

Introducción a MPLS

Javier Ger, Mariano Beiró

Facultad de Ingeniería UBA

26 de junio de 2012

Introducción a los Sistemas Distribuidos

<http://www.fi.uba.ar/materias/7543>

Temario

1 Introducción

- Diccionario y Acrónimos
- Usos actuales
- Antecedentes y RFCs

2 El protocolo MPLS

- Componentes Funcionales
 - Label Switched Router (LSR)
 - Label Switched Path (LSP)
 - Labels, PHP y UHP
- Header MPLS
- Procesamiento de paquetes

3 Protocolos de señalización

- RSVP vs. LDP
- RSVP
 - RSVP - Mensajes
 - RSVP - Extensiones MPLS
 - RSVP - LSP dinámicos
- LDP
 - Mensajes LDP
 - LDP - Parámetros de distribución
 - LDP - FEC

MPLS	<i>Multiprotocol Label Switch</i>
MPLS Labels	Etiquetas MPLS
ATM	<i>Asynchronous Transfer Mode</i>
VC	<i>Virtual Circuit</i> Circuito Virtual
PVC	<i>Permanent Virtual Circuit</i> Circuito Virtual Permanente
SVC	<i>Switched Virtual Circuit</i> Circuito Virtual Conmutado
TE	<i>Traffic Engineering</i> Ingeniería de Tráfico
SONET	<i>Synchronous Optical NETwork</i>
SDH	<i>Synchronous Digital Hierarchy</i>
ASIC	<i>Application-Specific Integrated Circuit</i> Hardware de propósito específico
IGP	<i>Interior Gateway Protocol</i>


Usos actuales

- Segmentar tráfico/servicios distintos sobre una infraestructura común
- Redes N-Play
- VPN para servicios de Capa 2 (ejemplos PW, VPLS) y Capa 3 (ejemplos VRF)
- Optimizar información de ruteo en los *Transit Router* o P (ejemplo: sesiones BGP entre routers no directamente conectados)
- Transición a IPv6 (ejemplo: topologías usando 6PE)

Antecedentes

- Actualmente MPLS es considerado, en muchos casos, como un medio para soportar múltiples servicios sobre una infraestructura IP común.
- Originalmente el objetivo fue reducir el retardo de tránsito.
 - Búsqueda en tablas de ruteo que se hacían por software y con una CPU central, en redes con ancho de banda creciendo.
- Adoptar ciertos conceptos de ATM. Estos switches trabajaban por hardware, con circuitos preestablecidos y con un encabezado de longitud fija (5 Bytes).
 - MPLS tiene un encabezado de longitud fija y genera un circuito predeterminado para el tráfico.
- Antes que MPLS fuera masivamente desplegado cambió la realidad.
 - Los routers empezaron a trabajar por hardware (ASIC - Application-Specific Integrated Circuit) para las búsquedas y envío de paquetes (igual de eficientes que los equipos ATM).
- Surge el concepto de Ingeniería de Tráfico (Traffic Engineering - TE) que le da un nuevo sentido a MPLS.

Antecedentes - TE

- TE podría ser definida como la capacidad de controlar el camino que los paquetes toman en una red.
- Hay formas originales muy rudimentarias (por ejemplo ruteo por origen) y confiar en las métricas del IGP tiene limitaciones (no es posible caminos distintos para flujos distintos).
- Otros modelos, con redes superpuestas operando en paralelo implican duplicación de recursos, tanto de redes como humanos, para su operación.
 - Un ejemplo sería una red IP con el core ATM. En este caso, el camino físico lo determinan los circuitos virtuales (VC) y no el IGP.
 - En el caso particular de ATM otra desventaja es la ineficiencia del protocolo respecto a la carga útil y los costos asociados al equipamiento de esta tecnología.
- Los fabricantes empezaron a soportar tecnologías de mayor capacidad en los routers directamente (ejemplo SONET) y, para mantener las ventajas del modelo mencionado, MPLS fue una opción interesante. 

Recomendaciones

- **www.ietf.org**
- RFC 2702, *Requirements for Traffic Engineering over MPLS*
- RFC 3031, *Multiprotocol Label Switching Architecture*
- RFC 3032, *MPLS Label Stack Encoding*
- Internet draft draft-ietf-isis-traffic-02.txt, *IS-IS Extensions for Traffic Engineering*
- Internet draft draft-katz-yeung-ospf-traffic-04.txt, *Traffic Engineering Extensions to OSPF*
- Internet draft draft-ietf-mpls-icmp-02.txt, *ICMP Extensions for Multiprotocol Label Switching*

Componentes Funcionales

LSR - *Label Switched Router*

- Label Switched Router (LSR): Cada router IP que soporta MPLS.
 - Los LSR entienden el encabezado MPLS y envían el tráfico a través del LSP.
- Hay distintos (4) tipos de LSR.
 - **Ingress Router.** Realiza la operación “Label Push” que permite que el tráfico IP nativo se encapsule en MPLS. Una vez hecho esto, el mismo se encamina hacia el Egress Router (downstream) a través del LSP.
 - **Transit Router.** Routers en el medio del LSP (entre 0 y 253 considerando el TTL). Realiza la operación de “Label Swap” (y decrementa el TTL). No utiliza información del encabezado IP.
 - **Penultimate Router.** Es un Transit Router con una función especial. Realiza la operación de “Label Pop” y remueve información MPLS del encabezado del paquete. El envío lo hace en base a la información MPLS, pero es un paquete nativo IP hacia el próximo salto del LSP luego de decrementar el TTL del mismo. Esto se llama PHP (Penultimate Hop Popping).
 - **Egress Router.** Equipo en el cual finaliza el LSP. Recibe el paquete del Penultimate Router y realiza una búsqueda en IP.

Componentes Funcionales

LSR - *Label Switched Router*

- Tanto los Ingress Router como los Egress Router se consideran Edge LSR y se los conoce también como Provider Edge (PE).
- A los Transit Router se los llama también Provider (P).

Componentes Funcionales

LSP - *Label Switched Path*

- Label Switched Path (LSP) - El camino creado por MPLS entre 2 extremos de la red.
- Unidireccionales, a diferencia de los VC que eran bidireccionales.
- Deben establecerse ambos para el funcionamiento correcto de los flujos de datos.
- Cada LSP sólo 1 Ingress y 1 Egress Router.

Componentes Funcionales

Labels, PHP y UHP

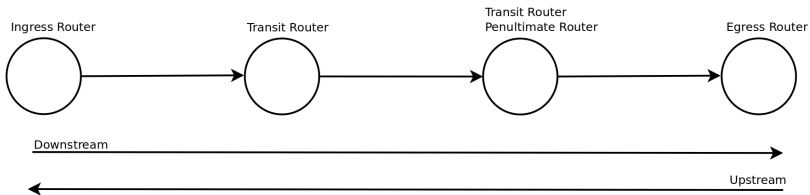
- Etiquetas (Labels) - El envío en una red MPLS se hace a través de labels o etiquetas.
- La asignación de las mismas es en sentido upstream.
- Los valores tienen significado local.
- Son parte del encabezado MPLS (shim header) de 32 bits. Está entre el encabezado IP y el encabezado de Nivel 2.
- Penultimate Hop Popping (PHP) - Realiza la operación de “Label Pop” y remueve información MPLS del encabezado del paquete. El envío se hace en base a la información del encabezado MPLS, pero es un paquete nativo IP hacia el próximo salto del LSP luego de decrementar el TTL del mismo.
- Permite mayor escalabilidad de la red ya que permite dividir el trabajo entre el Penultimate Router y el Egress Router.

Componentes Funcionales

Labels, PHP y UHP

- Hay fabricantes cuyo software trabaja por defecto con PHP en los LSP dinámicos.
- Ultimate Hop Popping (UHP) - Cuando el Egress Router realiza la operación de “Label Pop”. Es decir remueve la etiqueta de MPLS y realiza la búsqueda IP para el envío del paquete.
- Diferencia entre PHP y UHP está en el Egress Router y la etiqueta que le envía al Penultimate Router. Un valor de 3 implica que el Penultimate Router debe hacer PHP y enviar paquetes IP nativos.
- Un valor de 0 implica que el Penultimate Router debe hacer un “Label Swap” y enviar un paquete con encabezado MPLS.

Componentes Funcionales



Dada la bidireccionalidad de las comunicaciones de datos y la unidireccionalidad de los LSP, los mismos componentes deben estar presentes para el establecimiento del LSP en el sentido opuesto (el cual cambia también la referencia respecto a upstream y downstream).

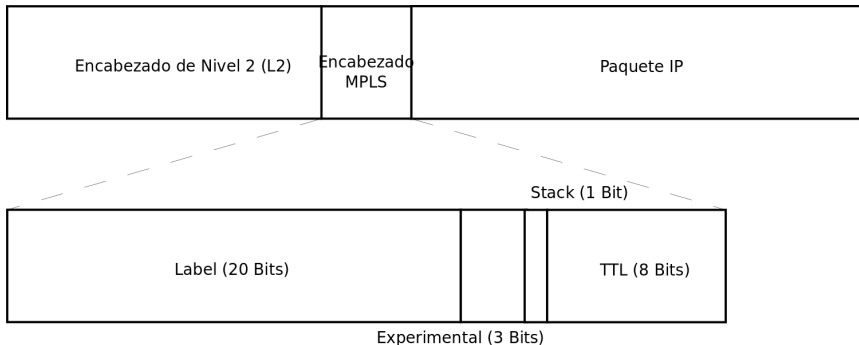
Header MPLS

Estructura

- Tiene 32 bits de longitud
 - Label (20 bits)
 - Experimental Bits (3 bits) - Se usa para identificar calidad de servicio.
 - Su nombre surge de la definición original en la RFC 3032 (actualizada por las RFC 3270, 5129 y 5462).
 - Stacking Bit (1 bit) - Valor 1 indica que sigue un paquete IP, valor 0 indica que hay otros encabezados MPLS.
 - Time To Live (8 bits) - Funcionalidad similar al TTL del encabezado IPv4 (Hop Limit en IPv6).
 - Opción de propagación de TTL entre el dominio IP Nativo y el dominio MPLS y viceversa.
- Cuando se incluye el encabezado MPLS se cambia el campo EtherType (0x8847 en el caso de IP unicast etiquetado y 0x8100 para IP unicast sin etiqueta).

Header MPLS

Estructura



Etiquetas

Valores reservados

- Tiene 20 bits de longitud
- Están en el rango de 0 a 15.
- 0 - IPv4 Explicit Null - Se hace un “Label Pop” y una búsqueda en la tabla de ruteo IPv4 para el envío del paquete.
- 1 - Router Alert Label - El paquete requiere un análisis particular (no es enviado directamente en función del valor de la etiqueta de ingreso).
 - Similar a la opción en el encabezado IP (RFC 2113 -The Router Alert option has the semantic “routers should examine this packet more closely”).
- 2 - IPv6 Explicit Null - Similar al valor 0 pero para IPv6.
- 3 - Implicit Null - No aparece nunca en el encabezado MPLS. Implica que se está trabajando con PHP.
- 4-15 - Reservados para uso futuro.

Tipos de LSP

- Estas etiquetas se asignaron como parte del establecimiento del LSP. El establecimiento del LSP puede ser estático o dinámico.
- LSP Estático. Cada router en el camino debe ser configurado explícitamente. Usa menos recursos pero cuenta con las desventajas de las definiciones estáticas (tolerancia a fallas, carga operativa, etc.)
- LSP Dinámico. Requiere señalización para su establecimiento ya que sólo se configura el ingress router.
 - En este contexto se entiende por señalización la existencia de un mecanismo dinámico de intercambio de información relacionada al dominio MPLS.
- LSP Estático PVC ATM
- LSP Dinámico SVC ATM
- Protocolos de Señalización. Resource Reservation Protocol (RSVP) y LDP (Label Distribution Protocol)

Procesamiento de paquetes

- En el ingress router se realiza una búsqueda IP (IP lookup) y se identifica como próximo salto un LSP. Se agrega un encabezado MPLS al paquete con la correspondiente etiqueta (“Label Push”).
- Routers en el medio del LSP realizan la operación de “Label Swap” y decrementan el TTL. No utilizan información del encabezado IP. Luego el Egress Router finaliza el LSP. Recibe el paquete del Penultimate Router y realiza una búsqueda en IP.
- Hay 2 casos diferentes de procesamiento según el trabajo del Penultimate Router: PHP y UHP

LSP Estáticos

- Se debe configurar en cada router a través del camino o que formará parte del LSP.
- Hay implementaciones de fabricantes que reservan ciertos valores de etiquetas para los LSP estáticos (es decir no se asignan para LSP dinámicos, aunque pueden trabajar con ellos si les son anunciados).
- No tolerante a fallas.
- Gran carga operativa.
- Diseño sencillo.

LSP Estáticos. Ejemplo

```
Ingress Router
[edit protocols mpls]
user@Ingress_Router# show
static-path inet {
  10.100.100.0/24 {
    next-hop 10.10.10.1;
    push 912;
  }
}
interface all;
```

```
Transit Routers
[edit protocols mpls]
user@Transit_Router1# show
interface all;
interface so-0/0/0.0 {
  label-map 912 {
    next-hop 10.10.20.2;
    swap 36;
  }
}
```

```
[edit protocols mpls]
user@Transit_Router2# show
interface all;
interface so-0/0/0.0 {
  label-map 36 {
    next-hop 10.10.30.3;
    swap 0;
  }
}
```

Egress Router
El Egress Router no necesita ninguna configuración particular, sólo habilitar MPLS en las interfaces, ya que realizará un Label Pop cuando reciba un paquete MPLS con una etiqueta 0 del Penultimate Router.

Protocolos de señalización

RSVP vs. LDP

- RSVP fue adaptado para MPLS, en cambio LDP fue diseñado específicamente para MPLS.
- Originalmente RSVP (IETF) fue diseñado para reservar recursos en la red (IntServ) pero nunca logró extender su uso.
- En cambio logró un despliegue importante con TE y establecimiento de LSP MPLS.

RSVP

Mensajes

- Usa flujos unidireccionales de señalización (y envío).
- El ingress router usa un PathMessage downstream (salto a salto) hasta el egress router. Cuando el mensaje alcanza el egress router la reserva real comienza y se realiza con un ResvMessage desde el egress router upstream hacia el ingress router.
 - Path Message
 - ResvMessage
 - Soft State
 - Refresh periódicos de cada router de Path y Resv para todos los Soft State activos.

RSVP

Mensajes

- También existen los mensajes PathTear (downstream) y ResvTear (upstream) para remover el Soft State y el camino establecido en el router que lo recibe.
 - PathTear
 - ResvTear
- PathError y ResvError no remueven el Soft State.
 - PathError
 - ResvError
- ResvConf permite confirmar a cada router , al egress router, que el ResvMessage fue recibido.
 - ResvConf

RSVP

Extensiones MPLS

- Permite incorporar información de etiquetas en mensajes Resv y Path.
- **Explicit Route Object (ERO)** - Permite que los mensajes Path vayan por camino específico de la red, incluso distinto al establecido por el IGP. El ingress router agrega el ERO al mensaje Path y los routers de tránsito lo cumplen.
 - ERO Loose Hops. El LSP debe transitar los nodos explicitados en el orden indicado con IGP estableciendo el camino entre ellos.
 - ERO Strict Hops. Informa el camino exacto del LSP. Los nodos mencionados deben estar directamente conectados (caso contrario mensaje de error y el LSP no se establece).
 - Se pueden soportar ambos en el mismo LSP.
- **Label Request Object (LRO)** - Parte del mensaje Path, permite a los routers asignar una etiqueta al LSP (lo guarda en el Soft State). Cuando llega el mensaje Resv del vecino downstream, la etiqueta es anunciada upstream en un RSVP Label Object.

RSVP

Extensiones MPLS

- **Label Object (LO)** - Parte del mensaje Resv, permite anunciar la etiqueta upstream.
- **Record Route Object (RRO)** - Tanto en mensajes Path como Resv. Registra los saltos intermedios y un router no envía el mensaje Resv si en la lista del RRO aparece una interface suya. Evita loops.
 - El ingress router genera un mensaje Path con el atributo RRO y cada router en el camino downstream agrega su interface de salida al RRO. Si hay loop se descarta el mensaje y se genera un mensaje PathError hacia el Ingress Router.
 - También el egress router genera un mensaje Resv con el atributo RRO y cada router en el camino upstream agrega su interface de salida al RRO. Si hay loop se descarta el mensaje y se genera un mensaje ResvError hacia el Egress Router.

RSVP

Extensiones MPLS

- **Session Attribute Object** - Parte del mensaje Path. Permite gestionar prioridad, affinity class, capacidades de local re-routing y preemption. Además se puede asignar un string ASCII para identificar el LSP en cada router de la red.
- **Traffic Specifier Object** (Tpspec Object) - Parte del mensaje Path. Contiene información como el ancho de banda del LSP (permite el establecimiento del LSP, no CAR), tamaños mínimos y máximos de paquetes del LSP.
- También se agrega mantenimiento del camino (LSP) y aspectos de escalabilidad.
 - Hellos para convergencia mas rápida y la opción de enviar varios Refresh en un mismo mensaje.

LSP dinámicos - RSVP

Ejemplo

```
Ingress Router
[edit protocols mpls]
user@Ingress_Router# show
label-switched-path Ingress-to-Egress {
to 10.100.100.100;
bandwidth 15m;
no-cspf;
primary via-Shiraz;
}
path via-Shiraz {
10.10.10.10 loose;
}
interface all;
```

Transit y Egress Routers

Tanto los Transit como el Egress Router no necesitan ninguna configuración particular, sólo habilitar MPLS y los protocolos de señalización en las interfaces correspondientes.

LDP

Mensajes

- Otro protocolo que permite señalar el establecimiento del LSP.
- Descubre vecinos adyacentes a través de los Hellos y luego establece una sesión TCP con ellos.
 - Los Hellos se envían periódicamente con dirección destino multicast (224.0.0.2, todos los routers en el segmento) hacia el puerto UDP 646.
 - Luego el router con el Router ID mas alto (Active Node) inicia una sesión TCP hacia el puerto 646 con dirección destino unicast (la IP origen del Hello o la dirección de transporte anunciada en el Hello).
- Descubrimiento de vecinos no adyacentes difiere sólo en la dirección destino de los mensajes Hellos. Se envían a la dirección IP unicast (debe ser definida) del vecino y luego el mecanismo de establecimiento es el mismo que para el caso de vecinos adyacentes.

LDP

Mensajes

- Una vez establecida la sesión TCP se intercambian mensajes de inicialización, keepalives y, finalmente, se inicia el intercambio de etiquetas/destinos.
 - Mensajes de Inicialización
 - Keepalives
 - LDP Address Mensajes
 - LDP Address Withdraw Mensajes

LDP

Parámetros de distribución

- **Unsolicited Downstream Distribution.** Los routers envían las etiquetas a todos los vecinos sin que se lo soliciten.
- **Independent Control Mode.** Los routers asignan y distribuyen las etiquetas de manera independiente (por su propia cuenta).
- **Liberal Label Retention Mode.** Los routers mantienen las etiquetas no usadas (aquellas enviadas por routers que no son el próximo salto o next-hop). Mejora los tiempos de convergencia.

LDP

FEC

- Surge el concepto de **Forwarding Equivalence Class (FEC)**.
- Una FEC se define como un conjunto de paquetes que son enviados de la misma manera, sobre el mismo camino y tienen el mismo tratamiento.
- Cada router asigna una determinada etiqueta a un FEC y lo anuncia a sus vecinos. Los vecinos hacen lo propio y de esta manera se construye el LSP.
- Por lo tanto, el envío se realiza asignando un paquete a un determinado FEC, determinando el next-hop del mismo y la correspondiente etiqueta/LSP.

Usos actuales

- Segmentar tráfico/servicios distintos sobre una infraestructura común
- Redes N-Play
- VPN para servicios de Capa 2 (ejemplos PW, VPLS) y Capa 3 (ejemplos VRF)
- Optimizar información de ruteo en los *Transit Router* o P (ejemplo: sesiones BGP entre routers no directamente conectados)
- Transición a IPv6 (ejemplo: topologías usando 6PE)