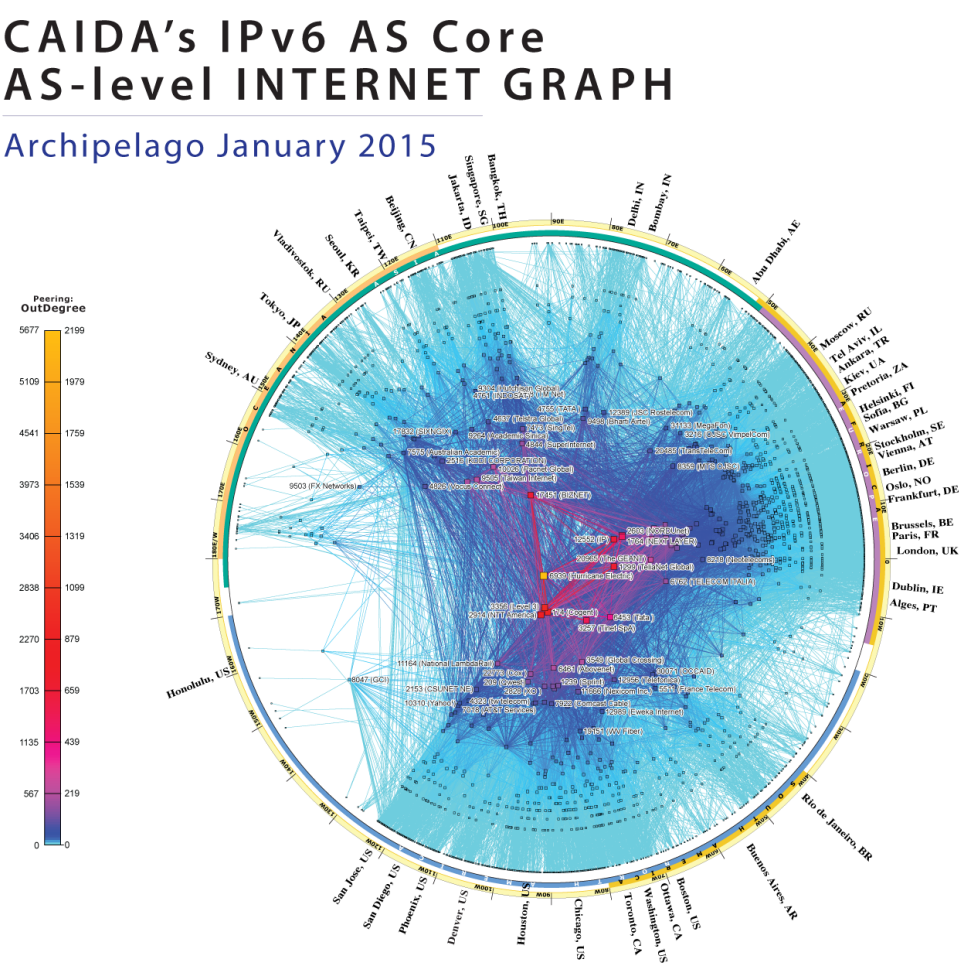
IPv6 en 2008



IPv4 vs IPv6 en 2013



La Internet Actual



fuente: http://www.caida.org

Referencias

- IPv6 Essentials (Integrating IPv6 into your IPv4 Network) Silvia Haggen. O'Reilly. 2nd Ed, 2006.
- Migrating to IPv6. Marc Blanchet. John Wiley and Sons, 2006.
- Tutorial de Jordi Palet Martínez: http://www.consulintel.es/html/ForoIPv6/Documentos/Tutorial de IPv6.pdf.
- TCP/IP Illustrated, Volume 1: The Protocols, 2nd Ed. W. Richard Stevens, Kevin R. Fall. Addison-Wesley Professional Computing Series, 2011.
- IPv6 for All: http://www.ipv6tf.org/pdf/ipv6forall.pdf.