



Ingeniería de Tráfico IP-TE y MPLS-TE

Ruteo IP y Tecnologías de Transporte

Federico Larroca

Basado en presentaciones de P. Belzarena y J.-L. Rougier





Introducción



- ¿Qué es la Ingeniería de Tráfico? RFC 3272 [Overview and Principles of Internet Traffic Engineering]:
 - “Internet traffic engineering is defined as that aspect of Internet network engineering dealing with the issue of performance evaluation and performance optimization of operational IP networks.”
- Tres elementos a entender:
 - Tráfico
 - Recursos
 - Performance (QoS)





Introducción

■ Ejemplo clásico: fórmula de Erlang-B

$$P_b = \frac{E^m / m!}{\sum_{i=0}^m E^i / i!}$$

E – Intensidad de tráfico
(Tráfico)

m – Número de circuitos
(Recursos)

P_b – Probabilidad de bloqueo
(Performance)

■ Responde preguntas como:

- Con mi instalación actual (m dado) y una intensidad de tráfico pico E , ¿qué P_b voy a obtener?
- Dado m ¿cuánto tráfico puedo soportar con una P_b menor a 0.001?
- Dado un E estimado ¿cuántos circuitos necesito para obtener una P_b menor a 0.001?



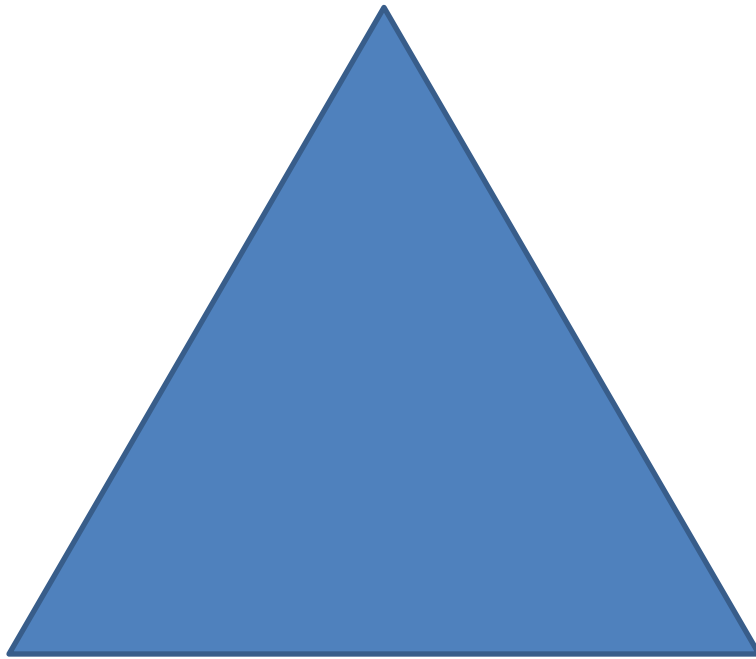


Introducción



- En resumen: TE requiere entender el triángulo

Performance



Tráfico

Recursos

Involucra varias áreas:

- Mediciones
- Caracterización
- Modelado
- Control



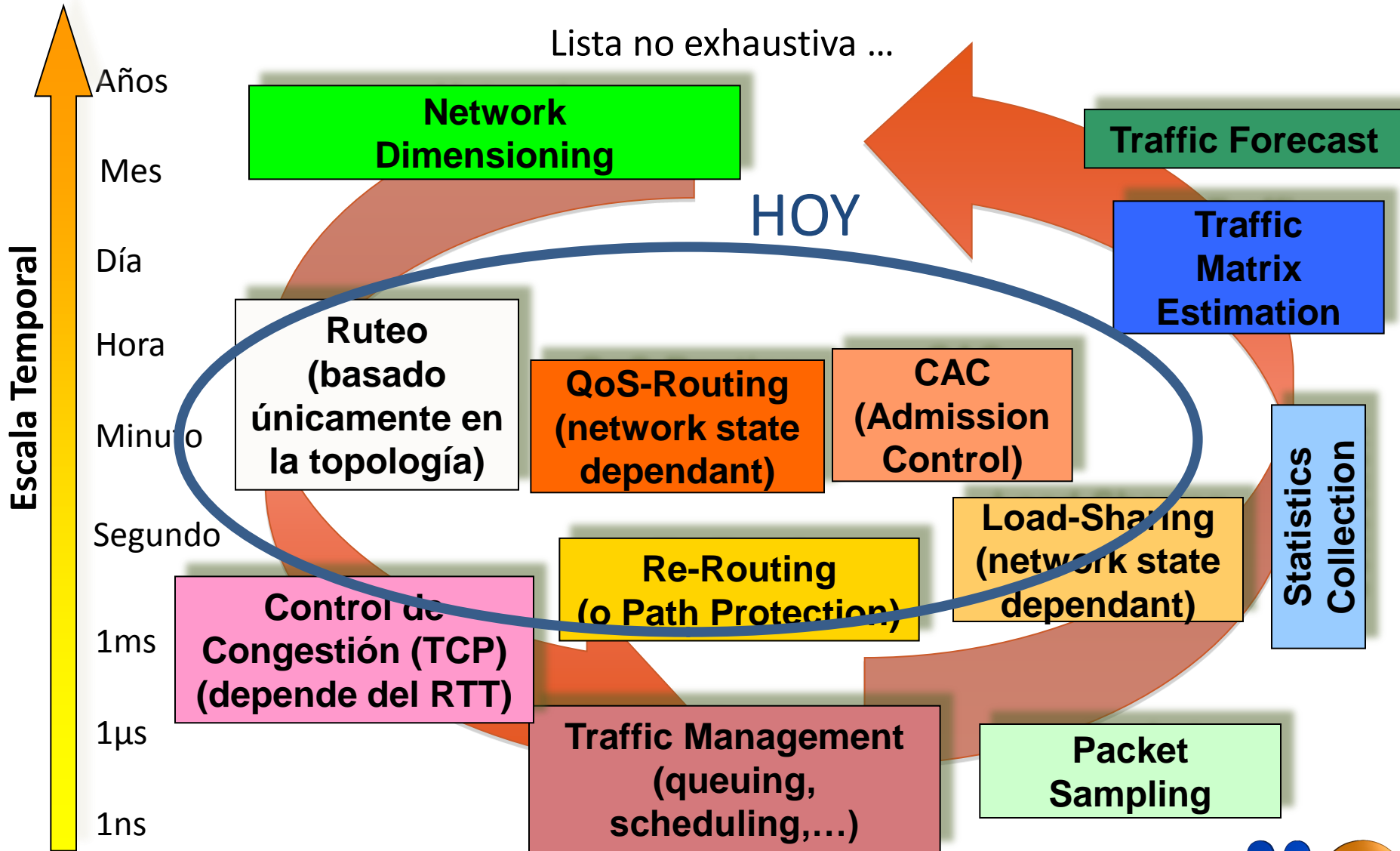
Introducción



- Problemas distintos a escalas temporales distintas:
 - Dimensionado de la red (Resource Planning)
 - Manejo de los recursos (Resource Management)
 - Layouts y topologías virtuales
 - Distribución del tráfico sobre la topología virtual
 - Ruteo y balanceo de carga (interno, externo)
 - Control de tráfico
 - Control de admisión
 - Control de Congestión (TCP)
 - Scheduling y manejo de buffers
 - Packet sampling (medidas)



Algunas Herramientas de TE





Introducción – TE ruteo



- Algunos posibles objetivos de TE a nivel ruteo:
 - **Mover el tráfico** del camino establecido por el IGP a un camino menos congestionado.
 - **Utilizar el exceso de ancho de banda** sobre los enlaces sub-utilizados. Maximizar la utilización de los enlaces y nodos de la red.
 - **Aumentar la confiabilidad** del servicio. Dividir el tráfico entre diversos enlaces minimiza el impacto de una falla simple. Asegurar capacidad necesaria para re-rutear tráfico en caso de fallas puntuales.
 - **Alcanzar requerimientos** impuestos





Introducción – TE ruteo



- Los objetivos se pueden clasificar en:
 - Orientados a tráfico:
 - Mejorar la QoS de los flujos de tráfico
 - Pérdidas
 - Retardos
 - Throughput
 - Orientado a recursos:
 - Optimizar el uso de los recursos
 - (También es orientado a tráfico) Evitar congestiones:
 - Más demanda de la que se puede soportar (utilizar otras técnicas de TE como control de congestión)
 - Evitar situaciones donde ciertos recursos se sub-utilizan y otros se sobre-utilizan

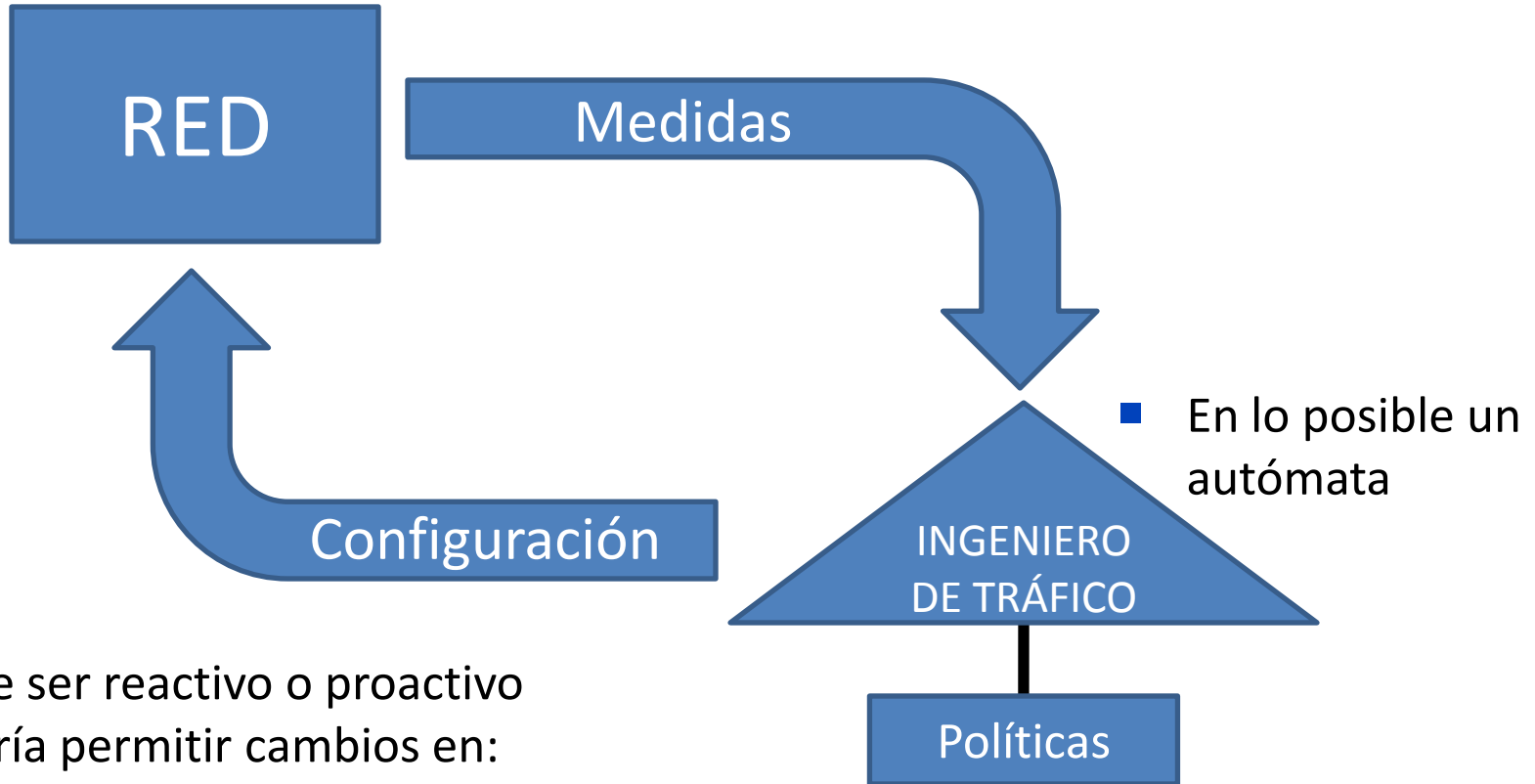




Introducción – TE ruteo



■ Esquema de TE:



- Puede ser reactivo o proactivo
- Debería permitir cambios en:
 - cómo se maneja el tráfico
 - configuración del ruteo
 - parámetros y atributos de los recursos



Agenda



- Introducción
- IP-TE
 - ECMP
 - IGP-WO
- MPLS-TE
 - Constrained Routes
 - RSVP-TE
- Comentarios finales ...

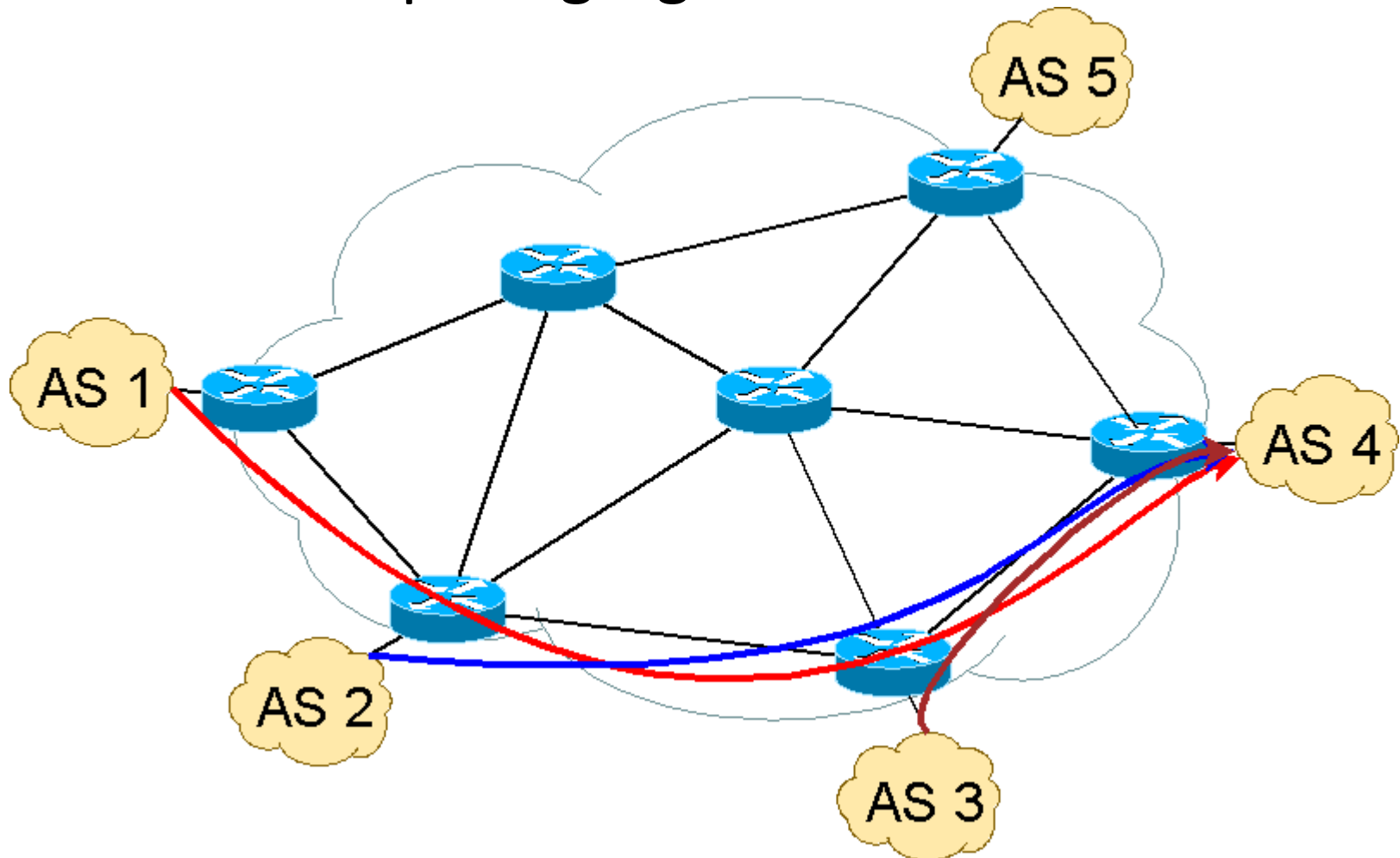




IP-TE



- Problema: Hiper-agregación en IP





IP-TE



- Problema: Hiper-agregación en IP
- Solución: Balanceo de carga (i.e. usar varios caminos)
 - Beneficio extra: resiliencia
- OSPF permite usar sólo en parte este potencial:
 - Equal Cost MultiPath (ECMP)
 - Balancear el tráfico en partes iguales por todos los caminos que tengan costo mínimo

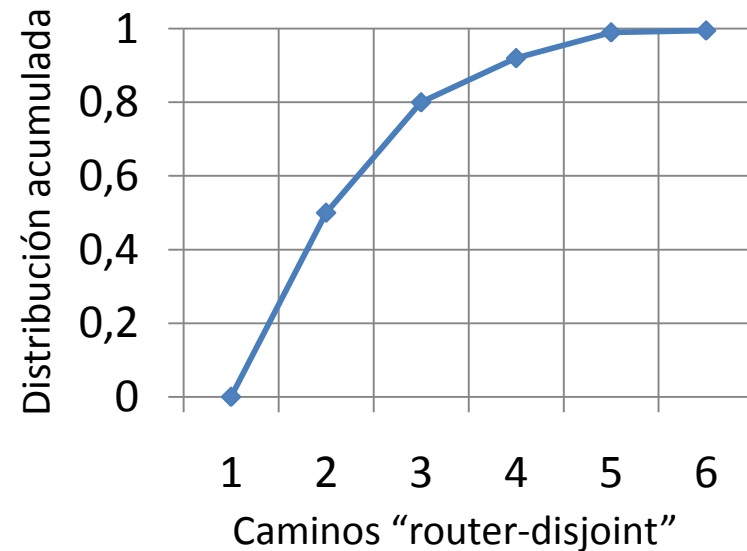


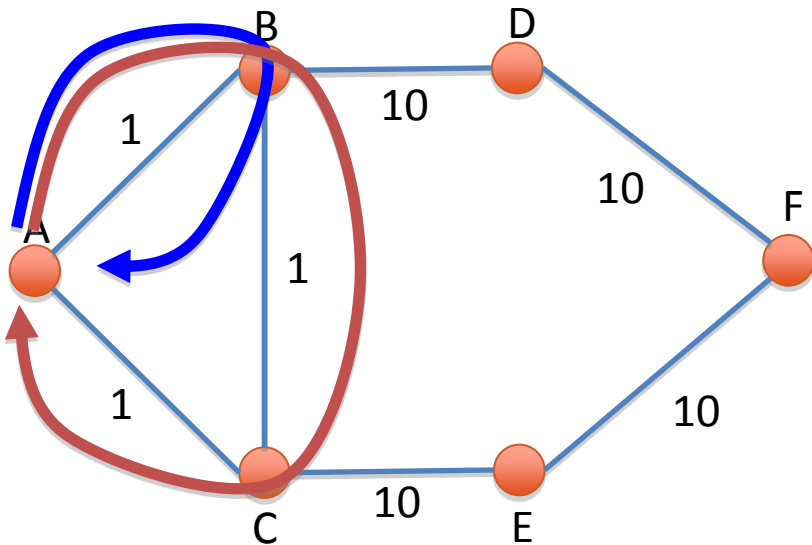
Ilustración: estimación de la diversidad de tráfico en un tier-1 [Teixeira 2003]





IP-TE: ECMP

- ¿Porqué sólo los caminos con igual costo?
- Ejemplo: agreguemos en la tabla de ruteo caminos con hasta 10% más que el mínimo



- ¡¡LOOPS!!

Nodo	Camino	Costo
A	ABDF	21
A	ACEF	21
B	BDF	20
B	BCEF	21
B	BACEF	22
C	CEF	20
C	CBDF	21
C	CABDF	22





IP-TE: ECMP



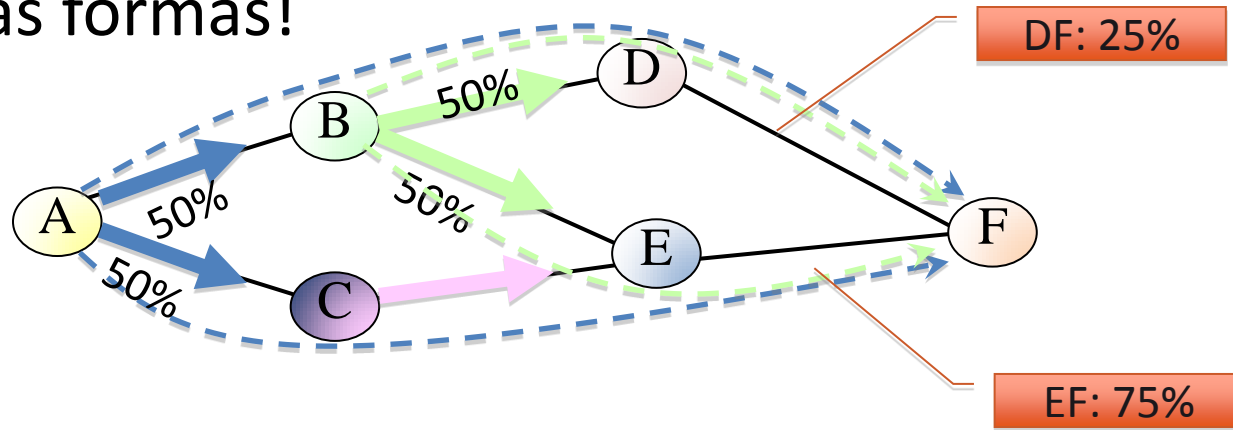
- ECMP genera otros problemas
 - ¿Cómo repartir los paquetes entre las interfaces?
 - Por playas de direcciones:
 - pésima distribución
 - Round-Robin por paquetes:
 - buena distribución
 - problema: si los caminos tienen distinto retardos los paquetes pueden llegar fuera de orden => mala performance de TCP
 - Por flujos: Hash-function sobre (src_addr, dst_addr, src_port, dst_port)
 - Paquetes llegan en orden y relativamente buena distribución





IP-TE: ECMP

- ECMP genera otros problemas
 - El balanceo se hace entre next-hops: puede resultar en una mala distribución de tráfico de todas formas!



- Se puede configurar una distribución no 50%-50% pero es complicado y engorroso (ver por ejemplo OSPF-OMP)



Agenda



- Introducción
- IP-TE
 - ECMP
 - IGP-WO
- MPLS-TE
 - Constrained Routes
 - RSVP-TE
- Comentarios finales ...





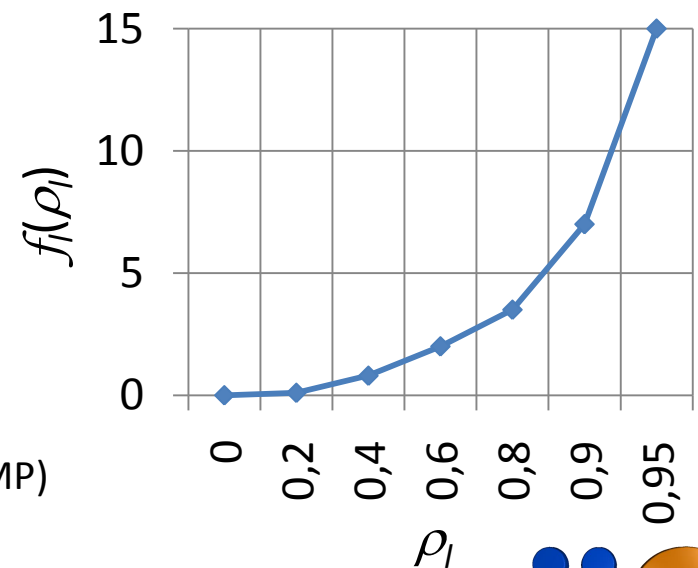
IP-TE: IGP-WO



- Otra técnica de Ingeniería de tráfico: IGP-WO (Weight Optimization)
 - Configurar los pesos OSPF de cada enlace de manera tal de lograr cierto objetivo (generalmente minimizar un costo)
- Ejemplo:
 - Inputs:
 - Tráfico promedio intercambiado entre cada par de routers (matriz de tráfico)
 - Topología completa de la red
 - Objetivo (minimizar el costo total):

$$\min_w \sum_l f_l(\rho_l)$$

ρ_l – utilización del enlace l de aplicar IGP-(ECMP)
con pesos w y la demanda del input





IP-TE: IGP-WO



- Dificultad del problema de optimización:
considerar OSPF-(ECMP)
 - Sólo se usa(n) los camino(s) más corto(s)
 - ECMP reparte de manera igualitaria
 - El ruteo es únicamente basado en el destino
- Imposible de resolverlo en tiempos razonables!
 - Se usan heurísticas.
- Buenos resultados (cercanos al óptimo global sin restricciones de ruteo)





IP-TE: IGP-WO



- Lado negativo de IGP-WO:
 - Ajuste continuos de los pesos OSPF generan problemas de estabilidad
 - Ejemplo: problema del Hot-Potato (i.e. la demanda de tráfico no es independiente de la configuración de OSPF)
- Posible extensión: dinamismo
 - A nivel de tráfico
 - más de una demanda de tráfico como input
 - A nivel topológico
 - fallas de routers/enlaces





Agenda



- Introducción
- IP-TE
 - ECMP
 - IGP-WO
- MPLS-TE
 - Constrained Routes
 - RSVP-TE
- Comentarios finales ...





MPLS-TE



- MPLS para TE se define (o discute) en el RFC 2702 [Requirements for Traffic Engineering Over MPLS]
- Terminología: Traffic Trunk (TT)
 - Agregado de tráfico perteneciente a una misma clase
 - En términos prácticos, un TT se puede caracterizar por:
 - Ingress-egress LSR
 - FEC(s) que se mapean a él
 - Características de comportamiento (parametrización, qué caminos quiere usar y cómo los mantiene)
- Ventajas de MPLS frente a OSPF:
 - Rutas explícitas
 - Mapeo de TT a LSPs posible
 - Agregación Y desagregación





MPLS-TE



- Se plantean tres problemas:
 1. Cómo mapear los paquetes a los FECs
 2. Cómo mapear los FECs a los troncales
 3. Cómo mapear los troncales a los recursos físicos
- El RFC se “limita” a 3.
- Se especifican tres funcionalidades:
 - Atributos asociados a los TTs
 - Definen su comportamiento
 - Atributos asociados a recursos
 - Limitan su uso por TTs
 - “Constraint-based Routing” (CBR)
 - Usando la información anterior elige los caminos para los TTs





MPLS-TE: Rutas Explícitas



- Los LSPs pueden ser usados para enviar tráfico por rutas independientes del algoritmo de IGP
 - Provee fuerte control sobre la función de ruteo
 - Habilita la implementación de políticas
 - Permite la optimización de la performance
- Es la función (principal) que habilita a hacer Ingeniería de Tráfico en MPLS
- Definición: secuencia de nodos lógicos entre un nodo de ingreso y uno de egreso
 - E.g.: Lista de direcciones IP
 - E.g.: especificar los primeros saltos y luego hop-by-hop





MPLS-TE: Rutas Explícitas



- Si el nodo ingreso quiere utilizar una ruta que...
 - ...no sigue el camino por defecto: debe utilizar un LDP que soporte la definición de rutas explícitas.
 - Distribution Mode: Downstream-on-demand
 - Dos protocolos posibles: (CR-LDP) y RSVP
 - ...provea un cierto nivel de QoS: debe utilizar un LDP que soporte la inclusión de requerimientos de tipos de servicio.
 - Distribution Mode: Downstream-on-demand
 - Los mismos dos protocolos: (CR-LDP) y RSVP-TE





MPLS-TE: Constrained Route



- Generalización de lo anterior: una ruta obtenida a partir de un conjunto de restricciones
 - E.g.: información de QoS del enlace (ancho de banda disponible, retardo, etc.), clases, prioridades, etc.
- El LSR de ingreso calcula una ruta satisfaciendo un conjunto de restricciones y de acuerdo al estado de la red
 - Una ruta explícita puede ser vista como una Constrained Route donde la restricción es la secuencia de nodos a atravesar.





MPLS-TE: Constrained Route



- Un ejemplo puede ser el de establecer una ruta que cumpla con:
 - Pasar por enlaces con un ancho de banda libre mayor a 10 Mbps
 - y
 - Pasar por enlaces de color “azul”
 - y
 - Pasar por enlaces con retardo menor a 200 ms



MPLS-TE



- Existen entonces dos posibilidades:

	Off-line provisioning	Constraint Based Routing
<i>Network load information</i>	network statistics	QoS aware routing distributes load information
<i>Path selection</i>	Optimal flow mesh calculated offline	New flows routed along best available path
<i>Service provisioning</i>	End-to-end path explicitly provisioned	Endpoints provisioned with constraints

RFC y lo que sigue...

Fuente : MPLS Forum 2000





MPLS-TE: Traffic Trunks



- Atributos de los TTs:
 - Parámetros de tráfico
 - rate pico, rate promedio, etc.
 - Indican los requerimientos de recursos del TT
 - Selección de caminos y administración
 - Caminos explícitos
 - total o parcial
 - obligatorio o no
 - Jerarquía o preferencia dentro de un conjunto de caminos
 - Afinidad con los recursos
 - Incluir o excluir ciertos recursos
 - Adaptabilidad
 - Permitir o no re-optimización
 - Balanceo de carga entre varios caminos
 - Cuánto se envía por cada camino



MPLS-TE: Traffic Trunks



- (Más) Atributos de los TT:
 - Prioridad
 - Útil en casos donde se use CBR (indica el orden en que establecen caminos los TT) y si se usa preemption
 - Preemption
 - Indica si un TT puede “sacar” a otro TT en un mismo camino
 - Resiliencia
 - Qué hacer con el TT cuando su camino falla
 - No enrutar
 - Re-enrutar si hay un camino con recursos suficientes
 - Siempre re-enrutar
 - Policing
 - Qué hacer con el TT cuando no cumple sus especificaciones (e.g. shaping)





MPLS-TE: Recursos



- Atributos de los recursos:
 - Maximum Allocation Multiplier (MAM)
 - Porcentaje del recursos disponible para reserva
 - Principalmente aplicable a ancho de banda (otro ejemplo: buffers)
 - Under y Over allocation
 - Clase del recurso
 - Colores
 - Principalmente aplicable a enlaces
 - Ejemplos de uso:
 - Políticas uniformes a un conjunto de recursos que no están en la misma región
 - Preferencia relativa para ser usados para rutear tráfico
 - Restringir el uso del recurso para rutear tráfico





MPLS-TE: CBR



■ “Constraint-Based Routing”

• Input:

- Atributos de los TT
- Atributos de los recursos
- Otra información topológica

• Output:

- Cada nodo computa rutas explícitas para los TTs que se originan en él
- Es decir, un LSP para cada TT tal que satisface sus requerimientos (expresado en su atributo de tráfico), y respeta las restricciones impuestas por la disponibilidad de recursos, políticas administrativas y otra información topológica.





MPLS-TE: CBR Hoy



- Implementación de CBR
 - Transportar información de la red mediante alguna extensión de protocolo IGP (e.g. OSPF-TE)
 - Encontrar los caminos que cumplen las restricciones mediante un algoritmo de ruteo con restricciones (e.g. CSPF)
 - Establecer los caminos mediante un protocolo que permita establecer caminos explícitos (e.g. RSVP-TE)
- Constrained Shortest Path First (CSPF)
 1. Eliminar los recursos que no satisfacen los requerimientos de la TT (en general reservas en el plano de control + atributos administrativos)
 2. Correr un algoritmo de “shortest path” en el grafo residual





MPLS-TE: OSPF-TE



- Especificado en RFC 3630 [Traffic Engineering (TE) Extensions to OSPF Version 2]
- Extensiones a OSPF: además de la info básica sobre estado y peso del enlace, OSPF-TE distribuye:
 - Máximo ancho de banda físico de la interfaz
 - Máximo ancho de banda reservable de la interfaz
 - Ancho de banda disponible por nivel de prioridad
 - Atributos del enlace
- Esta información se difunde:
 - Cuando el enlace es levantado o bajado
 - Cuando hay un cambio en la configuración
 - Periódicamente
 - Cuando hay un cambio en el ancho de banda reservado





Agenda



- Introducción
- IP-TE
 - ECMP
 - IGP-WO
- MPLS-TE
 - Constrained Routes
 - RSVP-TE
- Comentarios finales ...





RSVP-TE



- RFC 3209 [RSVP-TE: Extensions to RSVP for LSP Tunnels]
- ¿Qué extensiones hay con respecto a RSVP?
 - Mensaje RSVP **Path** incluye un objeto LABEL_REQUEST
 - Asocia una etiqueta a los nodos en el camino
 - Las etiquetas son asignadas downstream y distribuidas (enviadas upstream) por un objeto LABEL el mensaje RSVP **Resv**
 - También incluye un objeto EXPLICIT_ROUTE
 - Contiene información del camino que debe seguir el mensaje PATH para establecer el LSP
 - Puede estar definido explícitamente en su totalidad o no
 - También incluye un objeto SESSION_ATTRIBUTE
 - Especifica prioridades, afinidades, protección, etc.
 - El objeto SENDER_TSPEC se usa para especificar los parámetros de tráfico del TT





RSVP-TE



- Procesamiento del mensaje **Path** en cada router
 - Si la dirección de destino del mensaje pertenece a ese router
 - Finaliza el procesamiento del objeto EXPLICIT_ROUTE
 - Se envía upstream (el camino inverso del que llegó) un mensaje **Resv** con el objeto LABEL
 - En otro caso, se examina la siguiente entrada del EXPLICIT_ROUTE
 - Se consulta la tabla de ruteo y se determina el next-hop
 - Si la entrada en el EXPLICIT_ROUTE es del tipo “*strict*” se verifica que esté directamente conectado
 - Se reenvía el mensaje al next-hop correspondiente





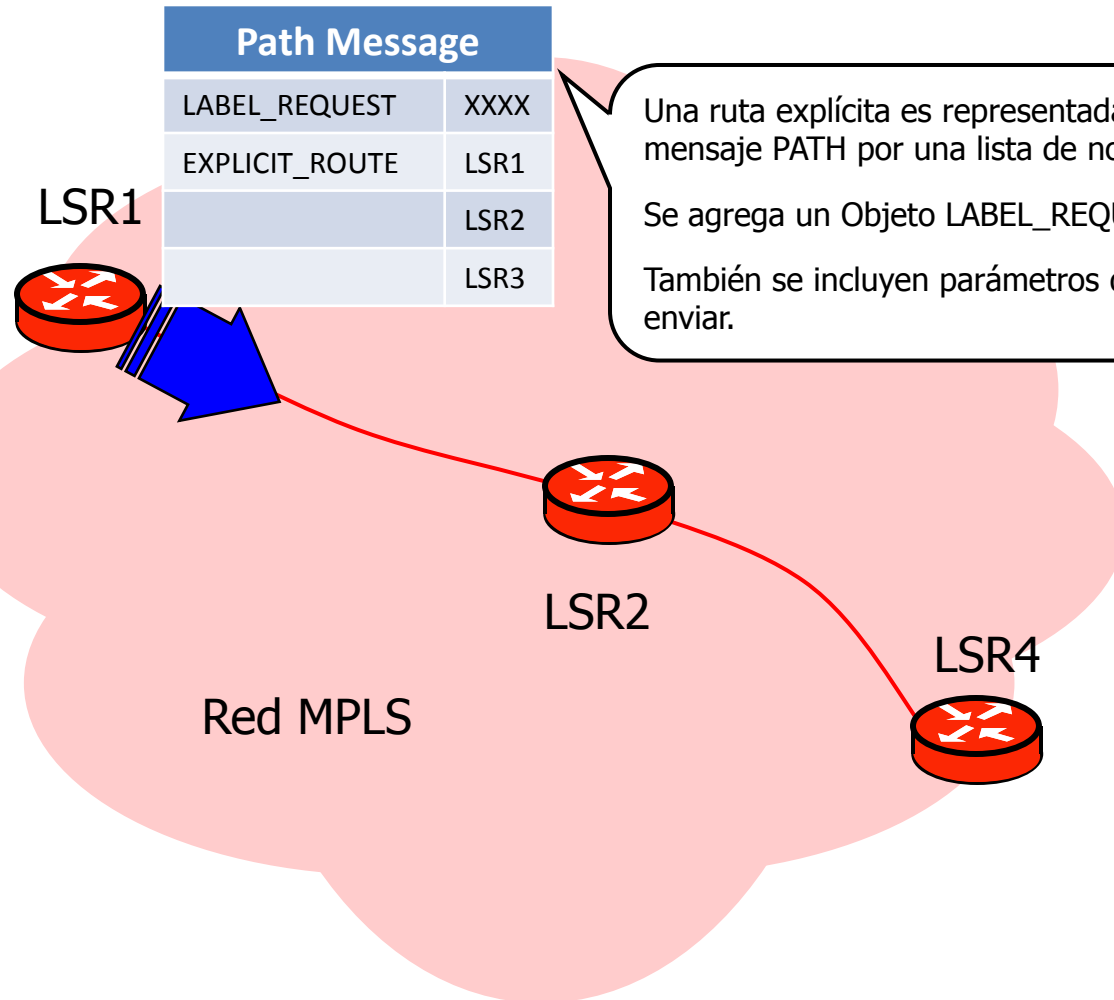
RSVP-TE



El LSR de ingreso determina que necesita establecer un LSP

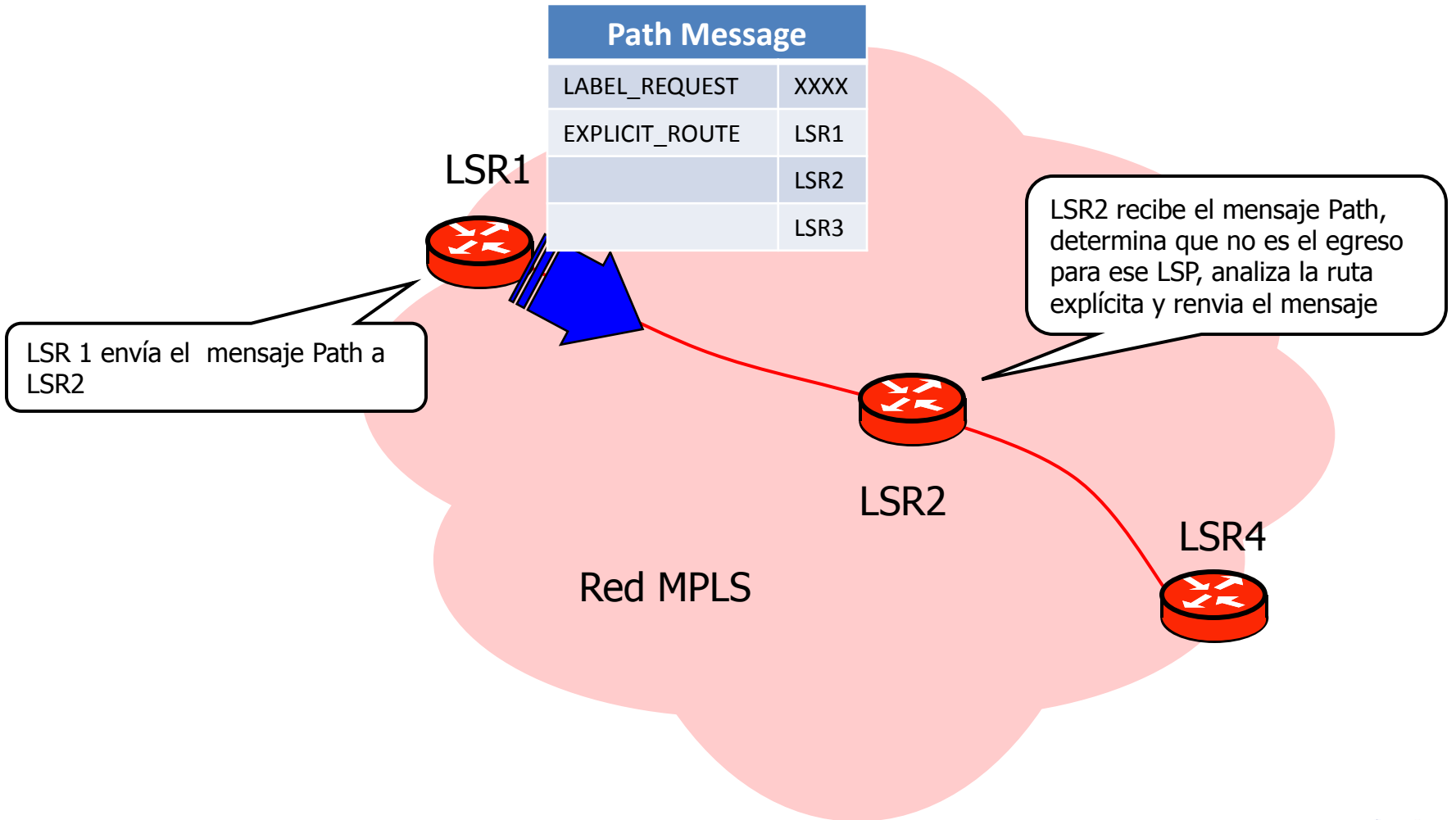
Una ruta explícita sujeta a las restricciones es calculada en un punto de ingreso de la red

Criterios incluyen pero no están limitados a información de ruteo, parámetros de tráfico y políticas administrativas



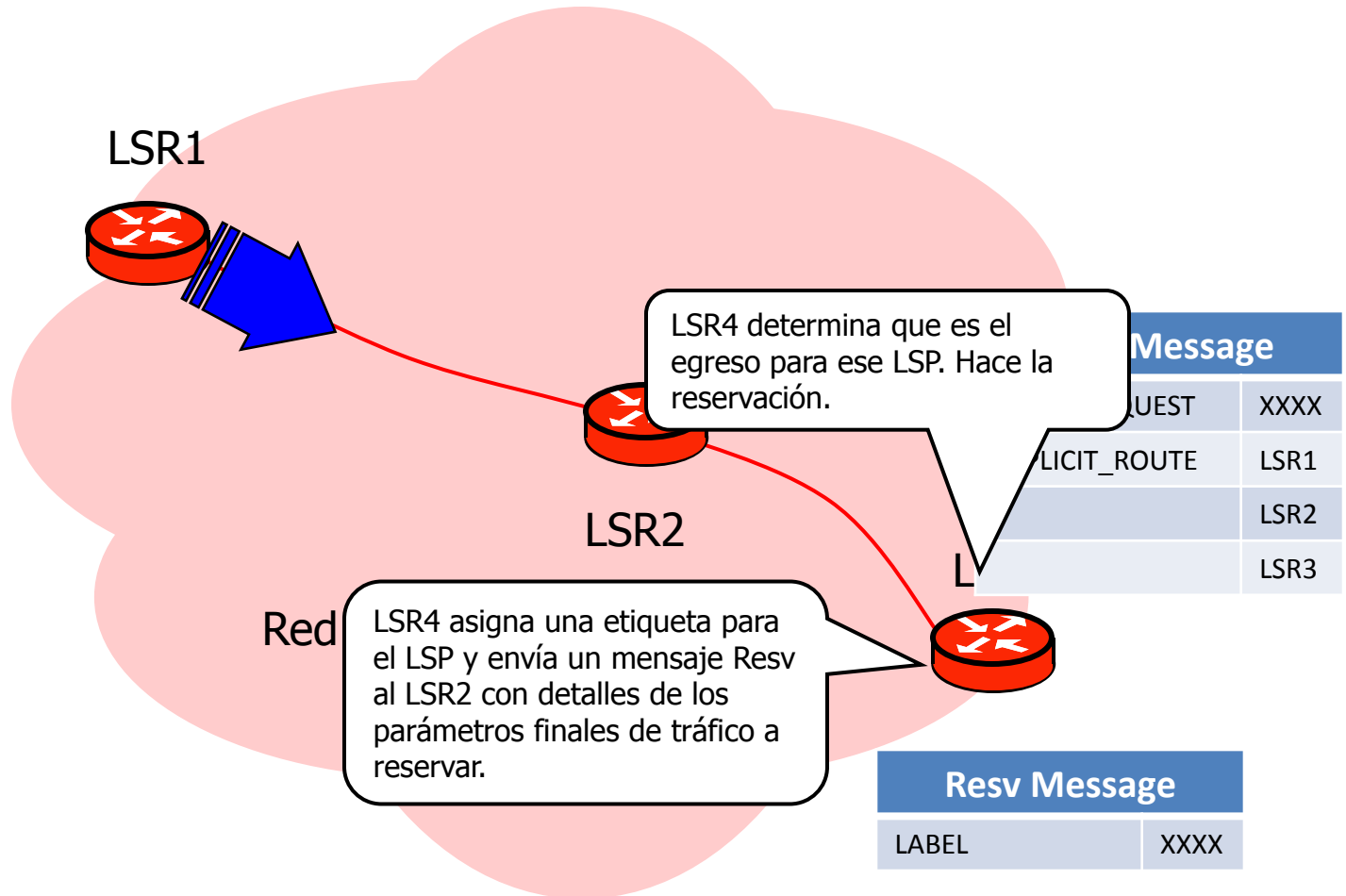


RSVP-TE



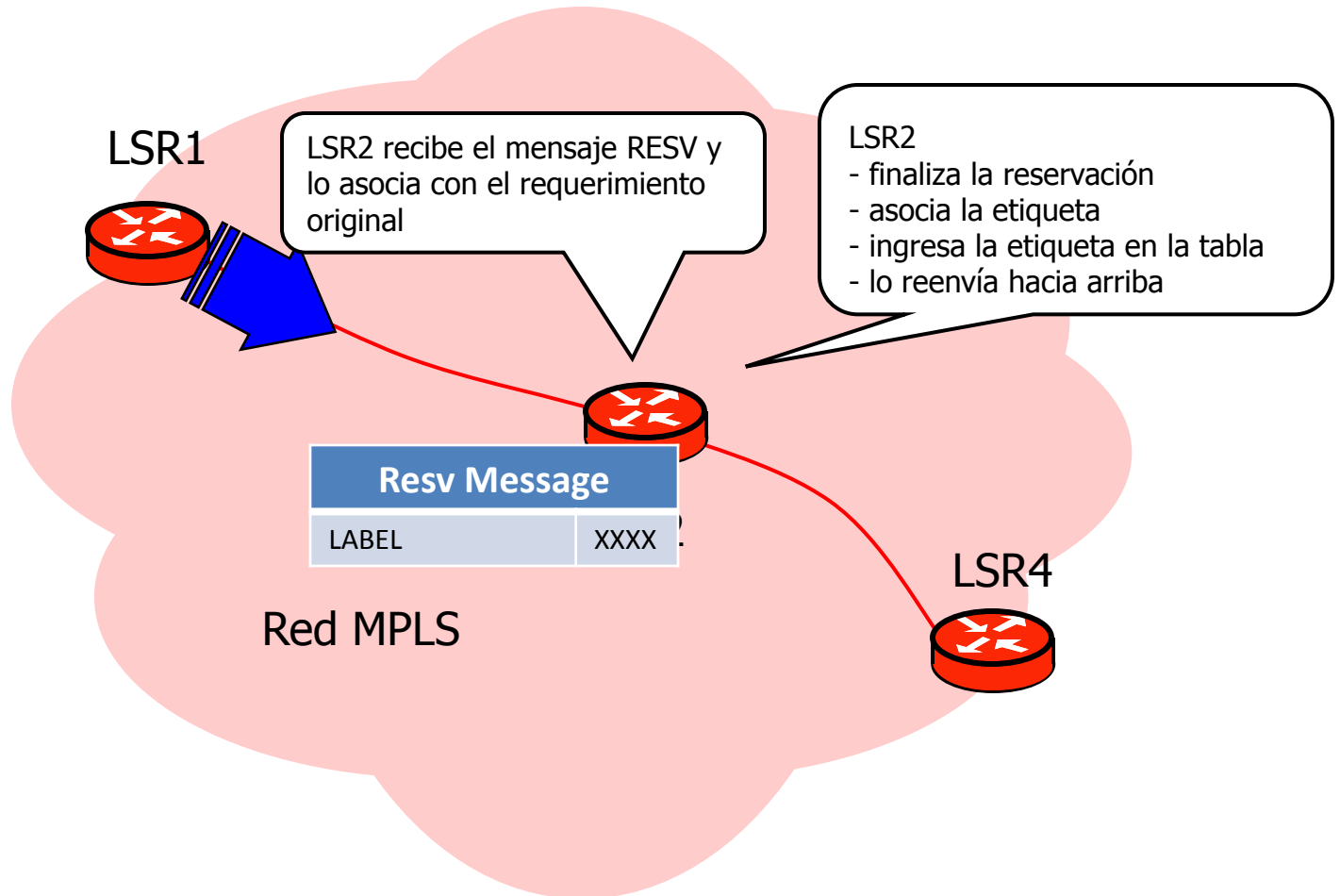


RSVP-TE





RSVP-TE





Agenda



- Introducción

- IP-TE
 - ECMP
 - IGP-WO

- MPLS-TE
 - Constrained Routes
 - RSVP-TE

- Comentarios finales ...





Comentarios Finales ...



- Si IGP-WO obtiene casi los mismos resultados que una red optimizada con MPLS-TE
 - ¿Para qué sirve MPLS-TE?
- Más que por la performance, es por sus funcionalidades:
 - Coloreo del grafo
 - Reserva de recursos (e.g. QoS por VPN)
 - Para el control de DiffServ (Diffserv-Aware Traffic Engineering, DS-TE)
 - Fast ReRouting
 - Etc...





Fast ReRouting



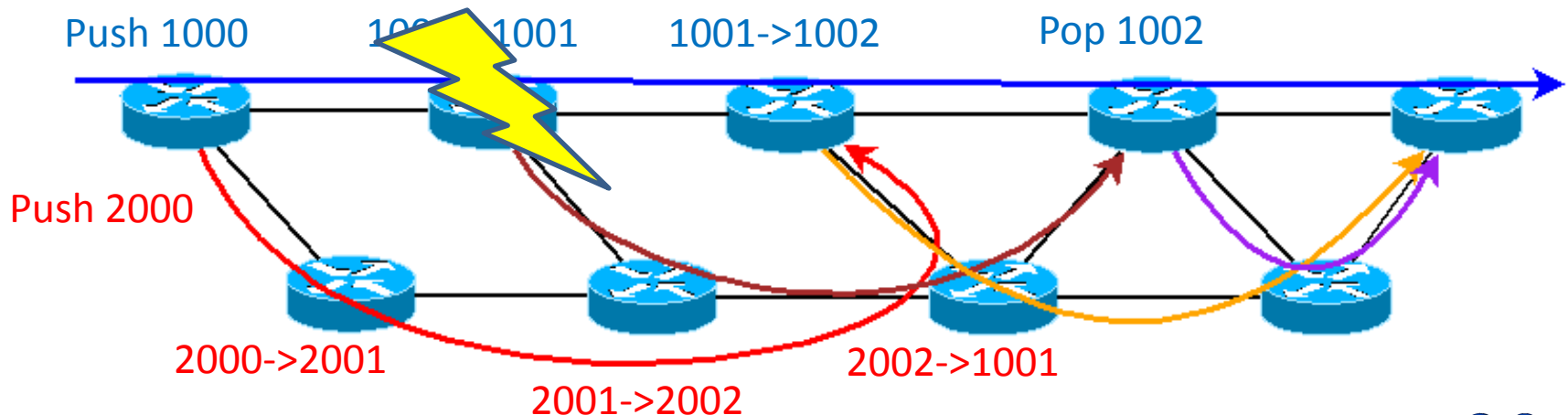
- Detección de fallas en OSPF:
 - Al detectar la falla se propaga la información, y se reconverge a la topología nueva
 - Problemas:
 - Reacción lenta (>1s)
 - Posibles inconsistencias en el transitivo
- Detección de fallas en MPLS:
 - Se usan mecanismos de capa 2 (e.g. SDH)
 - Se usan mensajes “hello” de RSVP
- Una vez detectada la falla sólo se ven afectados los LSPs que pasaban por ese enlace/nodo
- El router que detecta la falla re-dirige **localmente** los paquetes hacia el camino de backup
 - La notificación de la falla no necesita llegar al LER
- Referencia: RFC 4090 [Fast Reroute Extensions to RSVP-TE for LSP Tunnels]





Fast ReRouting

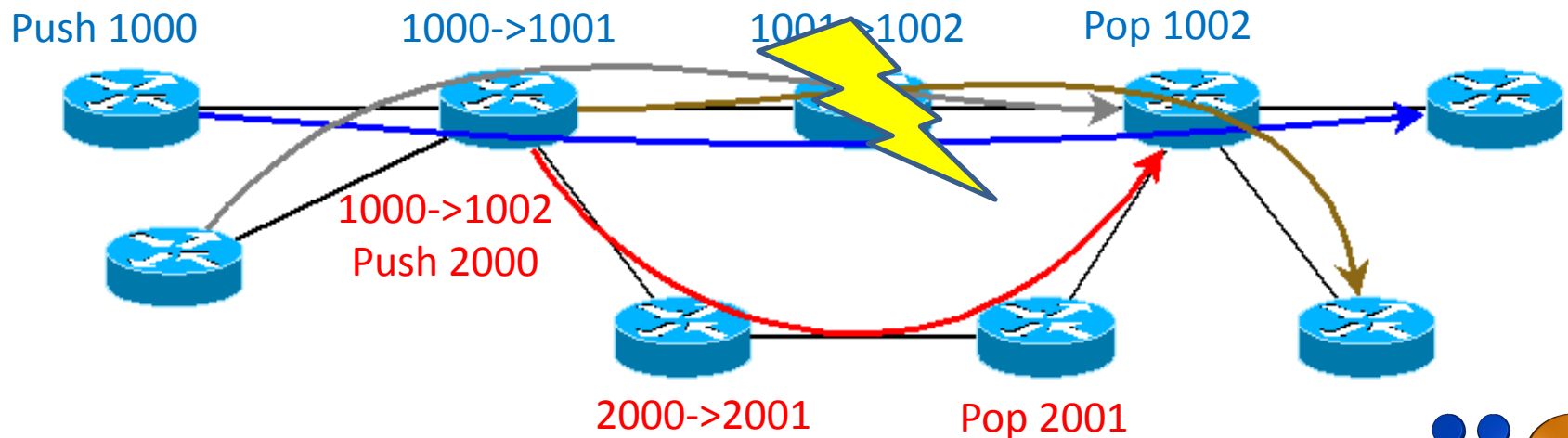
- Cada LSP se puede proteger de fallas en:
 - Nodos
 - Enlaces
 - Ambas
- Se puede proteger cada LSP de cada posible falla con un LSP alternativo: **one-to-one backup**





Fast ReRouting

- Cada LSP se puede proteger de fallas en:
 - Nodos
 - Enlaces
 - Ambas
- Se pueden proteger varios LSPs con un único LSP de backup: **Facility Backup**





MPLS y QoS



- Dos posibilidades (no excluyentes)
 - Usar DiffServ y MPLS:
 - La info de DiffServ TIENE que venir en el encabezado MPLS (MPLS no mira el encabezado IP) => Campo EXP
 - E-LSP
 - El tratamiento de cada paquete está dado EXCLUSIVAMENTE por el campo EXP del paquete
 - Tres bits de EXP => ocho posibilidades en total
 - L-LSP
 - La etiqueta determina la “PHB Schedule Class” y EXP determina la precedencia de descarte del paquete.
 - Usar DiffServ y MPLS-TE: DS-TE (DiffServ-Aware MPLS Traffic Engineering), RFCs 3564 y 4124
 - Idea básica:
 - Cada enlace tiene recursos globales (“global pool”) y subconjuntos de éstos (“subpools”).
 - Si se asocia cada clase a un subpool, se puede hacer CBR *por clase*

