

Máster en Ingeniería Informática

Redes de Nueva Generación

Profesor:

Dr. Juan Carlos Fabero Jiménez (UCM)



MPLS



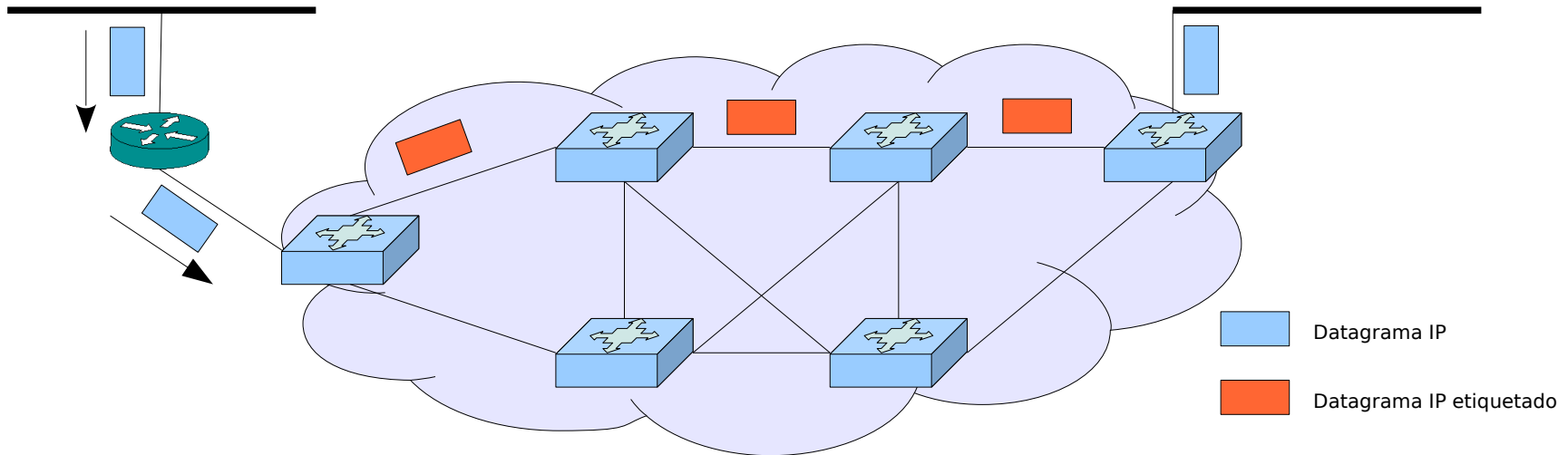
Introducción

Motivación

- **IP sólo ofrece un servicio no orientado a conexión.**
 - Cada encaminador toma las decisiones de encaminamiento para cada paquete, en función de la dirección destino.
 - Resulta difícil implantar criterios de calidad de servicio (QoS).
 - Se debe procesar la cabecera IP completa.
 - Sobrecarga considerable (menor para IPv6).
 - Estas tareas las realizan todos los encaminadores de la red, y de manera independiente.
 - La elección de la ruta se hace según “el prefijo más largo”, es decir, se elige la ruta más específica hacia el destino.
 - Difícil de implementar en hardware.
 - En software, requiere muchos accesos a memoria para encaminar un solo datagrama.

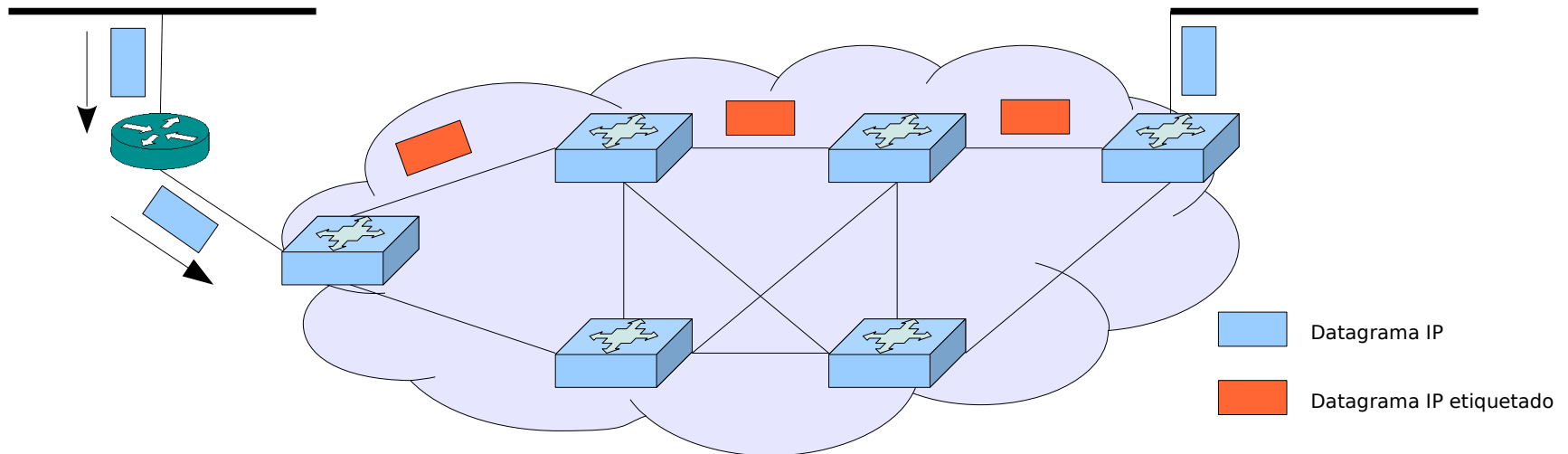
Multiprotocol Label Switching (RFC3031)

- Una red MPLS se compone de nodos LSR (Label Switching Routers)
- Cuando un paquete ingresa en el dominio MPLS, es clasificado en una FEC (Forwarding Equivalence Class).
- A cada FEC se le asocia una etiqueta.
- El encaminamiento se realiza según la etiqueta.



Multiprotocol Label Switching

- Cada FEC se define en función de la caracterización del tráfico que define a ese flujo.
- La asociación con la FEC se realiza sólo en el LSR de ingreso al dominio MPLS (LER, Label Edge Router).
- Las etiquetas tienen significado sólo local.



Multiprotocol Label Switching

■ Previo al reenvío:

- Debe establecerse la ruta LSP (Label Switching Path).
 - Especificar los parámetros QoS a lo largo de la ruta:
 - Recursos asignados al flujo.
 - Política de colas y de descartes.
 - Información de encaminamiento hacia el destino.
 - Mediante OSPF e iBGP.
 - Distribución de las etiquetas a lo largo de la LSP.
 - Mediante el protocolo LDP (Label Distribution Protocol, RFC5036).
 - Mediante una versión mejorada de RSVP.
- Cuando ingresa un datagrama IP en el dominio MPLS:
 - Se examinan sus requisitos QoS y se asocia con una FEC (y, por tanto, con una LSP).
 - Si la LSP no existe todavía, el LSR coopera con otros LSR para crearla.
 - Se añade al datagrama la etiqueta correspondiente con la FEC.

Multiprotocol Label Switching

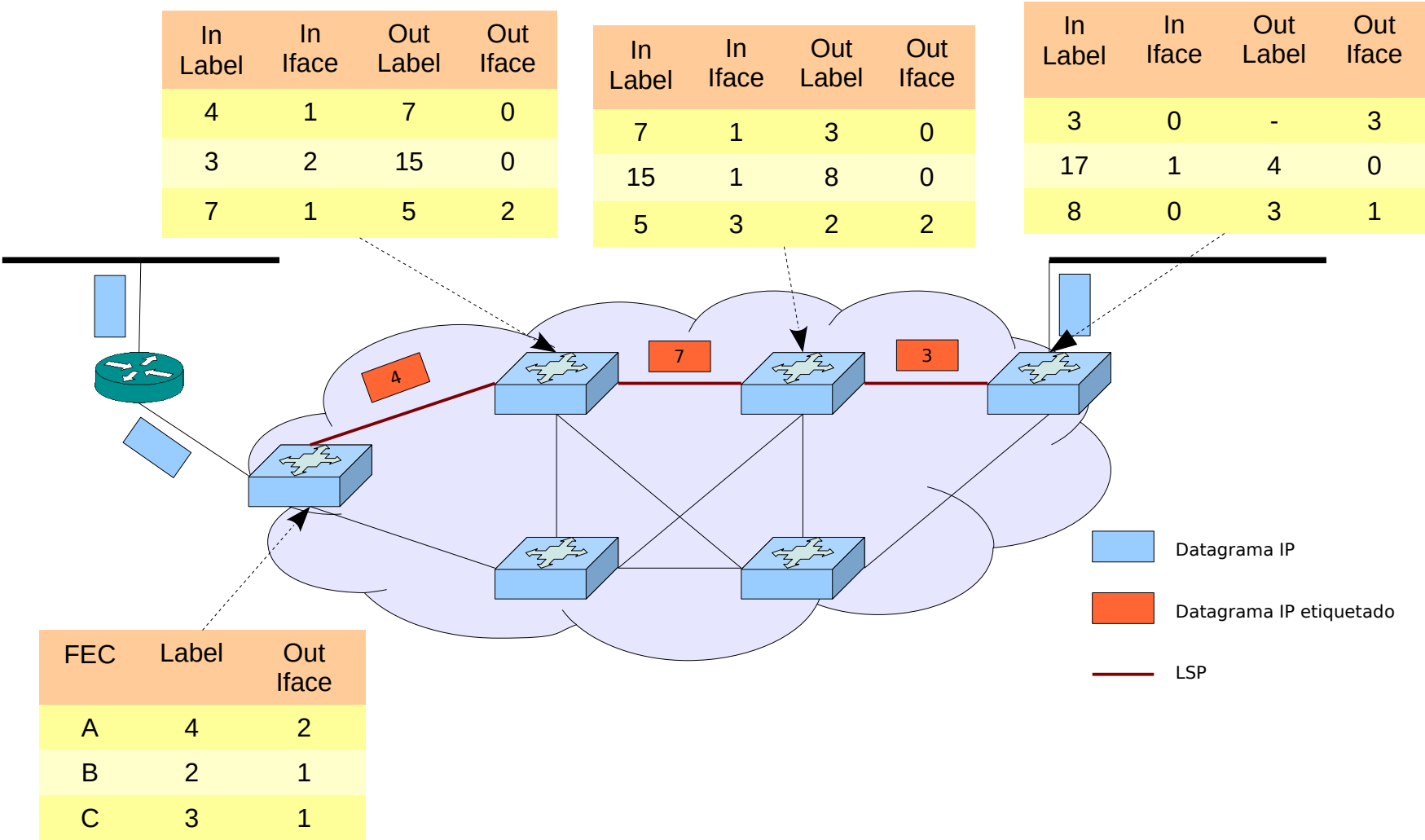
■ Durante el Reenvío:

- Cada LSR examina la etiqueta entrante.
 - Retira la etiqueta entrante.
 - Incorpora la etiqueta saliente.
 - Reenvía el datagrama al siguiente LSR dentro de la LSP.

■ Al salir del dominio MPLS:

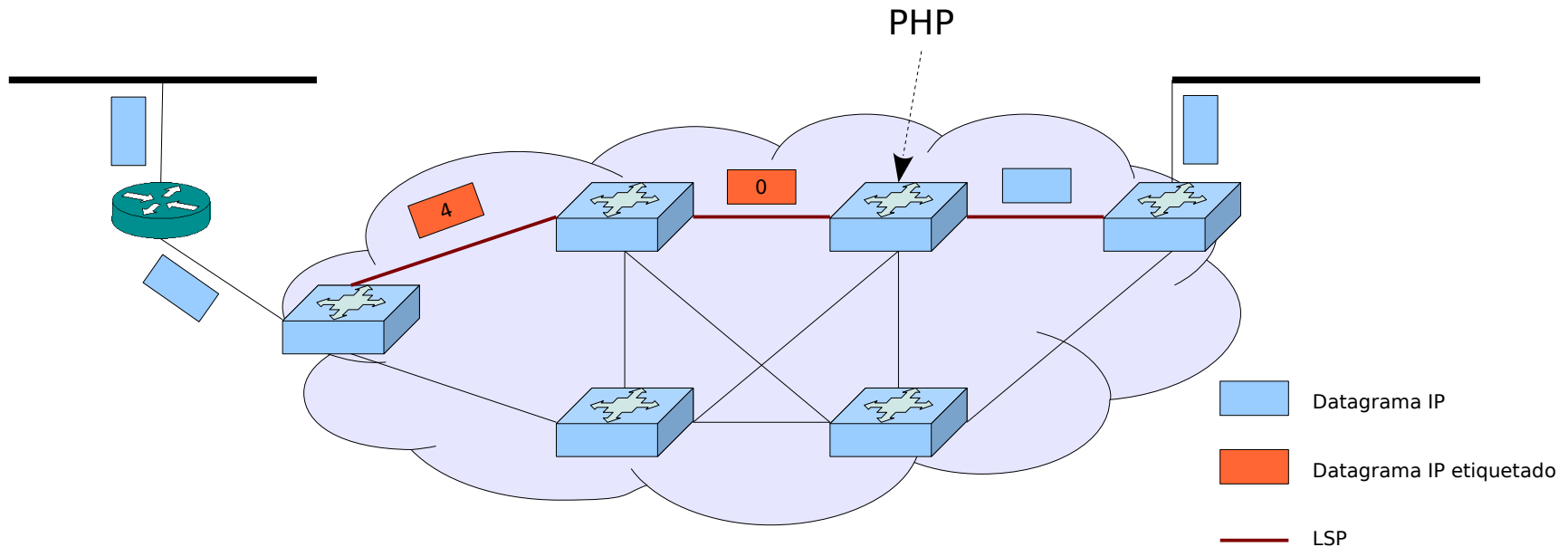
- El LSR saliente elimina la etiqueta y reenvía el datagrama IP sin etiquetar.

Multiprotocol Label Switching



Penultimate Hop Popping

- **El LSR anterior al LER de salida elimina la etiqueta MPLS.**
 - Reenvía el datagrama IP sin etiquetar.
 - El LER encamina el datagrama IP.
 - Evita tener que procesar dos veces el datagrama en el LER.



MPLS

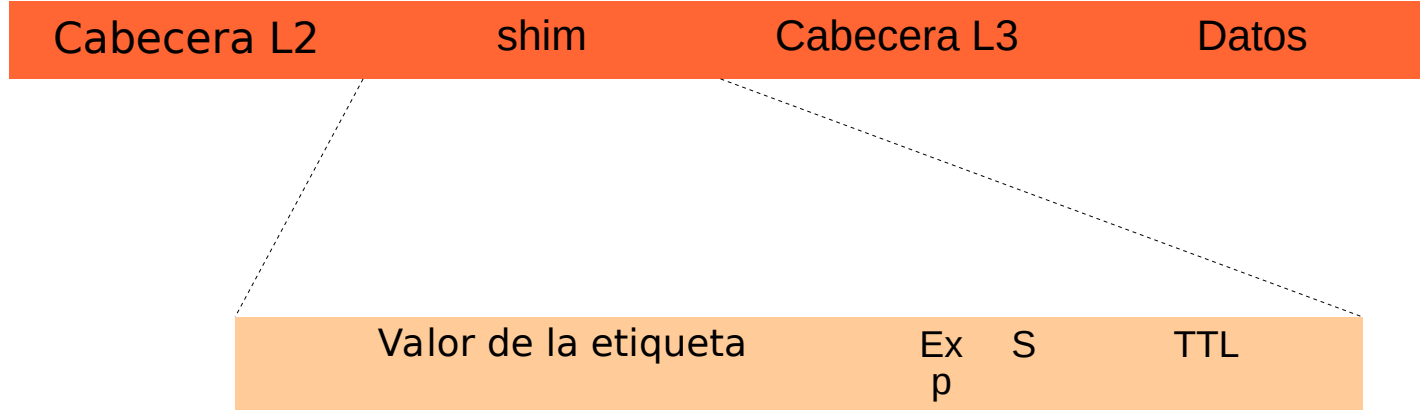


Etiquetas MPLS

Formato

■ Una etiqueta MPLS tiene 32 bits:

- Valor de la etiqueta (20 bits).
- Exp: uso experimental (3 bits). A partir del RFC 5462, Traffic Class.
- S: 1 para la etiqueta más antigua, 0 para las demás (apilado de etiquetas)
- TTL: tiempo de vida (8 bits)



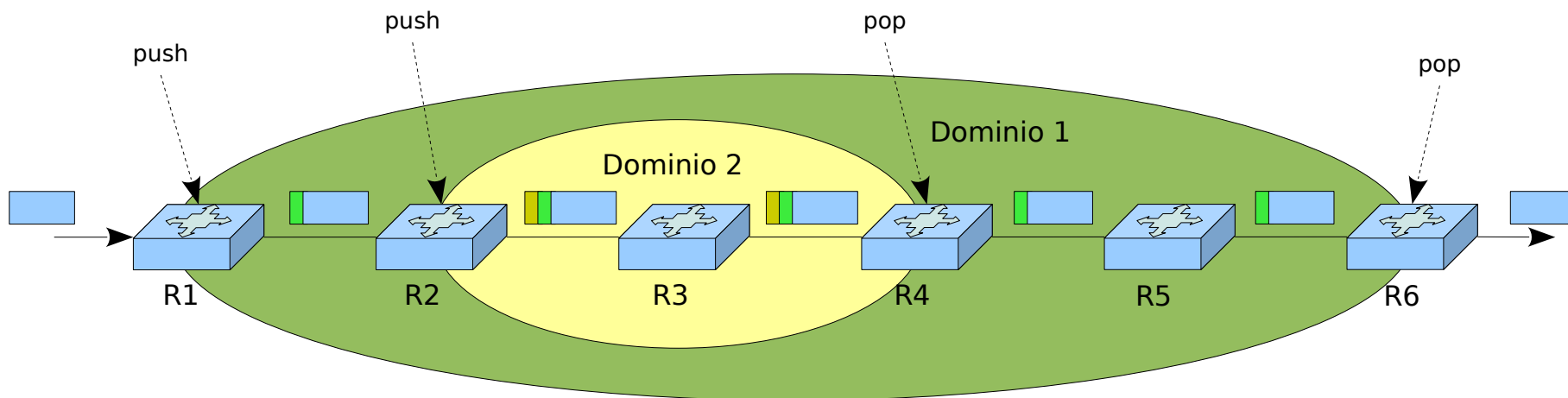
Tiempo de vida

- **Los encaminadores LSR no examinan la cabecera IP.**
 - No modifican el campo TTL/Hop Limit.
- **Cuando un datagrama ingresa en el dominio MPLS:**
 - Se añade la etiqueta MPLS de acuerdo a su FEC y se copia el campo TTL/Hop Limit de IP en el campo TTL de MPLS.
- **En cada reenvío:**
 - Se decrementa el campo TTL de la etiqueta.
 - Si TTL llega a 0, el paquete se descarta.
 - Si TTL es positivo, se reenvía el paquete al siguiente LSR.
- **Cuando el paquete sale del dominio MPLS:**
 - Se decrementa el TTL de la etiqueta.
 - Si TTL llega a 0, el paquete se descarta.
 - Si TTL es positivo, se elimina la etiqueta, se copia el valor en el campo TTL/Hop Limit y se reenvía el datagrama sin etiquetar.

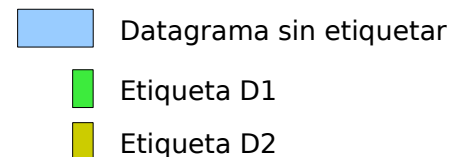
Apilado de etiquetas

■ MPLS permite apilar etiquetas.

- Un LSR puede añadir una etiqueta (push) aparte de la etiqueta ya existente, o eliminar una etiqueta (pop) de la pila.
- Se pueden crear túneles (dominios MPLS dentro de dominios MPLS).

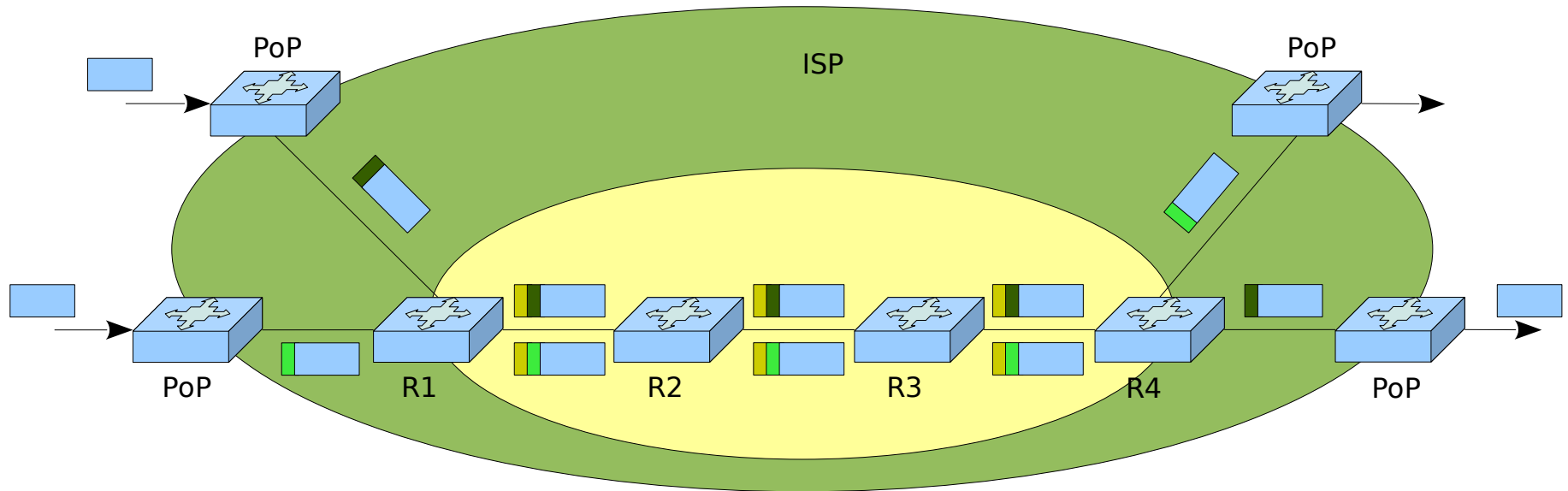


- R1 añade la etiqueta D1 acordada con R2
- R2 cambia la etiqueta D1 a la acordada con R4
- R2 añade la etiqueta D2 acordada con R3
- R3 cambia la etiqueta D2 a la acordada con R4
- R4 extrae la etiqueta D2
- R4 cambia la etiqueta D1 a la acordada con R5
- R5 cambia la etiqueta D1 a la acordada con R6
- R6 extrae la etiqueta D1






Apilado de etiquetas

■ Uso por parte de los ISP

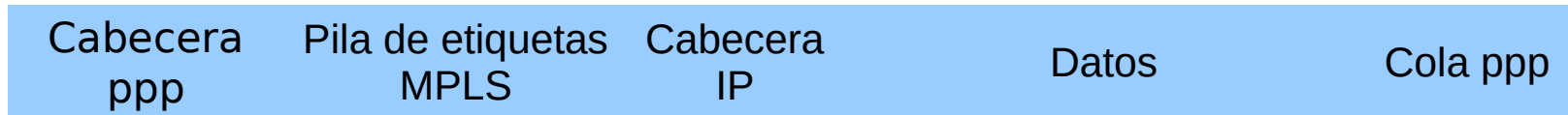


Se facilita el encaminamiento a través de una infraestructura común entre varios PoP (Point of Presence)

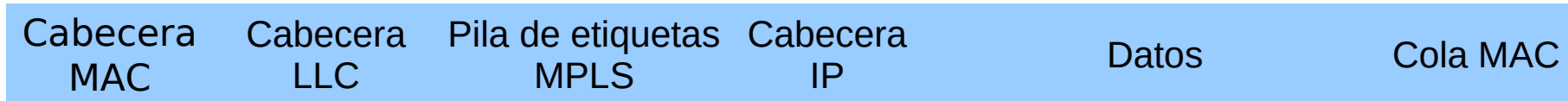
-  Datagrama sin etiquetar
-  Etiqueta D1
-  Etiqueta D2

Encapsulado MPLS

- **Las etiquetas MPLS se sitúan entre las cabeceras de enlace y de red.**
 - Nivel OSI “2.5”




Protocol=0x0281 (Unicast MPLS), 0x0283 (Multicast MPLS)



Type=0x8847 (Unicast MPLS), 0x8848 (Multicast MPLS)

MPLS



Ingeniería del Tráfico

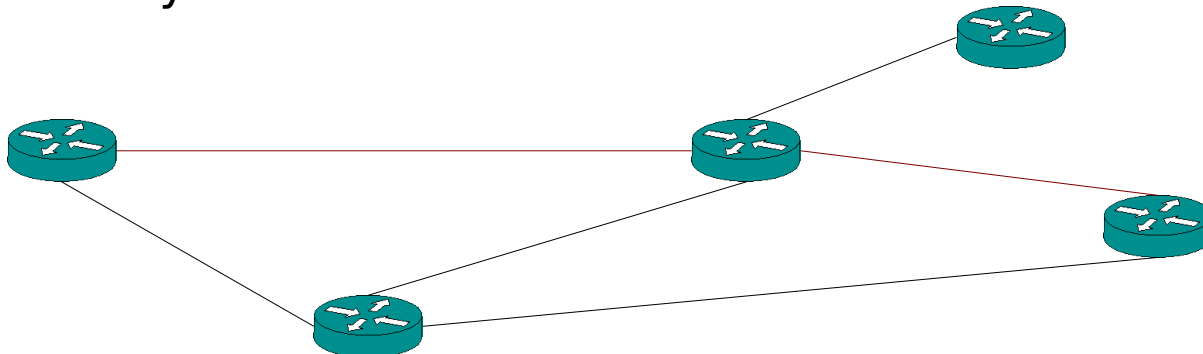
Ingeniería del tráfico

■ Consiste en:

- Definir rutas dinámicamente.
- Planificar entregas de tráfico según la demanda actual.
- Optimizar la utilización de la red.

■ Problema:

- Los algoritmos tradicionales de encaminamiento, como OSPF, sólo realizan de manera eficiente la primera de las tareas.
 - Los datagramas siguen siempre la misma ruta entre el origen y el destino, aunque esa ruta puede cambiar en caso de congestión.
 - No se distingue entre paquetes pertenecientes a distintos flujos de datos y con distintas necesidades.

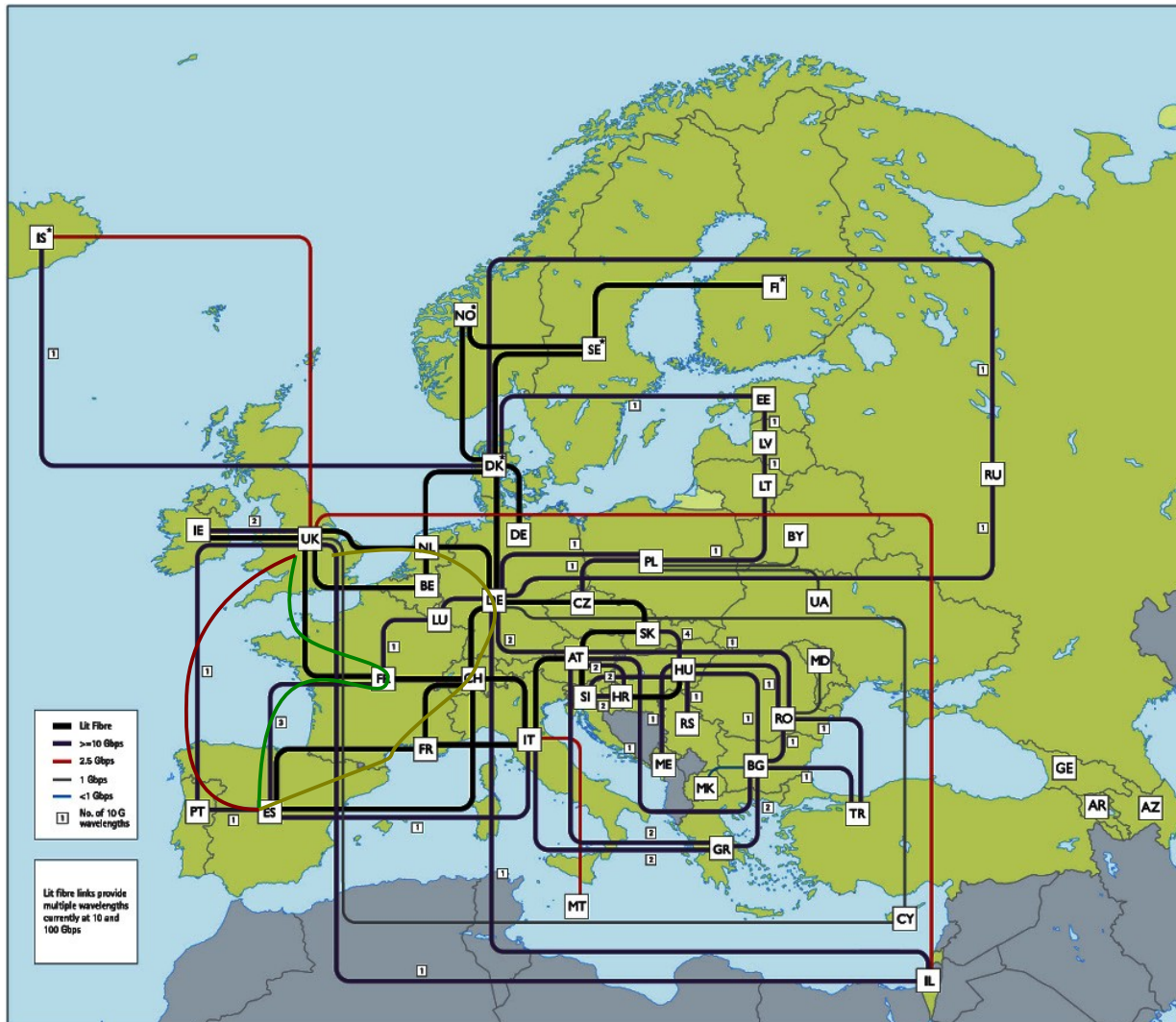


■ Solución MPLS:

- La LSP se establece según los requisitos del flujo:
 - Origen y destino.
 - Ancho de banda requerido.
 - Máxima latencia permitida.
 - ...
- La LSP se asocia con una FEC en el LER de ingreso.
- Se puede ver la LSP como un túnel entre el LER de ingreso y el de salida.
- Los protocolos de señalización se encargan de establecer la LSP.
 - LDP
 - RSVP-TE

Ingeniería del tráfico

■ ¿Cómo encaminar de Madrid a Londres?



Ingeniería del tráfico

■ Desde el punto de vista de MPLS:

- Cada LSP requiere un ancho de banda determinado.
- Se utiliza RSVP-TE para hacer un encaminamiento basado en restricciones.
 - Se elige el camino más corto que permite albergar el ancho de banda requerido por la LSP.
 - Si existe el camino, se distribuyen las etiquetas (señalización) para crear la LSP.
 - Se resta el ancho de banda de la LSP del ancho de banda disponible en los enlaces.
 - Una LSP posterior puede prohibirse por no disponer de ancho de banda suficiente.
 - Aunque, idealmente, será redirigida a través de un camino alternativo.
- Se pueden establecer prioridades entre las LSP, de manera que una LSP posterior más prioritaria podría “expropiar” a una LSP ya existente pero de menor prioridad.
 - La LSP expropiada debería ser redirigida por otro camino.

Ingeniería del tráfico

■ Problema:

- Cálculo del ancho de banda necesario.
 - Las redes IP son, por definición, de conmutación de paquetes.
 - El ancho de banda puede variar de manera drástica e impredecible.

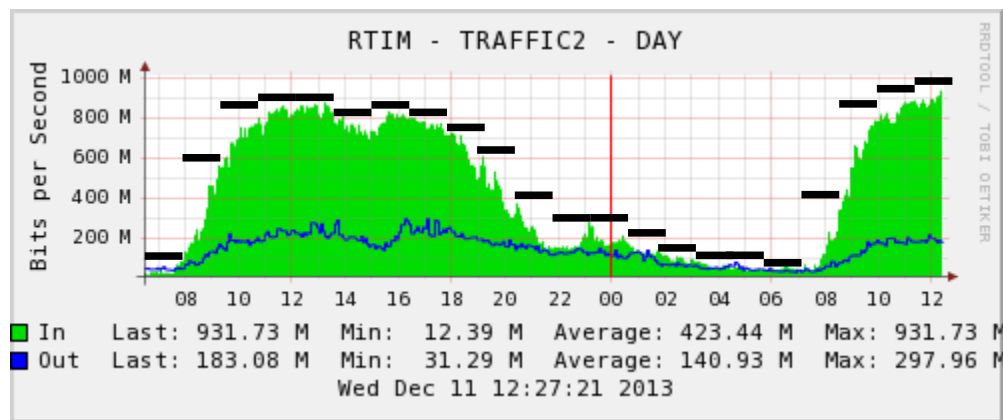
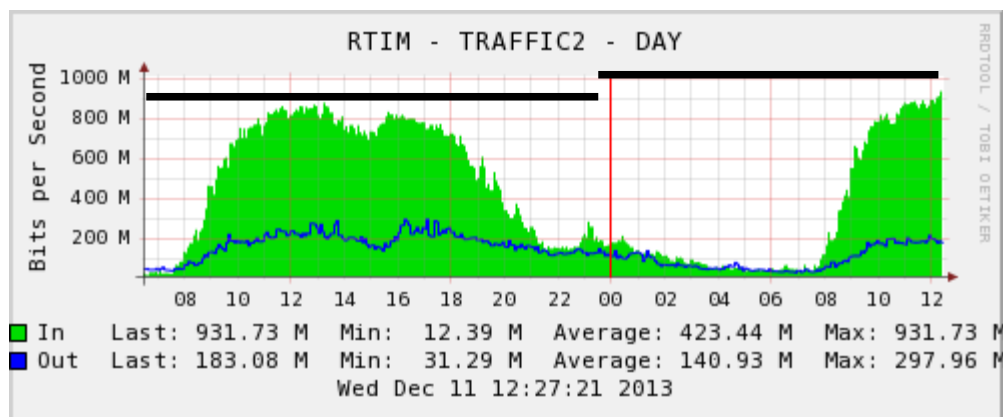
■ Soluciones:

- Predicción basada en modelos.
 - Es el primer mecanismo que se implantó.
 - Suele requerir el uso de herramientas de terceros.
- Medición en tiempo real.
 - El LSR mide el flujo de datos a través de la LSP.
 - Periódicamente se actualiza la reserva RSVP con los datos de ancho de banda real medidos por el LSR.

Ingeniería del tráfico

■ Medición en tiempo real.

- ¿Cada cuánto tiempo se actualiza el ancho de banda?



MPLS



Ajuste dinámico del ancho de banda

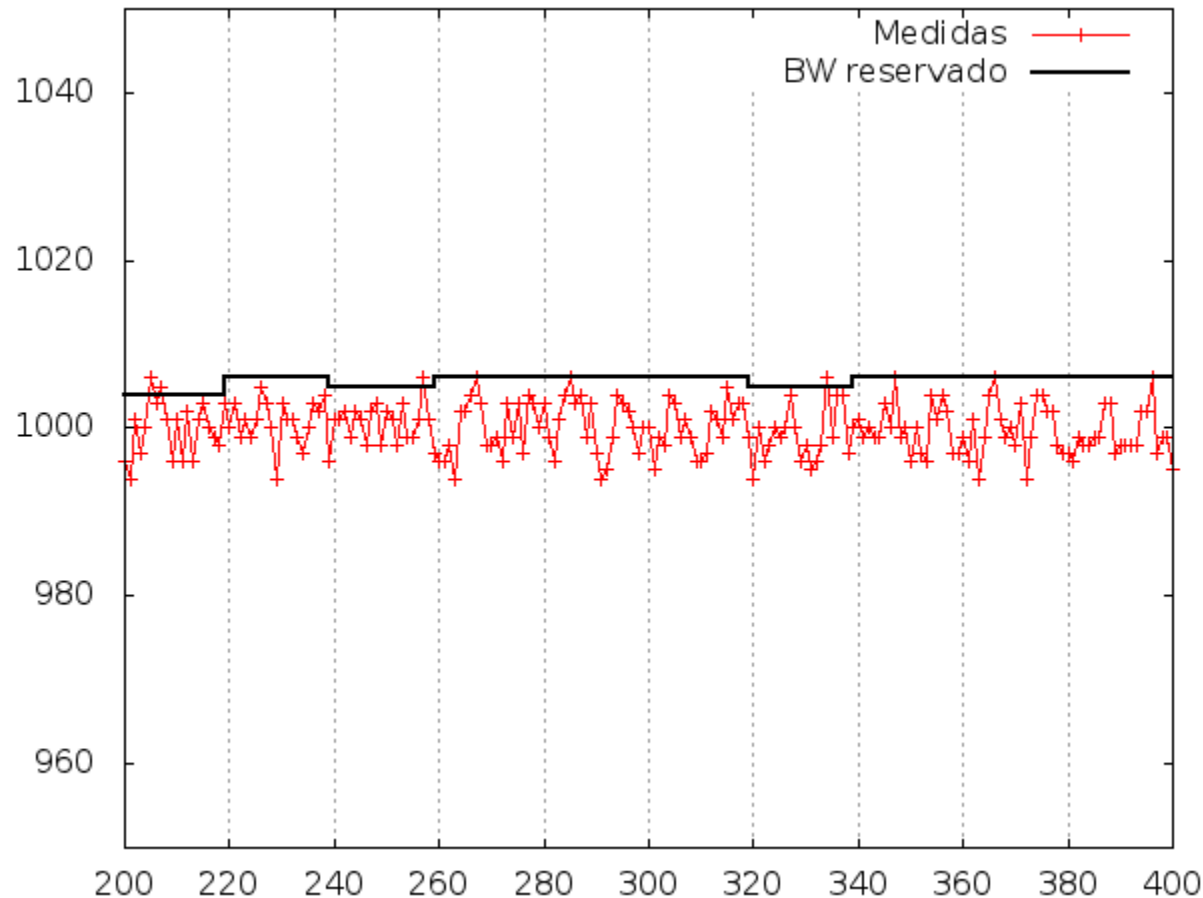
Ajuste dinámico del BW

- **El ajuste dinámico del ancho de banda sigue un algoritmo sencillo:**
 - Se toman muestras del ancho de banda requerido cada cierto intervalo de muestreo (p.e. cada 120 s)
 - Cada cierto tiempo (p.e. cada 600 s), se calcula el ancho de banda de la LSP como la máxima de entre las muestras del intervalo de actualización.
 - Si la diferencia entre el valor medido y el actual supera cierto umbral, se señala el nuevo ancho de banda para la LSP.

Ajuste dinámico del BW

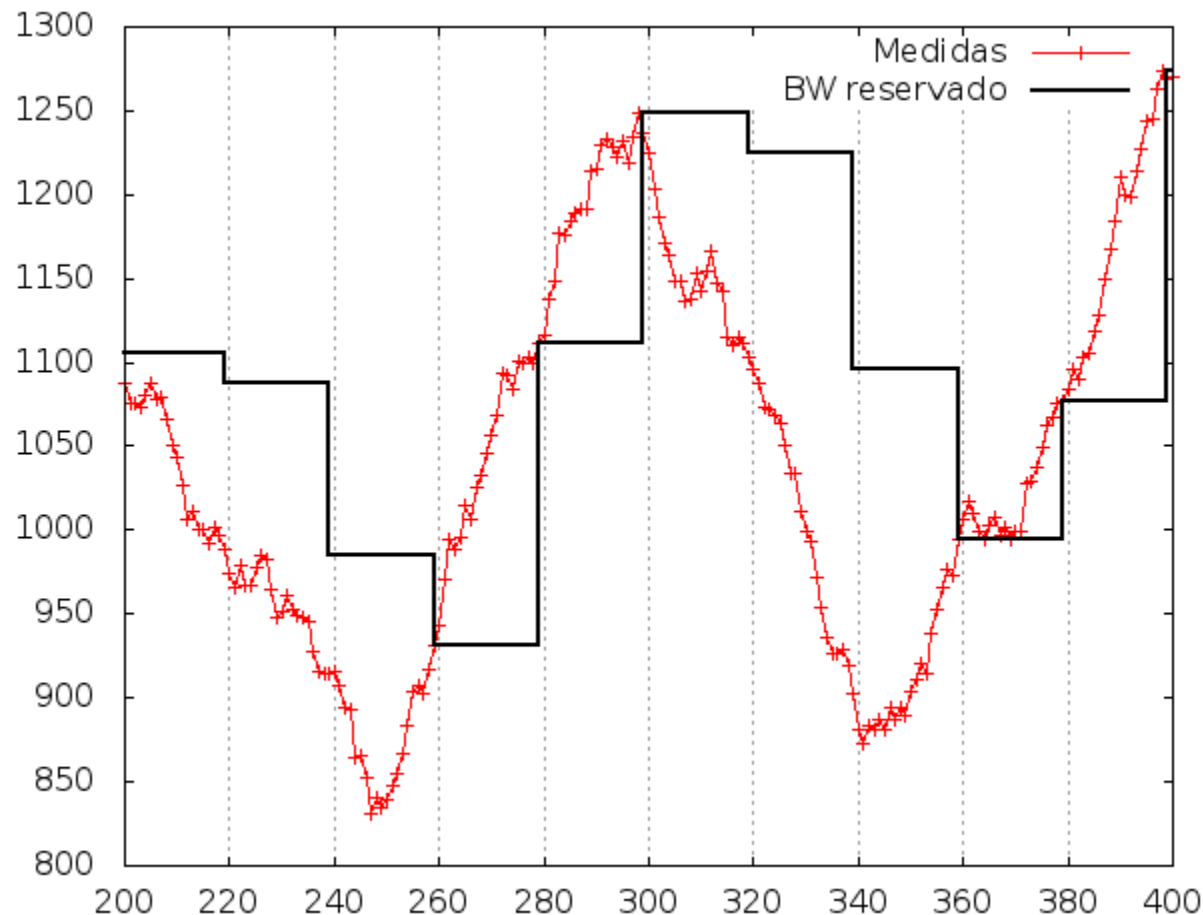
■ Medición en tiempo real.

- ¿Cuándo funciona bien?



Ajuste dinámico del BW

- **Medición en tiempo real.**
 - ¿Cuándo no funciona bien?



Ajuste dinámico del BW

■ Overflow

- Cuando un número de muestras ha superado el BW reservado, se actualiza el BW antes del intervalo de ajuste.
- Así se puede responder a situación de aumento repentino del ancho de banda necesario.
 - Por picos de tráfico en el flujo.
 - Por situaciones de aumento de tráfico por fallo de otros enlaces.
- Casi todos los fabricantes incorporan la detección de la situación de overflow.

Ajuste dinámico del BW

■ Underflow

- Se produce cuando el tráfico disminuye de manera brusca.
- Puede ser debido a diversas causas:
 - Variación en el flujo.
 - Durante la recuperación ante un fallo (make-before-break).
- No todos los dispositivos incorporan detección de underflow.

■ Make-before-break

- Cuando se va a cambiar una LSP, primero se señala la nueva antes de desmantelar la antigua.
 - Es necesario reservar ancho de banda para ambas simultáneamente, de manera temporal.
 - Cuando no se detecta el underflow, puede que no haya recursos suficientes en la red.

MPLS



Fast Reroute

Fast Reroute

- **Cuando se produce un cambio en la topología por fallo, es necesario buscar rutas alternativas.**
 - OSPF realiza una inundación controlada de LSA, y aplica el algoritmo SPF sobre la nueva topología.
 - BGP elimina las rutas que utilizaban el enlace fallido y espera a recibir nuevos anuncios por otros enlaces, o utiliza los recibidos anteriormente.
 - En ambos casos, la recuperación puede tardar desde unos cuantos segundos a varios minutos.
 - Mientras tanto, el tráfico se ve interrumpido o, al menos, seriamente afectado.
 - Pueden aparecer temporalmente bucles de encaminamiento.
 - MPLS, en cambio:
 - Calcula alternativas antes de que ocurra el fallo.
 - Esas alternativas se almacenan en la FIB, en espera de ser activadas.
 - Cuando se produce el fallo, la recuperación se hace en ms.

Fast Reroute

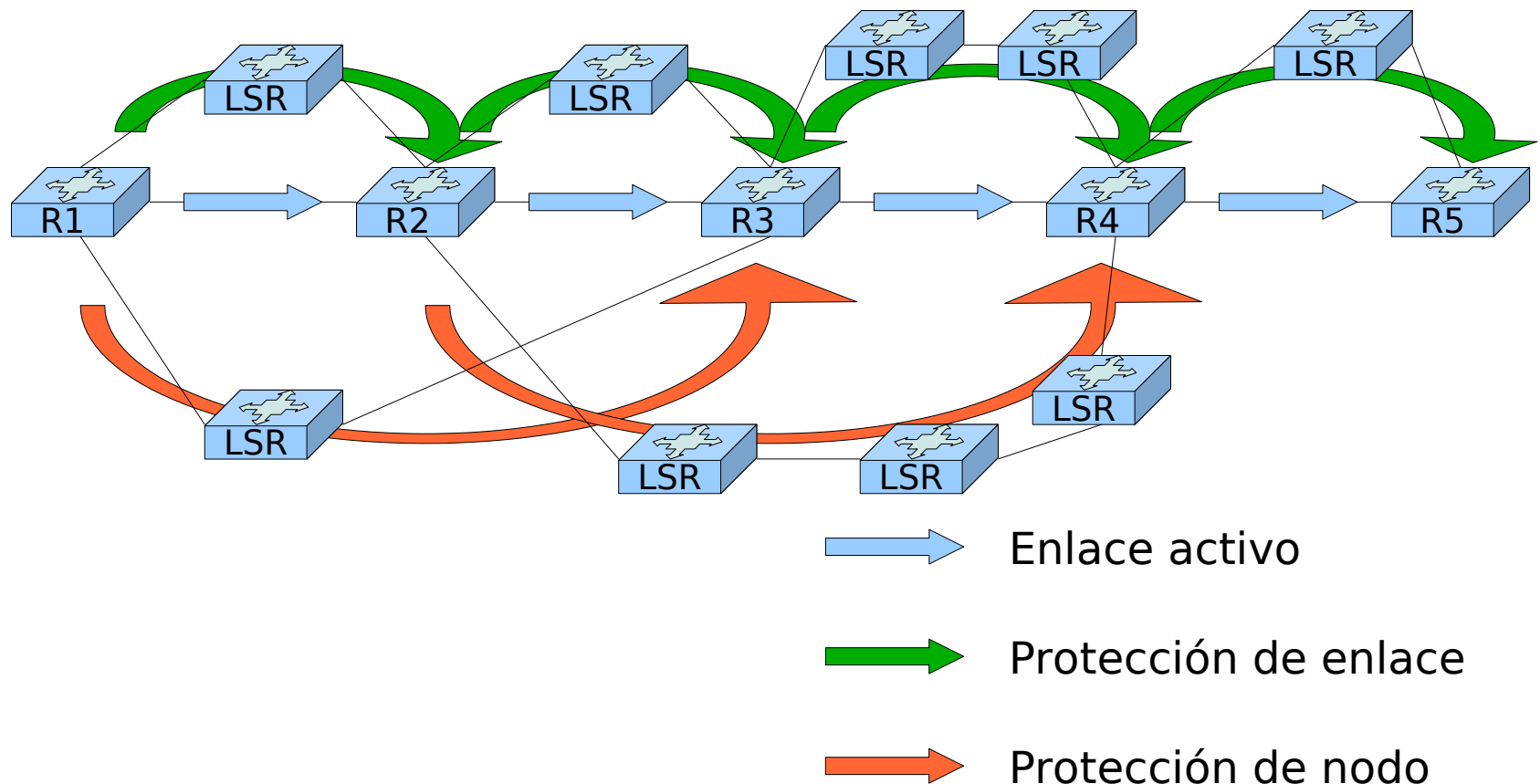
■ Protección:

- Qué tipos de fallos:
 - Frente a fallos en enlaces.
 - Se crea una LSP alternativa para cada enlace.
 - Frente a fallos en nodos.
 - Se crea una LSP alternativa entre nodos.
- Qué se hace con las LSP:
 - Uno a uno
 - Se crea una LSP alternativa para cada LSP activa.
 - La profundidad de etiquetas no aumenta.
 - Mucha sobrecarga.
 - Muchos a uno
 - Se crea una LSP alternativa para englobar varias LSP activas.
 - Aumenta en uno la profundidad de etiquetas.
 - Menos sobrecarga.

Fast Reroute

■ Protección de enlace y de nodo

- La LSP activa pasa por R1, R2, R3, R4 y R5.



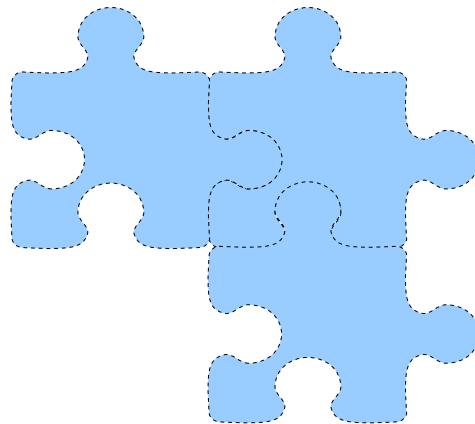
MPLS



IP sobre MPLS

IP sobre MPLS

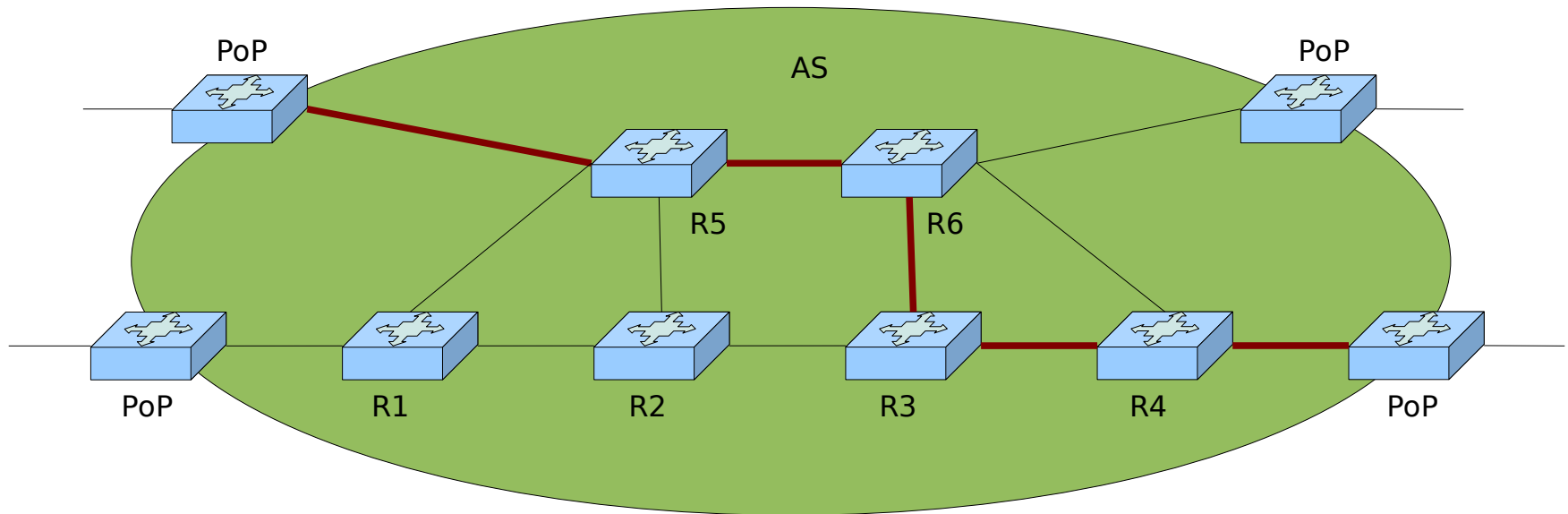
- **Cuando una red MPLS transporta datagramas IP, hay que tener en cuenta ciertos aspectos:**
 - Tratamiento del campo TTL/Hop Limit
 - En el LER de ingreso, se copia el campo TTL/Hop Limit del datagrama IP al campo TTL de la etiqueta.
 - En cada LSR se decrementa el campo TTL de la etiqueta.
 - En el LER de salida se copia el campo TTL de la etiqueta sobre el campo TTL/Hop Limit del datagrama.
 - Si el TTL llega a 0, se descarta el paquete...



MPLS. Resumen

■ Encaminamiento en troncales

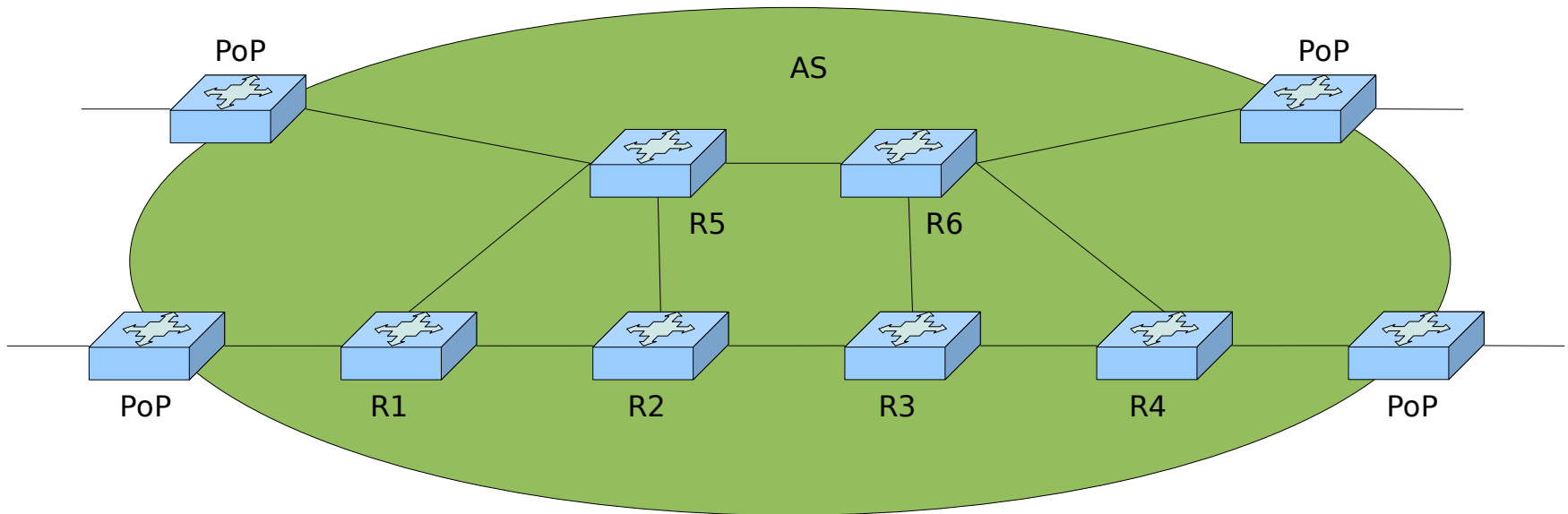
- Facilita encaminamiento en las grandes troncales.
- Reduce la sobrecarga de procesamiento.
- Rutas globales sólo necesarias en los LER (Label Edge Routers)
- Facilidad de recuperación ante fallos.



IP sobre MPLS

■ Grandes proveedores (troncales)

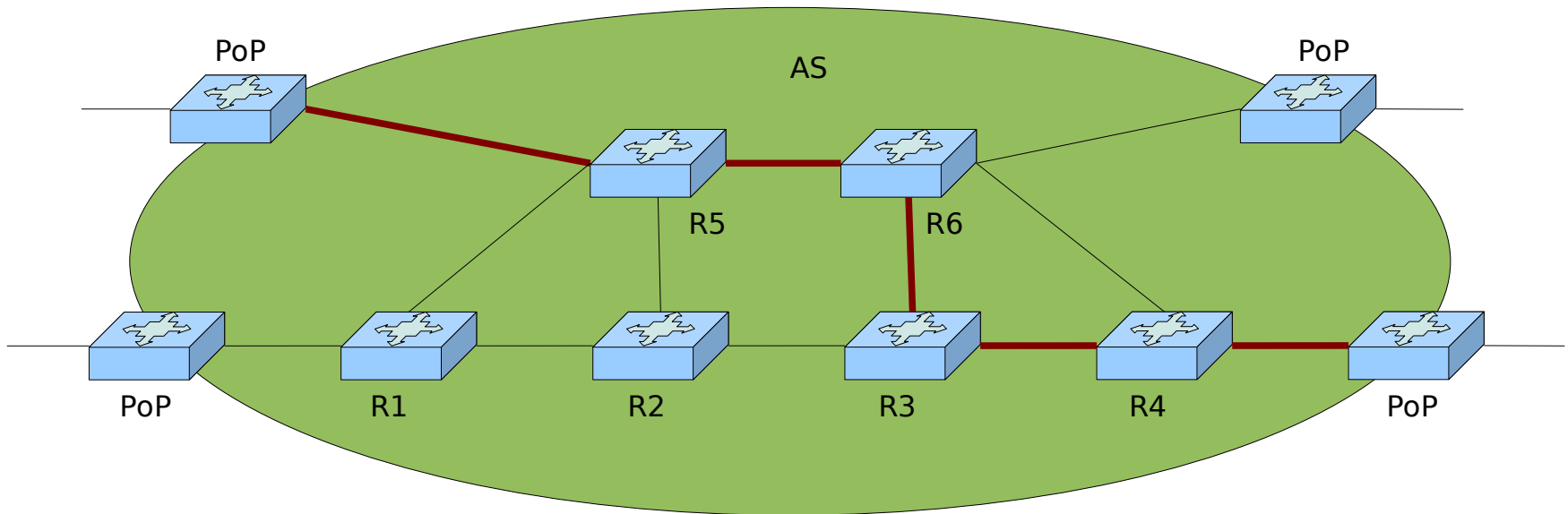
- Sus encaminadores internos (MPLS) no conocen rutas externas, sólo las aprendidas por IGP (p.e. OSPF).
- Los encaminadores LER ejecutan BGP y mantienen sesiones iBGP entre ellos, pero no inyectan rutas externas dentro de OSPF.



IP sobre MPLS

■ Grandes proveedores (troncales)

- ¿Qué sucede si un encaminador interno, pongamos R3, descarta un datagrama?
 - Debería enviar un ICMP al origen del datagrama.
 - Pero... ¡no conoce rutas externas!
 - Solución:
 - Se genera el ICMP y se envía por la LSP del datagrama original.



MPLS

LDP: Label Distribution Protocol

■ **Label Distribution Protocol**

- Definido en los RFC 5036, 6720, 6667, ...
- FRR implementa la especificación del RFC5036
- Se encarga de la coordinación de las etiquetas asociadas a una FEC a lo largo del dominio MPLS.

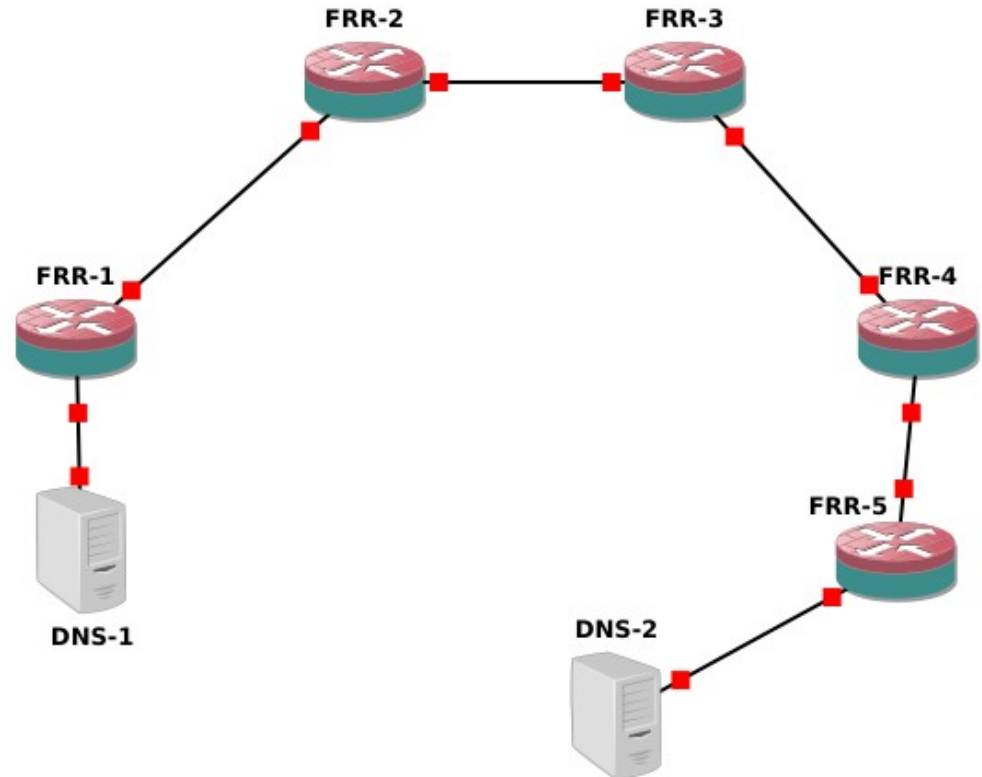
■ Label Distribution Protocol

- Soporte en Linux.
 - Necesario cargar el módulo **mpls_iptunnel** (en la máquina anfitriona).
 - En los contenedores FRR, añadir a /etc/sysctl.conf las entradas:
 - net.mpls.platform_labels=32
 echo net.mpls.platform_labels=32 >> /etc/sysctl.conf
 - net.mpls.conf.ethX.input=1
 con X desde 0 hasta 7.
 for i in {0..7} ; do echo net.mpls.conf.eth\$i.input=1 >>
 /etc/sysctl.conf ; done

LDP

■ Label Distribution Protocol

- FRR tiene soporte para LDP.
- Ejemplo:
 - Sea el siguiente AS.
 - Se usa OSPF como IGP.
 - Encapsulado MPLS.
 - Redes de usuario:
 - 172.16.1.0/24 para DNS1
 - 172.16.5.0/24 para DNS2



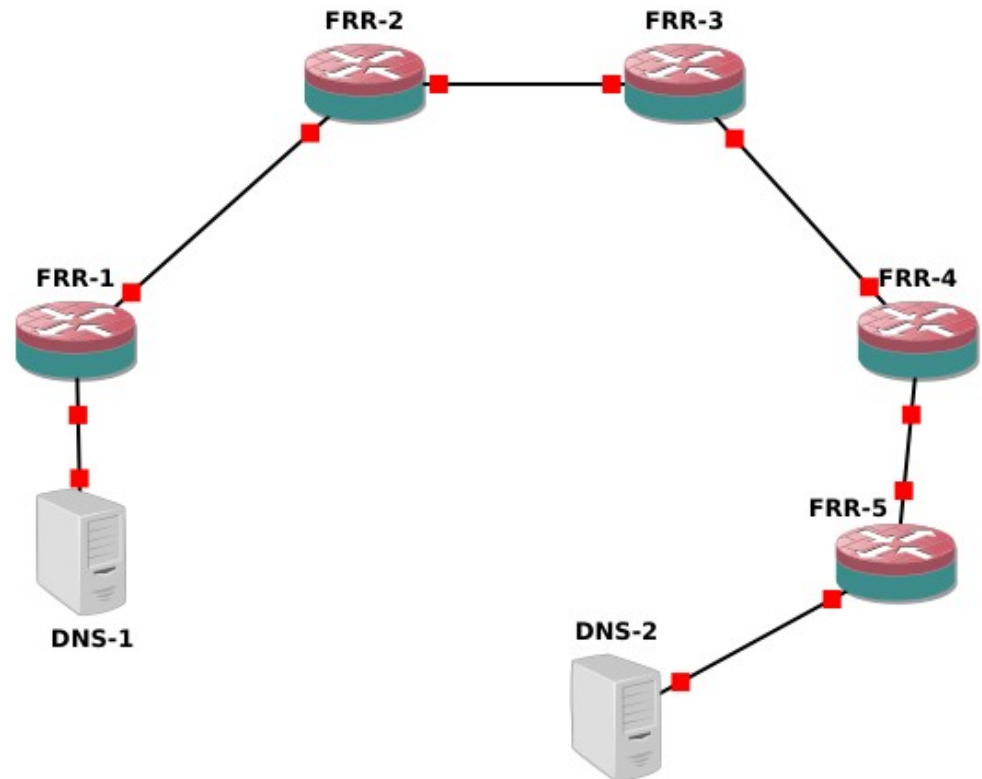
LDP

■ Label Distribution Protocol

- Configuración (en FRR-3, p.e.)

```
FRR-3# show running-config
```

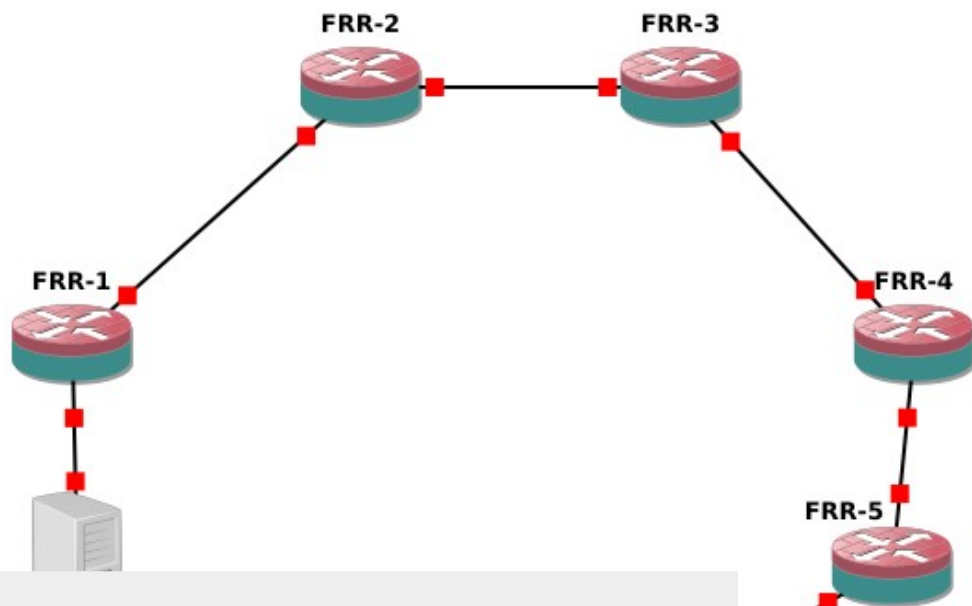
```
router ospf
network 10.0.23.0/24 area 0
network 10.0.34.0/24 area 0
!
mpls ldp
router-id 10.0.23.3
ordered-control
!
address-family ipv4
discovery transport-address 10.0.23.3
!
interface eth0
!
interface eth1
!
exit-address-family
!
```



LDP

■ Label Distribution Protocol

- *Binding* (en FRR-3, p.e.)



```
FRR-3# show mpls ldp binding
```

<u>AF</u>	<u>Destination</u>	<u>Nexthop</u>	<u>Local Label</u>	<u>Remote Label</u>	<u>In Use</u>
Ipv4	10.0.12.0/24	10.0.12.2	18	imp-null	yes
Ipv4	10.0.23.0/24	10.0.12.2	imp-null	imp-null	no
Ipv4	10.0.34.0/24	10.0.34.4	imp-null	imp-null	no
Ipv4	10.0.45.0/24	10.0.34.4	16	imp-null	yes
Ipv4	172.16.1.0/24	10.0.12.2	19	16	yes
Ipv4	172.16.5.0/24	10.0.34.4	17	17	yes

■ Label Distribution Protocol

- Encapsulado. Captura en el enlace entre FRR-3 y FRR-4.

The image shows a Wireshark packet capture window titled "*- [FRR-3 eth1 to FRR-4 eth0]". The capture filter is set to "icmp". The packet list shows five ICMP Echo (ping) packets. The selected packet, Frame 27, is an Echo (ping) reply from 172.16.5.2 to 172.16.1.2. The packet details pane shows the following structure:

- Ethernet II, Src: 56:c5:53:5b:77:bc (56:c5:53:5b:77:bc), Dst: 02:f1:2f:f1:48:2d (02:f1:2f:f1:48:2d)
- MultiProtocol Label Switching Header, Label: 17, Exp: 0, S: 1, TTL: 61
 - 0000 0000 0000 0001 0001 = MPLS Label: 17
 - 000. = MPLS Experimental Bits: 0
 - 1 = MPLS Bottom Of Label Stack: 1
 - 0011 1101 = MPLS TTL: 61
- Internet Protocol Version 4, Src: 172.16.1.2, Dst: 172.16.5.2

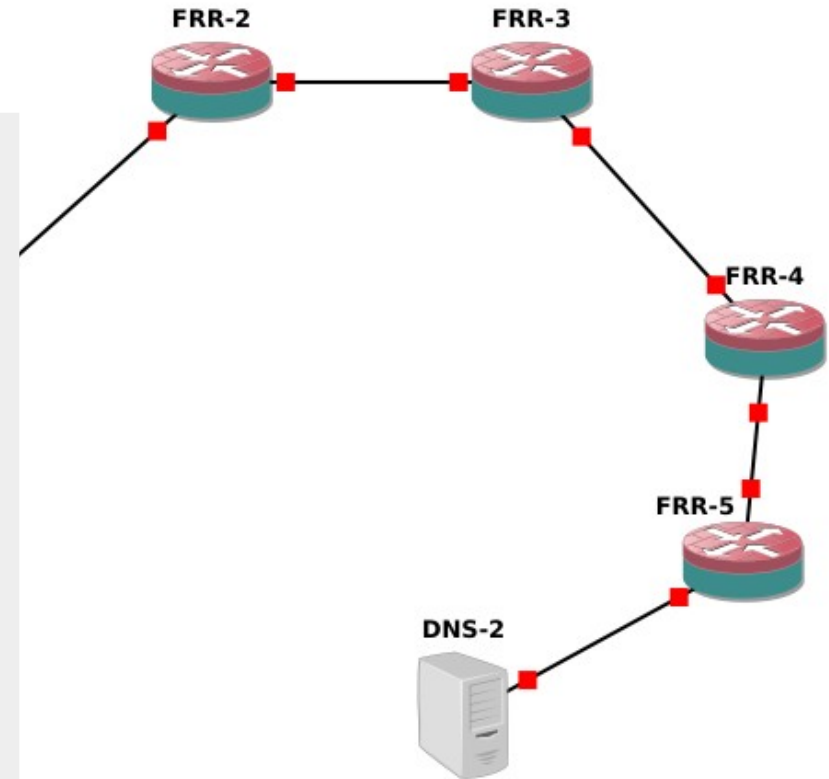
The packet bytes pane shows the raw data of the selected packet, with the ICMP Echo (ping) reply data visible.

LDP

■ Label Distribution Protocol

- Crear etiquetas sólo para rutas externas

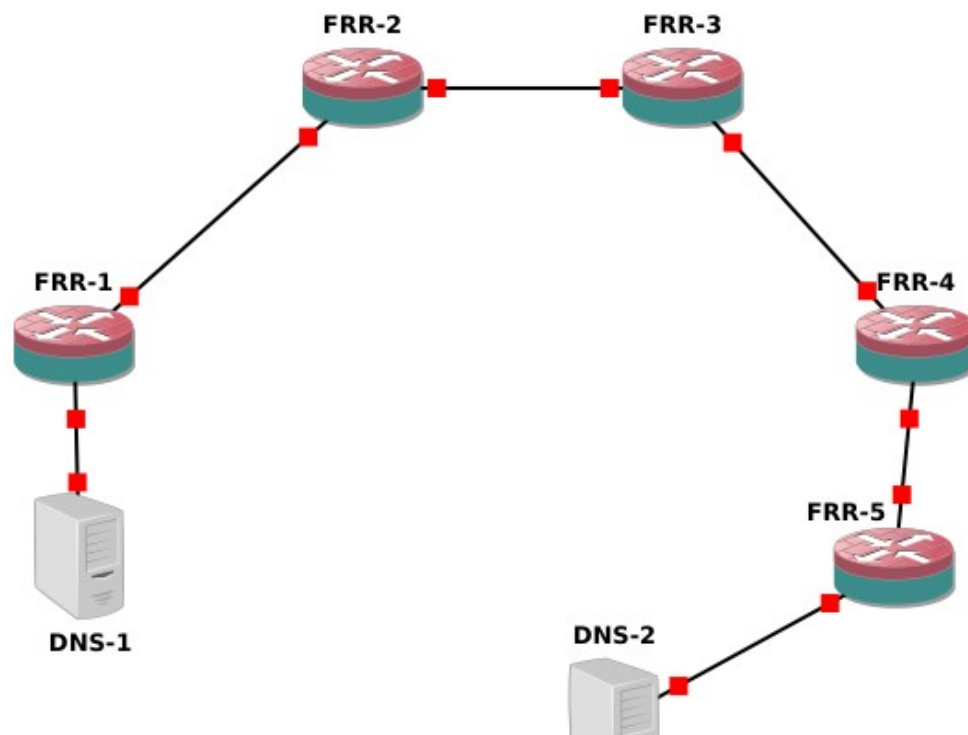
```
mpls ldp
router-id 10.0.12.1
ordered-control
!
address-family ipv4
discovery transport-address 10.0.12.1
label remote accept for external
label local allocate for external
!
interface eth1
!
exit-address-family
!
!
access-list external seq 10 deny 10.0.0.0/8
access-list external seq 20 permit any
```



LDP

■ Label Distribution Protocol

- Comprobar los LSP creados



```
FRR-1# show mpls ldp binding
```

AF	Destination	Nexthop	Local Label	Remote Label	In Use
ipv4	172.16.1.0/24	0.0.0.0	imp-null	-	no
ipv4	172.16.5.0/24	10.0.12.2	19	21	yes

LDP

■ Label Distribution Protocol

- Añadir un nuevo enlace entre FRR-2 y FRR-5
- Configurar el nuevo enlace en OSPF
- Añadirlo en mpls ldp
- Comprobar los LSP

