

Redes de Datos 2

Multiprotocol Label Switching (MPLS)

Temario

- Introducción
- Arquitectura y protocolos
- MPLS VPNs
- Ingeniería de Tráfico
- Calidad de servicio en IP y MPLS
- Otras aplicaciones de MPLS

Un poco de historia

Mediados de los 90', 2 tecnologías:

- ATM
 - tecnología de circuitos virtuales
 - Celdas de pequeño tamaño, tamaño uniforme
 - QoS garantizada
 - ALTA CAPACIDAD (para la época), facilitado por la búsqueda en la tabla de etiquetas
 - **CARO y complejo**. Difícil interacción con IP
- IP
 - Tecnología de datagramas
 - Paquetes de tamaño variable (+ eficiente en uso de la capacidad)
 - NO QoS
 - Capacidad moderada. Dificultad para realizar la búsqueda en la tabla de enrutamiento

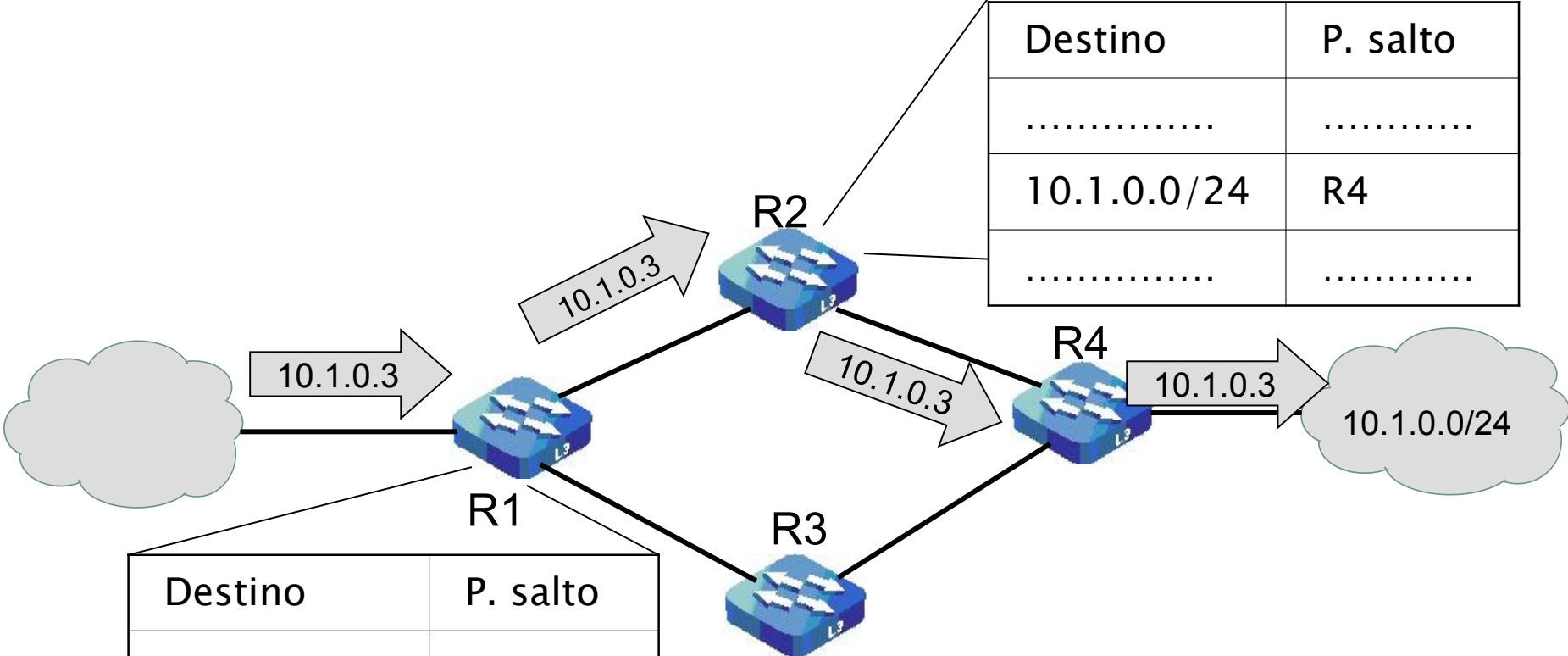
Historia (cont.)

- Idea: integrar las ventajas de ambos
- Diversos fabricantes plantean soluciones propietarias
- 1997: formación del grupo de trabajo MPLS en la IETF
- Objetivos iniciales:
 - Incrementar la velocidad de forwarding de los enrutadores
 - Facilitar la interacción IP/ATM
 - Estos objetivos ya no son relevantes

Redes de datagramas (IP)

- Cada paquete se encamina de forma independiente, en base a la dirección de destino
- Cada enrutador realiza la búsqueda en la tabla de enrutamiento (forwarding) para cada paquete y elige el próximo salto
 - Algoritmo “longest prefix match”
- Difícil implementar QoS, reserva de recursos, etc.

Encaminamiento IP



Destino	P. salto
.....
10.1.0.0/24	R4
.....

Destino	P. salto
.....
10.1.0.0/24	R2
10.1.0.0/16	R3
.....
0.0.0.0/0	R2

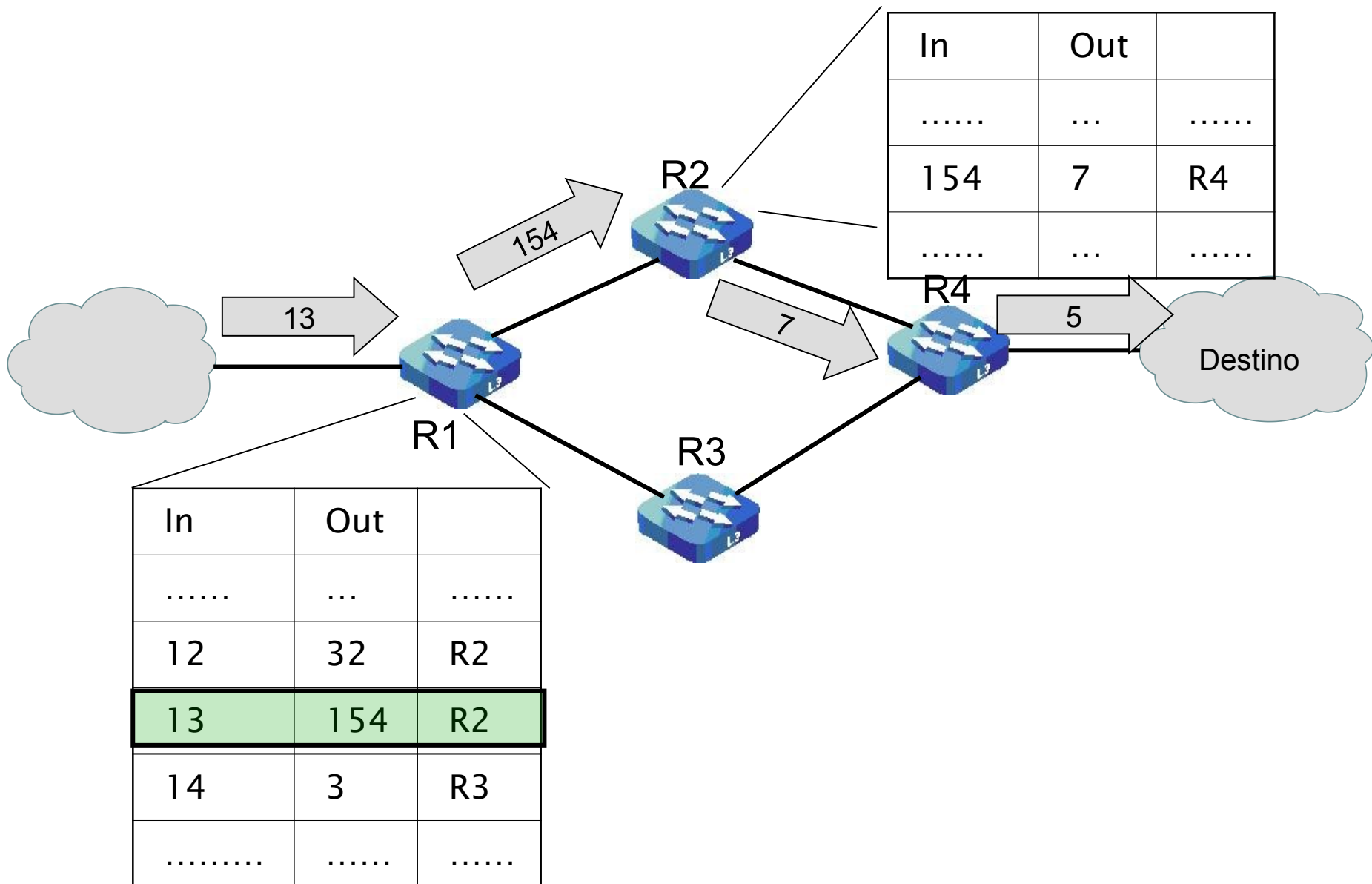
Redes de circuitos virtuales

- Redes de CV tradicionales se basan en tener un camino (circuito virtual) establecido antes de comenzar a enviar datos
 - Puede establecerse por señalización o administrativamente
 - La elección del camino lo hace la función de ruteo
- Se identifican 3 fases:
 - **Establecimiento** del CV
 - **Uso** del CV (envío de datos)
 - **Liberación** del CV

Identificadores de circuito virtual

- El camino se identifica por un número, genéricamente llamado “**identificador de Circuito Virtual**”
 - Distintas tecnologías lo llaman de otras maneras (**etiquetas**, DLCI, etc.)
- No es práctico tener un identificador de CV único para todo el circuito porque requeriría acuerdo y disponibilidad del mismo número en todos los nodos intermedios
 - Cada par de nodos acuerda un **identificador de CV local** para ese tramo
 - El CV se identifica **por la secuencia de los identificadores de todo el camino**
- Cada nodo debe mantener **tablas con los CVs establecidos**
- Cada nodo debe cambiar el identificador en el encabezado

Encaminamiento CV



Ventajas y desventajas de CV

- La búsqueda en la tabla de CV es mucho mas sencilla (y por tanto más rápida y barata)
- El camino se elige una vez, y se utiliza muchas (una por paquete)
 - Permite realizar reserva de recursos. Facilita QoS
- En principio todos los paquetes siguen el mismo camino
- Posibles dificultades ante la caída de nodos y enlaces
 - Debemos reconstruir los caminos
- Necesidad de establecer el CV para cada par origen-destino

Objetivos iniciales de MPLS

- Acelerar el forwarding de paquetes IP, utilizando el paradigma de los Circuitos Virtuales
- Desacoplar el enrutamiento y el forwarding
- Permitir una mejor integración de IP y ATM
- Servir como base para el desarrollo de nuevas aplicaciones y servicios
 - Ingeniería de tráfico
 - VPNs provistas por el ISP
 - Redes multiprotocolo (no solo IP)
 - Calidad de servicio

Realidades

- La búsqueda en la tabla de forwarding IP ya no es un cuello de botella
 - ASICs especializados
- ATM prácticamente ha desaparecido
 - Y nunca fue muy popular la integración mediante MPLS
- PERO las aplicaciones que facilita MPLS lo han hecho extremadamente popular en proveedores y algunas grandes empresas

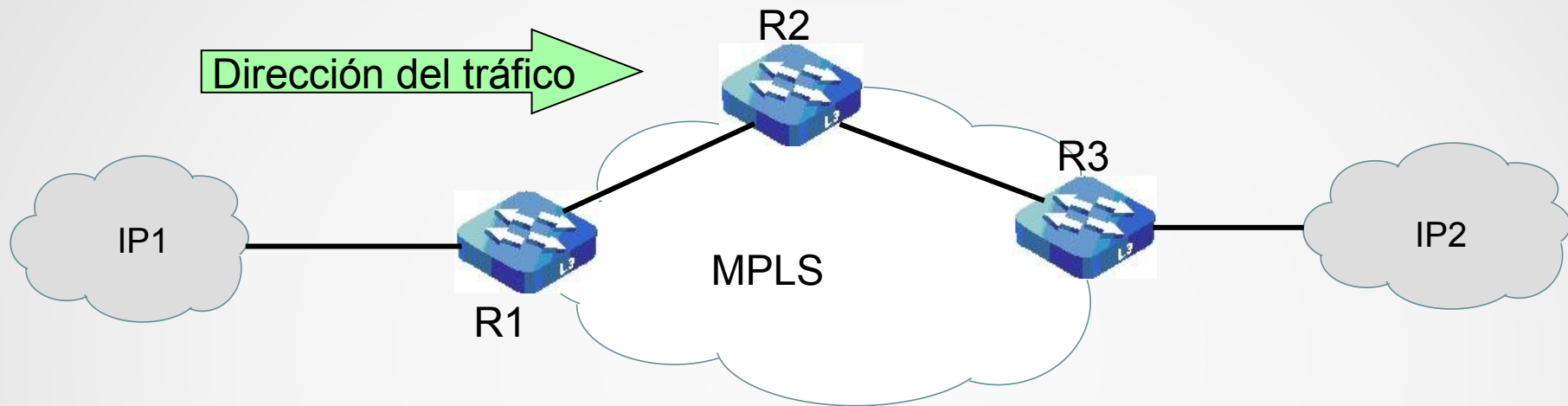
Multi Protocol Label Switching

- Definido en estándares de la IETF (múltiples RFCs)
- Arquitectura: RFC 3031, “Multiprotocol Label Switching Architecture”
- Por un lado, tendremos definiciones del formato de las etiquetas y el reenvío (forwarding)
 - Por ejemplo, RFC3032, “MPLS Label Stack Encoding”
- Por otro lado, los protocolos de señalización para definir los caminos y el significado de las etiquetas
- También definiciones de Operación y Mantenimiento, extensiones, generalizaciones...

Algunos nombres

- **Etiqueta (label)**: 2 posibles usos
 - Valor de identificador de Circuito Virtual (20 bits)
 - Coloquialmente, encabezado donde va el valor de la etiqueta (32 bits)
- **Upstream (“aguas arriba”)**: más cerca del origen
- **Downstream**: más cerca del destino
- **LSR: Label Switch Router**. Enrutador que basa el forwarding en la etiqueta MPLS
- **LER: Label Edge Router**. Enrutador en el borde de la red MPLS, encargado de recibir paquetes IP y reenviarlos con etiquetas
- **LSP: Label Switched Path**. Camino en la red MPLS (equivalente a un Circuito Virtual)

Nomenclatura



- LSR: R1, R2, R3
- LER: R1, R3
- R1 enrutador de ingreso
- R3 enrutador de egreso
- R1 upstream de R2
- R2 upstream de R3
- R3 downstream de R2
- R2 downstream de R1

¿Qué es un LSP (Label Switched Path)?

- Camino desde el enrutador de entrada al enrutador de salida, formado por una sucesión de valores de etiqueta
 - Similar a un circuito virtual tradicional
- Cada enrutador tendrá la relación entre la etiqueta de entrada y la etiqueta de salida
- El enrutador de ingreso le agregará una etiqueta
- Los enrutadores intermedios harán “label swap”, simplemente cambiando la etiqueta de entrada por otra de salida
- El enrutador de egreso retirará la etiqueta y encaminará el paquete normalmente

LSP

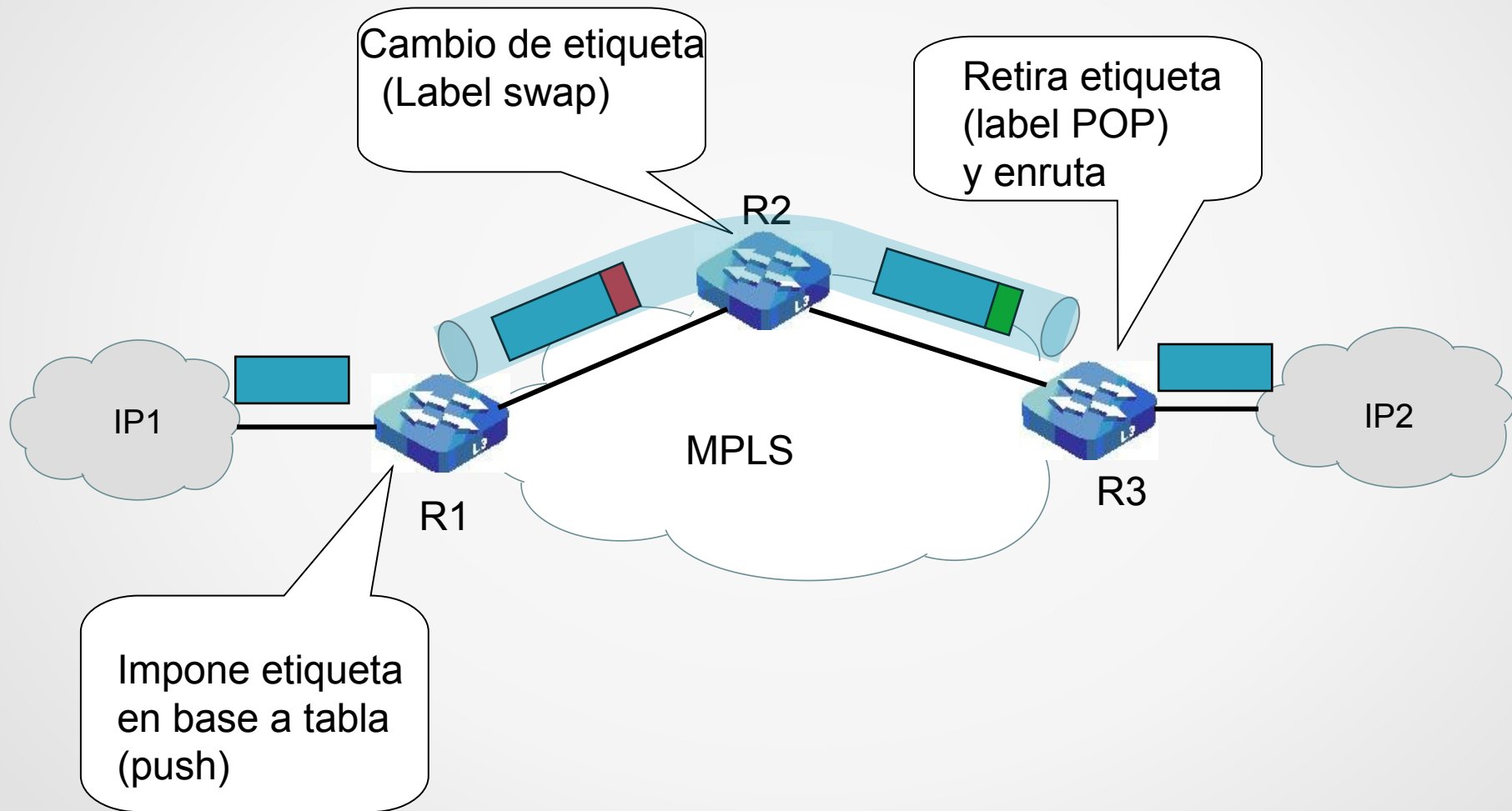


Tabla de etiquetas simplificada

- **Simplificadamente**, cada enrutador mantendrá una tabla con el mapeo de etiqueta de entrada a etiqueta de salida

Interfaz entrada	Etiqueta entrada	Interfaz Salida	Etiqueta Salida
i1	2001	o4	18
i2	2001	o3	99
i1	2002	o2	143
*	2003	o4	17

Espacio de etiquetas

- Las etiquetas son locales
- En general la asigna el que la va a recibir
- Dos políticas posibles de asignación de las etiquetas:
 - Espacio de etiquetas por interfaz: se asignan independientemente por cada interfaz
 - Espacio de etiquetas por plataforma: independientemente de la interfaz, el mismo valor de etiqueta tendrá el mismo comportamiento de salida

Funciones necesarias

- Las funciones primitivas a realizar sobre los paquetes son solo 3:
 - **Label push**: agregar una etiqueta
 - **Label swap**: cambiar una etiqueta por otra
 - **Label pop**: quitar una etiqueta

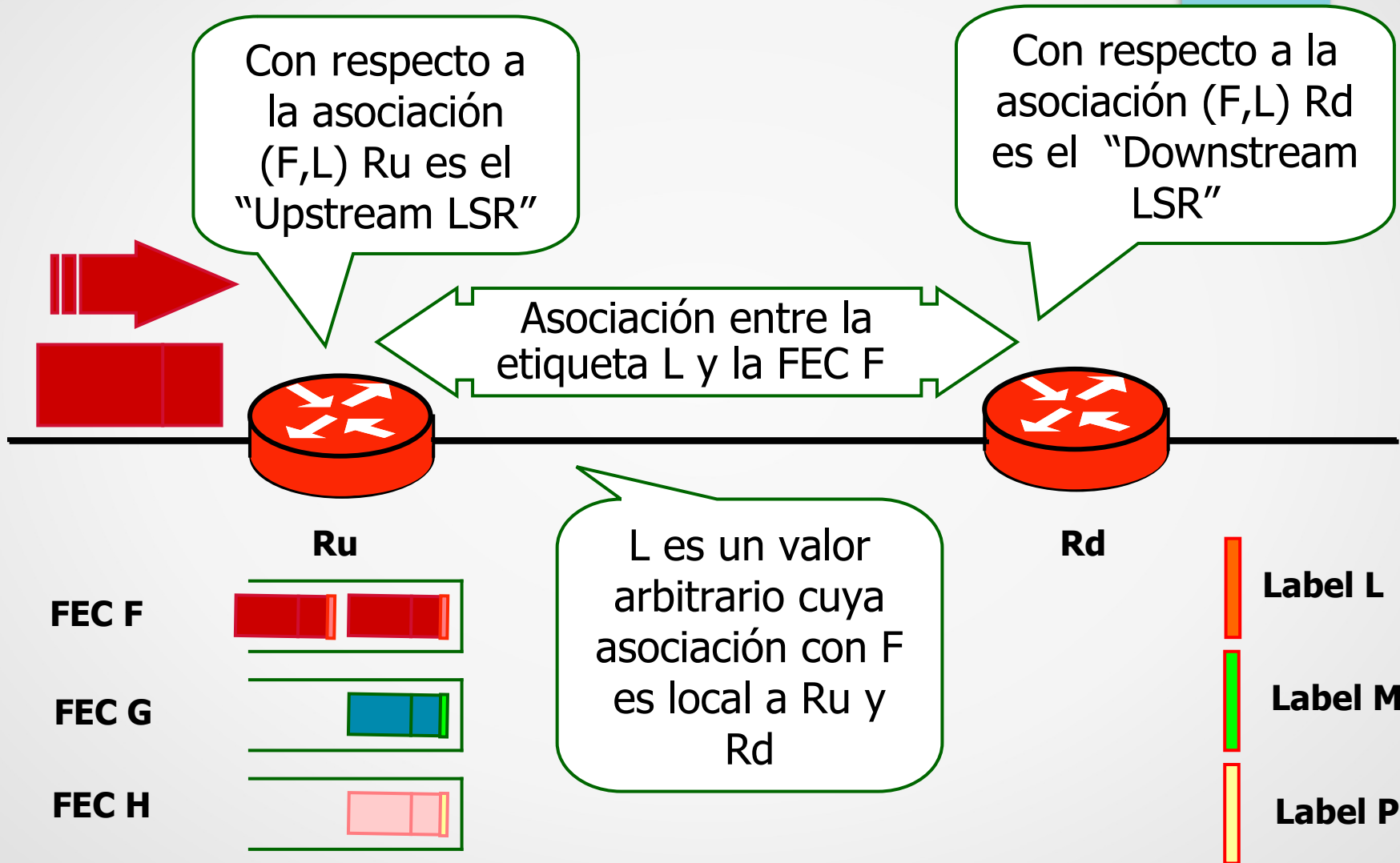
Clases de equivalencia (FEC)

- Se hace una partición de todos los paquetes en clases de equivalencia de encaminamiento, FECs (del Inglés, Forwarding Equivalence Class)
- Grupos de paquetes que serán reenviados de la misma manera
- Por ejemplo, en enrutamiento IP tradicional, las clases se determinan de acuerdo al algoritmo de “longest prefix match” sobre la dirección IP de destino
- Los paquetes de la misma clase seguirán todos el mismo camino

FEC

- El concepto de FEC provee gran flexibilidad y escalabilidad
- En MPLS la FEC se determina en el LER de entrada, no en cada enrutador (LSR)
- Paquetes con distinto destino pueden agruparse en la misma FEC
- Posibles criterios de FEC
 - IP Destino
 - IP Origen – IP Destino
 - IP Origen - IP Destino - puerto origen – puerto destino
 - “Todos los que pertenecen a la VPN de Cliente1”
 - Etc.

LSR. Asociación Etiqueta - FEC



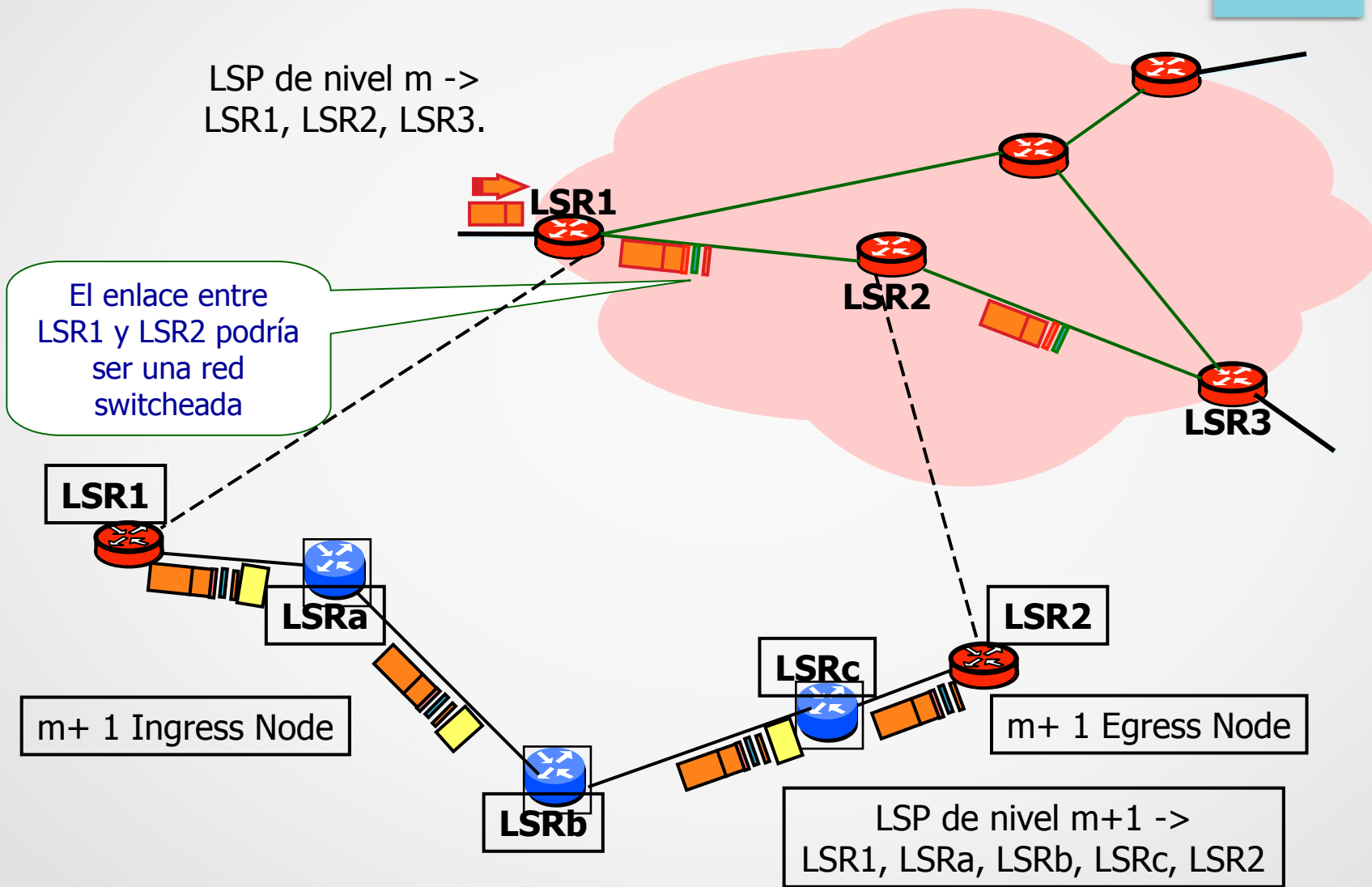
Stack de etiquetas

- Es útil tener un modelo más general en el cual un paquete pueda llevar un conjunto de etiquetas de modo LIFO (Last In First Out): "label stack".
- El procesamiento se basa siempre en la etiqueta externa
- Este mecanismo habilita:
 - MPLS jerárquico
 - Agregación
 - Aplicaciones de transporte y VPNs

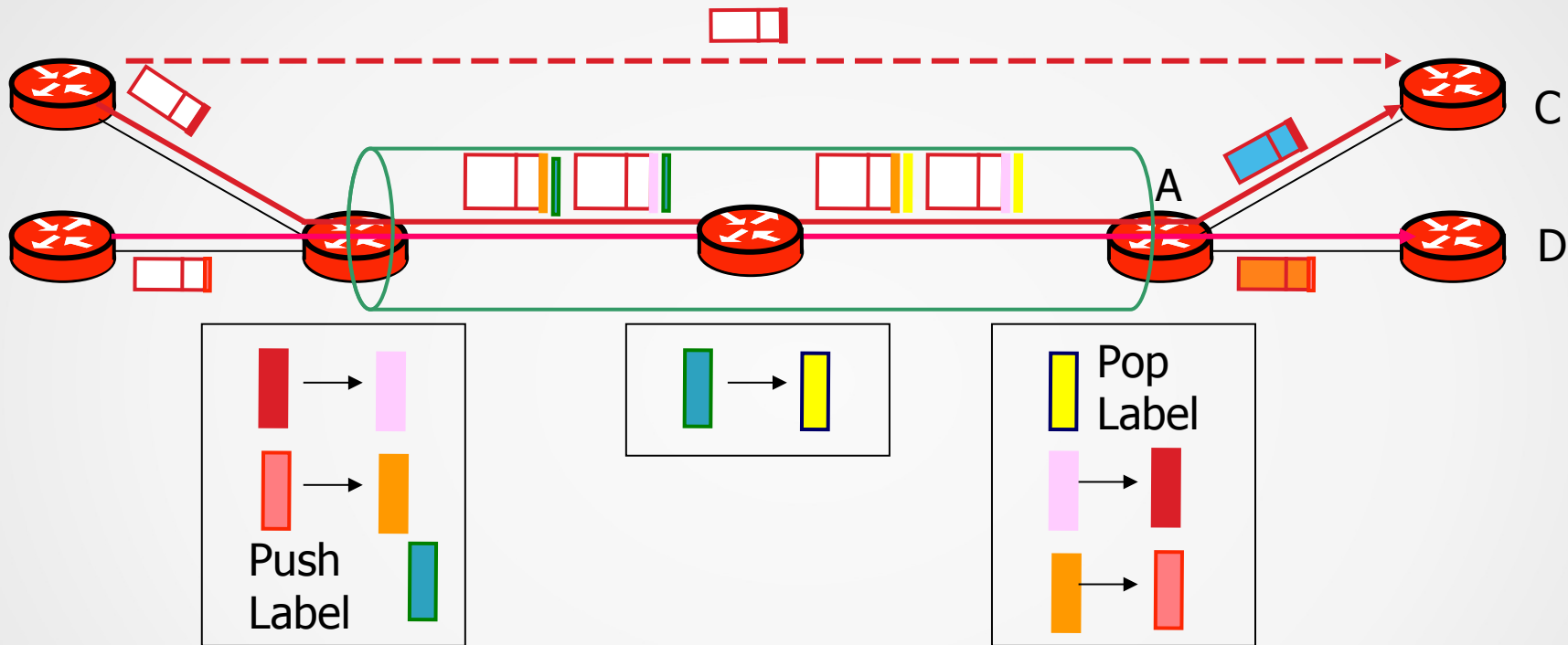


Bottom of Stack (bit S en 1)

Ejemplo de stack: MPLS Jerárquico



Ejemplo de stack: Agregación



Para enviar un paquete al LSR A no se necesita saber si su destino es C o D.

La agregación simplifica las tablas en el corazón de la red.

Mapeos

- Conceptualmente varias tablas indican qué hacer con el tráfico
 - Implementaciones pueden variar
- Tabla NHLFE: Next Hop Label Forwarding Entry. Indica las acciones a realizar sobre el paquete
- FTN: FEQ a NHLFE
 - Mapea la clase de equivalencia (FEC) de entrada, a la operación de salida
 - Para los paquetes que vienen sin etiquetar
- ILM: Etiqueta a NHLFE
 - Realiza el mismo mapeo para paquetes ya etiquetados

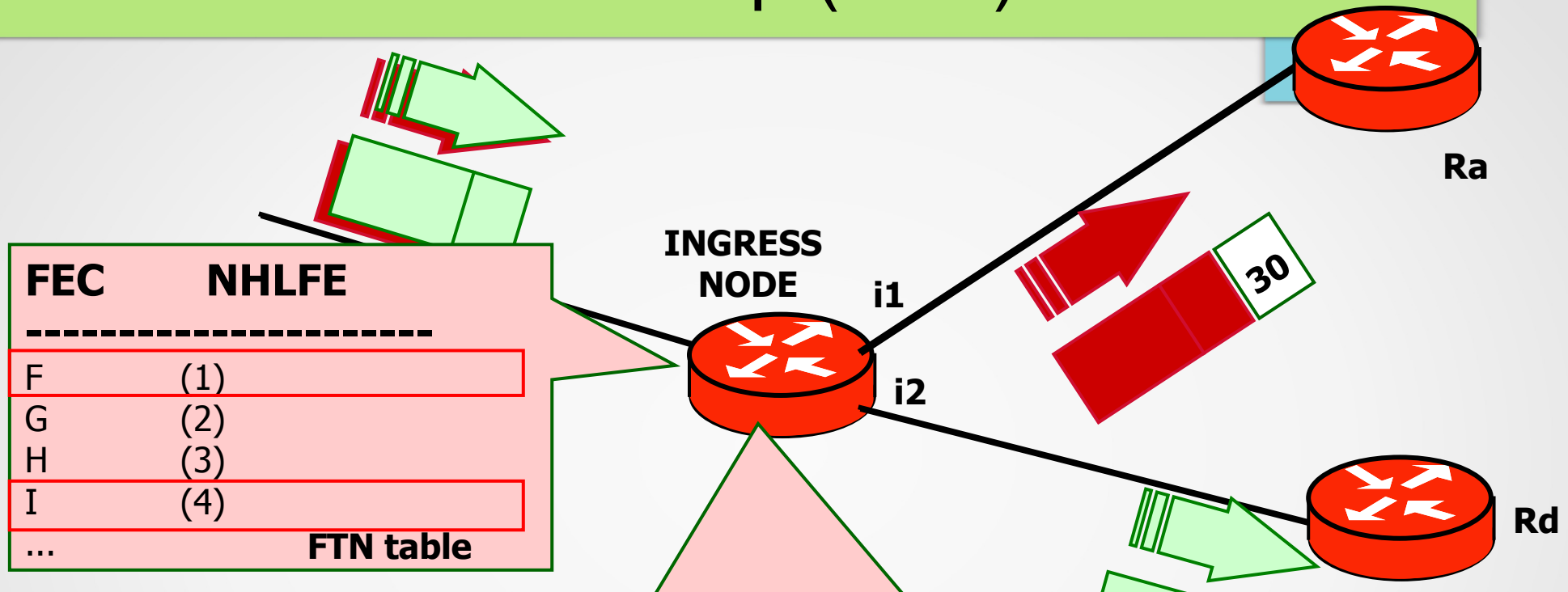
NHLFE: Next Hop Label Forwarding Entry

- Tabla que dice que hacer con el paquete: modificaciones a hacer, y a donde enviarlo
- Posibles acciones
 - Eliminar una etiqueta (pop)
 - Agregar una etiqueta (push)
 - Cambiar la etiqueta exterior por otra
- Información de reenvío
 - Próximo salto
 - Interfaz de salida
 - Etc.

NHLFE

NH.	Operation	Label	Interf.	Data Link
(1) Ra	Label Swap	30	i1	Frame Relay
(2) Rb	Label Swap	43	i1	Frame Relay
(3) Rc	Label Pop	--	i0	Ethernet
(4) Rd	Label Swap	56	i2	Ethernet
	Label Push	70		
...				

FEC-to-NHLFE Map (FTN)



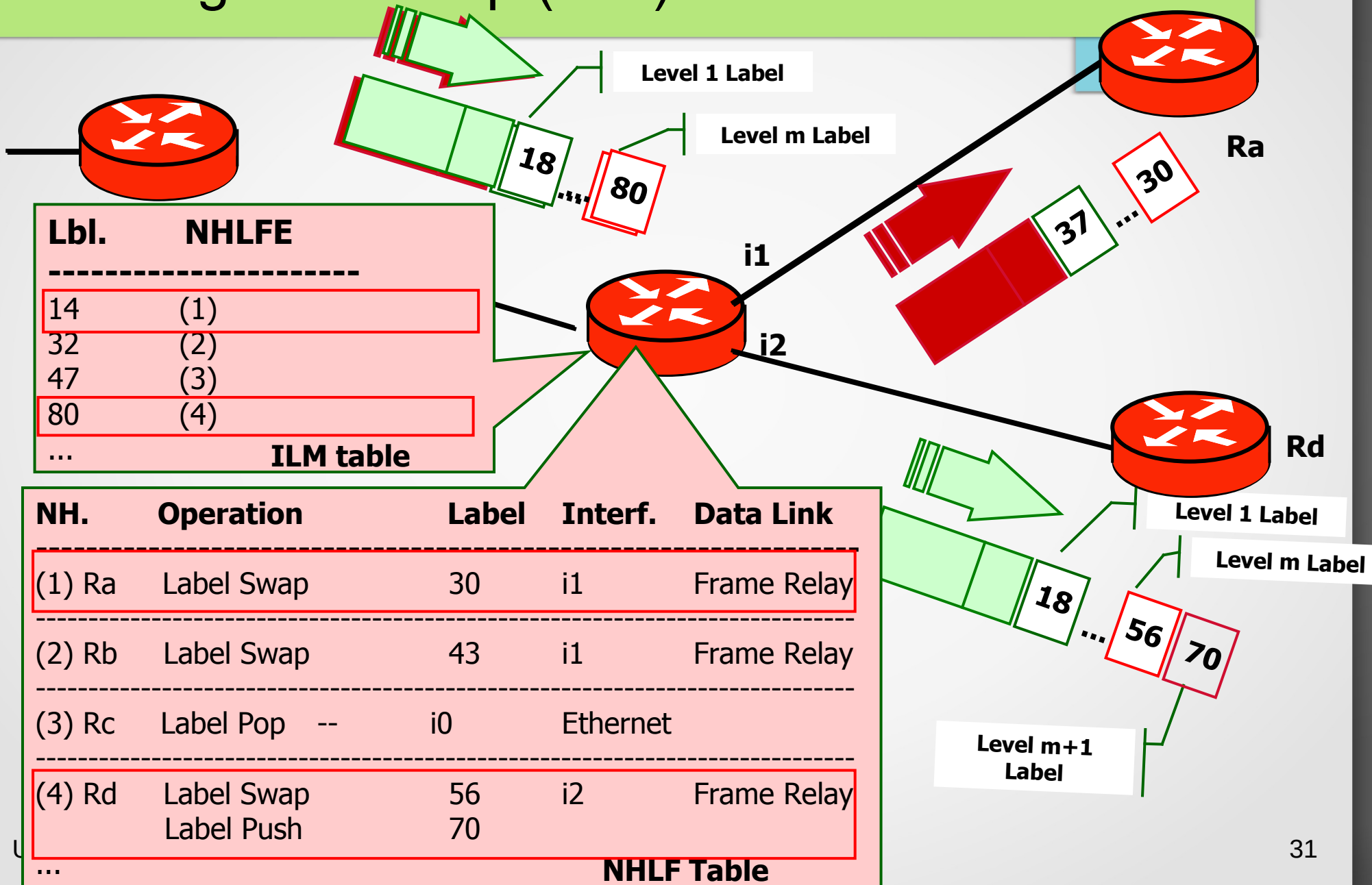
FEC	NHLFE
-----	-----
F	(1)
G	(2)
H	(3)
I	(4)
...	

FTN table

NH.	Operation	Label	Interf.	Data Link
-----	-----	-----	-----	-----
(1) Ra	Label Push	30	i1	Frame Relay
-----	-----	-----	-----	-----
(2) Rb	Label Push	100	i2	Frame Relay
-----	-----	-----	-----	-----
(3) Ra	Label Push	300	i1	Frame Relay
-----	-----	-----	-----	-----
(4) Rd	Label Push	200	i2	Frame Relay
U	...			

NHLF Table

Incoming Label Map (ILM)



Codificación de etiquetas

- RFC 3032
- Encapsulamiento genérico (frame mode)
 - Usado en Ethernet, PPP, etc.
 - Se agrega un encabezado (**shim header**) entre capa 2 y capa 3
- Cell mode (representación en VPI/VCI ATM)
 - En desuso

Formato del encabezado MPLS

- El stack de etiquetas es representado por una secuencia de "label stack entries". Cada entrada es representada por 4 bytes.
- El paquete de la capa de red sigue inmediatamente después del stack de etiquetas. La última entrada tiene el bit S seteado.
- Las entradas del stack de etiquetas aparecen después del encabezado de la capa de enlace pero antes de cualquier encabezado de la capa de red.

Codificación de Label Stack Entries



- Etiqueta (label): valor de 20 bits
- Exp: 3 bits, originalmente “experimentales”. Usados para QoS
- TTL: Time to live (8 bits). Contador de saltos, sirve para descartar paquetes en caso de loops
- Bit S: Bottom of Stack. Indica si es la última etiqueta del stack

Etiquetas especiales

- Los valores 0 a 15 están reservados
- 0 - “IPV4 Explicit NULL”
- 1 - “Router Alert”
- 2 - “IPV6 Explicit Null”
- 3 - “Implicit NULL” – se distribuye pero no aparece en el cabezal

Penultimate hop popping

- PHP: retiro de la etiqueta en el penúltimo nodo
- Sin PHP, el último nodo debería retirar la etiqueta, y luego hacer la búsqueda de acuerdo al siguiente encabezado
 - Opción pensada para optimización
- Se acuerda entre el último y penúltimo (se anuncia la etiqueta “implicit null”)
- No siempre se usa
 - Depende del fabricante y la configuración

Manejo del TTL

- El campo TTL cumple la misma función que el correspondiente de IP, descartar paquetes que quedaron en loop
- El TTL de MPLS se setea en el LER de ingreso
- Usualmente se copia del campo TTL de IP
 - Permite que el traceroute funcione aunque pasemos por una red MPLS
 - Puede evitarse este comportamiento (para que los nodos MPLS no sean visibles en el traceroute)
- Si el contenido no es IP, se pone un TTL genérico al ingreso a la red

Funciones de los distintos nodos

- LER de ingreso
 - Clasificar el paquete entrante en la FEC correspondiente
 - De acuerdo al mapeo FEC - NHLFE, agregarle las etiquetas que corresponda y enviarlo al próximo salto
- LER de tránsito
 - Realizar el swap de etiqueta correspondiente (mapeo ILM - NHLFE)
 - Enviar el paquete con tag al próximo salto
- LER de egreso
 - Retirar (al menos) el tag externo. Si lo que obtengo es un paquete tagueado, enviarlo de acuerdo a la tabla NHLFE. Si no está tagueado, reenviarlo utilizando la tabla de enrutamiento
 - Puede haber recibido el paquete sin tag si se realizó PHP (retiro de la etiqueta en el penúltimo salto)

Plano de control: distribución de etiquetas

- Podemos configurar los mapeos de forma estática
 - Muy trabajoso
- En general, protocolos de distribución de etiquetas
- Intercambiarán información de las asignaciones de cada FEC a un valor de etiqueta
- Casi todos los protocolos actuales precisan tener una red IP corriendo previamente
 - Ya que intercambian la información mediante IP

Protocolos de distribución de etiquetas

- LDP: Label distribution protocol
 - Propaga el mapeo de rutas internas a etiquetas
 - El encaminamiento obtenido coincide con el obtenido sin MPLS
- BGP
 - Carrying Label Information in BGP-4 (RFC ~~3107~~ - 8277)
 - Para rutas de VPNs capa 3: Multiprotocol BGP
- RSVP-TE
 - Utilizado para ingeniería de tráfico
 - Junto con OSPF-TE (o ISIS-TE)
 - Permite fijar caminos administrativamente, o de acuerdo a restricciones como ancho de banda

Asignación y distribución de etiquetas

- Las etiquetas siempre las genera el enrutador downstream (más cerca del destino)
 - Recibirá paquetes con las etiquetas que él asignó
- 2 modos de distribución de etiquetas:
 - Modo de control de LSP Ordenado
 - Si no soy el LSR de egreso, espero a recibir etiquetas del enrutador aguas abajo antes de enviar mis etiquetas
 - **Modo de control de LSP Independiente**
 - **En cuanto genero mis propias asociaciones de etiquetas, puedo enviarlas**

Modos de distribución: no solicitado vs. a demanda

- **En el modo no solicitado, distribuyo todas las asociaciones locales a todos los enrutadores**
 - **Unsolicited Downstream**
- En el modo a demanda (downstream-on-demand), solo envío las asociaciones que los demás enrutadores me solicitan
 - Utilizado principalmente en ingeniería de tráfico y ATM

Modos de retención

- **Modo liberal: guardo todas las asociaciones que me hayan enviado, aunque ahora no las vaya a usar**
 - Consume más memoria
 - Convergencia más rápida cuando los precise
- **Modo conservador: solo retengo las asociaciones que preciso en este momento**
 - Tendré que solicitarlas y esperar a recibirlas nuevamente cuando cambie la topología

Espacios de etiquetas

- **Espacio global: la etiqueta tiene un significado global, no importa la interfaz de entrada del paquete**
- Espacio por interfaz: cuando se genera una etiqueta para la FEC tomando en cuenta la interfaz de entrada
 - El mismo valor de etiqueta, para distintas interfaces de entrada, puede mapear a distinta FEC

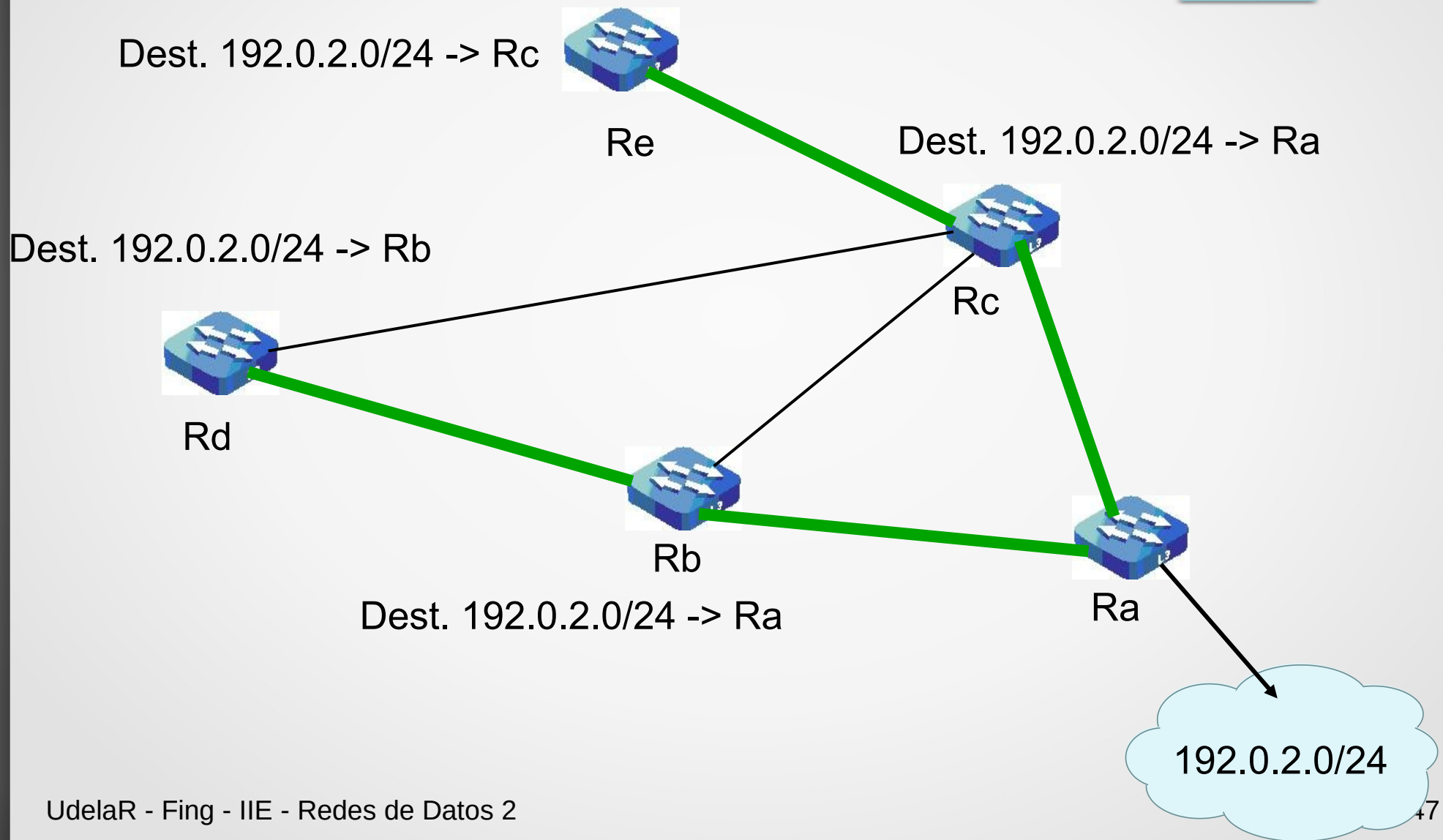
Resumen por protocolo (valores default)

Método de distribución	Control	Distribución	Retención	Espacio de etiquetas
LDP en modo frame	No ordenado	No solicitado	Liberal	Por plataforma
LDP en modo Celda	Ordenado	A demanda	Conservador	Por interface
RSVP-TE	Ordenado	A demanda	Conservador	Por plataforma

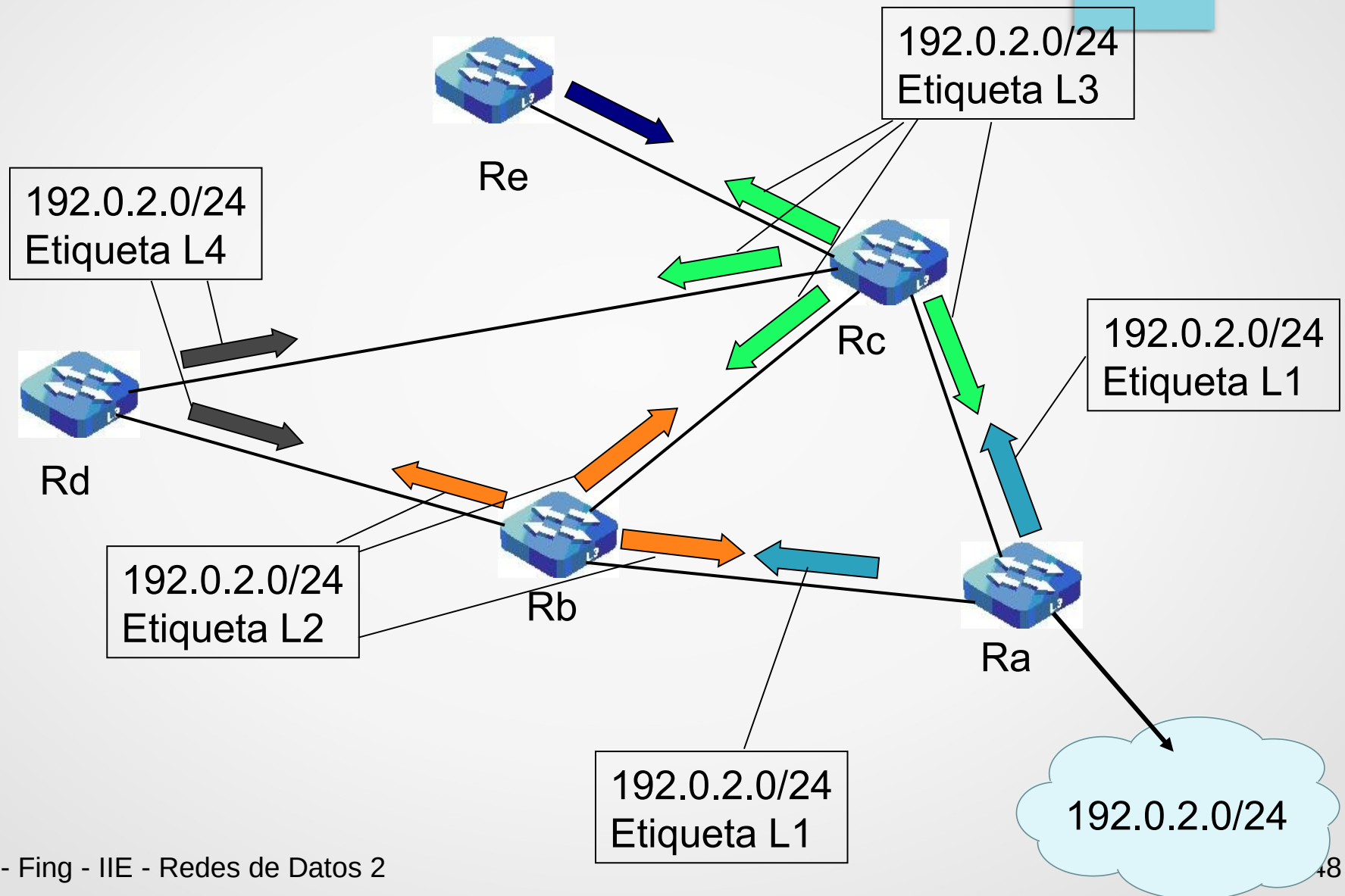
LDP: Label Distribution Protocol

- RFC 5036 (anteriormente RFC 3036)
- En el funcionamiento por defecto, LDP asigna una etiqueta a cada prefijo de la tabla de enrutamiento del IGP
 - Rutas estáticas, conectadas, aprendidas por OSPF, IS-IS, etc.
- Anuncia esta asignación a todos los vecinos LDP
- En la dirección downstream, utiliza la etiqueta que le anuncie el próximo salto de la tabla de enrutamiento
- Obtenemos el mismo camino que obtendríamos sin MPLS

Ejemplo: Caminos internos (IGP)



Ejemplo: LDP



Detalles sobre el ejemplo

- Ra posiblemente anuncie la etiqueta “implicit null”, de forma que Rb y Rc hagan PHP (penúltimo retira etiqueta)
- Aunque no los usen, todos los enrutadores guardarán todas las asociaciones (por si cambia el mejor camino)

4 funciones básicas de LDP

- Descubrimiento de vecinos
 - Paquetes Hello sobre UDP puerto 646
 - Similar (en función) a OSPF
- Establecimiento y mantenimiento de sesión
 - Sobre TCP puerto 646
- Anuncio de etiquetas
- Notificación

Descubrimiento de vecinos

- 2 tipos de sesiones:
 - Vecinos directamente conectados
 - Vecinos NO directamente conectados
 - Deben ser alcanzables por IP
 - No se descubren automáticamente por LDP
 - No los veremos ahora
- En el caso directamente conectado, se envían mensajes LDP Hello encapsulados en UDP, a la dirección 224.0.0.2 (all routers)
- Lo usual es que haya que configurar explícitamente las interfaces que van a intercambiar mensajes LDP
- Por defecto Hello cada 5 segundos, hold-time 15 seg.

Establecimiento de sesión

- En el Hello se indica la dirección a la cual establecer la sesión TCP (router-id)
 - Por defecto, elegirá una dirección de loopback del equipo (la mayor)
 - Puede configurarse explícitamente
- Esa dirección deberá ser alcanzable!!!
 - Si no se establece la sesión, verificar conectividad entre ambas direcciones (ambos router-id)
- Se elegirá uno como activo (IP mayor), este establecerá la sesión TCP

Establecimiento de sesión (cont.)

- Intercambio de parámetros (modo (no solicitado o a demanda), hold-time, etc.)
- Luego se enviarán keepalives periódicos
 - Independientes de los Hello por UDP
- Aunque exista más de un enlace entre dos vecinos, se establecerá una única adyacencia

Mensajes

- Detección de vecinos
 - *Hello*
- Mantenimiento de la sesión
 - *Initialization, Keepalive*
- Manejo de errores
 - *Notification*
- Manejo de mapeo de etiquetas e IPs de los vecinos
 - *Address, Address-withdraw*
 - *Label Request*
 - *Label Abort Request*
 - En modo On Demand
 - *Label Mapping*
 - *Label Withdraw*
 - *Label Release*

Intercambio de IPs de interfaces

- *Address*
 - El enrutador envía la lista de todas las direcciones IP en las interfaces donde LDP está activo
- *Address-withdraw*
 - Permite dar de baja una o varias direcciones IP previamente anunciadas
- Con estos 2 mensajes, cada LSR mantiene una lista de las IPs de las interfaces del vecino
- Necesario para saber si es el “next-hop” de una ruta

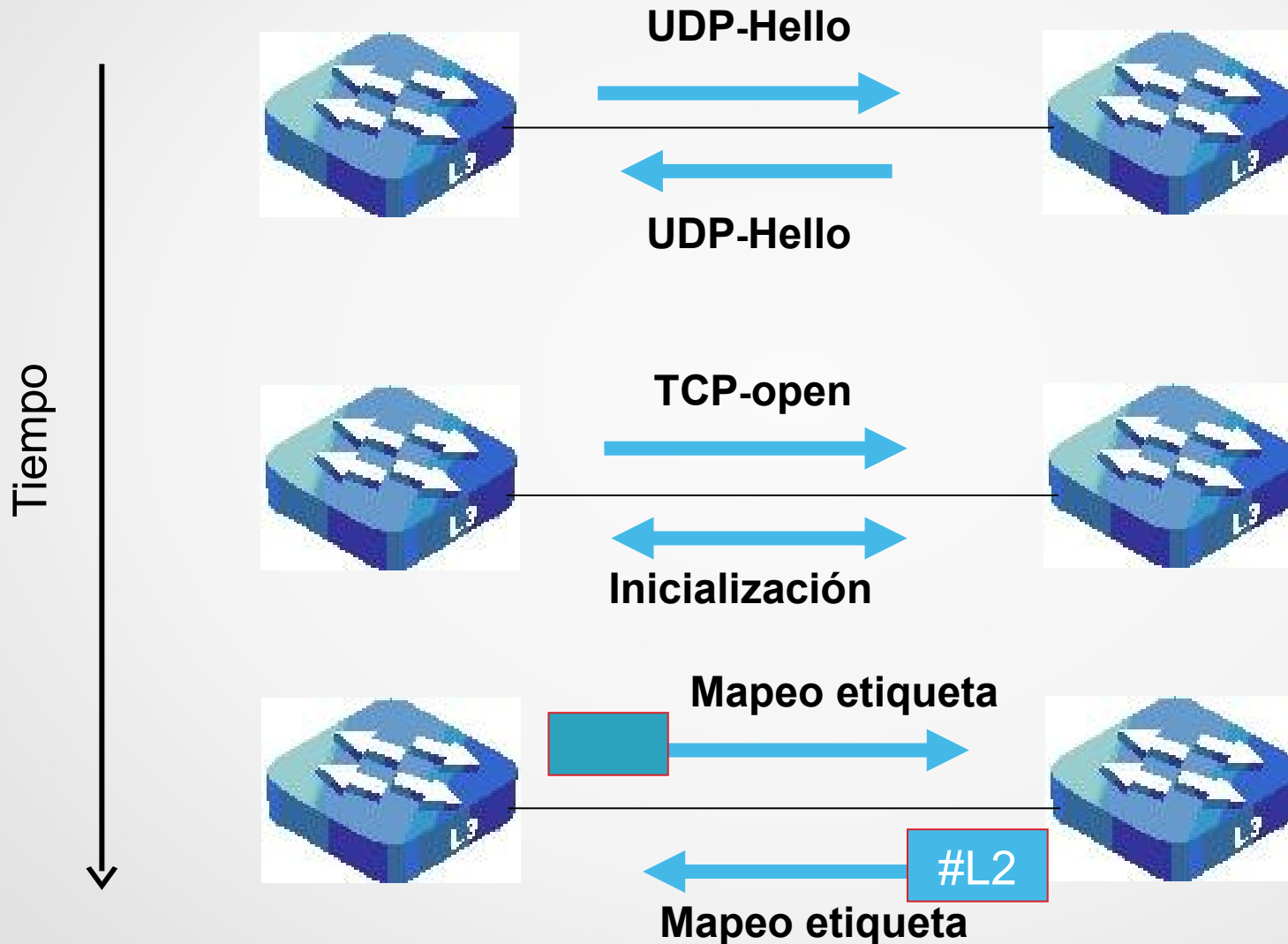
Mapeo de etiquetas (label mapping)

- El mensaje “*label mapping*” incluye una FEC, y un valor de etiqueta a asociar
- En LDP básico, la FEC corresponde con una red (una IP y una máscara o longitud de prefijo)
- Se avisa que el mapeo deja de ser válido con el mensaje “*label withdraw*”

Notificación

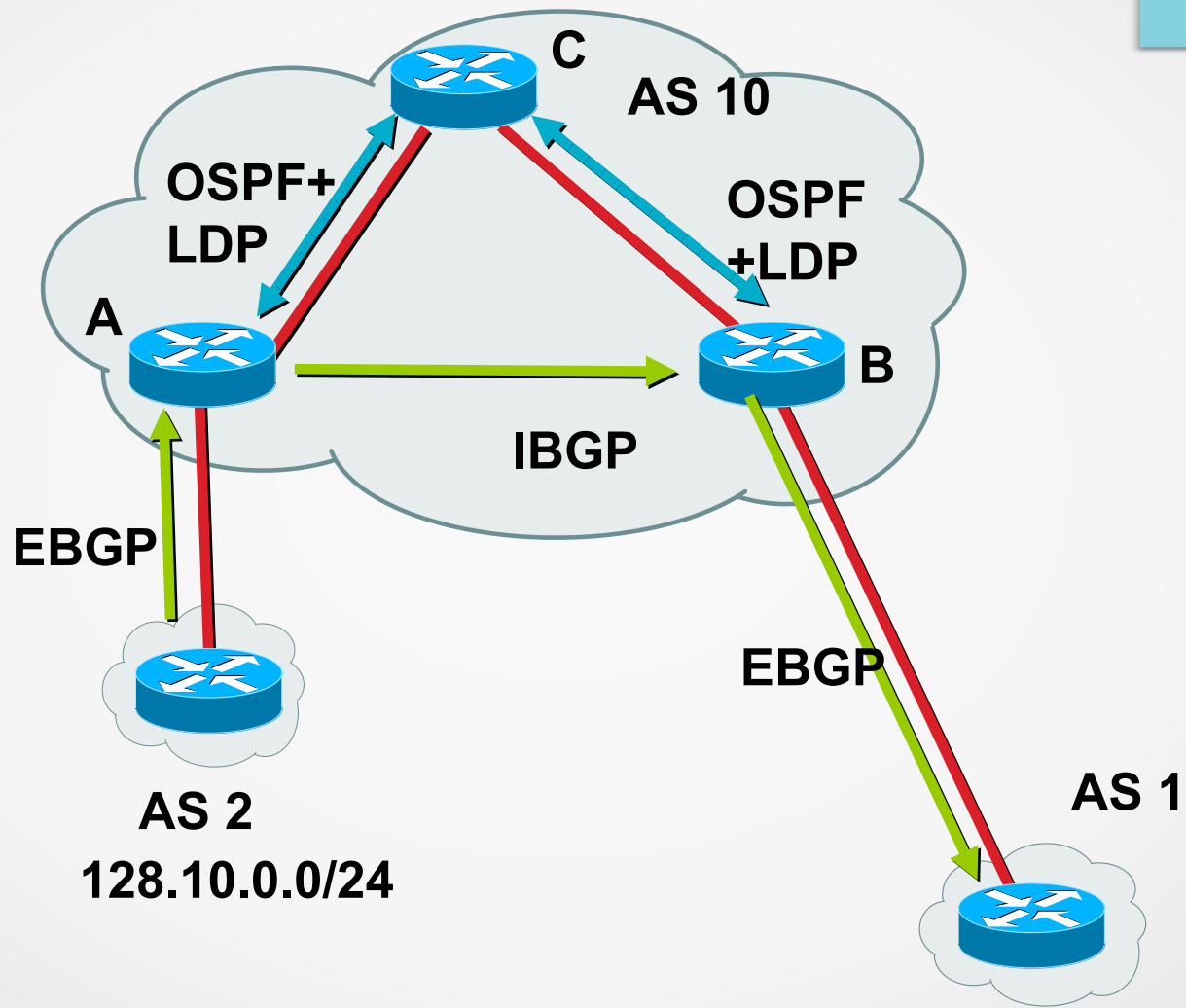
- Informe de problemas
- Errores
 - solo para errores fatales
 - Envía el mensaje y da de baja la sesión
 - Borra todos los mapeos de etiquetas aprendidos de ese vecino
- Notificaciones “informativas”
 - Warnings, sigue funcionando

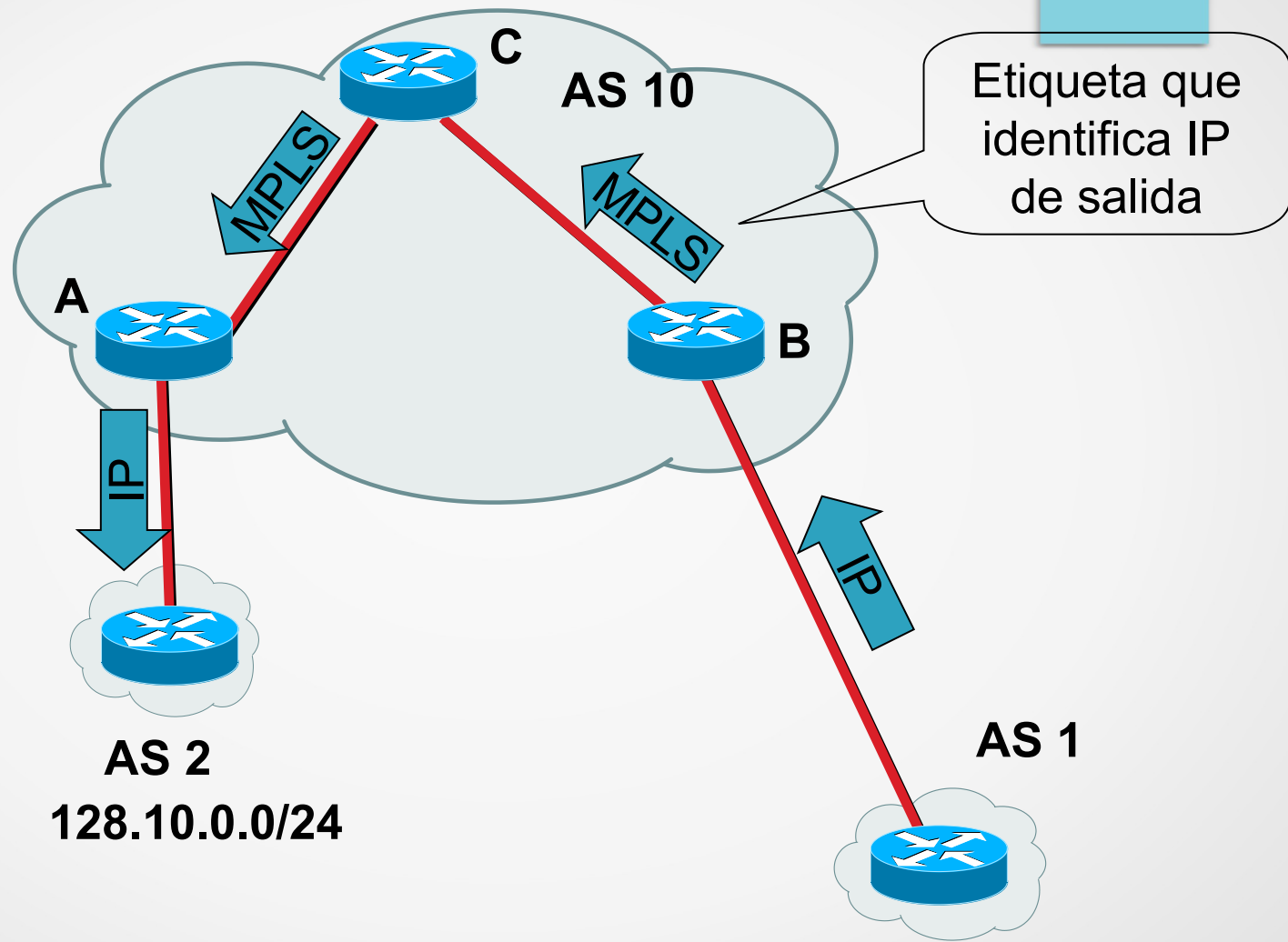
Esquema de funcionamiento



Interacción de MPLS/LDP con BGP

- Cuando un prefijo se propaga por BGP interno, se mantiene el próximo salto
- Podemos definir una FEC que sea “todos los paquetes que van al mismo próximo salto de BGP”
- A cada prefijo de BGP, le asociamos la etiqueta que corresponde con el next-hop
 - Muchos prefijos utilizarán la misma etiqueta
 - Esto se hace automáticamente
- Los enrutadores internos no precisarán hablar BGP, solo saber cómo llegar al next-hop

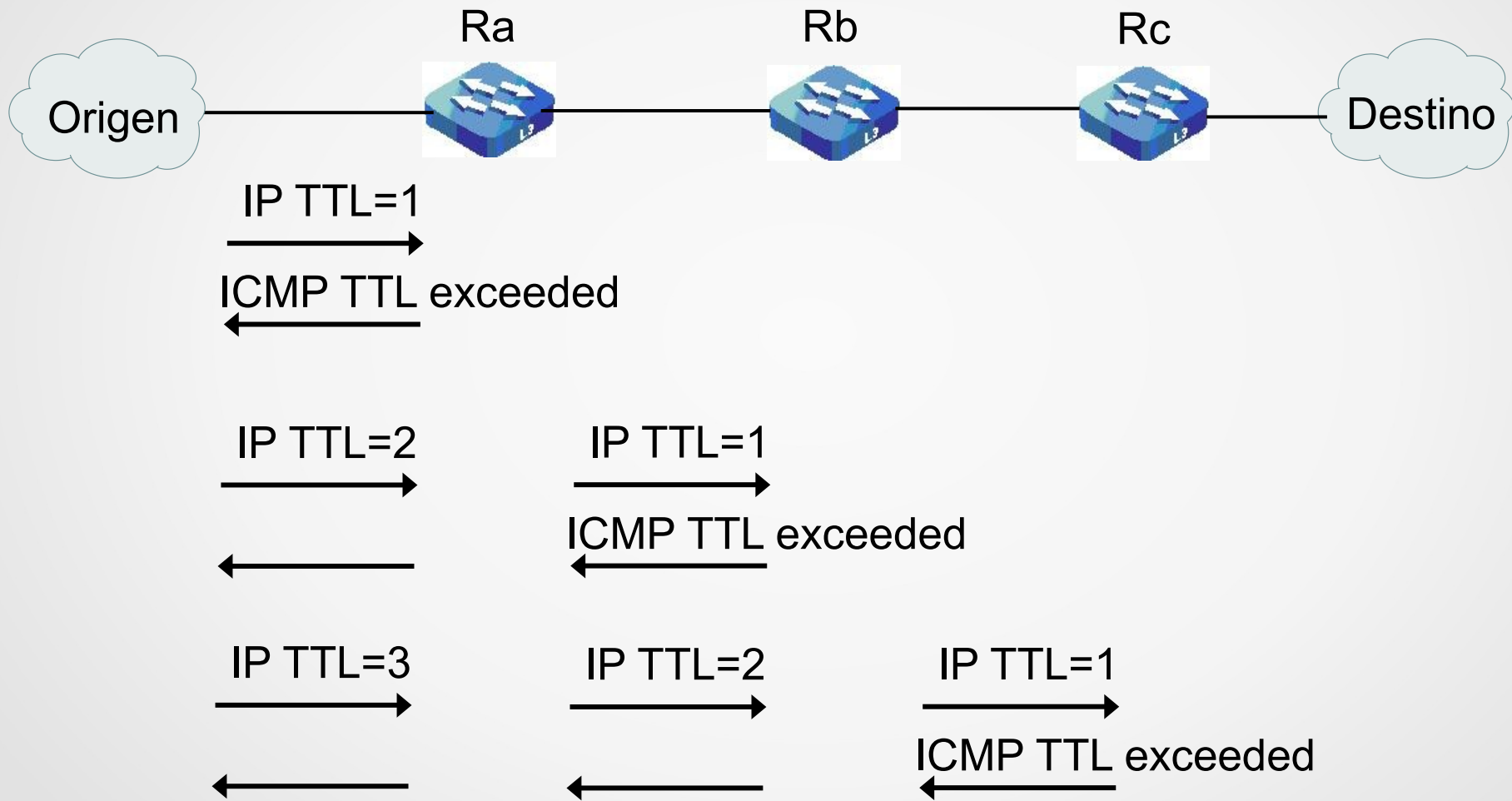




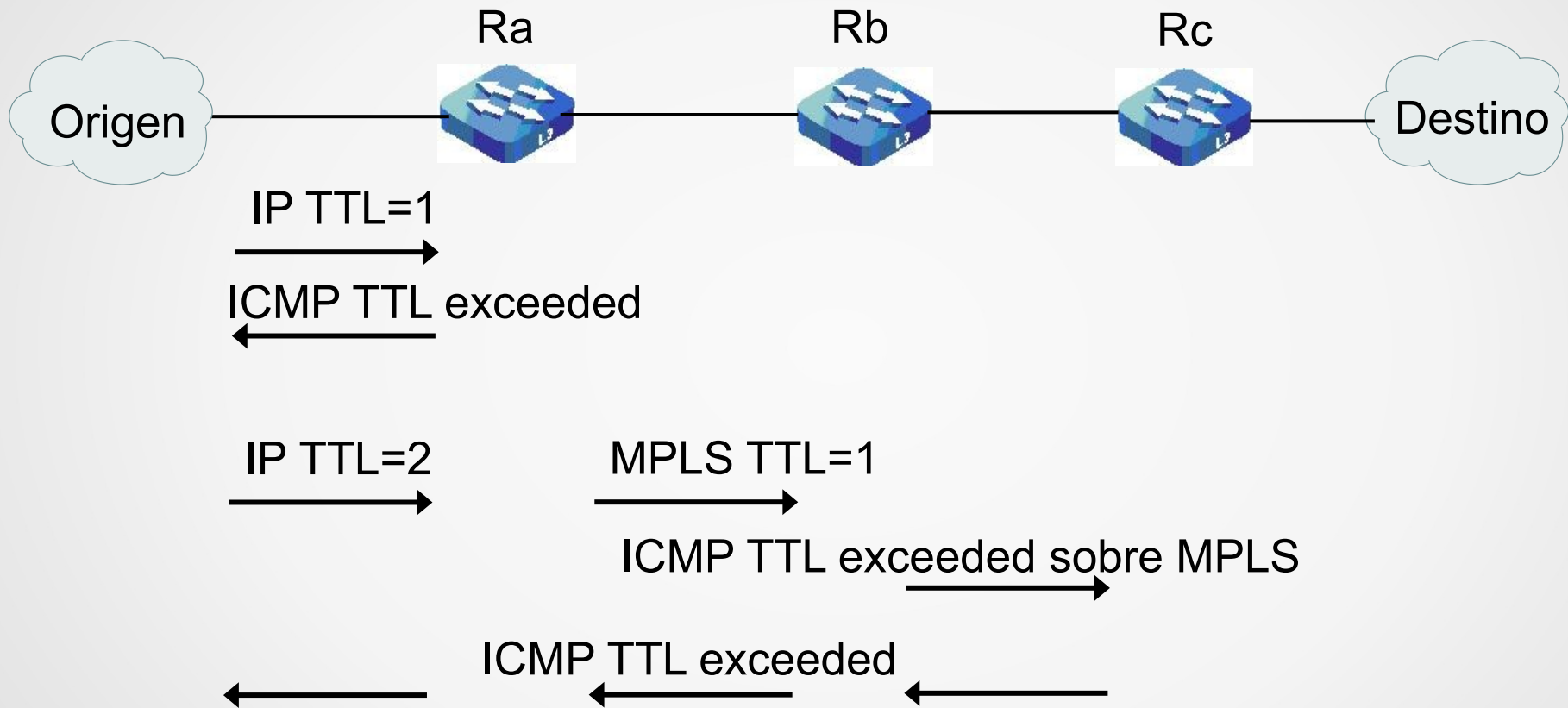
Traceroute en la red MPLS

- Importante herramienta de debug
- ¿cómo hacerla funcionar en MPLS?
 - Puede suceder que el router intermedio no tenga una ruta para llegar al origen del traceroute
 - Como el ejemplo que vimos de BGP
 - O cuando tengamos VPNs
 - Se envía el paquete ICMP por el LSP original, y el router de salida es el encargado de enviarlo al origen
 - Reutilizo el stack original de etiquetas (cambiando el TTL)
 - No es el mismo camino que seguiría en IP
 - En particular el retardo no es hasta el salto “n”
 - No todos los equipos lo soportan (lab no soporta)

Traceroute en IP



Traceroute en IP sobre MPLS



Formato de algunos mensajes LDP

Encabezado LDP.

Común a todos los mensajes

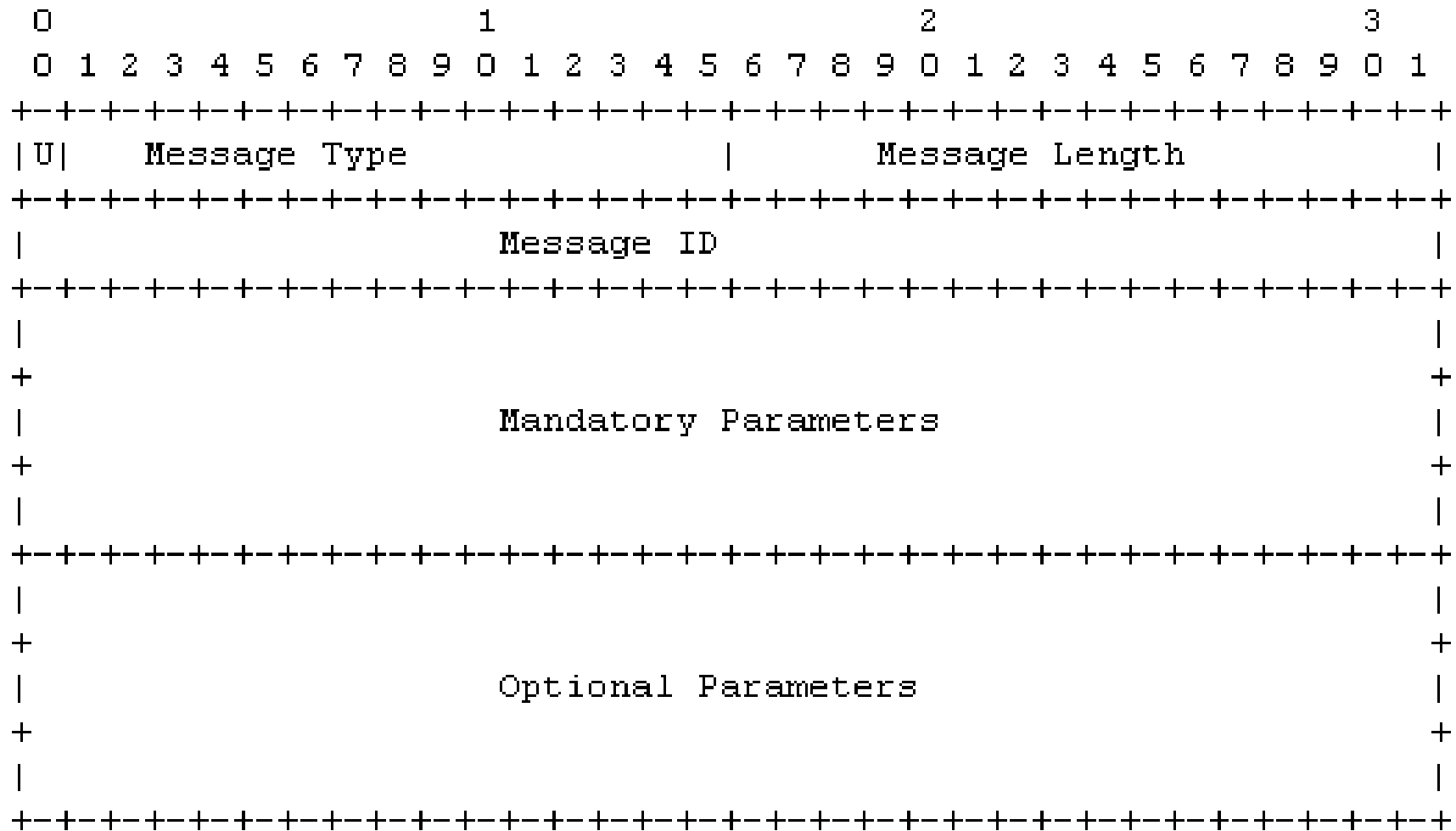
Version	PDU Length
LDP Identifier	
LDP Identifier (cont)	

Versión: 1

PDU Length: largo sin contar los campos Version y PDU Length

LDP Identifier: 6 octetos. Los 4 primeros son el Router-id, los últimos 2 octetos identifican un “espacio de etiquetas”, en general “0”

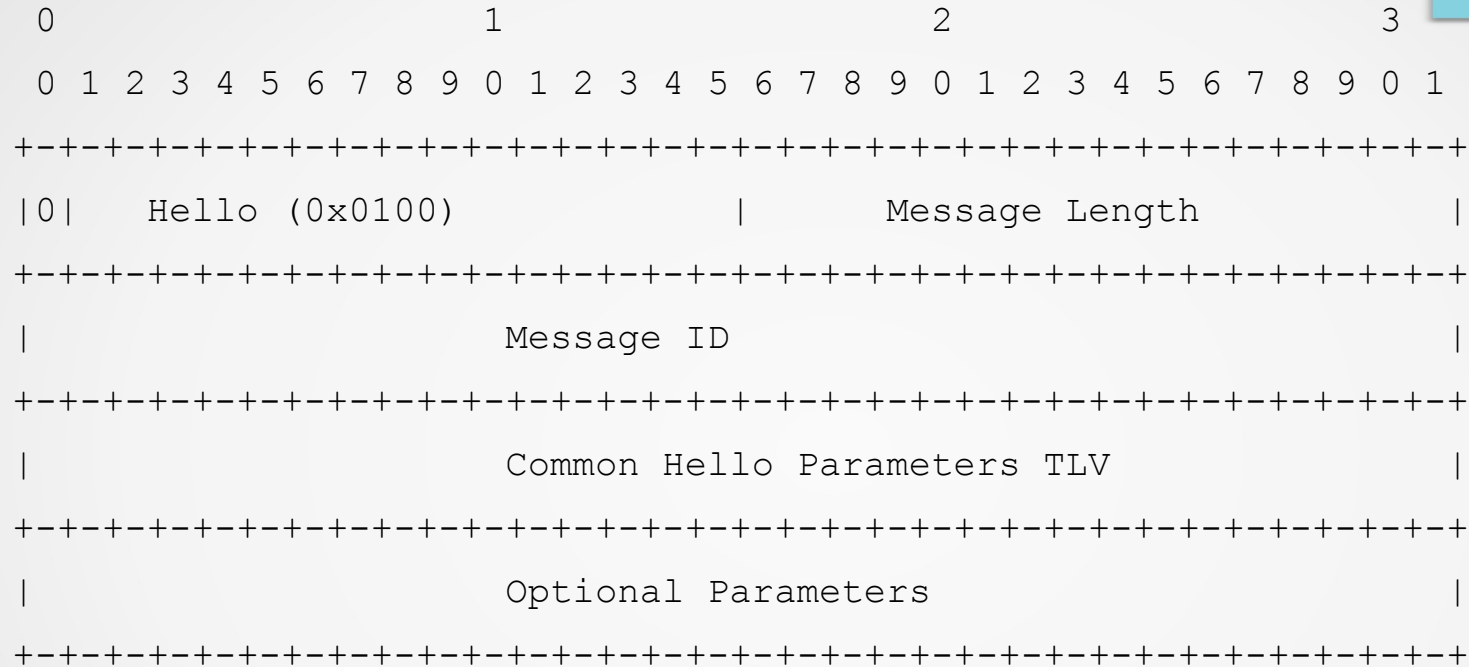
Formato de mensajes



Tipos de mensaje

Message Name	Type
Notification	0x0001
Hello	0x0100
Initialization	0x0200
KeepAlive	0x0201
Address	0x0300
Address Withdraw	0x0301
Label Mapping	0x0400
Label Request	0x0401
Label Withdraw	0x0402
Label Release	0x0403
Label Abort Request	0x0404
Vendor-Private	0x3E00- 0x3EFF
Experimental	0x3F00- 0x3FFF

Mensaje Hello



Message ID

32-bit value used to identify this message.

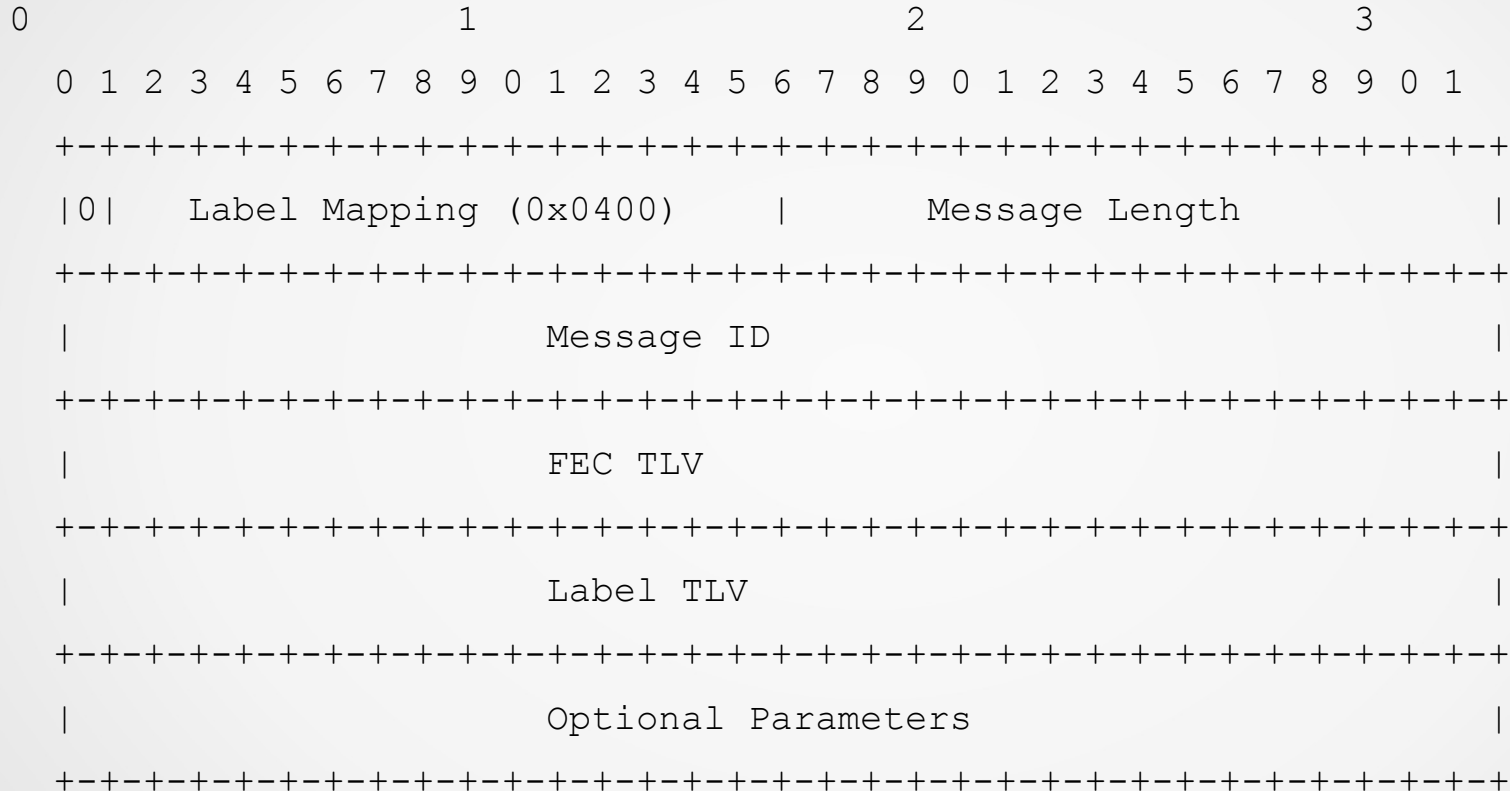
Common Hello Parameters TLV

Specifies parameters common to all Hello messages.

Mensaje Hello (cont.)

- Dentro de los parámetros comunes tenemos el Hold time, si es un “targeted hello”, y otras flags
- Dentro de los parámetros opcionales puede estar la IPv4 o IPv6 a la cual hay que establecer la sesión TCP, y el n° de secuencia de la configuración (para detectar cambios)

Ejemplo: Label mapping (binding)



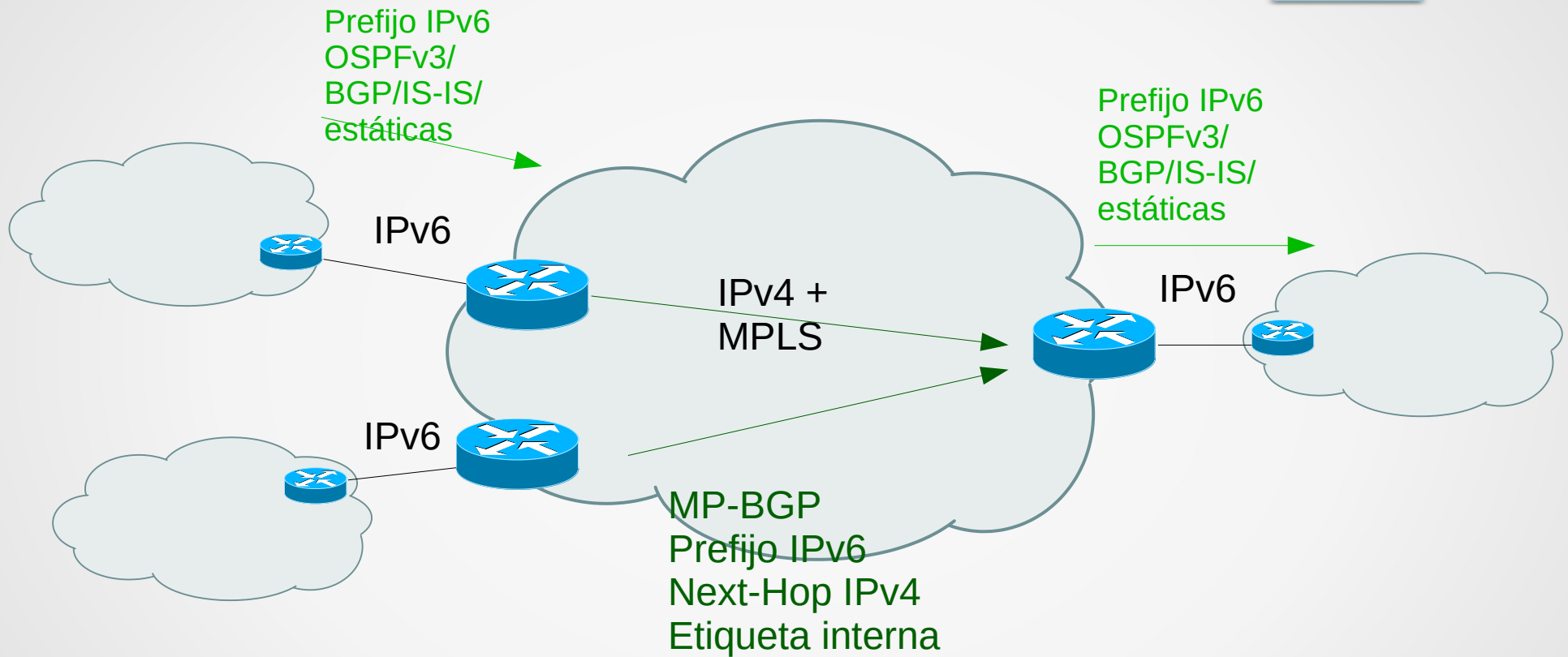
MPLS e IPv6

- RFC 7439 (Enero 2015): Gap Analysis for Operating IPv6-Only MPLS Networks
- Transporte de IPv6 (plano de datos): OK
- Señalización de servicios IPv6: OK
- Plano de control y Mgmt. sobre IPv6: NOK
 - LDP: Especificación no tiene suficiente detalle para garantizar interoperabilidad
 - RFC 7552: Updates to LDP for IPv6 (Junio 2015)
 - Otras faltas
- Implementaciones “recientes”
- Aún podemos encontrar problemas si trabajamos con una red “IPv6 only”

Transición: 6PE

- 6PE: Uno de los mecanismos de transición hacia IPv6
- RFC 4798: Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)
- Utiliza BGP multiprotocolo
 - AFI: IPv6, SAFI: Label
 - Se envía el prefijo IPv6, como next-hop la IPv4 del PE mapeada en IPv6 (::FFFF:x.y.z.w), y una etiqueta
- PE de entrada utiliza un stack de 2 etiquetas
 - Etiqueta exterior: Lleva a la IPv4 del PE destino
 - Etiqueta interior: la propagada por BGP
- Enrutadores P solo ven la etiqueta exterior

6PE: distribución de rutas



6PE: Tráfico

