UNIVERSIDAD CARLOS III DE MADRID

Escuela Politécnica Superior

PROYECTO FIN DE CARRERA

Ingeniería Técnica de Telecomunicaciones: especialidad de Telemática

Integración y optimización de redes MPLS: Un caso práctico

Autor: Álvaro González Carrasco
Tutor: Vicente Palacios Madrid
Julio de 2011

UNIVERSIDAD CARLOS III DE MADRID

Escuela Politécnica Superior

PROYECTO FIN DE CARRERA

Ingeniería Técnica de Telecomunicaciones: especialidad de Telemática

Integración y optimización de redes MPLS: Un caso práctico

Autor: Álvaro González Carrasco
Tutor: Vicente Palacios Madrid
Julio de 2011

Dedicado a mi hermana Patricia.

AGRADECIMIENTOS

Es el momento de acordarse de la gente que ha participado en mi vida universitaria, será más corto de lo previsto porque ya los ánimos no son los de antes pero bueno….

En primer lugar quisiera agradecer a Vicente por las facilidades que me has dado para la realización de este proyecto, dadas las circunstancias era justo lo que necesitaba.

A mi padre y mi madre, que han hecho de mi la persona capaz de realizar este proyecto y de muchas otras cosas que no se pueden agradecer en un párrafo.

A mi hermana que sabe de que va todo esto porque ha vivido lo mismo que yo.

A mi abuela que le hace mucha ilusión que acabe este proyecto.

A Almudena que ha sentido este proyecto como suyo y me ha prestado toda la ayuda que estaba en su mano en todo momento.

A Javi que gracias a su gusto y consejos estéticos, este proyecto es más bonito y ordenado.

A mi amigo David qué se ha leído la memoria entera.

A Álvaro, porque siempre ha estado muy pendiente de que acabe el proyecto, a Javi, Movilla y Alicia por sus sabias palabras en el momento adecuado.

A mis compañeros de clase, ya son más amigos que compañeros de clase y los nombro: Carlos, Raquel, Luis, Alex Alonso, Alex de Frutos, Valentín, Crispi, Raquel, Dani, Gemelo, Costo, Mena, Patri, Irene, Patillas y Mariu.

A mis compañeros de la beca que ralentizaron más aún la finalización del proyecto pero lo pasamos bien: Sergio, Movi, Mariano, Jeda, Gonzalo, Alfonso, Paula, Esther y Gonzalo.

Chesco también ralentizó la finalización del proyecto.

INDICE GENERAL

INDICE DE ILUSTRACIONES.

INDICE DE TABLAS

1 INTRODUCCIÓN

1.1 MOTIVACIÓN

El punto de partida de este proyecto es el supuesto de la unión de dos empresas lo cual tiene como consecuencia, entre otras, la necesidad de integrar las dos redes de comunicaciones de estas compañías. Esto tiene muchas implicaciones: desarrollar un plan de unificación del direccionamiento IP, integrar los sistemas, interconectar todas las sedes de las compañías que ahora pasan a ser una sola, etc…

Cada una de las compañías posee una filial que gestiona sus comunicaciones que a su vez también se fusionan en una sola.

Ambas empresas filiales poseen y gestionan una red troncal MPLS sobre la que prestan servicio a sus respectivas empresas matrices así como a otros clientes. Estas redes troncales proporcionan comunicación entre las distintas sedes que están distribuidas por toda la geografía.

Las dos empresas filiales gestionan cada una de las redes de telecomunicación que en el supuesto de este proyecto conoceremos como red de UFIN y red de DC.

La red de UFIN tiene un número menor de nodos y geográficamente está más concentrada. Además proporciona servicios de comunicaciones a la empresa matriz y a un número importante de clientes.

La red de DC tiene un número mayor de nodos y geográficamente es más dispersa. Prácticamente la totalidad de sus servicios los presta a la empresa matriz.

Ambas redes MPLS se han desplegado sobre la red de transmisión propia que ambas empresas poseen y que ha determinado, en gran parte, su distribución geográfica.

Puesto que es necesario dar conectividad a las dos redes y además la red de transporte de cada una de las empresas pasa a ser gestionada por una misma entidad, el paso lógico es el de la integración de las dos redes MPLS en una sola. El hecho de tener una sola red MPLS tiene las siguientes motivaciones:

1. Permitir la conectividad sobre la misma infraestructura de cualquier sede procedente de cualquiera de cada una de las dos compañías posibilitando la integración de las dos redes de comunicaciones de las empresas matrices que son los principales clientes.

2. Poseer una red con un mayor número de nodos y más extendida geográficamente permite tener más cobertura para ofrecer servicios a un mayor número de sedes de la empresa matriz y otros potenciales clientes.

1.2 OBJETIVOS

Partiendo de las motivaciones indicadas anteriormente, este proyecto persigue los objetivos que se describen a continuación, los cuales permitirán alcanzar la unificación de las dos redes MPLS.

1. Integración total de las dos redes. Esto implica la unificación de las dos redes a todos los niveles MPLS y MPLS VPN. Esto permite ofrecer cualquiera de los servicios posibles tanto de nivel 2 como de nivel 3 entre dos nodos cualesquiera de la red.

2. Homogeneización del diseño de la red. Los criterios de diseño que emplearon en cada una de las dos redes eran diferentes por lo que se persigue unificarlos para simplificar el funcionamiento y la operación de la red resultante. Esto aplica a diversos puntos: elección de IGP, protocolos de routing, parámetros de LDP, etc...

3. Optimización de la topología. Con el objetivo de ahorrar costes (eliminando equipamiento y enlaces redundantes) y de mejorar la disposición de los nodos (minimizando el número de saltos y mejorando el camino de los principales flujos de tráfico) se hará un rediseño de la topología de la red.

Un aspecto muy importante del proyecto y que ha condicionado la forma de actuar es que, puesto que estamos planteandolo como situación real, en cada tarea, una de las prioridades es minimizar el impacto que tendría cada acción para una indisponibilidad del servicio lo menor posible. Así, tras cada tarea, describiremos el impacto que se produciría en una red real.

Para alcanzar estos objetivos, el proyecto se desarrollará en diversas fases que se explican a continuación junto con el resto de la estructura de la memoria.

1.3 ESTRUCTURA

Para su mejor entendimiento, la memoria de este proyecto se ha estructurado de la siguiente manera:

En el capítulo **1. Introducción** se han descrito tanto las motivaciones para la realización de este proyecto como los objetivos y se fija el alcance de los mismos.

En el segundo capítulo **2. Estado del arte**, se describirán tanto el estado de las tecnologías empleadas como el impacto de las mismas en el proyecto. Se divide en dos bloques que son:

- MPLS: Aquí se describe tanto la arquitectura propia de MPLS como los protocolos necesarios para su funcionamiento: LDP y MP-BGP.
- Protocolos de routing comunes en entornos MPLS: En este apartado se describe el funcionamiento y características de OSPF e IS-IS que son los protocolos utilizados durante la vida real de este proyecto.

El capítulo **3. Herramientas** se describen los programas y aplicaciones utilizadas para la elaboración del proyecto, tanto de gestión de equipos utilizadas en el entorno de producción, como las que se han utilizado para la elaboración de la memoria.

El capítulo **4. Desarrollo del proyecto** contiene la parte más importante de este proyecto. Describe tanto los estados iniciales de ambas redes y las fases en las que se dividió la realización del proyecto así como los posibles problemas y su resolución. Estas fases son 6:

- La fase 0 es la definición e implementación de la topología física de red final. Al unificar dos redes independientes es necesario redefinir y optimizar la arquitectura de red considerando el número total de nodos y enlaces existentes. Para alcanzar la topología final es necesario la modificación de algunos enlaces troncales, lo que implica el alta y baja de nuevas conexiones entre varios nodos. Esto, de manera lógica forma otra fase, la fase 0.

  En el desarrollo temporal del proyecto, la fase 0 representa la definición de la nueva topología, pero su implementación, dado que nos encontramos en un supuesto que pretende acercarse a la realidad, se ejecutará de manera intercalada con el resto de fases, ya que, el establecimiento y eliminación de los enlaces físicos no se pueden hacer todos agrupados en el mismo periodo de tiempo, puesto que dependen de varios factores como pueden ser: disponibilidad de los medios de transmisión propios, equipamiento, plazos de operador público, obra de canalización, etc.. Esto se hace así por no retrasar el resto de tareas y en la memoria se han ido intercalando con fases según se han ido ejecutando de manera secuencial. Se ha decidido no agruparlas para que las capturas de los comandos que se incluyen en esta memoria, y que son tomadas directamente de los equipos en su momento temporal concreto, no sea necesario hacer supuestos sobre la situación de la red en cada momento.

- La fase 1 es la unión de ambas redes a nivel IP. En esta fase se establece conexión a nivel IP entre ambas redes pero a nivel MPLS. Con esta interconexión no se puede dar ningún servicio aún entre ambas redes.

- La fase 2 describe la homogeneización de ambas redes a nivel MPLS. Esto supone que sobre ese enlace se pueda establecer LDP y el IGP. Esto supone que sobre la red se pueden dar servicios de AToM pero no de VPN.

- La fase 3 es la homogeneización del MP-BGP unificando el AS y estableciendo los route reflector para todos los equipos de la red. Ya si tenemos una única red MPLS y se pueden ofrecer todos los servicios homologados en ambas redes.

- La fase 4 es la unificación de la versión de IOS de los nodos MPLS de la red. Si bien esta fase no es imprescindible, el tener una única versión de IOS en todos los nodos evita complejidad en su operación y mantenimiento.

- La fase 5 supone la unificación de la calidad de servicio de los enlaces troncales de la red. De cara a problemas de congestión es importante que haya criterios uniformes en toda la red.

El capítulo **5. Presupuesto,** es donde se detalla el gasto monetario asociado a cada tarea.

En el capítulo **6. Conclusiones,** se detallan, tras el análisis de los resultados del proyecto, las conclusiones obtenidas.

El en capítulo **7. Trabajos futuros**, quedan reseñadas algunas tareas posteriores que se pueden derivar de este proyecto, algunas son estéticas pero básicamente se refieren a modificar criterios que repercuten en la configuración de una red provocados por la fusión de filosofías distintas.

El octavo capítulo contiene las **Referencias Bibliográficas** empleadas para la realización de esta memoria.

Por último el capítulo **9. Glosario,** contiene los términos más frecuentes que se han empleado en esta memoria.

2 ESTADO DEL ARTE

2.1 MPLS

2.1.1 ANTECEDENTES AL MPLS Y EVOLUCIÓN

El enorme crecimiento de la red Internet ha convertido al protocolo IP en la base de las actuales redes de telecomunicaciones, contando con más del 80% del tráfico cursado. La versión actual de IP, conocida por IPv4 y recogida en la RFC 791, lleva operativa desde 1980. Este protocolo de capa de red (Nivel 3 OSI), define los mecanismos de la distribución o encaminamiento de paquetes, de una manera no fiable y sin conexión, en redes heterogéneas; es decir, únicamente está orientado a servicios no orientados a conexión y a la transferencia de datos, por lo que se suele utilizar junto con TCP (Nivel 4 de OSI) para garantizar la entrega de los paquetes.

A mediados de la década de los 90, la demanda por parte de los clientes de los ISP de aplicaciones multimedia con altas necesidades de ancho de banda y una calidad de servicio o QoS garantizada, propiciaron la introducción de ATM en la capa de enlace (Nivel 2 de OSI) de sus redes. En esos momentos, el modelo de IP sobre ATM satisfacía los requisitos de las nuevas aplicaciones, utilizando el encaminamiento inteligente de nivel 3 de los routers IP en la red de acceso, e incrementando el ancho de banda y rendimiento basándose en la alta velocidad de los conmutadores de nivel 2 y los circuitos permanentes virtuales de los switches ATM en la red troncal. Esta arquitectura, no obstante, presenta ciertas limitaciones, debido a: la dificultad de operar e integrar una red basándose en dos tecnologías muy distintas, la aparición de switches ATM e IP de alto rendimiento en las redes troncales, y la mayor capacidad de transmisión ofrecida por SDH/SONET y DWDM respecto a ATM.

Durante 1996, empezaron a aparecer soluciones de conmutación de nivel 2 propietarias diseñadas para el núcleo de Internet que integraban la conmutación ATM con el encaminamiento IP; como por ejemplo, Tag Switching de Cisco o Aggregate Route-Based IP

Switching de IBM. La base común de todas estas tecnologías, era tomar el software de control de un router IP, integrarlo con el rendimiento de reenvío con cambio de etiqueta de un switch ATM y crear un router extremadamente rápido y eficiente en cuanto a coste. La integración en esta arquitectura era mayor, porque se utilizaban protocolos IP propietarios para distribuir y asignar los identificadores de conexión de ATM como etiquetas; pero los protocolos no eran compatibles entre sí y requerían aún de infraestructura ATM.

Finalmente en 1997, el IETF establece el grupo de trabajo MPLS para producir un estándar que unificase las soluciones propietarias de conmutación de nivel 2. El resultado fue la definición en 1998 del estándar conocido por MPLS, recogido en la RFC 3031.

Se desarrolló como un protocolo de conmutación por etiquetas definido para funcionar sobre múltiples protocolos como Sonet, Frame Realy, ATM, Ethernet o cualquiera sobre el que pueda funcionar PPP. Las principales motivaciones para su desarrollo son la ingeniería de tráfico, la diferenciación de clases de servicio, y las redes privadas virtuales (VPN). En un principio, también proporcionaba una mayor velocidad puesto que los routers sólo deben mirar la etiqueta para conmutar y no leer la cabecera de la capa 3 para después decidir por dónde enrutar en función del destino y/u otros parámetros. Sin embargo, hay tecnologías que han conseguido aumentar la velocidad de los routers para consultar las tablas de enrutamiento (como ASIC).

Las ventajas que llevaron a desarrollar ATM: uso de conmutadores ATM más rápidos porque funcionaban con etiquetas, el poder ofrecer ingeniería de tráfico mediante circuitos virtuales... llevan a desarrollar MPLS unos años más tarde, basándose en la idea de las etiquetas, pero reduciendo la complejidad de las redes IP sobre ATM y mejorando la funcionalidad en algunos casos. IP sobre ATM conseguía aprovecharse de la velocidad que proporcionaban los conmutadores ATM para unir los routers IP, pero seguían siendo dos redes separadas (complejo de gestionar), y el número de circuitos virtuales aumenta mucho con el tamaño de la red. Varios fabricantes intentaron mejorar esto proponiendo soluciones mediante etiquetas que separasen las funciones de routing (encaminamiento, control de por dónde se envían los paquetes) de las de forwarding (reenvío en sí). El problema ahora es que eran incompatibles entre sí. MPLS es un intento de estandarizar estas soluciones.

MPLS aprovecha lo mejor de la capa 2, la rápida conmutación, sin perder de vista la capa 3, para no perder sus posibilidades. Esto se consigue separando de verdad la función de conmutación de la de enrutamiento. MPLS hace más viable la ingeniería de tráfico, permite enrutamiento rápido (porque en realidad hace conmutación, pero con información de enrutado), permite que los equipos de reenvío sean más baratos si sólo deben entender paquetes etiquetados, permite ofrecer QoS basándose en diferentes CoS (clases de servicio), hace más fáciles y flexibles las VPN (redes privadas virtuales), y además parece el primer paso para conseguir redes totalmente ópticas (ya que decidimos por dónde enviar el paquete según lo que diga la etiqueta y no hace falta procesar la cabecera de orden 3; es decir, aunque las decisiones del enrutado sean en el dominio eléctrico, la conmutación podría ser óptica).

MPLS utiliza los campos para etiquetas de ATM o Frame Relay, o añade una cabecera para el resto de protocolos entre la del nivel 3 y la del nivel 2. La diferencia con IP sobre ATM es que no tenemos una red diferente que nos proporciona conexión entre routers IP, sino que

los niveles están integrados, y las funciones de encaminamiento y reenvío separadas pero coordinadas. Hay una parte de control, que se encarga de las decisiones de encaminamiento, pero no construye una tabla en la que consultar la dirección IP de los paquetes que lleguen, sino que informa a la parte de reenvío, que construye una tabla con etiquetas; así no es necesario mirar la cabecera de la capa 3, y decidir para cada paquete, porque la decisión ya está tomada para cada etiqueta. El único router que tiene que hacer funciones de enrutamiento es el primero, que tiene que decidir que etiqueta coloca a cada paquete. Todos los paquetes que llevan la misma etiqueta forman un grupo que se denomina Forwarding Equivalent Class (FEC).

2.1.2 ARQUITECTURA MPLS

La denominación de *multiprotocol* fue posterior a la implementación de MPLS sobre routers Cisco. Antes del estándar se denominaba Tag Switching y solo permitía IPv4, después se implementó para el uso de otros protocolos como puede ser IPv6. La denominación de *Label Switching* es debido a que no se enruta en base a prefijos IPv4 u otro protocolo, sino que se conmuta en base a etiquetas. A continuación se describe para que se usan las etiquetas, como se usan y como se distribuyen en la red.

2.1.2.1 Etiquetas MPLS

Una etiqueta MPLS es un campo de 32 bits con una determinada estructura.

Ilustración 1- Formato etiqueta MPLS

Los primeros 20 bits son la etiqueta. Este valor oscila entre 0 y 2^{20}-1 o 1.048.575 sin embargo los primeros 16 bits no se emplean normalmente, tienen significado especial.

Los bits del 20 al 22 son los 3 bits experimentales (exp). Son usados exclusivamente para hacer calidad de servicio (QoS).

El bit 23 es el indicador de final de pila (Bottom of Stack – BoS). Su valor es 0 a menos que sea el final de la pila, en cuyo caso tomará valor 1. La pila de etiquetas es un conjunto de etiquetas que puede estar formado por una sola etiqueta o más. El número de etiquetas que pueden formar un paquete es ilimitado aunque lo normal es que no haya más de 4.

Los bits del 24 al 31 son los 8 bits que forman el TTL. Este TTL tiene la misma función que el TTL del paquete IP. En cada salto se va decrementando en una unidad y su principal función es que un paquete no esté dando vueltas por la red durante un tiempo ilimitado. Cuando alcanza el valor de 0 el paquete se descarta.

2.1.2.2 Apilado de etiquetas

Los routers con capacidad MPLS podrían necesitar más de una etiqueta para formar un paquete y enrutarlo a través de la red MPLS. Esto se realiza gracias al apilado de etiquetas. La primera etiqueta de la pila recibe el nombre de *top label* o etiqueta más externa y la última es

la *bottom label* o etiqueta más interna. Entre ambas puede haber cualquier número de etiquetas.

Ilustración 2- Pila de etiquetas

Nótese que en la pila de etiquetas de la figura, el bit BoS está puesto a 0 en todas las etiquetas menos en la *bottom label*, donde su valor es 1.

2.1.2.3 Codificación de MPLS

La pila de etiquetas MPLS se sitúa delante del paquete de nivel 3, esto es, antes de la cabecera del protocolo transportado, pero después de la cabecera de nivel 2.

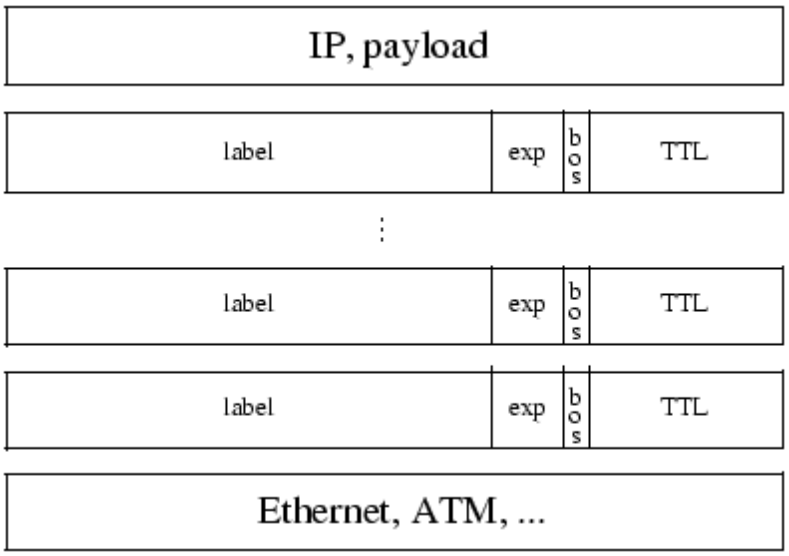

Ilustración 3- Encapsulación de paquetes

En el caso de nuestros equipos Cisco, el protocolo de capa 2 puede ser cualquiera que sea soportado por la versión de IOS: PPP, HDLC, Ethernet y demás… Suponiendo que el protocolo transportado es IPv4 y la encapsulación del enlace es PPP, la pila de etiquetas se situará después de la cabecera PPP pero antes de la cabecera IP. Debido a que la pila de etiquetas en la capa 2 está antes de la cabecera IP es necesario nuevos valores del campo *Data Link Layer Protocol*, indicando que lo que sigue a la cabecera del protocolo de nivel 2 es un paquete MPLS. El campo *Data Link Layer Protocol* es un valor que indica que tipo de tráfico está transportando la trama de nivel 2. En la siguiente tabla se muestra los nombres y valores del campo *Data Link Layer Protocol* en los diferentes protocolos de nivel 2.

Protocolo nivel 2	Nombre identificador protocolo nivel 2	Valor (hex)
PPP	PPP protocol field	0281
Ethernet / 802.3 LLC / SNAP encapsulation	Ethertype value	8847
HDLC	Protocol	8847
Frame Relay	NLPID (Network Level Protocol ID)	80

Tabla 1. –Códigos de MPLS en protocolos de nivel 2

2.1.2.4 MPLS y el modelo de referencia de OSI

Atendiendo al modelo de referencia de OSI, el nivel más bajo es el nivel físico, este nivel tiene que ver con propiedades físicas del cableado, aspectos mecánicos y características eléctricas. El nivel 2, nivel de enlace, tiene la responsabilidad de la conformación de las tramas. PPP, HDLC, Ethernet, etc…. son ejemplos de protocolos de capa 2. Este nivel sirve para establecer la comunicación entre dos máquinas pero no va más allá. De eso se encarga el nivel 3, nivel de red. IP es el protocolo más extendido. Tiene la responsabilidad de la comunicación extremo a extremo, el paquete no cambia en ningún momento a diferencia de los paquetes de nivel 2 que cambian por cada enlace que pasan.

Viendo esta clasificación, MPLS no es un protocolo de nivel 2 ya que es necesario que exista un nivel 2 y tampoco es de nivel 3 por lo mismo. Por eso y ya que la pila de etiquetas MPLS se sitúan entre el paquete de nivel 2 y 3, lo usual es verlo como un protocolo de capa 2,5.

2.1.2.5 Nodo MPLS (Label Switch Router - LSR)

Durante toda esta memoria haremos mención al concepto de nodo MPLS (también conocido como *Label Switch Router*). Un nodo MPLS solo es un router que soporta el protocolo MPLS, de hecho los equipos que componen ambas redes y que detallaremos ampliamente más adelante son routers comunes Cisco con funcionalidad MPLS. Esta funcionalidad MPLS se refiere a la capacidad de entender etiquetas MPLS y de recibir y transmitir paquetes etiquetados en los enlaces. Los tres tipos de nodos que existen en una red MPLS son:

1. Ingress LSR: Este router recibe un paquete que no está etiquetado aún, inserta la pila de etiquetas y lo manda por los enlaces.
2. Egress LSR: Este equipo recibe el paquete etiquetado, quita la etiqueta y lo manda por los enlaces.
3. Intermediate LSR: Reciben paquetes etiquetados, realizan operaciones sobre el paquete y lo reenvían por el enlace que corresponda.

4. Equipo conectado a la red MPLS: No es propiamente un componente de una red MPLS pero si nos referiremos ampliamente a ellos por lo que los enumeramos. Comúnmente se llaman CE o CPE.

   El Ingress y Egress LSR son dos tipos de equipos que también se denominan Provider Edge (PE). Aunque esta denominación es solo para entornos de VPN MPLS (que veremos más adelante) se utiliza de manera general. Un PE puede ser simultáneamente Ingress y Egress LSR. Lo mismo ocurre con los equipos Intermediate LSR que se denominan Provider (P).

   En la siguiente ilustración vemos un ejemplo de red MPLS con todos sus elementos.

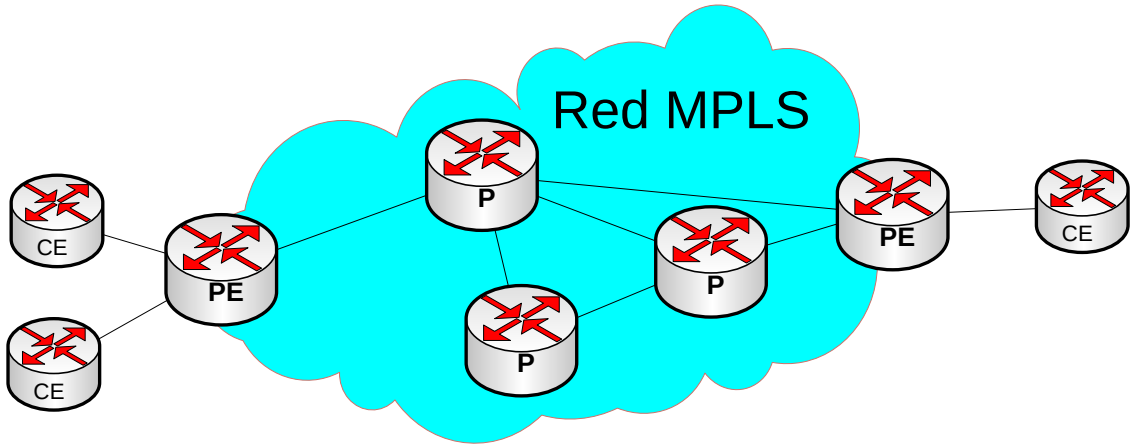

Ilustración 4- Elementos de una red MPLS.

2.1.2.6 Label Switched Path (LSP)

Un *Label Switched Path* o *LSP* es una secuencia de LSRs que conmutan un determinado paquete etiquetado a través de una red MPLS o parte de ella. Básicamente el LSP es un "camino" a través de la red MPLS. El primer LSR de un LSP es siempre un Ingress LSR y el último es un Egress LSR. Todos los routers intermedios del LSP son Intermediate LSR. En apartados posteriores veremos cómo se construye ese LSP.

Un detalle importante es que los LSPs no son reversibles, es decir, solo valen para un sentido de la comunicación.

2.1.2.7 Forwarding Equivalence Class

Un Forwarding Equivalence Class (FEC) es un grupo o flujo de paquetes que son enviados sobre un mismo camino y son tratados de la misma manera en lo relativo a cuestiones de conmutación. Todos los paquetes que pertenezcan al mismos FEC tienen la misma etiqueta pero no todos los paquetes que tienen la misma etiqueta pertenecen al mismo FEC. Como sus valores del campo EXP pueden ser distintos y su conmutación puede ser distinta, estos pueden pertenecer a distintos FEC. El router que decide que paquetes pertenecen a cada FEC son los Ingress LSR. Esto es lógico ya que el Ingress LSR clasifica y etiqueta los paquetes. Ejemplos de FECs son:

- Paquetes con dirección destino de nivel 3 que coinciden con un cierto prefijo
- Paquetes multicast que pertenezcan a un determinado grupo
- Paquetes con el mismo tratamiento de reenvío, basado en los campos de IP precedence o DSCP.
- Paquetes con la misma dirección de destino de nivel 3 que pertenezcan a un grupo de prefijos BGP, todos con el mismo siguiente salto de BGP. Esto quiere decir que todos los paquetes que entran por el mismo Ingress LSR y salen por el mismo Egress LSR tienen la misma etiqueta.

2.1.2.8 Distribución de etiquetas

La primera etiqueta la pone el Ingress LSR y pertenece a un LSP. El camino del paquete a través de la red MPLS está definido por el LSP. Todos los cambios se realizan sobre la etiqueta más externa. El Ingress LSR impone una o más etiquetas, los intermédiate LSR cambian la etiqueta externa, la de entrada, por otra y transmiten el paquete por el enlace de salida que corresponda. El Egress LSR del LSP quita todas las etiquetas del LSP y reenvía el paquete al CPE que corresponda.

Consideremos el ejemplo de IPv4 sobre MPLS, es el ejemplo más simple y común en una red MPLS. Todos los LSRs hablan un IPv4 Interior Gateway Protocol (IGP) como pueden ser OSPF, ISISI, EIGRP, etc.… El Ingress LSR mira la dirección IP destino del paquete, le pone la etiqueta y reenvía el paquete. El siguiente LSR recibe el paquete etiquetado, cambia la etiqueta más externa por otra y reenvía el paquete por el enlace correspondiente. El Egress LSR quita todas las etiquetas y envía el paquete IP por el enlace de salida adecuado. Para este proceso, todos los LSRs adyacentes deben estar de acuerdo en que etiquetas usar para cada prefijo IGP. Por lo tanto cada intermédiate LSR debe estar capacitado para decidir con que etiqueta de salida debe intercambiar la etiqueta de entrada. Esto significa que es necesario un mecanismo para que los routers sepan que etiquetas usar a la hora de encaminar un paquete. Las etiquetas son locales a cada par de routers adyacentes y no tienen sentido global para cruzar la red. Los routers adyacentes necesitan alguna forma de comunicación entre ellos para estar de acuerdo en que etiqueta usar para cada prefijo; los routers no saben que etiqueta de salida necesitan para sustituir cada etiqueta de entrada por si solos, es necesario un protocolo de distribución de etiquetas.

La distribución de etiquetas se puede hacer de dos formas:

- Distribución de etiquetas junto con la información de routing
- Utilizar un protocolo de routing especifico para la distribución de etiquetas

2.1.2.9 Distribución de etiquetas junto con la información de routing

Este método tiene la ventaja de que no es necesario otro protocolo distinto para la distribución de etiquetas aunque no es sencillo de implementar y mantener. La gran ventaja de este método es que está sincronizado el intercambio de prefijos y el de etiquetas, por lo que nunca se distribuirá por la red prefijos sin etiquetas ni etiquetas sin prefijos. Esto elimina la necesidad de de otro protocolo ejecutándose en el nodo MPLS con las ventajas de consumo de recursos en el router que esto supone. La implementación de un protocolo del tipo *vector*

distancia (como pudiera ser EIGRP) está altamente recomendado ya que cada router origina un prefijo desde su tabla de rutas, entonces el router asocia una etiqueta a cada prefijo que tenga.

Los protocolos del tipo *estado de enlace* (OSPF e IS-IS) no funcionan de esta manera. Cada router origina actualizaciones de estados de enlace que son reenviados sin modificar por todos los routers dentro de una misma área. El problema es que para que MPLS funcione, cada router necesita distribuir una etiqueta por cada prefijo que tenga el IGP en su tabla de rutas incluso si los routers no son los que originan este prefijo.

De todos los IGPs ninguno ha evolucionado para desarrollar este método, sin embargo, BGP puede transportar prefijos y etiquetas de manera eficiente. Pero BGP no es un IGP por lo que transporta prefijos externos y se usa principalmente para la distribución de etiquetas en las VPN MPLS.

2.1.2.10 Utilizar un protocolo de routing específico para la distribución de etiquetas

Este método tiene la ventaja de que la distribución de rutas se hace en un protocolo dedicado a ello. Independientemente del protocolo de routing que se use y tanto si tiene capacidad de distribución de etiquetas o no, es otro específico el que se encarga de esta tarea. La desventaja de esta opción es que son dos procesos ejecutándose en el nodo.

Para este cometido, el protocolo más usado es LDP, aunque no es el único, ejemplos son TDP, RSVP.

TDP es el precedente de LDP, fue el primer protocolo especifico para la distribución de etiquetas, desarrollado e implementado por Cisco. Después el IETF formalizó LDP. LDP y TDP son similares en el modo de funcionamiento pero LDP tiene más funcionalidades de TDP. Incluso en entornos Cisco, LDP fue sustituido por TDP que se ha quedado obsoleto.

2.1.2.11 Distribución de etiquetas con LDP

Para cada prefijo IGP en la tabla de rutas, el nodo crea una asociación local, es decir, asocia cada prefijo con una etiqueta. Entonces el nodo distribuye esta asociación a todos sus nodos vecinos. Estas asociaciones recibidas se denominan asociaciones remotas. Los vecinos entonces almacenan tanto las asociaciones remotas como las asociaciones locales en una tabla especial, la *Label Information Base* (LIB). Cada nodo tiene una sola asociación local por cada prefijo, y varias asociaciones remotas ya que lo lógico es tener varios vecinos.

Independientemente de las asociaciones remotas que reciba, el nodo debe seleccionar una sola etiqueta de salida para cada prefijo IP y decidir por enlace reenvía el tráfico. La tabla de rutas determina cual es el siguiente salto para cada prefijo IP. El nodo elige la asociación remota recibida del siguiente nodo en el camino. Este nodo es el siguiente salto en la tabla de rutas para ese prefijo. Así se va rellenando la *Label Forwarding Information Base* (LFIB) donde la etiqueta de la asociación local será siempre la etiqueta de entrada y la de salida será la de la asociación remota seleccionada gracias a la tabla de rutas. De esta forma, cuando un nodo recibe un paquete etiquetado, es capaz de intercambiar la etiqueta de entrada por la de salida obtenida del siguiente salto.

En la siguiente figura, se muestra como es este intercambio:

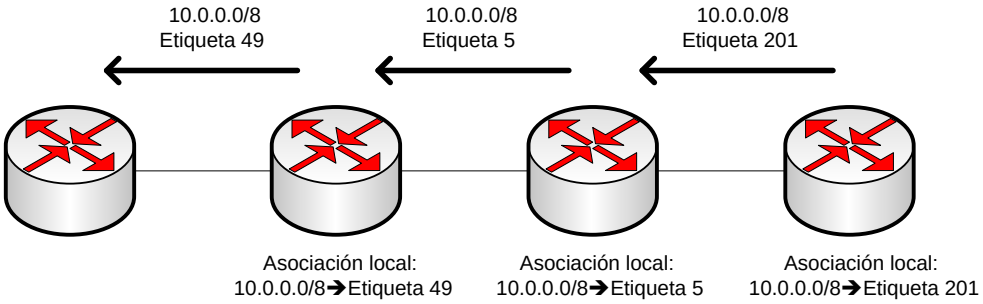

Ilustración 5- Intercambio de etiquetas por LDP para un prefijo

En la siguiente figura se muestra un paquete IP entrando en la red MPLS, se le pone la etiqueta 49 dado el prefijo IP destino y es enviado al siguiente nodo. El segundo nodo, intercambia la etiqueta de entrada (49) por la de salida (5) y envía el paquete hacia el tercer nodo. Este hace lo propio e intercambia la etiqueta de entrada (5) por la de salida (201) y reenvía de nuevo el paquete. Este proceso se repite hasta que sale de la red. Lo vemos gráficamente en la siguiente figura

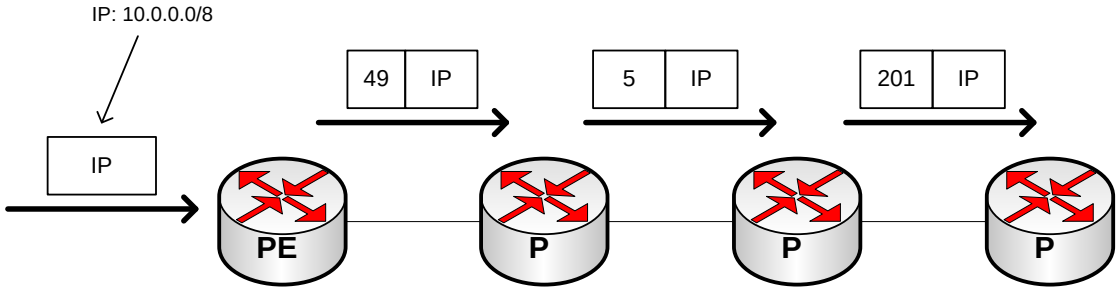

Ilustración 6- Etiquetado de un paquete IP en la red MPLS

2.1.2.12 Label Forwarding Instance Base

La LFIB es la tabla que se usa para reenviar paquetes etiquetados. Se va poblando con las etiquetas de entrada y salida para los LSPs. La etiqueta de entrada es la etiqueta de la asociación local (que previamente se pasó al vecino) en un nodo particular. La etiqueta de salida es la etiqueta de una asociación remota elegida por el nodo entre todas las asociaciones remotas recibidas. Todas estas asociaciones remotas se encuentran en la LIB. La LFIB elige solo una de todas las posibles etiquetas de salida de todas las asociaciones remotas en la LIB y la selecciona para la LFIB. La etiqueta remota elegida depende de que camino sea el mejor de todos los posibles. Para el caso de IP sobre MPLS, la etiqueta es asignada a un prefijo IP, sin embargo, la LFIB puede ser poblada con etiquetas que LDP no asigna. En el caso de ingeniería de tráfico las etiquetas son distribuidas por RSVP y si hablamos de VPN MPLS es BGP el que las asigna

2.1.2.13 MPLS Payload

La etiqueta MPLS no tiene campo de identificación de protocolo de red. Este campo está presente en todas las cabeceras de los protocolos de nivel 2 para indicar que protocolo de nivel 3 transportan. MPLS no tiene este campo porque no lo necesita, un nodo simplemente lee la etiqueta, la intercambia por otra y reenvía el paquete por el enlace adecuado.

Los nodos P no necesitan información de nivel 3, toda la información necesaria para conmutar un paquete se encuentra en la etiqueta más externa de la pila. Si la pila de etiquetas consiste en más de una etiqueta, las etiquetas más internas no podrían haber sido asignadas por el P y por tanto no podría tener conocimiento de ellas. Como los intermédiate LSR solo miran la etiqueta externa de la pila para tomar decisiones de encaminamiento, nunca supondrá un problema. Para realizar el encaminamiento, los Ps deben tener exclusivamente sus asociaciones locales y remotas para manipular la etiqueta externa.

El Egress LSR, el encargado de quitar la pila de etiquetas, si debe conocer que protocolo de red transporta MPLS porque debe encaminarlo más allá de la red. El Egress LSR es el que debe saber qué valor poner en el campo Network Level Protocol en la trama que envíe. Este nodo es el que puso la etiqueta (recibe el paquete con su asociación local) lo que significa que asignó la etiqueta para este FEC, por lo tanto el Egress LSR conoce que tipo de tráfico es mirando la etiqueta ya que fue creada por él mismo y conoce que tipo de tráfico es.

2.1.2.14 Espacio de etiquetas

En la siguiente figura podemos observar como el nodo A anuncia la etiqueta L1 para el FEC 1 al nodo B y la etiqueta L1 para el FEC 2 al nodo C pero solo si el nodo A puede distinguir de qué nodo recibe el paquete con etiqueta L1. En el caso del nodo B y C están directamente conectados al A. El hecho de que la etiqueta L1 es única por interfaz le da nombre al espacio de etiquetas: Espacio de etiquetas por interfaz. Si se usa este espacio, el paquete no es reenviado basándose exclusivamente en la etiqueta, sino que se observa el interfaz por el que se recibe para poder distinguir FECs.

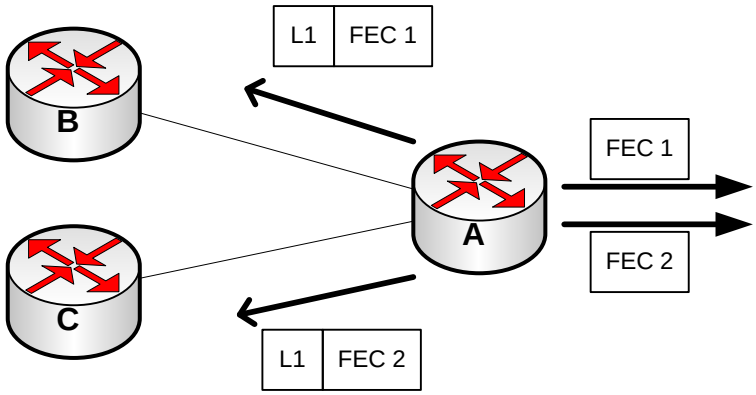

Ilustración 7- Espacio de etiquetas por interfaz

La otra posibilidad es que la etiqueta no sea única por interfaz, sino por nodo. Esto se denomina espacio de etiquetas por plataforma. En este caso el nodo A distribuye El FEC 1 con

la etiqueta L1 al nodo B y C como apreciamos en la siguiente figura. Si se distribuye una etiqueta para un FEC 2, debe ser distinta a L1. En este caso el reenvío siempre se hace exclusivamente atendiendo al valor de la etiqueta independientemente del interfaz por la que lo recibamos.

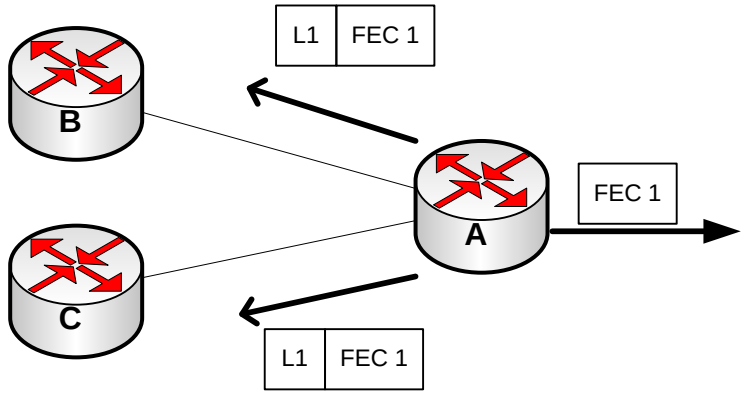

Ilustración 8- Espacio de etiquetas por plataforma

En Cisco, los interfaces LC-ATM tienen espacio de etiquetas por interfaz, el resto de interfaces tienen espacio de etiquetas por plataforma.

2.1.2.15 Modos de distribución de etiquetas

Un LSR puede usar diferentes modos de distribución a la hora de distribuir etiquetas a los demás LSRs:

1. Modos distribución de etiquetas
2. Modos retención de etiquetas
3. Modos control LSP.

Cada modo tiene sus características, vemos las ventajas de cada una.

- Modos de distribución de etiquetas

Hay dos modos para distribuir las asociaciones:

- Downstream-on-demand (DoD)
- Unsolicited Downstream (UD)

En el caso del modo DoD, cada nodo pregunta su siguiente salto (siguiente nodo) en el LSP una asociación de etiquetas para cada FEC. Cada nodo recibe una asociación por FEC de su siguiente salto para ese FEC.

En el modo UD, cada nodo distribuye una asociación a sus nodos adyacentes, sin que se lo soliciten. En este modo, un nodo recibe una asociación remota por cada nodo adyacente.

En el caso de DoD, la LIB muestra solo una asociación remota, mientras que en el caso de UD probablemente habrá varias. El modo de distribución de etiquetas usado depende del

tipo de interfaz e implementación. En routers Cisco todos los interfaces salvo LC-ATM usarán el modo UD. Todos los interfaces LC-ATM usan el modo DoD.

- Modos de retención de etiquetas.

Hay dos posibles modos que son:

- Liberal label retention (LLR)
- Conservative label retention (CLR)

En el modo LLR, un nodo mantiene en su LIB todas las asociaciones remotas recibidas. Una de esas asociaciones es la recibida del siguiente salto para un determinado FEC. La etiqueta de esa asociación remota recibida es la que figurará en la LFIB, pero ninguna de las etiquetas del resto de asociaciones remotas recibidas son puestas en la LFIB, de modo que no se usan para reenviar paquetes. Estas etiquetas se mantienen en la LIB ya que en cualquier momento la topología de la red puede cambiar (se puede caer un enlace o quitar un nodo por ejemplo) y entonces la LFIB se actualizaría más rápidamente con una etiqueta de la LIB.

El segundo modo de retención de etiquetas es el CLR. Un nodo que está en este modo no almacena todas las asociaciones remotas en la LIB, solo la que está asociada con el siguiente salto para un FEC particular.

En resumen, el modo LLR proporciona una rápida adaptación a cambios de red mientras que el modo CLR tiene la ventaja de que su LIB es más corta ya que almacena menos asociaciones lo que se traduce en optimización de memoria en el router.

En los routers Cisco, los interfaces LC-ATM emplean el modo CLR mientras que el resto de interfaces emplean el modo LLR.

- Modos de control LSP:

Hay dos modos que son:

- Modo de control independiente de LSP
- Modo ordenado de control LSP

El nodo puede crear una asociación local por FEC independientemente de los otros nodos. Esto se llama modo de control independiente de LSP. En este modo de control, cada nodo crea una asociación local para un particular FEC tan pronto como reconoce el FEC. Normalmente, esto se produce cuando añade un prefijo de un FEC a la tabla de rutas.

En el modo ordenado de control LSP, un nodo solo crea una asociación local para un FEC si reconoce que es el Egress LSR para el FEC o si el LSR ha recibido una asociación remota del siguiente salto para este FEC.

La desventaja del modo independiente es que algunos LSRs pueden empezar a conmutar paquetes antes de que el LSP se haya establecido completamente extremo a extremo; por lo tanto, el paquete no alcanzará el destino. Si el LSP no está completamente establecido el

paquete puede no ser conmutado o reenviado adecuadamente o puede ser incluso descartado.

Los routers Cisco usan el modo de control independiente.

2.1.3 REENVÍO DE PAQUETES ETIQUETADOS

En este apartado veremos cómo se reenvían los paquetes etiquetados a través de la red MPLS. El encaminamiento de paquetes en una red MPLS no solo se diferencia de una red IP en el análisis de la dirección IP por la consulta de una etiqueta en la LFIB, sino que se pueden realizar varias operaciones sobre las etiquetas. Estas son push, pop y swap.

Las operaciones consisten en:

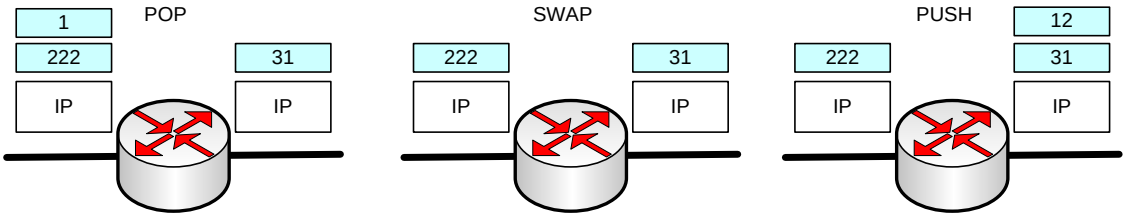

Ilustración 9- Operaciones POP, SWAP y PUSH

Comprobando la etiqueta externa del paquete etiquetado recibido y su correspondiente entrada en la LFIB, el nodo sabe como reenviar el paquete. El nodo determina que operación sobre la etiqueta debe ser aplicada, POP, SWAP o PUSH, y qué siguiente salto es debe ser reenviado el paquete.

Las operaciones que se pueden realizar sobre un paquete etiquetado son:

- SWAP consiste en cambiar la etiqueta externa de la pila por otra
- PUSH consiste en que la etiqueta externa es sustituida por otra (SWAP) y además se añaden una o más etiquetas a la pila.
- POP consiste en quitar la etiqueta externa. El paquete se reenvía con la pila de etiquetas restante o como un paquete sin etiquetar.
- UNTAGGED/NO LABEL: La pila completa es eliminada y el paquete se reenvía sin etiquetar.

2.1.3.1 Búsqueda por IP frente a búsqueda por etiqueta

Cuando un router recibe un paquete IP la búsqueda se hace en función de la dirección IP. En la IOS de Cisco significa que el paquete se busca en la tabla CEF. Cuando un router recibe un paquete etiquetado la búsqueda se realiza en la LFIB del router. El router sabe distinguir si el paquete recibido está etiquetado o es un paquete IP comprobando el campo de protocolo de la cabecera de nivel 2. Si un paquete es reenviado por CEF o por LFIB el paquete puede salir del router etiquetado o sin etiquetar.

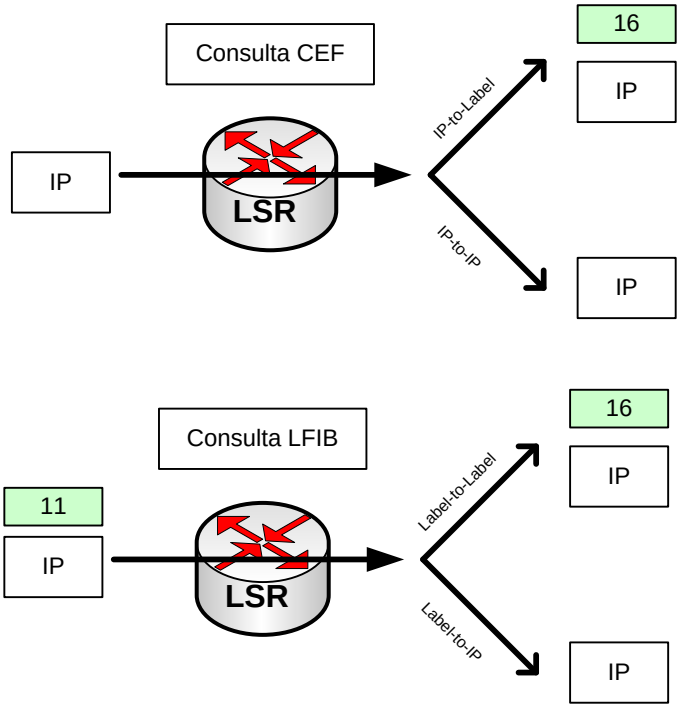

Ilustración 10- Consulta en tabla CEF o LFIB

Si un Ingress LSR recibe un paquete IP y lo reenvía etiquetado es un caso de IP-to-label. Si el LSR recibe el paquete etiquetado puede quitar todas las etiquetas y reenviarlo como un paquete IP o puede reenviarlo como un paquete etiquetado. El primer caso es label-to-IP y el segundo label-to-label.

```
MPLS-ASL#sh ip cef 192.168.1.12 detail
192.168.1.12/30, epoch 30
local label info: global/1363
nexthop 192.168.1.90 GE-WAN5/3 label 39
```

Los paquetes IP que llegan al LSR con destino 192.168.1.12/30 se envían por la interfaz GE-WAN5/3 después de haberles impuesto la etiqueta 39. El siguiente salto de ese paquete es la dirección 192.168.1.90. El reenvío IP-to-label se hace en el LSR que impone la etiqueta. En Cisco la conmutación CEF es el único modo de conmutación IP que puedes usar para etiquetar paquetes. Otros modos de conmutación IP como fast-switching no se puede usar porque la caché de fast-switching no guarda información de etiquetas. Como la conmutación CEF es la única que se soporta junto con MPLS se debe habilitar CEF cuando se habilita MPLS en el router.

El siguiente ejemplo muestra un extracto de la LFIB:

```
MPLS-ASL#sh mpls forwarding-table
Local   Outgoing        Prefix              Bytes Label     Outgoing       Next Hop
Label    Label or VC   or Tunnel Id          Switched      interface
16      No Label       l2ckt(13999)         88302         none           point2point
36      Pop Label      IPv4 VRF[V]          9529658011    aggregate/vpn10002:vrf1
```

```
37      Pop Label      IPv4 VRF[V]           597555473      aggregate/vpn10005:vrf1
40      No Label       10.2.60.0/24[V]       0              GE5/4.114      192.168.46.106
79      112            10.132.1.7/32         3465791970     GE5/1          10.132.1.17
81      explicit-n     10.132.1.5/32         64708175430    GE5/2          10.132.1.37
82      78             10.132.1.4/32         50196658414    GE6/1          10.132.1.26
83       explicit-n    10.132.1.3/32         1170536761272  GE6/1          10.132.1.26
85       explicit-n    10.132.1.52/30        0              GE6/1          10.132.1.26
90      84             10.132.2.128/26       150709263      GE5/1          10.132.1.17
92      91             10.132.1.248/30       217026         GE6/1          10.132.1.26
130     No Label       10.109.222.0/24[V]    3873412        Gi11/8         10.200.6.26
```

La etiqueta local es la etiqueta que este LSR asigna y distribuye a los demás, por tanto, este LSR espera que los paquetes etiquetados le lleguen con esta etiqueta como etiqueta externa de la pila. Si este LSR recibe un paquete con etiqueta externa 90 cambiará esta etiqueta por la 84 y lo reenviará por la interfaz GE5/1 (Operación SWAP). Este es un ejemplo de label-to-label.

Si el LSR recibe un paquete con la etiqueta externa 16 elimina todas las etiquetas y reenvía el paquete como un paquete IP porque el valor del campo outgoing label es *no label*. Este es un ejemplo de label-to-IP.

Si el nodo recibe un paquete con etiqueta 36 quita la etiqueta externa (operación POP) y reenvía el paquete como un paquete etiquetado o IP.

En el siguiente ejemplo vemos un caso de SWAP.

```
MPLS-ASL#sh mpls  forwarding-table 10.132.2.128 detail
Local  Outgoing      Prefix           Bytes Label   Outgoing   Next Hop
Label  Label or VC   or Tunnel Id     Switched      interface
90     84            10.132.2.128/26  151148265     GE5/1      10.132.1.17
       MAC/Encaps=14/18, MRU=1900, Label Stack{84}
       0015C711A6000015C70F65C08847 00054000
       No output feature configured
```

Lo explicado anteriormente especifica como un paquete etiquetado se reenvía al siguiente salto especificado después de una operación de etiquetas. La tabla de adyacencia CEF en cambio determina la encapsulación de la capa de enlace saliente. La tabla de adyacencia proporciona la información de capa 2 necesaria para reenviar el paquete al siguiente salto.

Ejemplo de tabla de adyacencia es:

```
MPLS-ASL#sh adjacency detail
Protocol Interface                 Address
IP       GigabitEthernet4/1        10.201.4.6(23)
147480694 packets, 17885488731 bytes
epoch 0
sourced in sev-epoch 2649
Encap length 14
001AA1EC1A4E0015C70F65C00800
```

```
L2 destination address byte offset 0
L2 destination address byte length 6
Link-type after encap: ip
ARP
IP       POS1/0/0                  point2point(9)
6633 packets, 661966 bytes
epoch 0
sourced in sev-epoch 8589
Encap length 4
0F000800
P2P-ADJ
P2P-ADJ
IP       Serial2/0/0.1/1/6/1:0.100 point2point(95)
15427341 packets, 2653464845 bytes
epoch 0
sourced in sev-epoch 8662
Encap length 4
184103CC
P2P-ADJ
```

2.1.3.2 Balanceo de carga de paquetes etiquetados

Si existen caminos múltiples de igual coste para un mismo prefijo IP, un router Cisco puede hacer balanceo de carga para los paquetes etiquetados como se muestra en el ejemplo a continuación:

```
MPLS-ASL#show mpls forwarding-table
Local        Outgoing      Prefix       Bytes Label   Outgoing    Next Hop
Label        Label or VC   or Tunnel Id  Switched    interface
79           112          10.132.1.7/32 3466978520    GE5/1       10.132.1.17
             75           10.132.1.7/32 755161472     GE6/1       10.132.1.26
```

La etiqueta local 79 tiene dos interfaces de salida, si los paquetes etiquetados son balanceados pueden tener la misma o diferente etiqueta saliente. Las etiquetas salientes son las mismas si los dos enlaces están entre un par de routers y los dos enlaces pertenecen al mismo espacio de etiquetas por plataforma.

Si existen múltiples LSR como siguiente salto, la etiqueta saliente para cada camino normalmente es diferente ya que los siguientes saltos asignan etiquetas independientemente como se puede comprobar en el ejemplo.

Si un prefijo es alcanzable a través de una mezcla de caminos etiquetados y sin etiquetar (IP) un router Cisco no considera los caminos sin etiquetar a la hora de hacer el balanceo de paquetes etiquetados. Esto es así porque en algunos casos el tráfico que atraviesa caminos sin etiquetar puede no alcanzar su destino. En el caso de que MPLS corra sobre una red IP los paquetes alcanzan su destino incluso si quedan sin etiquetar. Los paquetes se quedan sin etiquetar en los enlaces en los que MPLS no está habilitado y vuelven a ser etiquetados en el siguiente enlace en el que MPLS esté habilitado. En el punto en el que el

paquete queda sin etiqueta se produce una búsqueda por IP. Como en la red corre IP deberá ser capaz de entregar el paquete a su destino sin etiquetas, sin embargo, en algunos escenarios como las VPN MPLS o AToM un paquete que queda sin etiquetar en la red MPLS no llega a su destino final.

En el ejemplo de VPNs MPLS la carga de MPLS es IP pero los routers P normalmente no tienen tabla de rutas de las VPNs por lo que no pueden enrutar el paquete a su destino. En el caso de AToM la carga de MPLS es una trama de capa 2 por lo que si el paquete pierde la pila de etiquetas en un router P, el router P no tienen las tablas de encaminamiento de capa 2 para reenviar la trama. Por esta razón, en una red MPLS los paquetes etiquetados no son balanceados sobre caminos IP (sin etiquetar) y etiquetados. En general la inteligencia para reenviar el paquete transportado por MPLS está en el LSR del extremo (router PE) únicamente. Por lo tanto, un router P en la mayoría de los casos no puede reenviar un paquete sin etiquetar.

2.1.3.3 Etiqueta desconocida

En una situación normal, un nodo debe recibir solo paquetes etiquetados con una etiqueta externa conocida por él mismo ya que el LSR ha tenido que anunciar previamente esa etiqueta. Sin embargo es posible que algo vaya mal en la red MPLS y el nodo empiece a recibir paquetes etiquetados con una etiqueta externa que el LSR no encuentre en su LFIB. El LSR podría hacer teóricamente dos cosas:

1. Eliminar todas las etiquetas e intentar reenviar el paquete
2. Descartar el paquete

Los nodos Cisco descartan el paquete. Esto es lo correcto ya que este LSR no asignó la etiqueta externa y no sabe qué tipo de paquete hay detrás de la pila de etiquetas: no sabe si es IPv4, IPv6, una trama de nivel 2 u otra cosa. El LSR puede intentar averiguar que es realizando una inspección del paquete transportado pero entonces existirá el mismo problema descrito anteriormente: El LSR en el que el paquete o trama queda sin etiquetar probablemente no será capaz de encontrar el destino del paquete o de la trama. Incluso si el LSR intenta reenviar el paquete no está garantizado que este no sea descartado en otro nodo más adelante. La única opción correcta es descartar un paquete que llegue con una etiqueta desconocida.

2.1.3.4 Etiquetas reservadas

El rango de etiquetas de 0 a 15 están reservadas y un LSR no los puede usar para un caso normal de reenvío de paquetes. Los nodos asignan funciones específicas a cada una de estas etiquetas.

- Etiqueta 0: Explicit-null
- Etiqueta 1: Router alert
- Etiqueta 3: Implicit-null
- Etiqueta 14: OAM alert.

El resto de etiquetas no está definido aún.

2.1.3.4.1 Implicit-null

La etiqueta implicit-null es aquella que tiene valor 3. Un Egress LSR asigna la etiqueta implicit-null a un FEC si no quiere asignarle ninguna etiqueta a ese FEC por lo que solicita al LSR anterior que realice una operación POP. En el caso de una red IP sobre MPLS como una red en la que el protocolo LDP distribuye las etiquetas entre los LSR, el Egress LSR de Cisco asigna la etiqueta implicit-null a sus prefijos conectados y sumarizados. La ventaja de esto es que el si el Egress LSR asignara una etiqueta a estos FECs recibiría paquetes con una etiqueta externa por lo que tendría que realizar dos búsquedas: Primero, buscar una etiqueta en la LFIB para averiguar que la etiqueta tiene que ser eliminada y después realizar una búsqueda por IP. La primera de estas búsquedas sería innecesaria.

La solución para esta doble búsqueda es que el Egress LSR indique al penúltimo LSR (el anterior a él en el LSP) que debe enviar los paquetes sin etiqueta. El Egress LSR indica al penúltimo LSR que utilice implicit-null enviándole la etiqueta especial con valor 3 en vez de una etiqueta normal. El resultado es que el Egress LSR recibe un paquete IP y solo necesita hacer una búsqueda IP para reenviar el paquete. Esto mejora el rendimiento del Egress LSR. El uso de implicit-null al final de un LSP se llama *penultimate-hop-popping (PHP)***.** La entrada de la LFIB para ese LSP en el router PHP muestra Pop-label como etiqueta de salida.

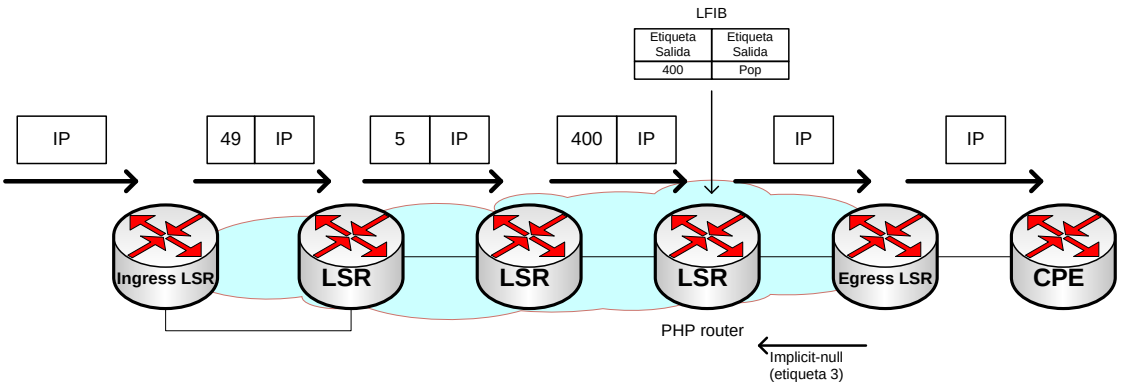

Ilustración 11- Uso de la etiqueta implicit-null

El uso de implicit-null no se limita al ejemplo anterior. En el caso de que el paquete tenga varias etiquetas en la pila entonces la etiqueta implicit-null usada en el Egress LSR indicará al PHP que quite una etiqueta y envíe el paquete etiquetado con el resto de etiquetas de la pila al Egress LSR. Entonces el Egress LSR no tendrá que realizar dos búsquedas de etiquetas. El uso de implicit-null no implica que todas las etiquetas de la pila deban ser eliminadas, sino solamente una de ellas. En cualquier caso el uso de la etiqueta implicit-null tenga que realizar dos búsquedas. Aunque la etiqueta con valor 3 indica el uso de implicit-null, nunca aparecerá como una etiqueta en la pila de etiquetas de un paquete MPLS. De ahí su nombre implicit-null

2.1.3.4.2 Explicit-null

El uso de implicit-null hace más eficiente el reenvío de paquetes, sin embargo, tiene un inconveniente: el paquete se envía con una etiqueta menos que cuando se recibió en el penúltimo LSR o sin etiqueta si se recibió con una sola. Además del valor de la etiqueta, esta también contiene los bits EXP. Cuando una etiqueta es eliminada, los bits EXP también se eliminan. Estos bits EXP se usa exclusivamente para calidad de servicio (QoS) por lo que esta información del paquete se pierde cuando la etiqueta externa se elimina. En algunos caso es necesario mantener esta información de QoS y entregarla al Egress LSR. En este caso implicit-null no se puede usar.

La etiqueta explicit-null es la solución a este problema, ya que el Egress LSR indica la etiqueta explicit-null con valor 0 al router PHP. El Egress LSR en este caso recibe paquetes etiquetados con una etiqueta externa con valor 0. El LSR no puede reenviar el paquete buscando el valor 0 en la LFIB porque puede estar asignado a múltiples FECs. El LSR elimina la etiqueta explicit-null y a continuación realiza otra búsqueda pero la ventaja es que ese nodo puede deducir la información de QoS del paquete recibido buscando los bits EXP de la etiqueta explicit-null

Es posible copiar el valor de los bits EXP al campo precedence o a los bits DiffServ cuando se realiza PHP para preservar la información de QoS. También si la pila de etiquetas tiene múltiples etiquetas y se elimina la externa se puede copiar los bits EXP al campo EXP de la nueva etiqueta externa.

2.1.3.4.3 Router alert label

Esta etiqueta tiene valor 1 y puede estar presente en cualquier parte de la pila de etiquetas excepto en la interna. Cuando la etiqueta de router alert es la externa indica al LSR que el paquete debe ser examinado más profundamente. En ese caso, el paquete no se reenvía mediante hardware sino que es examinado por un proceso software. Cuando el paquete se reenvía la etiqueta 1 se elimina, entonces se busca la siguiente etiqueta de la pila en la LFIB para decidir como reenviar el paquete. A continuación se realiza la acción necesaria sobre la etiqueta, POP, SWAP o PUSH y la etiqueta de router alert se pone de nuevo como etiqueta externa de la pila y se reenvía el paquete.

2.1.3.4.4 OAM Alert label

Esta etiqueta con valor 14 es la etiqueta Operation and Maintenance (OAM) alert label. Se describe por la recomendación Y.1711 de la ITU-T y la RFC 3429. Esta etiqueta se emplea básicamente para detección de fallos, localización y monitorización de rendimiento y diferencia los paquetes OAM de los paquetes de datos de usuario. Los routers Cisco no utilizan la etiqueta 14 aunque realiza MPLS OAM.

2.1.3.5 Etiquetas no reservadas

Exceptuando las etiquetas reservadas (rango de 0 a 15) es posible usar el resto de valores de etiquetas para el reenvío normal de paquetes. Las etiquetas de las 16 a 1.048.575 se usan para este reenvío de paquetes. En los routers Cisco el rango es de 16 a 100.000 que es

más que suficiente para etiquetar todos los prefijos IGP aunque si se quiere etiquetar los prefijos BGP este número puede resultar insuficiente. Se puede cambiar el rango de etiquetas con el comando *mpls label range [min max]*.

2.1.3.6 Comportamiento del TTL en paquetes etiquetados

El tiempo de vida (TTL) es un mecanismo muy popular gracias a IP. En la cabecera IP es un campo de 8 bits que indican el tiempo de vida que le queda a un paquete antes de ser descartado. Cuando un paquete IP es enviado, su TTL es normalmente 255 y se va decrementando en una unidad en cada salto. Si el TTL llega a 0 el paquete es descartado. En cualquier caso, el router que descarta el paquete IP envía un mensaje ICMP (tipo 11 y código 0) al equipo originador del paquete.

Aplicando este concepto a la red MPLS, en la etiqueta también hay un campo TTL. Hay un mecanismo que traslada este TTL IP al TTL MPLS al entrar en la red y del TTL MPLS al TTL IP al salir de la red. Esto asegura que no habrá ningún paquete que esté eternamente dando vueltas por la red entrando y saliendo de la red MPLS si hubiera un bucle.

2.1.3.7 Comportamiento del TTL en los casos IP-to-label

En MPLS, el uso del TTL es igual que en el caso de IP. Cuando un paquete IP entra en la red MPLS, a través de un Ingress LSR, el valor del TTL IP, después de ser decrementado en 1, se copia al campo TTL MPLS de la etiquetas agregadas (pueden ser una o varias). En el Egress LSR, se quita la etiqueta y es de nuevo la cabecera IP la que encamina pero con el valor del TTL copiado de la etiqueta MPLS al TTL IP pero decrementado en una unidad. De todas formas, en routers Cisco se ha implementado un mecanismo que evita posibles bucles de routing mediante el cual no se copia el TTL MPLS al TTL IP si es TTL MPLS es mayor que el TTL IP del paquete. Si el TTL MPLS se copiara a la cabecera IP el TTL IP que es más pequeño sería sobrescrito por un valor más alto. Si el paquete IP fuera reinyectado en la red MPLS debido a un bucle de routing el paquete podría no caducar nunca porque el TTL nunca llegaría a 0.

El siguiente ejemplo muestra un ejemplo del comportamiento por defecto de copia y propagación de TTL entre la cabecera IP y la etiqueta MPLS y viceversa.

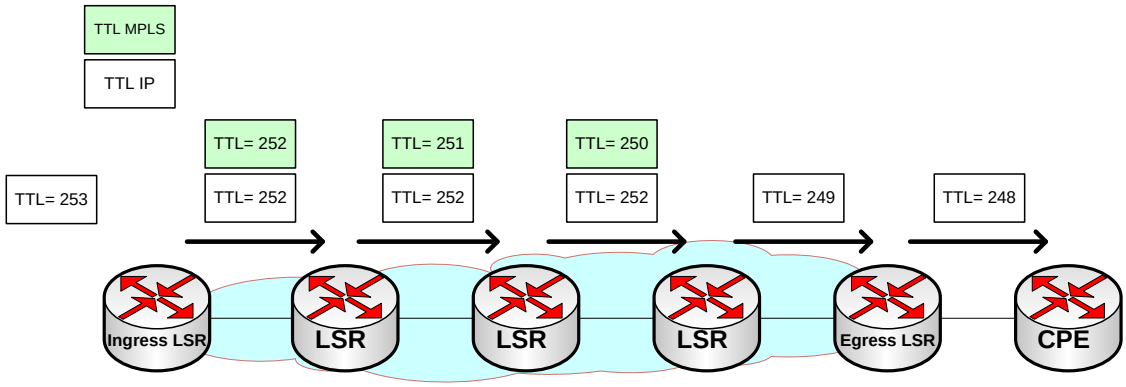

Ilustración 12- Comportamiento del TTL entre IP y MPLS.

2.1.3.8 Comportamiento del TTL en el caso Label-to-label

Si la operación realizada sobre un paquete es SWAP, el TTL de la etiqueta de entrada se copia decrementado en la etiqueta intercambiada. Si la operación realizada es PUSH de una o más etiquetas, el TTL MPLS recibido se copia decrementado en la etiqueta intercambiada y en todas las añadidas. Si la operación es POP, el TTL MPLS de entrada es copiado decrementado a la nueva etiqueta externa a menos que el valor sea mucho mayor que el TTL de la nueva etiqueta externa, en cuyo caso no se copia. En la siguiente ilustración lo vemos para las 3 operaciones básicas.

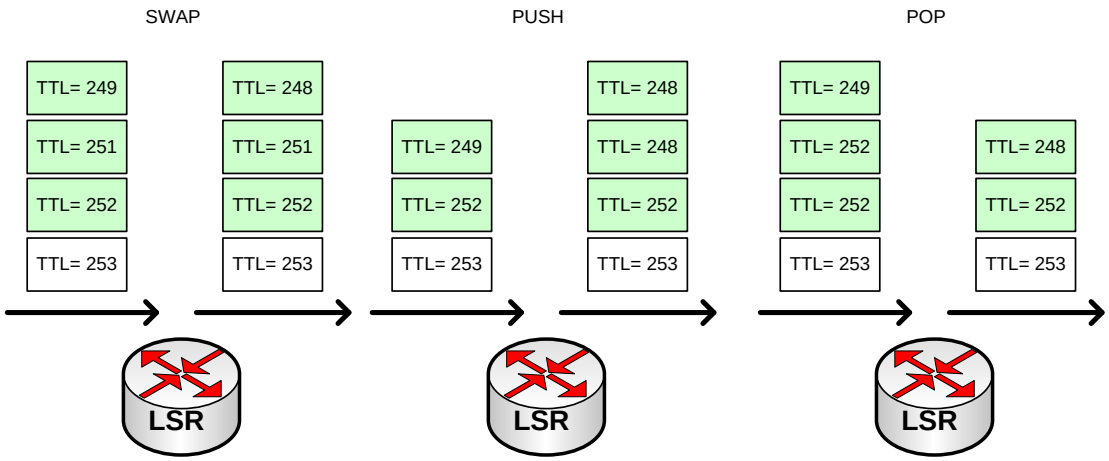

Ilustración 13- Comportamiento del TTL en la red MPLS

Un intermédiate LSR nunca cambia el campo TTL IP, solo mira o cambia el campo TTL de la etiqueta externa de la pila de etiquetas.

2.1.3.9 Expiración del TTL

Cuando un paquete etiquetado es recibido con el TTL a 1, el LSR lo descarta y envía un mensaje ICMP de tiempo de vida expirado (tipo 11, código 0) a la dirección IP origen del paquete (emisor del paquete); es el mismo comportamiento que se podría dar en un router IP cuando un paquete expira. La peculiaridad es que el paquete ICMP no es enviado al origen ya que el LSR podría no tener un camino IP hacia el origen del paquete, dado esto, el mensaje ICMP es enviado a lo largo del LSP original que el paquete iba siguiendo.

En las siguientes dos ilustraciones se ve la diferencia entre la respuesta ante un paquete expirado en una red IP y un paquete expirado en una red MPLS.

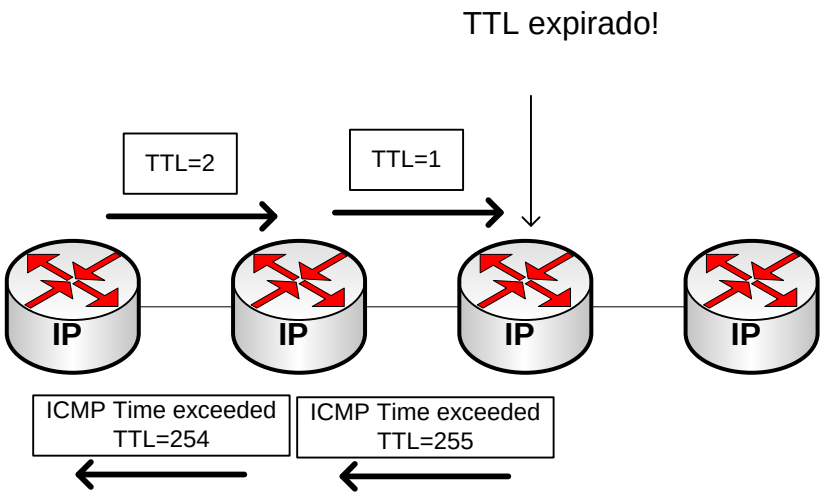

Ilustración 14- Comportamiento del TTL expirado en una red IP

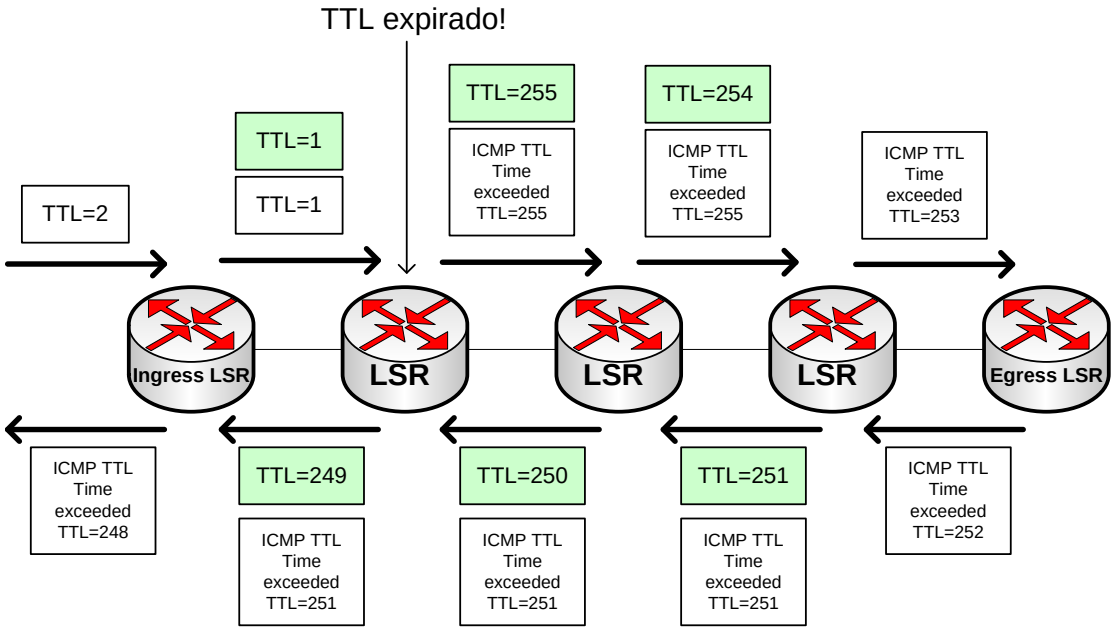

Ilustración 15- Comportamiento del TTL expirado en una red IP-MPLS

La razón por la que enviar el mensaje de ICMP a través del LSP que iba a seguir el paquete expirado es que en muchos casos, el nodo que genera el paquete ICMP no sabe cómo llegar hacia el originador del paquete. De hecho, el intermédiate LSR podría estar cerca del generador del paquete y no saberlo. Un caso puede ser el de la VPN MPLS. En este escenario, los routers P son incapaces de de encaminar el mensaje CIMP hacia el originador del paquete ya que los routers P no tienen tablas de rutas en las VPNs. Entonces el P envía el mensaje ICMP siguiendo el LSP original y es el Egress LSR (que si maneja tablas de rutas) el encargado de reencaminar el mensaje ICMP hacia el origen.

También es especialmente importante que el router P donde expira el TTL conozca el protocolo de nivel 3 que transporta, examina si es IPv4 o IPv6, si es así puede generar el paquete ICMP y enviarlo a través del LSP, si no es ni IPv4 ni IPv6 descarta el paquete. En el caso de que transporte AToM, como el protocolo transportado es de nivel 2 es incapaz de detectar un campo TTL y va a descartar el paquete siempre.

2.1.3.10 MPLS MTU

La unidad máxima de transferencia (MTU) es un parámetro clásico del mundo IP. Indica el tamaño máximo de un paquete IP que puede ser enviado por un enlace sin ser fragmentado. Los enlaces en redes MPLS también tienen una MTU específica, pero para paquetes etiquetados. Cogemos el caso de una red MPLS/IP. Los paquetes IP tienen una o más etiquetas. Esto implica que los paquetes etiquetados son ligeramente más grandes que los paquetes IP debido a que por cada etiqueta que se añada el paquete, son 4 bytes más. En resumen, si *n* es el número de etiquetas, *n**4 bytes son añadidos al tamaño del paquete cuando se etiqueta.

Aquí explicaremos que el parámetro MPLS MTU se refiere a un paquete etiquetado.

En un interfaz Ethernet, la MTU por defecto es 1500, por lo tanto habría que fragmentar cualquier paquete Ethernet que llegase con ese tamaño al añadirle aunque fuese una sola etiqueta. Para ello, en routers Cisco existe el concepto de MPLS MTU que te permite especificar como de grande será un paquete en un enlace. Por ejemplo si se sabe que el máximo de etiquetas por enlace es de dos, se puede establecer la MPLS MTU en 1508 (1500 eth + 4*2 etiqueta). Así todos los paquetes etiquetados de 1508 bytes (contemplando etiquetas) pueden ser enviados por el enlace sin fragmentarlos. Por defecto, en un enlace la MTU MPLS es igual a la MTU.

2.1.3.11 MPLS MRU

La unidad máxima de recepción (MRU – Maximum receive unit) es un parámetro que usan los routers Cisco. Indica al LSR como de grande puede ser un paquete recibido de un cierto FEC para que pueda ser reenviado. Este valor es un valor que se asigna por FEC (por prefijo) y no por interfaz. Esto es lógico si pensamos en el siguiente ejemplo.

Un router tiene todos sus interfaces configurados con una MTU de 1500 bytes. Esto significa que el paquete más grande que puede ser recibido y transmitido por todos sus interfaces sería de mil 1500 bytes. Ahora imaginemos que los paquetes pueden ser etiquetados con un máximo de 2 etiquetas (2 etiquetas es típicamente lo que se añaden en redes VPN MPLS y AToM). Visto esto establecemos la MPLS MTU en 1508 bytes para permitir la transmisión de estos paquetes sin fragmentar. Pero, sin embargo, la operación que vamos a realizar al paquete de entrada para un determinado FEC fuese POP, el paquete recibido podría haber sido 4 bytes más grande (de 1512 bytes) ya que le vamos a quitar una etiqueta con la operación POP y lo vamos a reenviar con un tamaño de 1508 bytes. En el caso contrario si la operación que vamos a realizar es PUSH, el paquete máximo de recepción no puede pasar los 1504 bytes ya que le vamos a añadir una cabecera que lo dejará en el máximo, 1508 bytes.

Con esto hemos comprobado que la operación que se realice sobre el paquete es determinante en el tamaño de la MRU y como hemos comprobado, ya que la operación depende del prefijo (FEC) recibido, la MRU cambia acorde al FEC.

El siguiente ejemplo muestra el ejemplo en nuestra red con una MTU configurada de 1900 bytes. Para el prefijo 192.168.255.6 la MRU es el máximo ya que no va a añadir ni quitar cabeceras.

```
MPLS-ASL#sh mpls forwarding-table 192.168.255.6 detail
Local  Outgoing      Prefix           Bytes Label   Outgoing   Next Hop
Label  Label or VC   or Tunnel Id     Switched      interface
100    58            192.168.255.6/32  0             GE5/3      192.168.1.90
       MAC/Encaps=14/18, MRU=1900, Label Stack{58}
       00169D4E5E800015C70F65C08847 0003ª000
       No output feature configured
```

2.1.3.12 Fragmentación de paquetes MPLS

Si un nodo recibe un paquete etiquetado que es demasiado grande para ser reenviado por un enlace, fragmenta el paquete de manera similar que se hace en IP. Si un paquete etiquetado es recibido y el nodo ve que la MTU de salida es inferior al tamaño del paquete lo que hace es quitar la pila de etiquetas, fragmentar el paquete, poner de nuevo la pila de etiquetas (después de realizar la operación PUSH, POP o SWAP que aplique) y reenviar el paquete por el enlace de salida. Solo si la cabecera IP tiene el bit don't fragment (DF) activado descarta el paquete, pero en ese caso, devuelve un mensaje ICMP de error "Fragmentantion needed and do not fragment bit set" (ICMP tipo 3, código 4) al originador del mensaje. Al igual que en el caso de TTL expirado, el paquete sigue el LSP correspondiente hasta que el Egress LSR lo devuelve al originador del mensaje (Recordemos que solo el Egress LSR sabe encaminar el paquete hasta el origen).

Como la fragmentación es un problema muy común en la red, los equipos IP modernos utilizan un método que se conoce como Path MTU Discovery que consiste en que los paquetes se mandan con el bit DF activado, cuando se recibe un mensaje ICMP de paquete demasiado grande, se baja el tamaño de los paquetes hasta que se dejan de recibir estos mensajes.

2.1.4 PROTOCOLO DE DISTRIBUCIÓN DE ETIQUETAS (LDP)

2.1.4.1 Visión general de LDP

Para encaminar los paquetes a lo largo de un LSP a través de una red MPLS, todos los nodos deben de ejecutar un protocolo de distribución de etiquetas e intercambiarse asociaciones (etiqueta - prefijo de red). Cuando todos los nodos tienen la etiqueta para un particular FEC, los paquetes pueden ser enviados por el LSP, lo que significa que cada LSR sabe encaminarlo en base a la conmutación de etiquetas. Las operaciones sobre las etiquetas las realizan los nodos de acuerdo con la LFIB. La LFIB, que es la tabla de reenvío de paquetes etiquetados y es poblada con las asociaciones que se encuentran en la LIB. La LIB se completa gracias a las asociaciones recibidas por LDP, RSVP, MP-BGP o asignaciones estáticas. Ya que

RSVP solo se usa para hacer ingeniería de tráfico y MP-BGP distribuye las etiquetas solo para las rutas de BGP, LDP se encarga de distribuir las etiquetas para las rutas interiores. Por lo tanto, todos los nodos directamente conectados deben establecer sesiones LDP entre ellos. Los pares LDP intercambian los mensajes de asociaciones de etiquetas a través de su sesión LDP. Una asociación de etiquetas está ligada a un FEC. El FEC es un conjunto de paquetes que están ligados a un cierto LSP que se envían sobre ese LSP a través de la red MPLS.

Aquí solo hablaremos de asociaciones de etiquetas para prefijos IP del IGP.

LDP tiene 4 funciones principales:

1.- Descubrimiento de los LSRs que ejecutan el protocolo LDP
2.- Establecimiento y mantenimiento de las sesiones
3.- Anuncio de asociaciones de etiquetas
4.- Notificación

Cuando dos nodos están ejecutando LDP y comparten uno o más de un enlace entre ellos deben descubrirse mutuamente mediante mensajes *hello.* El segundo paso es establecer una sesión entre ellos mediante una conexión TCP. A través de ella, LDP anuncia los mensajes de asociación de etiquetas entre dos pares LDP. Estos mensajes de asociación de etiquetas se usan para anunciar, cambiar o quitar asociaciones de etiquetas. LDP proporciona los medios para notificar al vecino LDP mensajes de aviso y de error enviando mensajes de notificación.

2.1.4.2 Descubrimiento de los nodos con LDP habilitado

Los LSR que ejecuten LDP envían mensajes *hello* por todos los enlaces que tengan LDP activado. Estos *hello* son mensajes UDP que son enviados sobre los enlaces “todos los routers de esta red”, a la dirección de multicast 224.0.0.2. El puerto UDP usado es el 646. El nodo que recibe este *hello* en un cierto interfaz es entonces consciente de de la presencia de un router LDP por ese interfaz. Estos *hello* tienen un tiempo de espera, si no se recibe ningún *hello* de ese LSR antes de que expire el tiempo de espera, el nodo borra ese otro nodo de la lista de vecinos LDP descubiertos. Si los mensajes de *hello* son enviados y recibidos en un interfaz se establece una adyacencia LDP sobre ese enlace entre los dos LSR que están ejecutando LDP.

Podemos ver las adyacencias de LDP de la siguiente manera:

```
MPLS-ASL#show mpls interfaces
Interface              IP            Tunnel   BGP Static Operational
POS1/0/0               Yes (ldp)     Yes      No  No     Yes
GE-WAN5/1              Yes (ldp)     Yes      No  No     Yes
GE-WAN5/2              Yes (ldp)     Yes      No  No     Yes
GE-WAN5/3              Yes (ldp)     No       No  No     Yes
GE-WAN6/1              Yes (ldp)     Yes      No  No     Yes
```

Los valores por defecto para el holdtime y el interval entre *hello* es de 15 y 5 segundos respectivamente. Si dos vecinos LDP tienen diferentes valores del *holdtime* configurados, el más pequeño de los dos será el valor real de *holdtime* para esta fuente. Si el tiempo de *holdtime* expira para un enlace, ese enlace es borrado de la lista de fuentes LDP descubiertas.

Si el último de los enlaces de las fuentes de LDP descubiertas para un vecino LDP se elimina, la sesión LDP se pierde. Hay que tener mucho cuidado con los valores que se configuran para el *holdtime*. Si se establece un valor muy bajo y se pierde uno de estos paquetes debido a, por ejemplo, alta congestión, se caería la adyacencia. En el caso contrario, si este valor es muy alto y se cae el enlace, la red tardaría mucho en darse cuenta de este evento y serían muchos los paquetes etiquetados perdidos.

Los nodos que tienen LDP activo tienen un identificador LDP, o LDP ID. Este LDP ID es un campo de 6 bytes que consiste en 4 bytes identificando el nodo y 2 bytes identificando el espacio de etiquetas que usa el LSR. Como vimos en el apartado de arquitectura, si los dos últimos bytes son 0, el espacio es por plataforma y en cualquier otro caso es por interfaz. Si se da el caso de espacio por interfaz se utilizan múltiples LDP IDs, donde los primeros 4 bytes tienen el mismo valor pero los dos últimos indican el espacio de etiquetas. Los primeros 4 bytes del campo son una dirección IP tomada del router mediante una operación interna. Si existe interfaces de loopback, la dirección más alta de uno de ellos es usada como LDP ID. Si no hay interfaces de loopback se selecciona la más alta de las direcciones IP de asignadas a los interfaces. También se puede forzar este parámetro de manera manual.

En el siguiente ejemplo vemos el LDP ID de un nodo MPLS:

```
MPLS-ASL#sh mpls ldp discovery
Local LDP Identifier:
10.132.1.2:0  <- IP loopback más alta : espacio de etiquetas.
Discovery Sources:
Interfaces:
POS1/0/0 (ldp): xmit/recv
LDP Id: 10.132.1.6:0
GE-WAN5/1 (ldp): xmit/recv
LDP Id: 10.132.1.1:0
GE-WAN5/2 (ldp): xmit/recv
LDP Id: 10.132.1.5:0
GE-WAN5/3 (ldp): xmit/recv
LDP Id: 192.168.255.3:0
GE-WAN6/1 (ldp): xmit/recv
LDP Id: 10.132.1.3:0
Targeted Hellos:
10.132.1.2 -> 10.132.1.1 (ldp): active/passive, xmit/recv
LDP Id: 10.132.1.1:0
10.132.1.2 -> 10.132.1.4 (ldp): active/passive, xmit/recv
LDP Id: 10.132.1.4:0
10.132.1.2 -> 10.132.1.3 (ldp): active/passive, xmit/recv
LDP Id: 10.132.1.3:0
10.132.1.2 -> 10.132.1.6 (ldp): active/passive, xmit/recv
LDP Id: 10.132.1.6:0
```

En sombreado vemos el LDP ID que se compone [IP de loopback más alta: espacio de etiquetas]. En los LSRs de Cisco, la dirección IP que sea la LDP ID tiene que estar en la tabla de rutas de los nodos vecinos para que se pueda establecer la adyacencia.

2.1.4.3 Establecimiento y mantenimiento de sesiones LDP

Si dos LSRs se han descubierto intercambiando LDP *Hellos*, intentan establecer una sesión LDP entre ellos. Un LSR intenta abrir una conexión TCP (puerto 646) al otro nodo. Si la conexión TCP se establece, ambos LSR negocian parámetros para la sesión LDP intercambiando mensajes de inicialización, estos parámetros incluyen cosas como:

- Valor de temporizadores
- Método de distribución de etiquetas
- Rangos VPI/VCI en caso de enlaces ATM
- Rangos DLCI en caso de enlaces Frame Relay

Si ambos nodos están de acuerdo en los parámetros de sesión mantendrán la conexión TCP entre ellos.

Una vez se ha establecido la sesión, es mantenida por, o la recepción de paquetes LDP o mensajes periódicos de keepalive. Cada vez que el vecino LDP recibe un paquete LDP o un keepalive, el temporizador se resetea para ese vecino. El temporizador de keepalive o Holdtime para sesión LDP se puede configurar también y oscila entre 15 y 2.147.483 segundos, por defecto está a 180 sg.

El siguiente ejemplo nos muestra estos parámetros en una adyacencia entre nodos:

```
MPLS-ASL#show mpls ldp neighbor 10.132.1.3 detail
Peer LDP Ident: 10.132.1.3:0; Local LDP Ident 10.132.1.2:0
TCP connection: 10.132.1.3.23580 - 10.132.1.2.646; MD5 on
Password: not required, neighbor, in use
State: Oper; Msgs sent/rcvd: 7860/7863; Downstream; Last TIB rev sent 197
Up time: 4d16h; UID: 4; Peer Id 3;
LDP discovery sources:
GE-WAN6/1; Src IP addr: 10.132.1.26
holdtime: 15000 ms, hello interval: 5000 ms
Targeted Hello 10.132.1.2 -> 10.132.1.3, active, passive;
holdtime: infinite, hello interval: 10000 ms
Addresses bound to peer LDP Ident:
10.132.1.3      10.132.1.54     192.168.1.86    10.132.1.33
10.132.1.26
Peer holdtime: 180000 ms; KA interval: 60000 ms; Peer state: estab
Clients: Dir Adj Client
```

En sombreado vemos los puertos TCP utilizados para la comunicación (23580 y 646) y los valores de holdtime (180 sg) y de keepalive (6 sg).

Aquí vemos los temporizadores de descubrimiento y de sesión en un nodo MPLS.

```
MPLS-ASL#show mpls ldp parameters
Protocol version: 1
Session hold time: 180 sec; keep alive interval: 60 sec
Discovery hello: holdtime: 15 sec; interval: 5 sec
Discovery targeted hello: holdtime: 90 sec; interval: 10 sec
```

```
Accepting targeted hellos; peer acl: ACL_THELLO_LDP
Downstream on Demand max hop count: 255
LDP for targeted sessions
LDP initial/maximum backoff: 15/120 sec
LDP loop detection: off
```

2.1.4.4 Anuncio de asociaciones de etiquetas.

El anuncio de asociaciones de etiquetas es la tarea principal de LDP. En la arquitectura de MPLS vimos 3 modos de comportamiento de los LSRs. Cada modo tiene 2 posibilidades por lo que se pueden dar 6 combinaciones posibles:

1.- Unsolicited Downstream (UD) frente Downstream-on-demand (DoD).

2.- Liberal Label Retention (LLR) frente a Conservative Label Retention (CLR).

3.- Independent LSP control frente a Ordered LSP control.

No importa qué modo esté funcionando, el objetivo es anunciar las asociaciones de etiquetas. En el modo UD, los vecinos LDP distribuyen las asociaciones de etiquetas sin solicitárselas entre ellos. Sin embargo, una asociación de etiquetas es un conjunto de (LDP ID, etiqueta) por prefijo. Un router LDP recibe múltiples asociaciones de etiquetas por prefijo, exactamente uno por vecino LDP. Esas asociaciones son almacenadas en la LIB del router. Sin embargo, solo un vecino es el siguiente salto en un LDP para un prefijo determinado. Por supuesto si se hace balanceo de carga, puede haber más de un siguiente salto para un LDP.

El nodo siguiente salto para un FEC determinado es encontrado mirando el siguiente salto para un prefijo en la tabla de rutas. Solo la asociación remota con el nodo siguiente salto deberían ser usados para rellenar la LFIB. Esto significa que solo una etiqueta de todas las anunciadas por los LSR vecinos debe ser usada como etiqueta de salida en la LFIB para este prefijo. El problema es que los anuncios de las asociaciones son [LDP ID, etiqueta] sin la dirección IP de los interfaces. Esto quiere decir que para encontrar la etiqueta de salida para un prefijo particular, debes mapear al LDP ID la dirección IP del interfaz en el nodo siguiente salto del LSP. Estas direcciones IP son anunciadas por el vecino LDP con mensajes de direcciones.

En el siguiente ejemplo vemos todas las direcciones IP asociadas a cada LDP ID de los nodos vecinos:

```
MPLS-ASL#show mpls ldp neighbor detail
Peer LDP Ident: 10.132.1.1:0; Local LDP Ident 10.132.1.2:0
TCP connection: 10.132.1.1.646 - 10.132.1.2.53519; MD5 on
Password: not required, neighbor, in use
State: Oper; Msgs sent/rcvd: 9461/9443; Downstream; Last TIB rev sent 209
Up time: 5d15h; UID: 1; Peer Id 0;
LDP discovery sources:
Targeted Hello 10.132.1.2 -> 10.132.1.1, active, passive;
holdtime: infinite, hello interval: 10000 ms
GE-WAN5/1; Src IP addr: 10.132.1.17
holdtime: 15000 ms, hello interval: 5000 ms
```

```
Addresses bound to peer LDP Ident:
10.132.1.1      10.132.1.253    10.132.1.42     10.132.1.21
10.132.1.17
Peer holdtime: 180000 ms; KA interval: 60000 ms; Peer state: estab
Clients: Dir Adj Client
……
```

Cada nodo asigna una etiqueta local a cada prefijo en la tabla de rutas. Esta es la asociación local. Estas asociaciones locales son almacenadas en la LIB de cada router. Cada una de estas etiquetas y prefijos son anunciadas a cada vecino LDP. Estas asociaciones de etiquetas son las asociaciones remotas de los vecinos LDP y almacenadas en su LIB.

Para una mayor claridad, el siguiente ejemplo relaciona la tabla de rutas, con la LIB, la LFIB y los vecinos LDP.

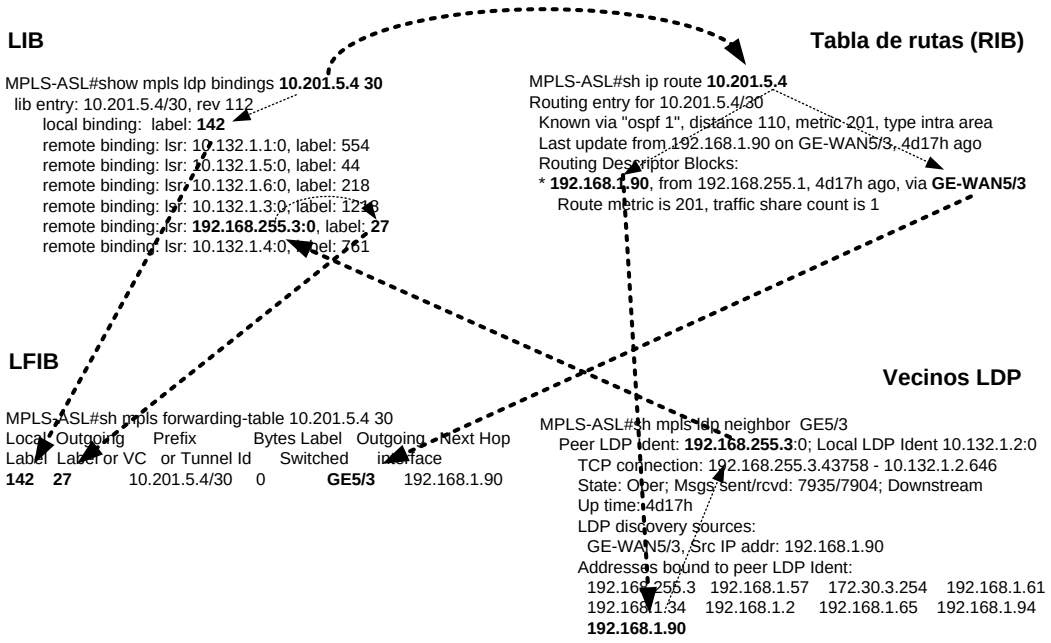

Ilustración 16- Relación entre LIB, LFIB, Vecinos LDP y Tabla de rutas.

Vemos que a la hora de rellenar la LFIB con una nueva asociación, la etiqueta local es escogida directamente de la LIB, en este caso es la 142. Para encontrar la etiqueta de salida el proceso es un poco más complejo. Hay que cotejar el prefijo de la LIB (10.201.5.4) en la tabla de rutas para encontrar el siguiente salto, de ahí obtenemos el interfaz de salida (GE-WAN5/3) pero aún necesitamos la etiqueta de salida. Como en la tabla de rutas tenemos el siguiente salto (192.168.1.90) comprobamos de entre los vecinos LDP, el LDP ID (192.168.255.3) del nodo que contiene esa dirección (192.168.1.90) que es nuestro siguiente salto. Con este dato, de nuevo en la LIB comprobamos la etiqueta correspondiente a este vecino con el LDP ID buscado, en este caso es la 27 que pasamos a completar la nueva asociación en la LFIB con ella.

Es importante reseñar que LDP asigna etiquetas locales a todos los prefijos IGP y anuncia estas asociaciones a todos sus vecinos LDP. En este caso el concepto de Split horizon no existe. Un nodo asigna una etiqueta a un prefijo y se lo manda a su siguiente salto LDP aunque ese LSR posea el prefijo (si es la red de interconexión). En los siguientes ejemplos se ve más claro, el nodo de ASL le pasa la asociación local al nodo de CPII de la dirección que es loopback de CPII.

La dirección 10.132.1.3 es la IP de loopback del nodo de CPII.

```
MPLS-CPII#show ip interface brief | include  Loopback0
Loopback0              10.132.1.3      YES NVRAM  up                    up
```

Comprobamos que CPII le pone la etiqueta local de imp-null ya que él es el último salto.

```
MPLS-CPII#sh mpls ip binding 10.132.1.3 32
10.132.1.3/32
        in label:     imp-null
        out label:    58         lsr: 192.168.255.8:0
        out label:    19         lsr: 10.132.1.7:0
        out label:    194        lsr: 10.132.1.4:0
        out label:    116        lsr: 10.132.1.1:0
        out label:    44         lsr: 10.132.1.6:0
        out label:    112        lsr: 10.132.1.2:0
```

En la LIB comprobamos que recibe las asociaciones remotas de los demás nodos, incluyendo el vecino físico ASL. Y de nuevo, el nodo de ASL establece una asociación local con el prefijo de la dirección de loopback de CPII y se la vuelve a enviar:

```
MPLS-ASL#sh mpls ip binding 10.132.1.3 32
10.132.1.3/32
        in label:     112
        out label:    116        lsr: 10.132.1.1:0
        out label:    57         lsr: 10.132.1.5:0
        out label:    44         lsr: 10.132.1.6:0
        out label:    exp-null   lsr: 10.132.1.3:0     inuse
        out label:    105        lsr: 192.168.255.3:0
        out label:    194        lsr: 10.132.1.4:0
```

Después en la LFIB del nodo de CPII comprobamos que no hay interfaz de salida ya que es propia y que no etiqueta el paquete.

```
MPLS-CPII#sh mpls forwarding-table 10.132.1.3
Local  Outgoing      Prefix           Bytes Label   Outgoing    Next Hop
Label  Label or VC   or Tunnel Id     Switched      interface
None   No Label      10.132.1.3/32    0             aggr-punt
```

2.1.4.5 Baja de las etiquetas

Cuando un vecino LDP anuncia una asociación de etiquetas, los LSRs vecinos que la reciben la mantienen hasta que la sesión LDP cae o hasta que la etiqueta se da de baja. La etiqueta puede darse de baja si la etiqueta local cambia. Esa etiqueta local puede cambiar si, por ejemplo, el interfaz con un cierto prefijo asociado se cae pero otro nodo lo sigue anunciando. Por lo tanto, la asociación local de etiquetas para un prefijo cambia de implicit-null a una etiqueta no reservada. Si llega a ocurrir esto, la etiqueta implicit-null es inmediatamente dada de baja por medio del envió de un mensaje de "baja de etiquetas" a los vecinos LDP. La nueva etiqueta se anuncia en un mensaje de "mapeo de etiquetas".

2.1.4.6 Sesiones Targeted LDP

Normalmente, las sesiones LDP se dan entre nodos directamente conectados. En una red en la cual las rutas IGP necesitan ser etiquetadas es más que suficiente ya que la conmutación de etiquetas de paquetes es salto a salto. Sin embargo, algunas veces es necesario una sesión LDP remota (targeted LDP session) que es una sesión LDP entre nodos no directamente conectados. Ejemplos de redes en las que es necesario targeted sessions son las redes AToM y al hacer Ingeniería de tráfico en VPN MPLS. En el caso de las redes AToM, es necesario sesiones LDP entre todos los PEs de la red. Para el caso de la ingeniería de tráfico en las VPN MPLS, cuando se hace un túnel, es necesario una sesión remota entre el nodo de entrada y el de salida del túnel.

Un vecino remoto LDP puede mejorar el tiempo de convergencia de etiquetas comparado con dos vecinos LDP directamente conectados si hay enlaces cayendo y levantando. Esto es porque si se cae un enlace entre dos vecinos LDP directamente conectados, la sesión LDP cae, pero si tenemos una targeted session y un camino alternativo para conseguir que los paquetes TCP, necesarios para establecer la sesión, lleguen de un LSR a otro, esta se mantendrá.

2.1.4.7 Control de anuncio y recepción de etiquetas

En algunos escenarios o a la hora de hacer ingeniería de tráfico, puede surgir la necesidad de controlar el anuncio de etiquetas, a nivel de nodo, se puede controlar qué etiquetas y a que vecinos LDP (entendemos remotos y directamente conectados) le pasamos esas etiquetas.

De la misma manera, podemos controlar qué etiquetas y de que LSR queremos recibirlas o denegarlas

2.2 VPN MPLS

Las VPN MPLS o redes privadas virtuales MPLS, es la más popular y usada implementación de la tecnología MPLS. Su popularidad ha crecido exponencialmente desde que fueron inventadas. Aunque muchos proveedores de servicios las han implementado como sustitutos de sus antiguas redes ATM o Frame Relay, muchas grandes compañías las están desarrollando dada su escalabilidad y la capacidad de dividir redes en redes más pequeñas, lo

cual es muchas veces útil en empresas de gran tamaño, donde con la misma infraestructura tienes que dar servicio a departamentos individuales.

2.2.1 Definición de una VPN

Una VPN es una red que emula redes privadas virtuales sobre una infraestructura común. Una VPN puede ofrecer comunicación en las capas 2 y 3 del modelo de referencia de OSI.

La característica fundamental de una VPN es que todas las ubicaciones conectadas a esa VPN deben poder utilizar infraestructura común con otras ubicaciones de otra VPN y tener el tráfico completamente separado. Si hablamos de VPNs de nivel IP se amplían mucho las posibilidades como puede ser ofrecer conectividad entre VPNs distintas en incluso conectividad a internet entre ellas. Las VPN MPLS son posibles porque el proveedor de servicios dispone de una red MPLS por debajo, que desvincula el plano de control del plano de tráfico lo cual es imposible con una red IP tradicional.

2.2.2 Modelos de VPN

Las VPNs existen antes del desarrollo de MPLS. Las más populares fueron Frame Relay y ATM ofreciendo servicios de nivel 2. El proveedor tiene una Frame Relay o ATM red de core y ofrece conectividad de nivel 2 a los clientes. Esto era comúnmente conocido como modelo overlay. Existían las redes peer-to-peer en las que se ofrecía conectividad de nivel 3 lo cual supone un manejo bastante complejo de listas de distribución y acceso. No eran muy populares a causa de esto.

2.2.3 Modelo de VPN MPLS

Es importante familiarizarse con la terminología de VPN MPLS, aunque lo venimos haciendo durante todo el capítulo de MPLS dado que es bastante frecuente que se extienda más allá del ámbito de VPN MPLS.

Un router Provider Edge (PE) está directamente conectado al customer edge (CE) que es un router de cliente de nivel 3. Un router Provider (P) es un router que no está conectado a ningún equipo de cliente. En una VPN MPLS los P y PE tienen funcionalidad MPLS lo que significa que tienen capacidad de intercambio de etiquetas entre ellos.

Un CE tiene conexión directa de nivel 3 con un router PE. Un router CE no tiene capacidad MPLS. Un router C es un equipo de cliente que no tiene conexión directa con un router PE. Los routers C y CE no necesitan tener capacidades de MPLS.

En la siguiente ilustración vemos todos los elementos de una VPN.

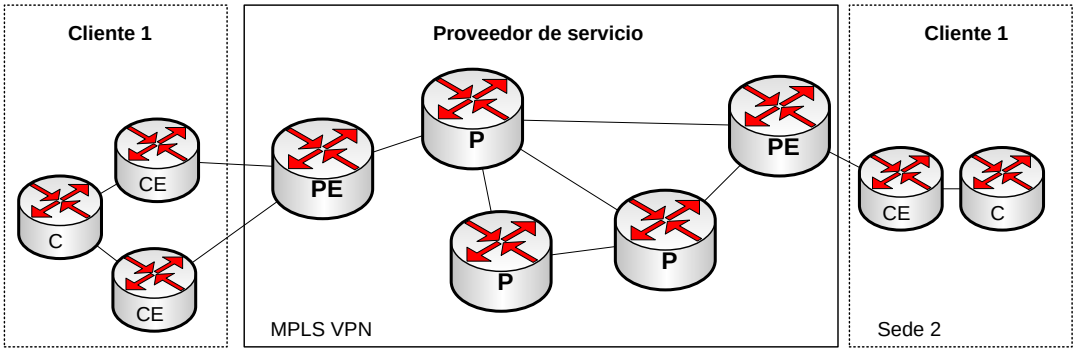

Ilustración 17- Elementos de una VPN MPLS.

Debido a que tanto los routers CE como los PE interactúan a nivel 3, es necesario que hablen entre ellos un protocolo de routing dinámico (o rutas estáticas). El CE solo tiene un equipo conectado fuera de su ubicación, el PE. El CE no tiene conectividad física directa con ningún otro CE. El nombre de este modelo se llama peer-to-peer ya que el CE y el PE tienen una conexión de nivel 3.

Una VPN debe ser privada, por ello los clientes pueden tener su propio plan de direccionamiento, puede usar, tanto direccionamiento público como privado e incluso se puede repetir direccionamiento entre clientes. Si los paquetes fuesen reenviados como paquetes IP en los nodos P habría un problema de routing. Si no se les permitiese a los clientes tener su propio direccionamiento, este debería ser asignado por el proveedor de servicios. Suponiendo esto, los paquetes podrían ser reenviados atendiendo a su dirección IP destino en cada router de la red del proveedor. Esto significa que tanto los nodos P como los nodos PE deberían tener una tabla de rutas completa con el direccionamiento de cada cliente y esa tabla podría ser muy grande. El único protocolo de routing capaz de manejar semejante tabla es BGP por lo que tanto nodos P como nodos PE deberían hablar iBGP entre ellos. Llegados a este caso no sería un esquema válido debido a que no es un entorno privado para cada cliente.

Otra solución sería que tanto LSRs P como LSRs PE manejasen tablas de rutas distintas para cada cliente. Debería haber tantos procesos de routing como VPNs de cliente hubiera configuradas en la red. Esta no es una solución muy escalable ya que cada vez que un nuevo cliente se diese de alta en la red habría que configurar en cada nodo (tanto P, como PE) un proceso de routing. Además, al entrar un paquete a la red a través de un PE, ¿Cómo se podría identificar a que VPN pertenece? La solución pasaría por modificar el paquete IP añadiéndole un campo de identificación de VPN. Entonces los nodos P deberían mirar además del campo IP destino el campo de VPN para reenviar adecuadamente el paquete.

Una solución escalable es que los routers P no tuviesen consciencia de VPN lo que les liberaría de la carga de tener información de las rutas para cada VPN. Precisamente esto es la solución que ofrece MPLS. Los paquetes IP de cada cliente son etiquetados en la red MPLS para conseguir una VPN privada para cada cliente. Además, los routers P no necesitan conocer la tabla de rutas gracias a la utilización de dos etiquetas MPLS. Por lo tanto, BGP no es necesario en los routers P. Las rutas para cada VPN solo se manejan en los nodos PE al igual

que solo hay concepto de VPN en los PEs lo que hace que las VPN MPLS sean una solución escalable.

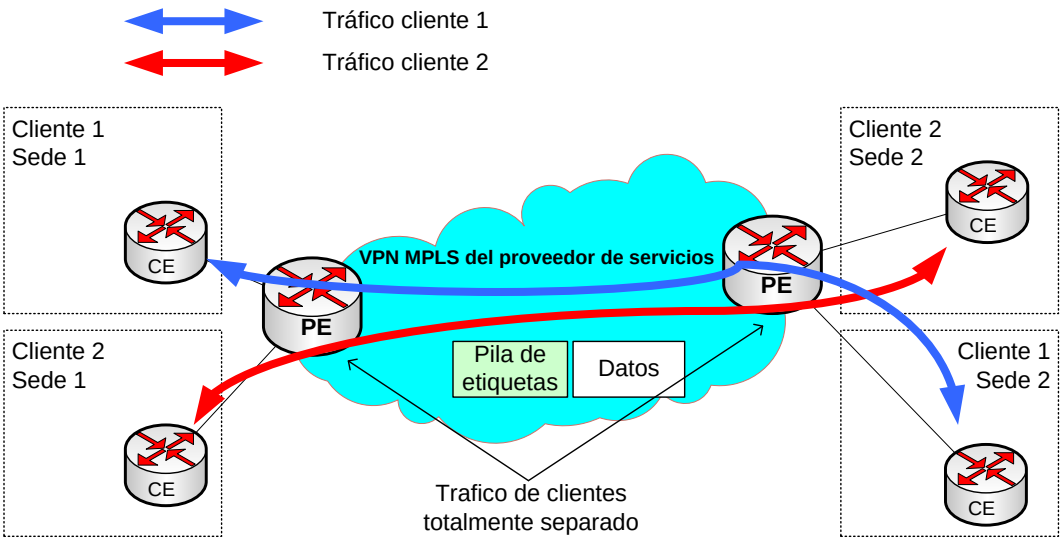

Ilustración 18- Modelo de VPN MPLS.

2.2.4 Arquitectura de VPN MPLS

Los conceptos básicos para comprender el funcionamiento de una VPN en MPLS son: VRF, route distinguisher (RD), route target (RT), propagación de rutas en MP-BGP y el reenvío de paquetes etiquetados.

2.2.5 Virtual Routing Forwarding (VRF)

Una VRF es una instancia de enrutamiento y reenvío en la VPN. Es el nombre que recibe la combinación de la tabla de routing de la VPN, la CEF de la VRF y los protocolos de routing IP asociados en el router PE. Un nodo PE tiene una instancia de VRF para cada VPN asociada.

En la siguiente ilustración podemos ver como un nodo PE tiene su tabla de rutas global IP y también una tabla de routing VRF por cada VPN conectada al PE.

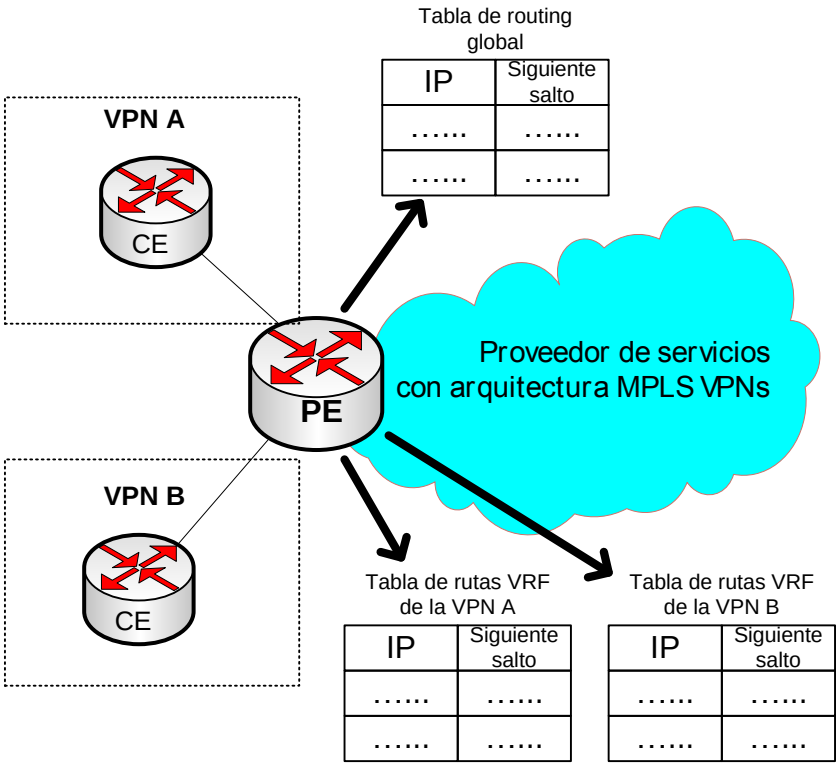

Ilustración 19- VRFs en un nodo PE.

Como la tabla de rutas debe estar separada y ser privada para cada cliente dentro de un nodo PE, cada VPN debe tener su propia tabla de rutas. Esta tabla de rutas privada se llama tabla de rutas VRF. El interfaz del PE que conecta con el CE puede pertenecer solo a una VRF por lo que todos los paquetes recibidos en la interfaz de esa VRF se identifican inequívocamente como pertenecientes a esa VRF. Puesto que hay una tabla de rutas separada por VPN también hay una tabla CEF específica por VPN para reenviar esos paquetes en el router PE. Esta tabla se llama tabla CEF VRF. Al igual que con la tabla de rutas global y la tabla CEF global, la tabla CEF VRF deriva de la tabla de rutas VRF. Una interfaz solo se puede asociar a una VRF y no a varias, pero una VRF puede estar asociada a varios interfaces.

La tabla de rutas VRF tiene prefijos aprendidos mediante protocolos dinámicos y rutas estáticas al igual que cualquier tabla de rutas estándar. Los conceptos de métrica, distancia, siguiente salto, etc.… tampoco cambian.

En equipos Cisco, CEF es el único método de conmutación soportada para el reenvío de paquetes IP de interfaces que pertenezcan a VRFs, por tanto, CEF debe ser habilitado de manera global en el router PE y en todas las interfaces de VRF.

La configuración y el comportamiento de VRFs sobre equipos sobre los que corre MPLS es el mismo que para cualquier otro equipo Cisco que soporte esta funcionalidad. El concepto de VRF no es específico de equipos MPLS y su uso es independiente: El uso de VRFs no implica MPLS y viceversa aunque en MPLS es habitual el uso de las VRFs.

2.2.6 RD

Los prefijos de la VPN se propagan a través de la VPN sobre MPLS mediante Multiprotocol-BGP. El problema es que cuando BGP transporte estos prefijos sobre la red deben ser únicos y si el cliente tiene direccionamiento IP solapado el routing podría ser erróneo. Para solucionar este problema se crea el concepto de RD que convierte los prefijos IP en únicos. Cada prefijo de cada cliente recibe un identificador único RD para distinguir el mismo prefijo de distintos clientes. El prefijo deriva de la combinación del prefijo IP y del RD y se llama prefijo VPNv4. El MP-BGP transporta los prefijos VPNv4 entre los routers PE.

El RD es un campo de 64 bits pero no indica a que VRF pertenece el prefijo. La función del RD no es ser un identificador de VPN ya que algunos escenarios de VPN más complejos pueden requerir más de un RD por VPN. Cada instancia de VRF en un nodo PE debe tener un RD asignado.

El valor del campo del RD puede tener dos formatos: *ASN:nn* o *DirecciónIP:nn* donde *nn* representa un número. El formato más usado comúnmente es *ASN:nn* donde *ASN* es el número de sistema autónomo. Normalmente ASN es el número de sistema autónomo asignado por IANA al proveedor de servicio y nn es el número que el proveedor de servicio asigna unívocamente a la VRF. El RD no impone semántica y se usa solamente para identificar de manera única las rutas de la VPN. Esto es necesario para evitar solapamiento IP entre clientes. La combinación del RD y el prefijo IP proporciona un prefijo VPNv4 de 96 bits de longitud.

Esto es un ejemplo:

```
RD: 64987:140014
Prefijo IPv4: 10.200.3.4/30
Prefijo VPNv4: 64987:140014:10.200.3.4/30
```

Un cliente puede usar diferentes RDs para una misma ruta IP. Cuando una sede de la VPN se conecta a dos PEs, las rutas de la sede VPN pueden tener dos RDs diferentes dependiendo en que PE se reciben las rutas. Cada ruta IP tendrá en ese caso dos RDs diferentes asignados y tendrá dos rutas VPNv4 totalmente diferentes, esto permite a BGP verlas como diferentes rutas y aplicarles diferentes políticas.

2.2.7 ROUTE TARGET (RT)

Si los RDs se usaran solo para identificar la VPN, la comunicación entre sedes de distintas VPNs sería problemática y a veces esto es necesario p.e cuando dos clientes necesitan acceder a un mismo recurso (DMZ, servidor, segmento de red, etc.…) Una sede de un cliente A no podría comunicarse con una sede de un cliente B porque los RDs no coincidirían. El concepto de sedes de distintos clientes con comunicación entre si se llaman extranet VPN. El caso más sencillo de comunicación entre sedes de un mismo cliente (de la misma VPN) se conoce como intranet VPN. La comunicación entre sedes se controla mediante otra funcionalidad de la VPN MPLS llamada Route Target (RT).

Un RT es una comunidad extendida de BGP que indica que rutas deben ser importadas de MP-BGP a la VRF. Exportar un RT significa que a cada ruta VPNv4 exportada se le añade una comunidad BGP extendida (esto es el RT) cuando esta ruta se redistribuye de la tabla de rutas VRF al MP-BGP. Importar un RT significa que para cada ruta VPNv4 recibida de MP-BGP se comprueba si su comunidad extendida (RT) coincide con alguna de las asociadas a alguna VRF. Si coincide el prefijo se incluye en la tabla de rutas VRF como una ruta IP. Si no coincide el prefijo es rechazado.

La siguiente ilustración muestra como los RTs controlan que rutas se importan en cada VRF desde los PEs remotos y con que RTs se exportan los prefijos VPNv4 hacia los PEs remotos. Más de un RT puede ser asociado a un prefijo VPNv4. Para que la importación hacia la VRF se permita solo es necesario que un RT del prefijo VPNv4 coincida con alguno de los RTs importados en esa VRF.

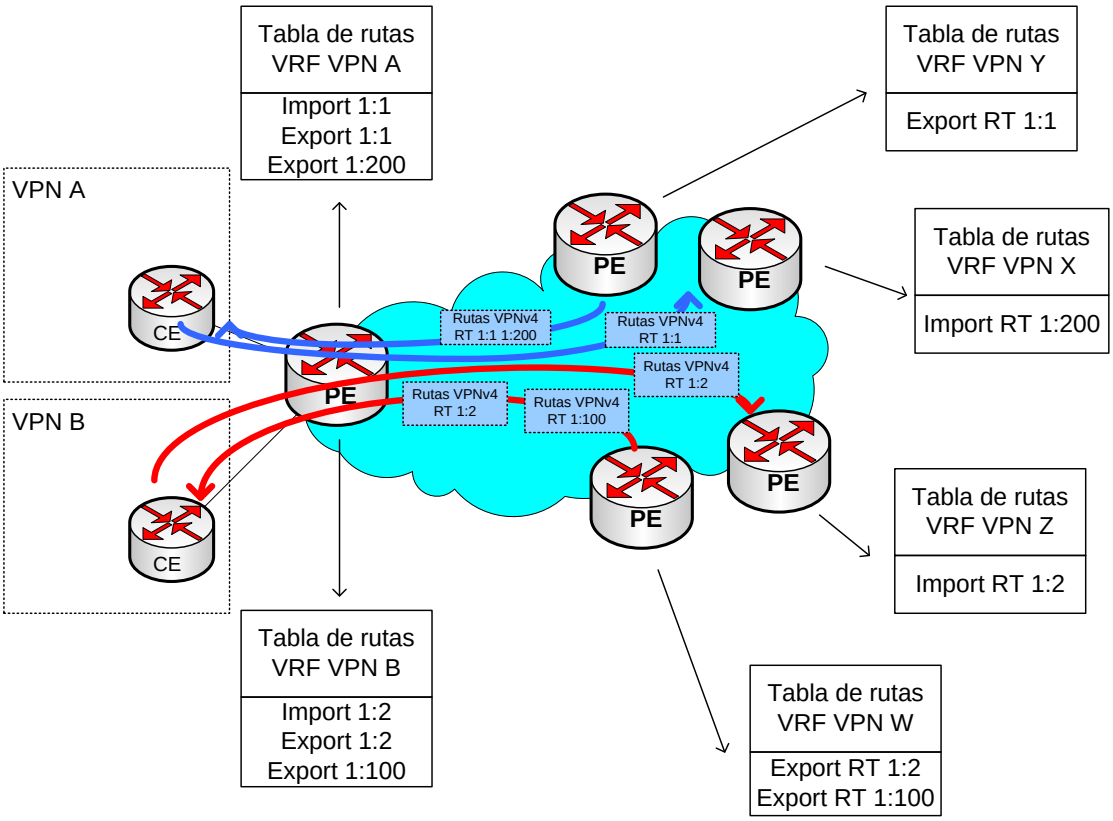

Ilustración 20- Funcionamiento de los Route Targets

El siguiente ejemplo es una configuración de VRF donde se reflejan los elementos que hemos visto:

```
ip vrf vpnA                          → Nombre de la VRF
 description Cliente_A               → Descripción
```

```
 rd 64987:140014                          → Route Distinguisher; formato ASN:nn
 route-target export 1:1                  → Comunidad extendida exportada
 route-target import 1:1                  → Comunidades extendidas importadas
 route-target import 1:200
```

Cuando una VPN se configura con varias sedes en la misma VRF, sin que estas sedes tengan que comunicar con sedes de otra VPN solo es necesario configurar un RT para importar y exportar en todos los routers PE que tengan una sede conectada a esa VRF. Es el caso más simple, intranet, como es el siguiente ejemplo.

```
ip vrf vpn14:vrf13
 description Cliente_A
 rd 64987:140013
 route-target export 64987:149999
 route-target import 64987:149999
```

Cuando las sedes que pertenecen a una VPN de un cliente necesitan tener comunicación con sedes de la VPN de otro cliente, extranet, los RTs a los que exporta un cliente deben ser importados por el otro y viceversa. El siguiente ejemplo lo muestra:

VRF Cliente A

```
ip vrf vpn14:vrf14
 description Cliente_A
 rd 64987:140014
 route-target export 64987:149999
 route-target import 64987:149999
 route-target import 64987:158888
```

VRF Cliente B

```
ip vrf vpn15:vrf15
 description Cliente_B
 rd 64987:150015
 route-target export 64987:158888
 route-target import 64987:158888
 route-target import  64987:149999
```

También se podría no querer que todas las rutas de dos VRFs se intercambiasen absolutamente todas las rutas ya que entonces no tendría sentido el concepto de separar el tráfico. Esto puede ser útil, como en el caso de ambas redes cuya unión es objetivo de este proyecto, para la importación y exportación de redes de gestión. El proveedor de servicios configura en la red una VPN de gestión, las redes de esta VPN son exportadas al resto de VPNs y tan solo importa las direcciones de gestión de los equipos. Así se consigue que en la VPN de gestión solamente circule tráfico cuyo origen y destino solo puede ser el segmento de gestión y los equipos a gestionar.

2.2.8 PROPAGACIÓN DE RUTAS VPNv4 EN UNA VPN MPLS

Las VRFs separan las rutas de cliente en los nodos PE, pero absolutamente todos los prefijos son transportados a través de la red MPLS. Potencialmente pueden ser cientos de miles de rutas ya que pueden ser numerosas las VPNs de cliente configuradas. Para este transporte de rutas, BGP es el protocolo ideal ya que está probado y es estable para el manejo de grandes tablas de rutas, por eso es el protocolo estandarizado para internet. Gracias a la transformación de prefijos IP en prefijos VPNv4 (RD + prefijo IP), todas las rutas se pueden transportar de manera segura a través de la red.

El nodo PE recibe rutas IP desde el CE mediante un IGP o mediante eBGP. Estas rutas IP de una VPN determinada se insertan en una tabla de rutas VRF. Esta VRF depende de la que esté configurada sobre el interfaz del PE que conecta con el CE que inyecta las rutas. Estas rutas IP se convierten en rutas VPNv4 una vez que los prefijos se asignan al RD correspondiente, es entonces cuando entran en el proceso de MP-BGP. BGP se encarga de distribuir estas rutas VPNv4 hacia todos los PEs en esa VPN. El que la ruta VPNv4, después de separarse del RD, sea puesta en la tabla de VRF como rutas IP o no depende de si los RTs permiten la importación a esa VRF. Esas rutas IP son entonces anunciadas al router CE mediante un IGP o eBGP que esté corriendo entre el PE y el CE.

Para comprender todos estos procesos, en la siguiente ilustración se ven los pasos que se establecen para que se produzca comunicación IP entre dos CEs a través de una VPN MPLS.

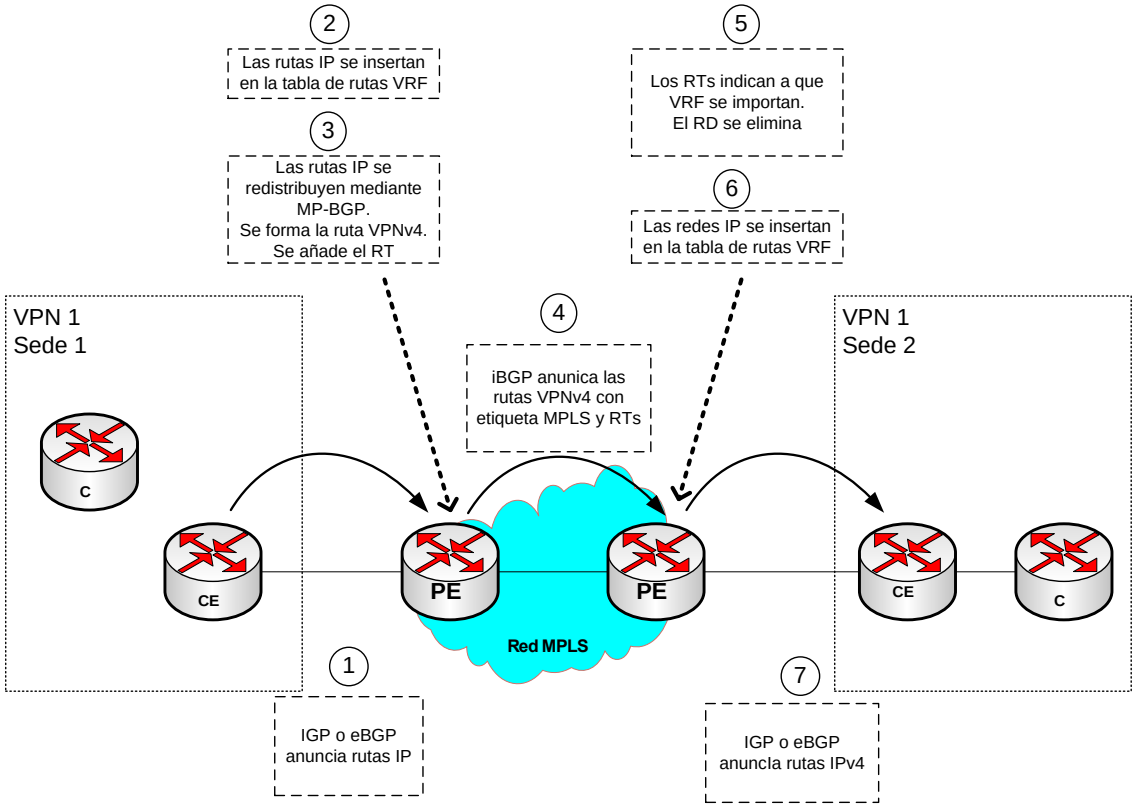

Ilustración 21- Propagación de rutas en una VPN MPLS paso a paso

2.2.9 REENVÍO DE PAQUETES EN UNA RED VPN MPLS

Los paquetes no pueden ser reenviados como paquetes puros IP entre dos sedes. Los routers P no pueden reenviarlos porque no tienen información alguna de VRFs. MPLS soluciona este problema etiquetando los paquetes. La manera más habitual es hacerlo con LDP entre todos los routers P y PE así todo el tráfico IP es reenviado basado en etiquetas. También se puede usar RSVP con extensiones para ingeniería de tráfico pero LDP es lo más común. Los paquetes IP son reenviados basándose en etiquetas desde el Ingress PE hasta el Egress PE. Un nodo P nunca tiene que tiene que consultar la cabecera IP. Esta es la manera en que los paquetes se conmutan entre el Ingress PE y el Egress PE. Esta etiqueta se llama *etiqueta IGP*, ya que es la etiqueta que se asocia a un prefijo IP en la tabla de routing global de los routers P y PE y es anunciada por el IGP.

Para resumir, el tráfico VRF a VRF tiene dos etiquetas en una red VPN MPLS. La etiqueta externa es la etiqueta IGP y es distribuida mediante LDP o RSVP entre todos los routers P y PE salto a salto. La etiqueta más interna es la *etiqueta VPN* que es anunciada por MP-BGP de PE a PE. Los routers P consultan la etiqueta IGP para reenviar los paquetes hacia el nodo PE correcto. Los Egress PE usan la etiqueta de VPN para reenviar el paquete al CE correcto.

En el siguiente ejemplo podemos ver como es el reenvío de paquetes en una red VPN MPLS. Los paquetes entran en el router PE en la VRF asociada al interfaz de entrada como un paquete IP. Es reenviado a través de la red VPN MPLS con dos etiquetas. Los routers P reenvían el paquete mirando la etiqueta externa. Esta etiqueta externa es intercambiada en cada nodo P. Las etiquetas son eliminadas en el Egress PE y el paquete es enviado como un paquete IP sobre el interfaz que corresponda a la VRF adecuada hacia el CE. El CE se encuentra mirando la etiqueta VPN.

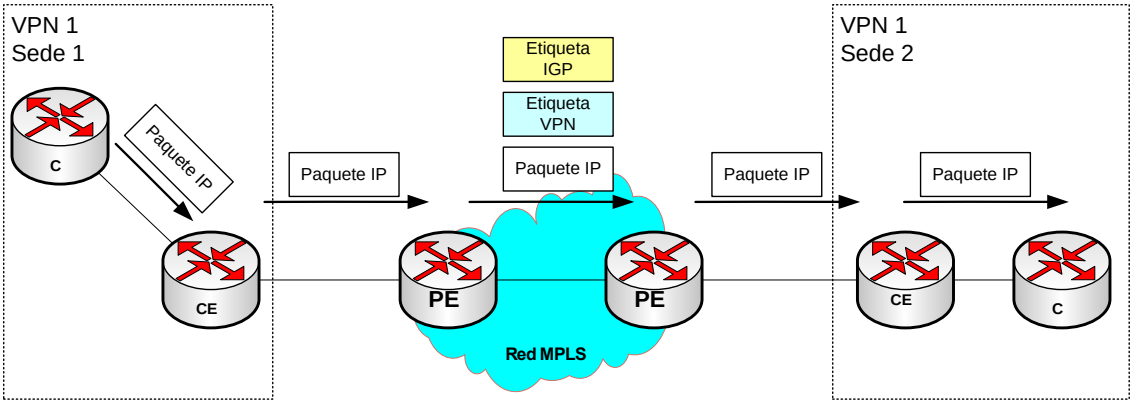

Ilustración 22- Formato de paquetes en una red VPN MPLS

2.2.10 BGP

Antes de seguir avanzando en conceptos de encaminamiento en una red VPN MPLS es necesario conocer un poco acerca de las características de BGP necesarias en una VPN MPLS.

BGP ha sido durante muchos años el estándar de protocolo de routing entre dominios. Es el protocolo que hace que internet funcione. Los proveedores de servicios intercambian rutas mediante BGP. Se interconectan con otros proveedores de servicios mediante eBGP y en su red hablan iBGP. BGP es un protocolo de rutas que está totalmente adaptado para transportar cientos de miles de rutas. También es un protocolo que permite implementar políticas flexibles y extendidas. Son estas características las que le convierten en un buen candidato para transportar rutas VPN MPLS. Como mencionamos antes, lo que realmente transporta son prefijos VPNv4.

BGP-4 está descrito en la RFC 1771, pero en esta RFC solo se describe el uso de BGP para transportar rutas IP pero BGP puede hacer mucho más que transportar rutas IP. La RFC 2858, "Multiprotocol extensions por BGP-4" fue escrita para extender BGP y que fuese capaz de transportar otra información de routing además de IP. Un equipo BGP permite a sus vecinos BGP que extensiones multiprotocolo para BGP-4 soporta utilizando anuncios de capacidades. Los vecinos BGP se mandan entre si las capacidades que soportan. Las capacidades que dos vecinos compartan, son las que pueden usar. La RFC 3392 "Capabilities Advertisement with BGP-4" describe el funcionamiento de los anuncios de capacidad.

Cuando un nodo que tiene BGP configurado manda un mensaje de "open" a sus vecinos BGP, puede incluir un parámetro de capacidad opcional listando todas las capacidades del equipo. Todos los vecinos BGP pueden hacer lo mismo.

Las extensiones multiprotocolo para BGP-4 definen dos nuevos atributos: Multiprotocol Reachable NLRI y Multiprotocol Unreachable NLRI. Estos atributos anuncian o dan de baja rutas. Ambos tienen dos campos: el Address Family Identifier (AFI) y el Subsequent Address Family Identifier (SAFI). La combinación de estos atributos describe exactamente qué tipo de rutas BGP se transportan.

La siguiente tabla indica algunos códigos AFI y su descripción:

Número	Descripción
0	Reservado
1	IPv4
2	IPv6
11	IPX
12	AppleTalk

Tabla 2. –Códigos AFI y descripción

La siguiente tabla muestra códigos SAFI y su descripción:

Número	Descripción
1	NLRI para reenvió unicast
2	NLRI para reenvío multicast
3	NLRI para reenvío unicast y multicast
4	NLRI para reenvío de IPv4 y etiquetas
128	NLRI para reenvío VPN etiquetado

Tabla 3. –Códigos SAFI y descripción

Para soportar el comportamiento multiprotocolo de BGP en tecnología Cisco, existe el concepto de familia de direcciones (address family). Las cuatro familias de direcciones que actualmente se soportan son: IPv4, IPv6, VPNv4 (VPN IPv4) y VPNv6 (VPN IPv6). Las siguientes familias de direcciones que se pueden combinar con las anteriores son unicast, multicast y VRF.

2.2.11 COMUNIDAD EXTENDIDA BGP: RT

El atributo comunidad es un atributo opcional transitivo que se describe en la RFC 1997. La comunidad extendida es también un atributo opcional transitivo. Se creó para extender el rango de comunidades y tiene una estructura reforzada sobre el atributo comunidad de BGP. Muchos atributos comunidad extendida son definidos, pero solo uno es necesario para VPNs MPLS: el route target. Este atributo indica a los PEs si una ruta debe ser importada a una VRF.

Veamos el siguiente ejemplo:

```
MPLS-ASL#show ip bgp vpnv4 rd 64987:170001 10.200.19.190
BGP routing table entry for 64987:170001:10.200.19.190/32, version 5962
Paths: (1 available, best #1, table vpn17:vrf1)
  Advertised to update-groups:
     21          22          24          25          26          27          28
     1
  Local
    10.200.17.10 from 0.0.0.0 (10.132.1.2)
      Origin incomplete, metric 1, localpref 100, weight 32768, valid,
sourced, best
      Extended Community: RT:64987:179999 RT:64987:999999998
      mpls labels in/out 446/nolabel
```

Aquí vemos que la ruta VPNv4 64987:170001:10.200.19.190/32 posee los RTs 64987:179999 y 64987:999999998. Solo las VRFs que estén configuradas para importar al menos uno de esos RTs, insertarán la ruta IP 10.200.19.190/32 en su tabla de rutas VRF.

2.2.12 TRANSPORTE DE ETIQUETAS CON BGP

BGP anuncia los prefijos VPNv4 en la VPN MPLS. Esto no es suficiente para que el Egress PE sea capaz de identificar correctamente hacia que CE debe enviar el paquete que le llega ya que debe distinguir a que VPN pertenece ese tráfico. Para ello necesita basarse en una etiqueta. El Egress PE asigna una etiqueta a cada prefijo VPNv4 llamada etiqueta VPN. El Egress PE debe anunciar esta etiqueta junto con el prefijo VPNv4 a los posibles Ingress PE. La codificación de la etiqueta con el prefijo se describe en la RFC 3107 "carrying label information in BGP-4". Esta etiqueta se envía junto con el prefijo VPNv4 y se anuncia mediante BGP usando el atributo de la extensión de multiprotocolo BGP. Esta etiqueta está contenida en el campo NLRI. El AFI tiene valor 1 y el SAFI es 128 en el caso de VPN MPLS para IPv4.

2.2.13 ROUTE REFLECTOR (RR)

Un route reflector es un nodo que habla BGP y refleja rutas de otros equipos que hablan BGP. Los RR surgieron cuando las redes se hicieron muy extensas ya que el iBGP requiere que todos los equipos que hablen BGP tengan una configuración mallada entre ellos. Esto es posible cuando el número de equipos es bajo pero da problemas cuando las redes pasan de un cierto tamaño. En una red con *n* nodos hablando iBGP, cada nodo tiene *n-1* vecinos por lo que tendrá que establecer *n*(n-1)/2* sesiones BGP en total. Para solucionar esto se crearon los route reflectors y las confederaciones de BGP. Los nodos se agrupan y dentro de cada grupo se definen uno o más equipos con el papel de RR. Cada equipo que forma el grupo establece una sesión BGP únicamente contra el/los RR/s de su grupo y no con el resto de equipos que lo forman. Los route reflector únicamente reenvían o reflejan las rutas BGP que reciben. Si se quiere emplear RRs con VPN MPLS los route reflector deben reflejar prefijos VPNv4 los cuales transportan etiquetas. El RR solo cambia la etiqueta si se convierte en el siguiente salto para las rutas cosa que no es habitual. Los RRs que se convierten en el siguiente salto para una ruta iBGP deben reenviar el tráfico para esas rutas. Esto puede hacer que una gran cantidad de tráfico tenga que atravesar unos pocos RRs. Los RRs no deben asumir la carga del reenvío de tráfico sino que únicamente deben reflejar las rutas BGP.

Los RR difieren en otra característica del resto de PEs en una VPN MPLS. Nunca descartan rutas VPNv4. Un PE tiene este comportamiento para ahorrar memoria pero un RR no puede descartarlas ya que no tiene conocimiento de que RTs permiten o deniegan el resto de nodos. Para aliviar la carga, se pueden implementar grupos route reflector para dividir la carga entre varios RRs o Grupos RR. Cada RR o Grupo RR redistribuirá un subgrupo de rutas VPNv4.

2.2.14 GRUPO RR

No es necesario para un RR o grupo de RRS tener todas las rutas VPNv4 en la tabla BGP. Se pueden subdividir las rutas VPNv4 en grupos y permitir varios RRs o Grupos RR transportar cada uno de los grupos de rutas. Esto incrementa la escalabilidad de la red. Para realizar esto se debe especificar una lista de comunidades extendidas en cada Grupo RR. Estas comunidades extendidas especifican que RTs quieres que el Grupo RR permita o deniegue.

En la siguiente figura vemos una red MPLS con dos RRs con Grupos RR configurados para diferentes conjuntos de rutas. Uno filtra las rutas con los RTs pares y el otro los RTs impares.

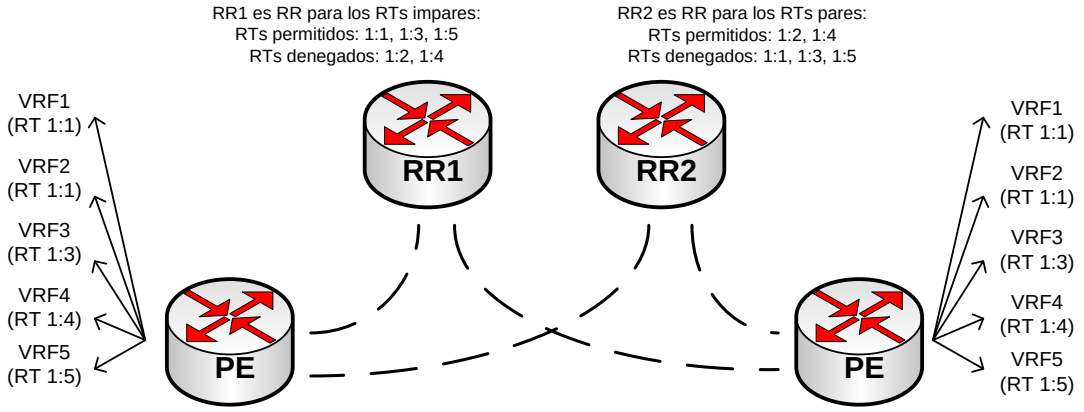

Ilustración 23- Ejemplo de red VPN MPLS con Grupos RR

2.2.15 Selección de rutas BGP

Si un CPE está conectado a dos PEs (o dos CPEs de la misma sede a dos PEs), se recibirá la misma ruta por dos nodos por lo que BGP debe seleccionar la mejor ruta. El proceso de selección de la mejor ruta es en este caso, idéntico al caso de IPv4 solo que ahora los prefijos no son de 32 bits (dirección IP) sino de 96 bits (VPNv4). De cualquier manera, dado este caso, un paquete en un Ingress PE con destino esta red de cliente tendrá dos rutas con diferente next-hop (los dos Egress PE). Entonces el Ingress PE selecciona mediante los mecanismos de BGP la mejor ruta de las dos y la inserta en su tabla de rutas VRF.

2.2.16 Reenvío de paquetes

En este apartado veremos un ejemplo específico donde se muestra como es el paso de un paquete IP en una red VPN MPLS.

La primera característica de estas VPNs es el multiprotocol BPG; debe establecerse una sesión entre el Egress y el Ingress PE que distribuyen los prefijos VPNv4 y la etiqueta VPN.

La segunda característica es el uso de un protocolo de distribución de etiquetas, aquí asumimos que es LDP.

Entre los CEs y los PEs es necesario un protocolo de routing para rellenar la tabla de rutas de la VRF que luego serán distribuidas mediante MP-BPG.

Las siguientes figuras son útiles para comprender todos estos procesos. La primera de ellas muestra como es el anuncio de la red 10.109.2.1/32 y de las etiquetas correspondientes dentro de la red MPLS .Las flechas de la parte baja de la ilustración son el anuncio del siguiente salto. La dirección siguiente salto de BGP es siempre el Egress PE, el cual se anuncia como prefijo IGP por LDP. La red de cliente 10.102.2.1/32 se anuncia CE-PE mediante un protocolo de routing cualquiera y es el Egress PE el que transforma ese prefijo en un prefijo IPv4

añadiéndole el RD correspondiente, en este caso 64987:170001 y lo envía al Ingress PE vía MP-BGP.

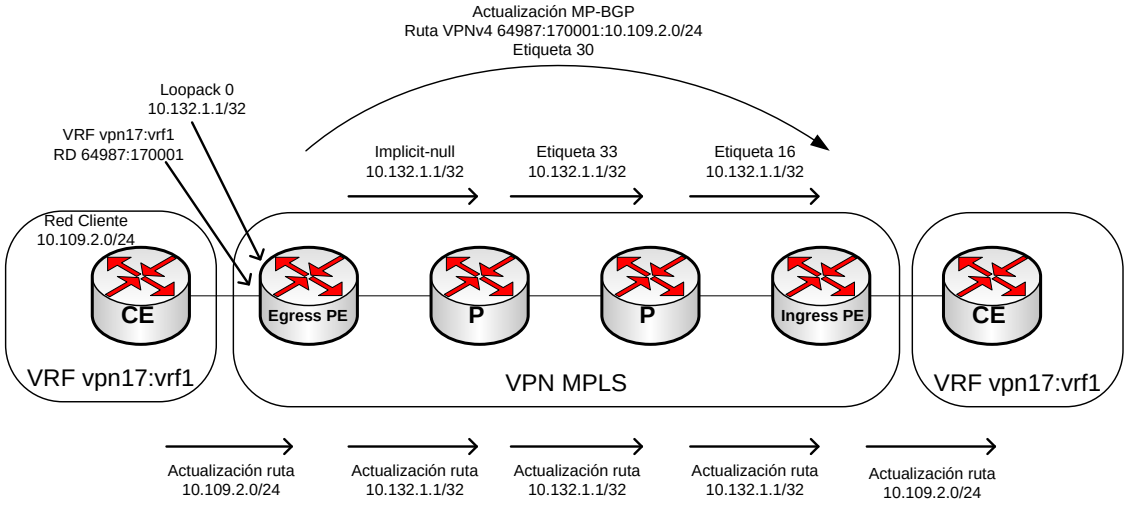

Ilustración 24- Vida de un paquete IP en una red VPN MPLS: enrutamiento y anuncio de etiquetas.

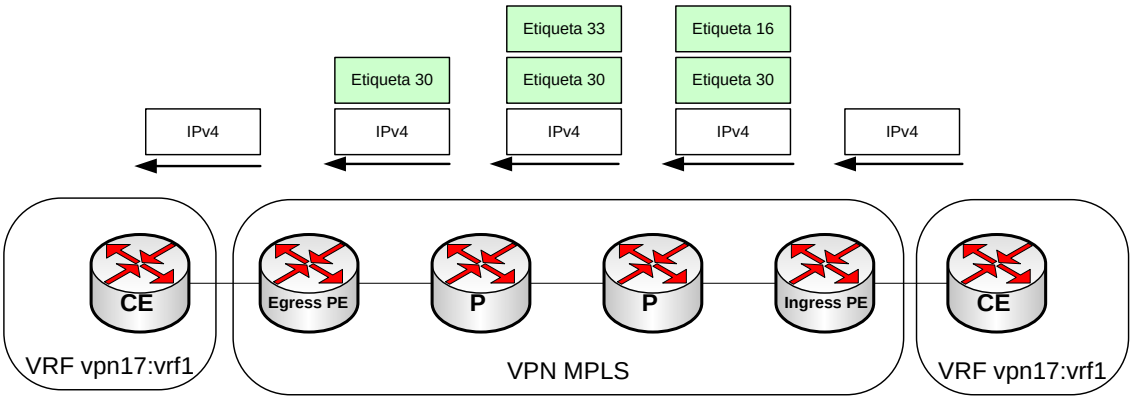

Ilustración 25- Vida de un paquete IP en una red VPN MPLS: reenvío de paquetes.

Cuando un paquete IP entra en un Ingress PE, este mira su dirección IP en la tabla CEF de la VRF vpn17:vrf1; el Ingress PE sabe en qué tabla mirar ya que el interfaz por el que entra el paquete está asociado exclusivamente a una VRF. La entrada específica en la tabla CEF de la VRF normalmente indica que dos etiquetas son necesarias.

Primero, el Ingress PE inserta delante del paquete IP la etiqueta de VPN 30 para su anuncio mediante MP-BGP. Esta será la etiqueta más interna de la pila de etiquetas: Entonces el Ingress PE le añade a la pila de etiquetas la etiqueta IGP como etiqueta externa. Esta etiqueta está asociada unívocamente con el prefijo IP 10.109.2.1/32 y está relacionada con el siguiente salto BGP que será, normalmente, la dirección de loopback del Egress PE para ese prefijo. El prefijo es anunciado salto a salto a través de toda la red y la etiqueta va cambiando su valor en cada nodo P. El Ingress PE le asocia la etiqueta 16.

El paquete IPv4 sale del Ingress PE con dos etiquetas en la pila. La etiqueta más externa (etiqueta IGP) es intercambiada en cada nodo por el que pasa el paquete. Normalmente, dado que es el comportamiento por defecto en equipos Cisco, el PHP se da entre el último P y el Egress PE, de esta manera la etiqueta IGP es extraída en el último router P por lo que llega al Egress PE con una única etiqueta, la etiqueta VPN. El Egress PE comprueba esta etiqueta VPN en su LFIB y toma decisiones de encaminamiento. Ya que la etiqueta de salida es "No label" el resto del paquete pierde sus etiquetas MPLS y es reenviado como un paquete IP hacia el CE. El PE no tiene que comprobar la dirección IP destino del paquete IP si la etiqueta es "No label" ya que la información de siguiente salto la encuentra en la etiqueta VPN en la LFIB.

En los siguientes ejemplos se puede comprobar cómo es el anuncio mediante LDP y MP-BGP y su uso en la tabla CEF de la VRF y en la LFIB.

```
MPLS-FII#show ip cef vrf vpn17:vrf1 10.109.2.0 255.255.255.0 detail
10.109.2.0/24, epoch 27
  recursive via 10.132.1.1 label 30
    nexthop 10.132.1.33 POS6/0/0 label 16
```

```
MPLS-FII#show ip bgp vpnv4 rd 64987:170001 10.109.2.0
BGP routing table entry for 64987:170001:10.109.2.0/24, version 1487089
Paths: (2 available, best #2, table vpn17:vrf1)
  Advertised to update-groups:
     3          4          5          6          7          13         24
  65123
    10.132.1.1 (metric 701) from 192.168.255.18 (192.168.255.18)
      Origin incomplete, metric 0, localpref 100, valid, internal
      Extended Community: SoO:64987:170004 RT:64987:179999
      Originator: 10.132.1.1, Cluster list: 0.0.2.128
      mpls labels in/out nolabel/247
  65123
    10.132.1.1 (metric 701) from 192.168.255.9 (192.168.255.9)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      Extended Community: SoO:64987:170004 RT:64987:179999
      Originator: 10.132.1.1, Cluster list: 0.0.2.128
      mpls labels in/out nolabel/30
```

```
MPLS-MS#show mpls forwarding-table labels 30
Local  Outgoing      Prefix            Bytes Label   Outgoing   Next Hop
Label  Label or VC   or Tunnel Id      Switched      interface
30     No Label      10.109.2.0/24[V]  24673951339   Gi10/1     10.200.17.110
```

2.3 PROTOCOLOS DE ROUTING COMUNES EN ENTORNOS MPLS.

En este apartado veremos los protocolos de routing más comunes en entornos MPLS. Dada la estrecha relación de BGP con MPLS lo vimos dentro del apartado de MPLS. Otros dos

protocolos muy comunes en arquitecturas MPLS son OSPF e IS-IS. Ambos, debido a sus características, pueden ser utilizados como protocolos IGP así como protocolos de routing CE-PE.

Antes de ver ambos protocolos vamos a hacer una primera diferenciación dentro de los protocolos de routing. La mayor parte de los protocolos de encaminamiento pueden clasificarse en dos grandes grupos:

1.- Protocolos de tipo Vector Distancia (DS).

2.- Protocolos de tipo Estado de Enlace (LS).

2.3.1 PROTOCOLOS DE ROUTING DE TIPO VECTOR DISTANCIA.

Un protocolo de tipo Vector Distancia se llama así por la información que anuncia a sus vecinos que es del tipo [distancia, dirección]. La distancia es la métrica asociada al camino para llegar a la red de destino y la dirección se mide en términos de "siguiente salto". Este algoritmo de operación fue introducido por Bellman – Ford - Fukerson de la universidad de Princeton.

Las características básicas de funcionamiento son:

1.- Actualizaciones periódicas: Dependiendo del protocolo de routing en cuestión variará el tiempo entre actualizaciones.

2.- Envío de la tabla de rutas completa. Esto tiene el inconveniente de que se puede producir el efecto de Split-Horizon (Consiste en que tu generas una ruta que pierdes al caer un enlace pero la recibes del vecino, cuando en realidad la del vecino es la tuya que le enviaste anteriormente). Se evita si no envías a un vecino rutas que has aprendido de él.

3.- Envío de las actualizaciones a una dirección de broadcast (Excepto en RIPv2).

4.- Temporizadores de invalidez de rutas: Las rutas dejan de ser válidas cuando pasado un tiempo no han sido actualizadas. No hay mensajes para dar de baja rutas.

Como hemos descrito, un router le pasa a sus vecinos directamente conectados la tabla de rutas completa, a su vez, estos pasas estas rutas a sus vecinos añadiéndoles una unidad al coste.

En la siguiente ilustración vemos como se distribuyen y como se almacenan en las tablas de routing las rutas en protocolos de Vector Distancia

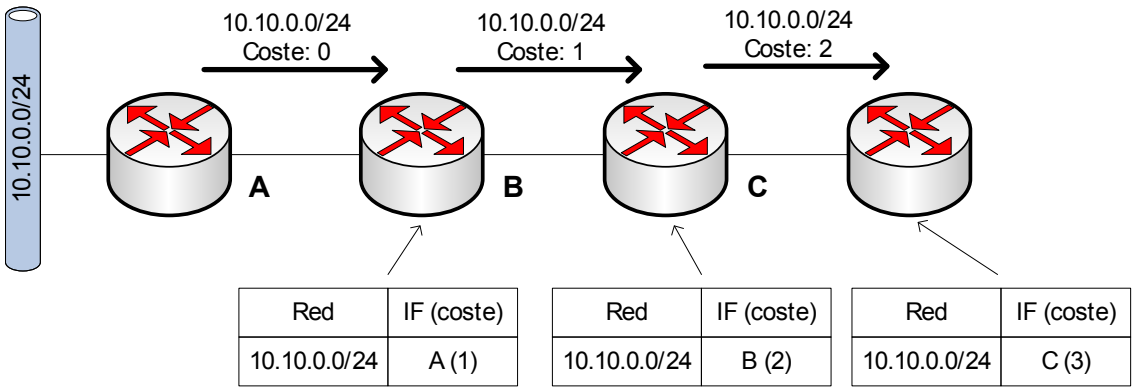

Ilustración 26- Propagación de rutas y relleno de tablas en escenarios de tipo Vector Distancia

Con este comportamiento podemos llegar a varias conclusiones. Cada router tiene una visión limitada de de la red, para un determinado destino tan solo conoce la distancia a la que se encuentra y el interfaz por donde se alcanza.

También tiene un tiempo de convergencia lento ya que cuando un router recibe un cambio tiene que recalcular la tabla de rutas antes de anunciar dicho cambio.

Tiene un alto consumo de recursos de la red ya que envía periódicamente toda su tabla de rutas per por el contrario tiene un bajo consumo de recursos del router ya que no hay base de datos que mantener ni algoritmos asociados.

Ejemplos de protocolos de routing comunes basados en vector distancia son RIP (todas sus versiones) y EIGRP.

2.3.2 Protocolos de routing de tipo Estado de Enlace

Tal y como afirma Doyle, un protocolo de tipo Vector Distancias es a una señal en la carretera lo que un protocolo de tipo Estado de Enlace es a un mapa de carreteras. En un protocolo de tipo Estado de Enlace cada router tiene un conocimiento total de la topología de la red. Cada router genera una unidad de datos con la información de cada uno de sus enlaces y el coste asociado a los mismos. Se conocen con el nombre de *Shorted path first* (el camino más corto primero).

Su operación se basa en el mantenimiento de tres bases de datos:

1.- Base de datos de vecinos

2.- Bases de datos topológica

3.- Tabla de rutas

Estas bases de datos se crean en orden:

BD Vecino ➔ BD Topológica ➔ Tabla de rutas

Antes de nada, cada router tiene que descubrir a sus vecinos mediante el intercambio de mensajes *Hello*. Tras conocerse, los routers se ponen de acuerdo para sincronizar sus bases de datos.

Cada router envía a sus vecinos la información de sus enlaces y también el estado de los enlaces recibido de otros routers. Así se mantiene la base de datos topológica. Una vez se ha recibido esta información cada router ejecuta su algoritmo de cálculo de rutas (SFP) y actualiza la tabla.

Los protocolos de Estado de Enlaces tienen las siguientes características:

- Responden rápidamente a los cambios de la red
- Envían actualizaciones específicas cuando se produce un cambio en la red
- Envían actualizaciones periódicas (refrescos del estado de los enlaces), en largos intervalos de tiempo (30 minutos)

Estos protocolos generan una actualización de routing sólo cuando se produce un cambio de topología en la red. Cuando el estado de un enlace cambia, el equipo que lo detecta genera un anuncio de estado del enlace (LSA). Este LSA se propaga a todos los equipos vecinos usando una dirección de multicast. Cada equipo toma una copia de ese LSA, actualiza su base de datos de estado de los enlaces y reenvía el LSA a sus vecinos. La inundación de LSAs asegura que todos los equipos actualicen sus bases de datos antes de actualizar sus tablas de rutas para reflejar la nueva topología.

La base de datos de estado de los enlaces se usa para calcular el mejor camino a través de la red. Los routers determinan el mejor camino aplicando el algoritmo de Dijkstra sobre la base de datos de estado de los enlaces para calcular el árbol de SPF (Shortest Pach First). Los mejores caminos de este árbol se seleccionan y se incluyen en la tabla de rutas.

Los protocolos de Estado de Enlaces recogen información de los demás routers en la red o del área de la red a la que pertenezcan y una vez que han recogido esta información, cada router de manera independiente calcula el mejor camino para cada uno de los destinos usando el algoritmo de Dijkstra. Cada router posee información de toda la red, por lo que conoce más información que aquellos que emplean protocolos de vector distancia. Por tanto realizan decisiones más precisas para el enrutamiento y son más robustos ante la posibilidad de que un equipo que pueda introducir información errónea.

Estos protocolos almacenan toda esta información recogida en las diversas tablas indicadas anteriormente. Los recursos de memoria necesarios para mantener estas tablas son una desventaja de los protocolos de Estado de Enlaces. Ya que la tabla de topología es idéntica en todos los routers del área y contiene información de todos los routers y enlaces de la misma, cada router es capaz de seleccionar el mejor camino, basado en coste, para alcanzar cada red del área. Este es un beneficio sobre la limitación de “routing por rumores” que

presentan los protocolos de vector distancia. Con protocolos de vector distancia los routers confían en las decisiones de routing de sus vecinos y no tienen una visión total de la topología de la red. En cambio los de Estado de Enlaces tienen una visión completa de la red y toman las decisiones de routing independientemente y basándose en una visión precisa de toda la topología de la red.

2.3.3 OSPF.

OSPF es un protocolo de encaminamiento basado en el algoritmo Enlace-Estado (LSA-Link State Algorithm). Está descrito en la RFC 2328.

Su principio de funcionamiento es el siguiente: un router no intercambia distancias con sus vecinos, en lugar de eso, comprueba activamente el estado de sus enlaces con cada uno de sus routers vecinos y envía esta información a dichos vecinos, los cuales, a su vez, propagarán esta información por toda la red. Entre esa información propagada se pasa el coste de sus enlaces. Este valor es configurable y es en base a retardo, tasa de transferencia, coste monetario de la línea u otros factores. En base a la información recibida, cada router mantiene una base de datos que refleja la topología conocida del sistema del cual forma parte. Esa topología solo varía en función de los cambios en los enlaces.

Los routers se intercambian mensajes de Hello por todos sus interfaces (flooding) para descubrir a los routers vecinos, se identifican mediante un router ID que es configurable. Este mensaje de saludo contiene una lista de todos los identificadores de vecinos que este router ha recibido mediante otros mensajes Hello. De esta forma, todos los routers se tienen información sobre quiénes son sus vecinos y los vecinos de sus vecinos. De esta forma además cada router sabe si sus mensajes Hello han sido procesados por el resto de routers.

La topología construida tiene los siguientes elementos:

1.- Vértices o nodos, de dos tipos:

- Router

- Red, que a su vez puede ser de dos tipos

  - Red de tránsito: los datos ni se originan ni se consumen en los extremos de esta red

  - Red final: no es una red de tránsito.

2.- Aristas, de dos tipos:

- Aristas que conectan dos vértices routers cuando estos routers están conectados entre sí por un enlace punto a punto

- Aristas que conectan un vértice router a un vértice red cuando el router está conectado directamente a esa red.

La siguiente ilustración es un ejemplo de red basada en los conceptos OSPF:

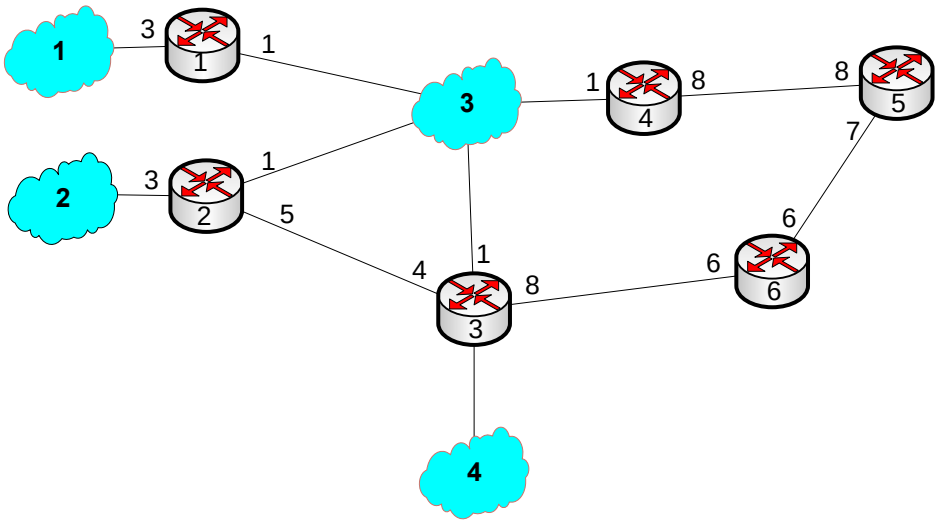

Ilustración 27- Gráfico de una red

El gráfico que se construirá siguiendo las reglas de OSPF es como se muestra en la siguiente ilustración:

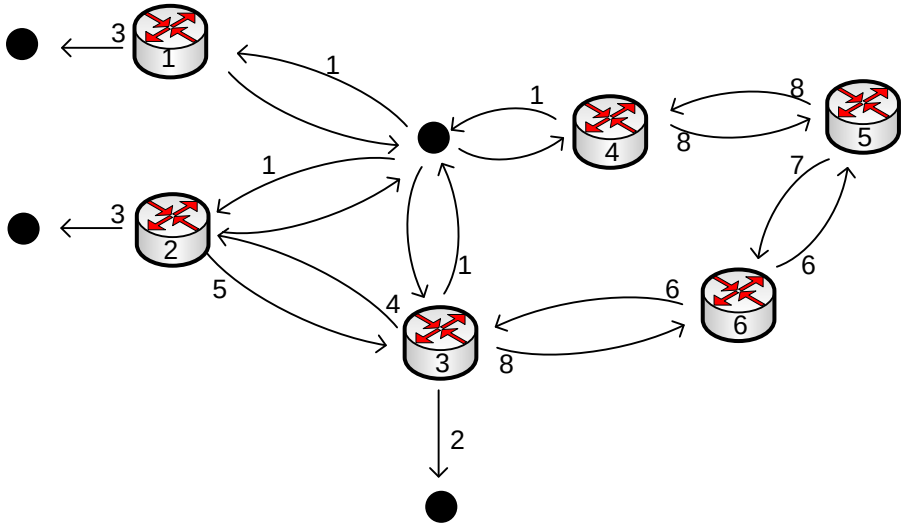

Ilustración 28- Gráfico de una red con la topología de OSPF

Estas reglas son:

- Dos routers interconectados mediante una línea punto a punto se representan mediante 2 aristas, una en cada dirección
- Cuando múltiples routers están interconectados a través de una red LAN, se representa la red como un vértice y todos los routers conectados bidireccionalmente a ese router.
- Si un router está conectado únicamente a esa red, la unión se representa en una única dirección

- Se representa un coste asociado unívocamente a cada arista. Si una arista no tiene coste etiquetado, este es 0. Esto ocurre en las aristas con origen un vértice de red.

En cada router se mantiene una base de datos con esta información. A partir de estos gráficos, cada router ejecuta un algoritmo, el algoritmo de Dijkstra para calcular el camino de menos coste para todas las redes de destino.

En la siguiente ilustración vemos como quedaría el árbol formado por el router 6 para alcanzar las redes destino. Solo existe un camino para alcanzar cada red.

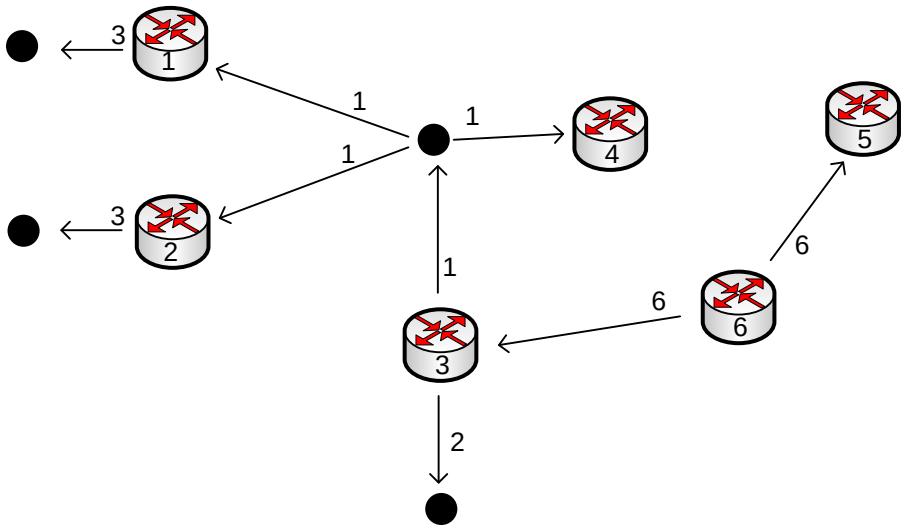

Ilustración 29- Gráfico de una red con la topología de OSPF

Una vez llegado a esta situación, los routers intercambian mensajes de estado de enlace, que se denominan LSAs, cuando se producen cambios en algún enlace, así, cada router podrá recalcular su árbol ejecutando de nuevo el algoritmo de Dijkstra.

2.3.3.1 Áreas OSPF

Los protocolos de Estado de Enlaces requieren una estructura de red jerárquica, y esto es respetado por OSPF.

En redes grandes el número de enlaces es alto y por tanto el número de potenciales caminos a un destino es elevado. En ese caso los cálculos del algoritmo de Dijkstra comparando todas las posibles rutas pueden ser muy complejos y tomar un tiempo significativo.

Los protocolos de Estado de Enlaces normalmente reducen los cálculos del algoritmo de Dijkstra dividiendo la red en áreas. El número de routers en un área y el número de LSAs que se inundan sólo dentro de ella es pequeño, lo que significa que la base de datos de topología o de estado de enlaces de un área es reducida. Por lo tanto, el cálculo del algoritmo de Dijkstra es más sencillo y requiere un tiempo menor.

Los protocolos de Estado de Enlaces, y entre ellos OSPF, usan una jerarquía con dos capas de áreas de la siguiente manera:

- Áreas de tránsito: son áreas OSPF cuya principal función es el transporte rápido y eficiente de paquetes IP. Estas áreas de tránsito interconectan otros tipos de áreas OSPF. Por lo general los usuarios finales no se encuentran en estas áreas. El área 0 en OSPF, también conocida como backbone, es por definición un área de tránsito.

- Áreas regulares: son áreas OSPF cuya principal función es conectar usuarios y distintos recursos. Estas áreas habitualmente se agrupan de manera funcional o geográfica. Por defecto un área regular no permite que el tráfico de otra área use sus enlaces para alcanzar otras áreas: todo el tráfico de otras áreas debe atravesar áreas de tránsito, como el área 0. Las áreas de tránsito o áreas no backbone, puede ser a su vez de diversos tipos: área "stub", área "totally stub" o área "not-so-stubby" (NSSA)

- Áreas stub: no aceptan información de rutas externas al sistema autónomo. Si un router necesita rutas a redes externas al sistema autónomo, emplea la ruta por defecto. Las áreas stub no pueden contener ASBRs y suelen tener un solo ABR.

- Áreas totally stub: no aceptan rutas externas al sistema al sistema autónomo ni rutas sumarizadas de otras áreas del sistema autónomo. Si un router necesita enviar un paquete a una red externa al área, emplea la ruta por defecto. Las áreas totally stub no pueden contener ASBRs.

- Áreas not-so-stubby (NSSA): estas áreas ofrecen los beneficios de las áreas stub o totally stub (minimizar la información de routing) pero a diferencia de éstas pueden contener ASBRs. Emplean LSAs especiales, de tipo 7. Éstos son generados por el ASBR y el ABR los convierte en LSA de tipo 5 para que se propague por el sistema autónomo.

  El empleo de áreas en OSPF permite:

  - minimizar las entradas en la tabla de rutas
  - localizar el impacto de un cambio de topología dentro de un área
  - detener la inundación de LSAs dentro de un área
  - requiere un diseño jerárquico de la red

El empleo de áreas supone un compromiso entre tener información de toda la red y almacenar toda la información en cada uno de los routers, lo cual no es escalable en redes muy grandes. Los routers de un área mantienen información detallada de los enlaces de su área y solamente información general o sumarizada de los routers y enlaces de otras áreas.

Cuando hay un cambio de topología la información se inunda solo a los routers del área. Manteniendo una estructura jerárquica y limitando el número de routers en un área, un sistema OSPF autónomo puede escalar a tamaños muy grandes.

Una estructura jerárquica implica que todas las áreas deben estar conectadas al área 0, y por tanto todo el tráfico interarea debe atravesar el área 0. Además el número óptimo de routers por área se recomienda que no sea mayor de 50-100 routers.

2.3.3.2 Routers en OSPF

Los routers del área 0 se denominan routers backbone. Los routers borde de área (ABR o Area Border Router) conectan el área 0 con otras áreas y tienen las siguientes funciones:

- separar las zonas de inundación de LSAs
- ser el punto primario para la sumarización de direcciones de un área
- funcionar habitualmente como fuente de rutas por defecto
- mantener la base de datos de estado de enlaces de cada área a las que pertenece

El diseño ideal supone que cada ABR esté conectado únicamente a dos áreas: el área backbone y otra área.

Los routers en OSPF se clasifican en cuatro tipos:

- Internal Routers: son routers cuyas interfaces pertenecen todas a la misma área y por tanto tienen idénticas LSDB
- Backbone Routers: son routers en el perímetro del área de backbone que tienen al menos una interfaz conectada al área 0. Estos routers mantienen la información de routing usando los mismos procedimientos y algoritmos que los Internal Routers
- ABRs: son routers cuyas interfaces pertenecen a diferentes áreas, manteniendo por tanto LSDB separadas para cada área a la que se encuentran conectados. Estos equipos enrutan el tráfico destinado o procedente de otras áreas. Los ABRs son puntos de salida del área, por lo que la información de routing desinada a otras áreas sólo puede ser transmitida a través del ABR del área local.
- Los ABRs pueden ser configurados para sumarizar la información de routing de las LSDBs de las áreas asociadas. Además los ABRs distribuyen la información de routing hacia el backbone y a su vez los routers de backbone envían la información a otros ABRs.
- ASBRs: son routers con al menos una interfaz en una red externa de interconexión, es decir, de otro sistema autónomo, que pertenece a una red no OSPF. Los ASBRs pueden importar información de redes no OSPF a la red OSPF y viceversa (redistribución de rutas)

Un router puede ser de más de un tipo de los indicados anteriormente.

Un router tiene LSDBs independientes para cada área a la que está conectado y dos routers de una misma área mantienen idénticas LSDBs para esa área.

2.3.3.3 Adyacencias OSPF

Los routers descubren a sus vecinos mediante el intercambio de paquetes de hello y declaran a sus vecinos como operativos tras comprobar ciertos parámetros u opciones en estos paquetes.

En los enlaces punto a punto los vecinos se declaran completamente adyacentes (fully adjacent). En los enlaces LAN los vecinos establecen adyacencia con los router DR y DBR y mantienen estado bidireccional con los otros routers (DROTHERS).

Una vez declaran a sus vecinos operativos, los routers sincronizan sus bases de datos de estado de enlaces (LSDB) intercambiando LSAs y hecho esto declaran establecida la adyacencia.

Las actualizaciones de routing y la información de topología se intercambian sólo entre routers adyacentes.

En enlaces LAN, se elige uno de los routers como “designated router” o DR y otro como “backup designated router” o BDR. Todos los routers de la LAN forman adyacencia completa con estos dos routers y les envían LSAs únicamente a ellos. El DR reenvía las actualizaciones de un vecino a todos los demás vecinos de esa LAN.

Una de las principales funciones del router DR es conseguir que todos los routers de la misma LAN tengan bases de datos idénticas. Además envía esta base de datos a los nuevos vecinos que establezcan adyacencia. De esta manera se intercambia la información necesaria de una manera eficiente.

2.3.3.4 Cálculo de rutas en OSPF

Los routers determinan la mejor ruta para cada destino aplicando el algoritmo SPF de Dijkstra sobre la base de datos de Estado de Enlaces de la siguiente manera:

- cada router dentro de un área tiene idéntica base de datos de Estado de Enlaces
- Cada router se coloca a sí mismo en la raíz del árbol que construye
- El mejor camino se calcula según el menor coste del total de los enlaces a cada destino específico.
- Las mejores rutas se incluyen en la base de datos de forwarding

Para OSPF el comportamiento por defecto es calcular el coste de una interfaz según el ancho de banda de la misma aunque este coste también puede ser definido manualmente en cada interfaz.

2.3.3.5 Tipos de LSAs

Los LSAs se usan para formar la base de datos de estado de enlaces (LSDB) y en conjunto sirven para describir la topología de toda la red o área OSPF.

Los distintos tipos de LSAs se describen a continuación.

- LSA tipo 1: Router Link Advertisements.

Cada router genera estos LSAs para cada área a la que pertenece. Estos LSAs describen el estado de los enlaces del router: incluye una lista de los enlaces directamente conectados, identificados por su prefijo IP. Se envían únicamente a un área particular y no atraviesa ABRs (es decir no salen de la propia área). Uno de los campos que se indican en la cabecera de estos paquetes es el identificador del router que lo genera (link-state ID).

- LSA tipo 2: Network Link Advertisements.

Los routers DR generan estos LSAs para redes multiacceso, los cuales describen el conjunto de routers asociados para una red particular. Se genera un LSA de este tipo para cada red de tránsito dentro de un área: una red de tránsito es aquella que tiene al menos dos routers OSPF directamente conectados. Incluyen una lista de los routers conectados a la red de tránsito, así como la dirección IP de dicha red. Se inundan al área que contiene esa red, no atraviesan ABRs, y el identificador que se especifica es la dirección IP de la interfaz del router DR que lo genera.

- LSA tipo 3 y 4: Summary Link Advertisements.

Son generados por routers ABR y describen las siguientes rutas interáreas:

  - Tipo 3: describen rutas de redes y de redes agregadas. Se emplean para enviar información de un área fuera del área que la origina (interárea): especifica todas las redes de esa área. Son enviados por el ABR del área que lo genera y son inundados a través de todo el sistema autónomo, atravesando las distintas áreas. Por defecto OSPF no sumariza las distintas subredes, pero esto es configurable. Por defecto los LSAs de tipo 3 se anuncian al área backbone para cada subred definida, lo cual puede causar problemas de inundación, por lo que es recomendable la configuración de la sumarización en los ABRs.

  - Tipo 4: describen rutas a ASBRs. Se emplean para anunciar un ASBR, y la ruta hasta él, a todas las demás áreas del sistema autónomo. Los genera el ABR del área origen y se inundan por todo el sistema autónomo: son regenerados por los ABRs que atraviesa. Contiene únicamente el router ID del ASBR.

Estos LSAs se inundan a través del área backbone a otros ABRs. No se envían en cambio a áreas del tipo “totally stubby”

- LSA tipo 5: Autonomous System External Link Advertisements.

Son generados por los ASBRs y describen rutas hacia destinos externos al sistema autónomo. Son generados por el ASBR e inundados a todo el sistema autónomo, a cualquier área excepto a las áreas de tipo “stub”. Incluyen el router ID del ASBR y este permanece inalterado a través de todo el sistema autónomo. Los LSAs de tipo 4 son necesarios para alcanzar el ASBR.

Puesto que se inundan a todo el sistema, la no sumarización de estas rutas puede provocar problemas de inundación, por lo que no es recomendable.

- LSA tipo 6: Multicast OSPF LSAs.

  Son LSAs especializados que se emplean para aplicaciones multicast.

- LSA tipo 7: Defined for Not-So-Stubby Areas.

  Son LSAs usados en areas NSSA (not-so-stubby area)

- LSA tipo 8: External Attributes LSA for BGP.

  Son LSAs especializados usados para la interconexión de redes OSPF y BGP

- LSA tipo 9, 10 y 11: Opaque LSAs.

Son LSAs destinados a futuros desarrollos de OSPF. Cisco los emplea por ejemplo en redes MPLS con OSPF.

2.3.4 ISIS

En los últimos años la popularidad del protocolo IS-IS (Intermediate System-to-Intermediate System) ha aumentado considerablemente, al extenderse su uso entre los proveedores de servicio. La simplicidad y estabilidad de IS-IS lo hace robusto en redes de gran tamaño. IS-IS se encuentra en muchos ISPs y en muchas redes que soporten protocolos OSI. Se trata de un protocolo de estado de enlaces, el cual permite una rápida convergencia y gran escalabilidad. Es también un protocolo muy flexible, que ha sido extendido para soportar funcionalidades como ingeniería de tráfico en MPLS,

Entre las características de de IS-IS destacan:

- Routing jerárquico
- Comportamiento classless
- Rápida transmisión de la nueva información
- Rápida convergencia
- Alta escalabilidad
- Configuración flexible de los temporizadores
- Implementación de routing multiárea de Cisco
- Implementación de route-leaking de Cisco
- Implementación del bit de overload de Cisco

IS-IS es un protocolo de routing dinámico de intradominio de OSI, especificado por la ISO 10589. El protocolo está diseñado para funcionar en CLNS (Connectionless Network Service) de OSI. Los datos son transportados siguiendo el protocolo especificado en la ISO 8473,

Para soportar dominios de routing grandes, se emplea una jerarquía de dos niveles. Un dominio grande se divide administrativamente en áreas y cada sistema reside en una de estas áreas. El routing dentro del área se denomina routing de Nivel 1 mientras que el routing entre áreas se denomina routing de Nivel 2. Un Sistema Intermedio (IS) de Nivel 2 mantiene información de los caminos hasta las áreas de destino. Un IS de Nivel 1 mantiene información de routing dentro de su propia área. Para un paquete cuyo destino es otra área, un IS de Nivel 1 envía dicho paquete al IS de Nivel 2 más cercano dentro de su propia área, sea cual sea el área de destino. El paquete a continuación es enviado mediante routing de Nivel 2 al área de destino, dentro de la cual empleará el routing de Nivel 1 para llegar a su destino. Cabe destacar que la selección de un camino de salida del área basada en el routing de Nivel 1 hasta el IS de Nivel 2 más cercano puede resultar en un enrutamiento no óptimo de los paquetes (routing subóptimo).

Existe una extensión propuesta en la RFC 2966 que permite que la información de routing disponible en el Nivel 2 esté disponible también para el routing de Nivel 1. Este mecanismo soluciona el problema del routing subóptimo a costa de la escalabilidad en entornos de IS-IS integrados. Sin embargo no es aplicable a entornos CLNS.

En medios multiacceso de broadcast (LAN), se elige un Sistema Intermedio Designado (DIS) el cual dirigirá la inundación de información en el medio. Este sistema DIS es análogo al router designado (DR) en OSPF, aunque su funcionamiento, incluyendo el proceso de elección y las adyacencias son considerablemente distintos. El DIS es elegido por prioridad: será aquel con la prioridad más alta. Esta prioridad se configura por interfaz. En caso de empate, el router con MAC más alta será elegido DIS.

2.3.4.1 CLNS

CLNS de OSI es un servicio del nivel de red similar a IP. Una entidad CLNS se comunica con otra mediante el protocolo CLNP (Connectionless Network Protocol).

En la arquitectura OSI se definen sistemas: los routers son Sistemas Intermedios (IS) y los hosts son Sistemas Finales o End Systems (ES). Los ES no tienen información de routing, sino que descubren los routers o IS escuchando los hellos de los sistemas intermedios (ISH) y enviando el tráfico a cualquier router al azar. A su vez los ES envían paquetes de hello (ESH). No eligen los routers DIS para enviarles el tráfico, por lo que el routing óptimo se consigue mediante la redirección del tráfico.

Los IS descubren los ES escuchando sus hellos (ESH) y los IS envían hellos a su vez a los ES.

En estos entornos no existe ARP ni ICMP ni protocolos de routing interdominio (IDRP), pero el protocolo ES-IS (End System to Intermediate System) proporciona las mismas funciones para los IS y ES. El protocolo ES-IS se define en la ISO 9542.

IS-IS es un protocolo IGP para el enrutamiento en OSI. Los paquetes de IS-IS no se encapsulan en CLNS o IP, sino que se encapsulan directamente sobre el protocolo de la capa de enlace.

2.3.4.2 IS-IS Integrado o Dual

IS-IS puede emplearse como IGP para soportar IP además de OSI. Esto permite emplear un solo protocolo de routing en entornos puramente OSI, puramente IP y entornos duales. IS-IS integrado se extendido en entornos puramente IP como protocolo empleado en las redes de proveedores de Internet (ISP). Se define en la RFC 1195.

2.3.4.3 Funcionamiento de IS-IS

De forma general, el funcionamiento de IS-IS es el siguiente:

- Los routers sobre los que corre IS-IS envían paquetes de hello por todas las interfaces donde está habilitado IS-IS para descubrir los vecinos y establecer adyacencias.
- Los routers que comparten un enlace se convierten en vecinos IS-IS si sus paquetes de hello contienen información que cumple los criterios para formar adyacencia. Estos criterios se diferencian ligeramente dependiendo del tipo de medio usado (punto a punto o broadcast). Los criterios principales son la coincidencia en la autenticación, tipo de IS y tamaño de MTU
- Los routers construirán un paquete LSP basándose en las interfaces locales en las que se haya configurado IS-IS y en los prefijos aprendidos de otros routers adyacentes.
- Los routers generalmente envían LSPs a todos los vecinos adyacentes excepto al vecino del que se ha recibido el mismo LSP. Existen varias maneras de inundar esta información y diversos escenarios posibles.
- Los routers construyen la base de datos de estado de los enlaces a partir de los LSPs recibidos.
- Cada IS calcula un árbol de camino más corto (SPT) y a partir de este árbol construye la tabla de rutas.

En IS-IS los routers establecen adyacencias con otros routers con lo que tienen enlaces punto a punto, mientras que en entornos LAN los routers las establecen con los DIS designados. El DIS genera para ello un LSP adicional, llamado LSP de pseudonodo. El DIS es el responsable de controlar la inundación de información en la LAN y de mantener la sincronización.

El routing en IS-IS consiste en la ejecución de cuatro procesos:

Recepción

Es el punto de entrada de toda la información. Este proceso pasa la información de usuario y repostes de errores al proceso de reenvío y la información de routing y los paquetes de control al proceso de actualización

Actualización

Genera la información de los enlaces locales que se enviará a los routers adyacentes y recibe, procesa y reenvía la información de enlaces recibida de los routers adyacentes. Controla las bases de datos de estado de enlaces de Nivel 1 y Nivel 2 e inunda los LSPs de Nivel 1 y Nivel 2 al área.

Cada LSP tiene un tiempo de vida restante, checksum y número de secuencia.

El tiempo de vida restante es un contador decreciente de 1200 segundos. El routers que origina el LSP debe refrescar el LSP periódicamente para evitar que llegue a cero. Si llegara a cero el LSP expirado se almacena otros 60 segundos en la base de datos antes de ser eliminado.

Si un LSP recibido tiene un cheksum incorrecto, se elimina y se inunda el LSP con un tiempo de vida de cero. Esto hace que el originador del LSP lo envíe de nuevo. Esto es diferente de OSPF, donde sólo el originador puede eliminar el LSP.

Decisión

Este proceso ejecuta el algoritmo SPF sobre la base de datos de estado de enlaces y crea la base de datos para el reenvío. Procesa la información de siguiente salto y los caminos de mismo coste, creando un conjunto de adyacencias que se usarán para el balanceo de carga.

Reenvío

Obtiene la información del proceso de recepción y emplea la base de datos de reenvío para reenviar los paquetes de datos hacia su destino. También redirige el balanceo de carga y genera los reportes de error.

2.3.4.4 Áreas y dominios de routing

Un dominio de routing en IS-IS es similar a un sistema autónomo de BGP: es un conjunto de áreas bajo una misma administración que implementa políticas de routing en dicho dominio.

Área de backbone

IS-IS no tiene un área de backbone como el Área 0 de OSPF. El backbone de IS-IS es un conjunto contiguo de routers con capacidad de Nivel 2 cada uno de los cuales puede pertenecer a un área diferente.

Áreas

En IS-IS un router pertenece únicamente a un área y la frontera entre áreas es un enlace que conecta routers que pertenecen a distintas áreas. Esto se diferencia de OSPF, donde existen routers de borde de área que pertenecer a más de un área. La razón es que en IS-IS un router generalmente tiene únicamente una dirección NSAP (Network Service Access Point) mientras que un router IP tiene típicamente múltiples direcciones IP. La dirección NSAP es la dirección de la capa de enlace en los paquetes de CLNS.

IS-IS tiene una jerarquía de dos niveles. El backbone lo forman los routers de Nivel 2 contiguos. Los routers de Nivel 1 de Nivel 2 pertenecen a las distintas áreas. Los routers pueden ser de Nivel 1, de Nivel 2 o de ambos niveles. En Cisco la configuración por defecto es de ambos niveles a la vez. Los routers de Nivel 2 conectan todas las áreas del dominio de routing. Los routers de Nivel 2 anuncian su dirección NSAP a los demás routers de Nivel 2 del backbone. Todos los routers de Nivel 1 y hosts en un área deben tener una NSAP con la misma dirección de área.

Router de Nivel 1

Este router conoce únicamente la topología de su propia área, en la cual tiene vecinos de nivel 1 o de Nivel 1 y 2. Mantiene un base de estado de los enlaces de Nivel 1 con toda la información de routing intra área. Usa el router de Nivel 2 más próximo de su área para enviarle los paquetes cuyo destino está fuera del área, lo cual puede producir routing sub-óptimo.

Para informar a un router de Nivel 1 de cuál es el router de Nivel 2 más próximo, se sigue el siguiente mecanismo: el router de Nivel 1 y 2 con conexión a otra área, envía un LSP de nivel 1 con el bit “attached bit”. Todos los IS de nivel 1 de esa área reciben ese LSP y sabrán donde deben enviar los paquetes destinados fuera del área. Si los routers emplean IS-IS integrado, la ruta por defecto se instalará automáticamente en los routers de Nivel 1 apuntando al router de Nivel 1 y 2 más próximo, el cual originó el LSP con el “attached bit”. Un router de Nivel 1 y 2 no conectado a otra área puede también detectar que un vecino de Nivel 2 únicamente está conectado a otra área y generar el LSP con “attached bit” en nombre de su vecino de Nivel 2.

Si existe más de un punto para salir del área, el router más cercano se determina en función del coste. Si hay dos caminos de igual coste, se hará balanceo de carga.

Router de Nivel 2

Un router de nivel 2 puede tener vecinos en la misma área o en otras diferentes. Almacena información de routing inter área en su base de datos de estado de enlaces de Nivel 2. El router de nivel 2 tiene información de otras áreas pero no tiene información de Nivel 1 de su propia área. En OSI, un router debe conocer la topología de su propia área, por lo que el router de Nivel 2 no debe configurarse cuando se enruta tráfico OSI. Si el tráfico en el área es IP exclusivamente, todos los routers pueden configurarse como Nivel 2.

Router de Nivel 1 y 2

El router de Nivel 1 y 2 puede tener vecinos de cualquier área. Tiene dos bases de datos de enlaces: la de Nivel 1 para el routing intra área y la de Nivel 2 para el routing inter área. Ejecuta dos SPFs por lo que se requiere mayor memoria y capacidad de procesamiento.

En entornos de IS-IS integrado, un router de Nivel 1 y 2 traspasa información de subredes de Nivel 1 al Nivel 2. Estas subredes pueden además ser sumarizadas. En Cisco, por defecto la configuración de los routers es de Nivel 1 y 2.

En la siguiente ilustración vemos como se pueden separar las áreas y la función de cada router.

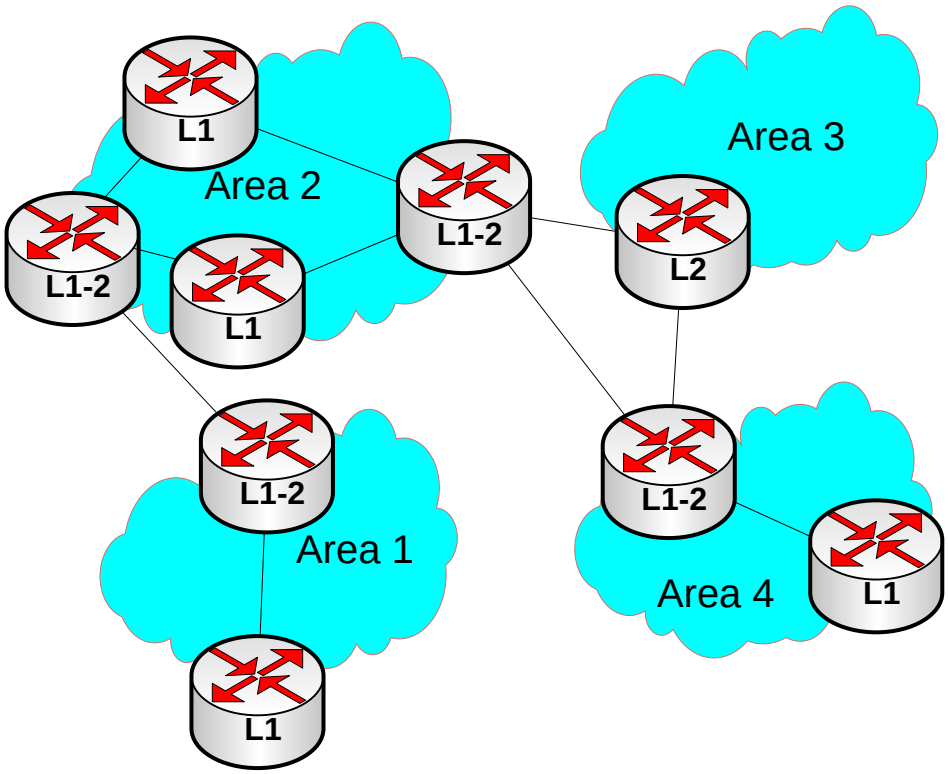

Ilustración 30- Esquema funcional de IS-IS y sus componentes

3 HERRAMIENTAS

Para el desarrollo de este proyecto han sido dos las herramientas con usos concretos que se describen a continuación:

3.1 SECURE CRT -VERSION 5.1.2 (BUILD 274)

Esta herramienta es una aplicación que gestiona las conexiones remotas con los equipos, ofrece posibilidad de utilizar varios protocolos con sus extensiones como son Telnet, SSH y RLogin. La totalidad de los equipos implicados en la realización de este proyecto son accesibles mediante una conexión estándar de telnet al puerto 23.

Una de las ventajas que ofrece esta aplicación es la gestión de sesiones contra los equipos que hay en la red, no es necesario insertar la dirección IP del equipo para poder acceder a él, se guardan accesos directos con las sesiones adecuadamente configuradas para poder acceder.

También se puede configurar scripts que facilitan la ejecución de las tareas, el único que tenemos corriendo es el de login y usuario, que mediante una combinación de teclas introduce login y password en la consola automáticamente.

También ofrece un buen sistema de logs ya que absolutamente todas las sesiones iniciadas por cada usuario quedan registradas en formato de texto. Esto es bastante útil tanto en la elaboración del proyecto ya que nos permite poder ejecutar los comandos sin tener que copiar al instante la información relevante como a la hora de resolver incidencias.

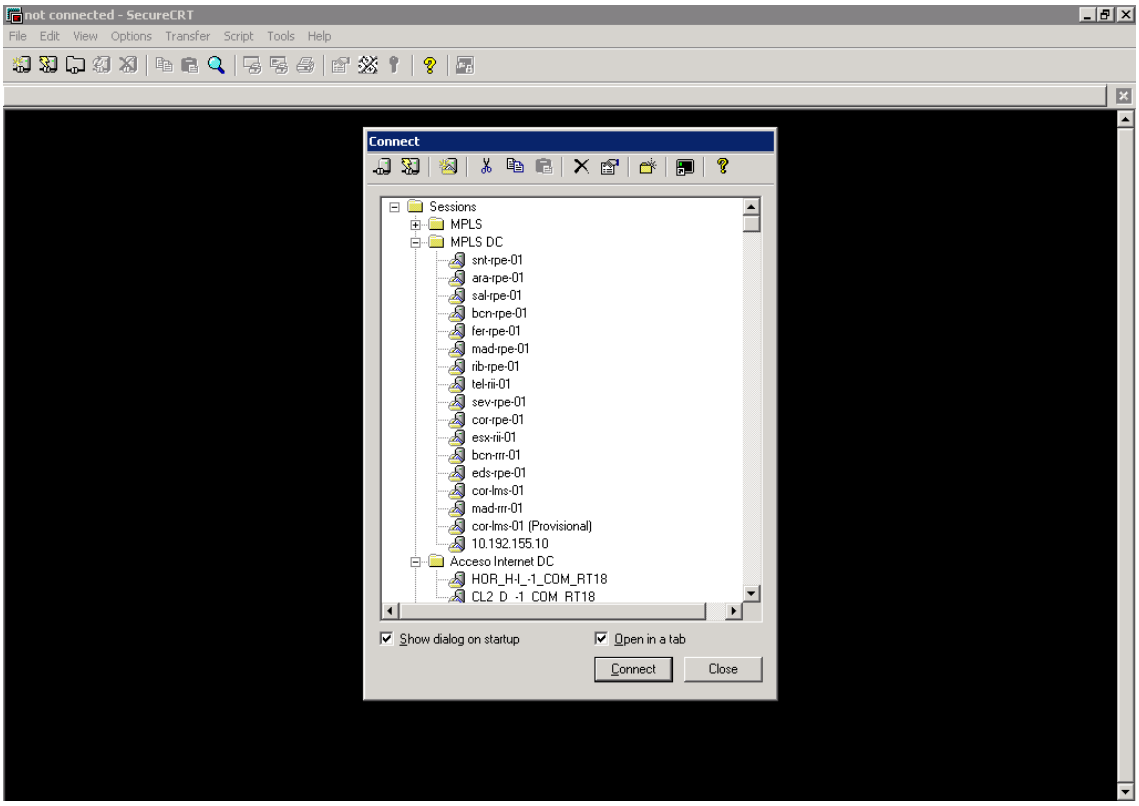

Ilustración 31- Apariencia de Secure CRT

3.2 MICROSOFT VISIO 2003

Para la elaboración de planos de red se ha utilizado esta aplicación. Si bien no es una herramienta de gestión (otras si ofrecen estas capacidades) con ella se puede plasmar con bastante nivel de detalle lo que es un escenario de red.

4 DESARROLLO.

4.1 SITUACIÓN INICIAL

En un principio, se parte de dos redes MPLS totalmente aisladas entre sí e independientes tanto en concepción como en desarrollo de servicios a través de ella. Ambas redes poseen, por ejemplo, distintas concepciones de calidad de servicio, en una se utiliza IP precedence y en la otra DSCP + IP precedence.

También hay diferencias en cuanto a la filosofía de uso. En la red de UFIN, como norma general, el CPE debe de ser propiedad y mantenimiento de UFIN. El primer equipo no gestionado por UFIN debe de ser el C y nunca el CPE. Esto repercute directamente en la concepción de la calidad de servicio en la red MPLS puesto que el tráfico es marcado con un valor de IP Precedence en el CPE y la red MPLS no tiene que volver a remarcar y confía en lo entregado. En el caso de la red de DC no es así, siempre se hace un remarcado de tráfico puesto que el CPE puede ser gestionado por el cliente y por tanto el tráfico no es confiable.

En cuanto a los servicios, en ambas redes el 95% del tráfico es tráfico de VPN MPLS por lo que son redes de filosofía parecida. En el caso de la red de UFIN, además, se dan servicios de nivel 2, concretamente VPLS y EoMPLS. Por ello, veremos un apartado en el que se extiende la capacidad de dar servicios de nivel 2 a la red de DC.

En el siguiente punto, veremos el hardware que componen ambas redes.

4.1.1 HARDWARE

4.1.1.1 Familia Cisco C7600

La familia de Routers C7600 (también denominado Cat6500) está basada en la integración evolucionada de la arquitectura de la Familia de Routers C7500 y la Familia de Switches Cat6500 con las siguientes características:

- Integración de Arquitectura de N2 y N3 en una Plataforma única.
- WAN Port Adapters Familia C7500.
- LAN Modules Familia Cat6500.
- OSM Line Cards.
- E-FlexWAN Carrier Card.

Los Routers C7613 son Chasis de 13 slots con 2 Fuentes de alimentación 4000 W DC con unas dimensiones (Alto 82.3 cm, Ancho 42.5 cm, Profundo 44.7 cm) y 40.82 Kg de peso, ocupando 19 RU en rack. Con las dos fuentes en funcionamiento se proporciona redundancia y balanceo de carga ya que cada una de las fuentes suministra la mitad de la energía necesaria y sólo en caso de avería de una de ellas se asume el suministro total por parte de la otra, es decir cada fuente es capaz de asumir el consumo total del equipo.

Cada una de las fuentes 4000W-DC puede funcionar en un modo de funcionamiento 2700W ó 4000W en función de las conexiones de alimentación que tengan. Se instalarán todos los equipos con sus fuentes en modo de funcionamiento 4000W. El consumo de cada equipo depende del tipo y el número de tarjetas que tenga instaladas.

En la siguiente figura correspondiente a un equipo Cisco 7613 se pueden ver los slots distribuidos de forma horizontal, que se empiezan a numerar de arriba hacia abajo, las fuentes de alimentación en la parte inferior del equipo y los ventiladores en el lateral izquierdo.

Ilustración 32- Chasis de routers Cisco C7613, C7609 y C7606

Ilustración 33- Fuente de alimentación DC

En relación a la situación de los slots en cada chasis para introducir las tarjetas, por criterios de ventilación se puede ver en la siguiente ilustración que el flujo de aire atraviesa por igual a todos los slots, por lo que no hay una recomendación, por lo que salvo la obligatoriedad de poner en los slots 7 y 8 las procesadoras, el resto de los slots se han puesto empezando de arriba abajo (Empezando por el 1). Sin embargo, la arquitectura del C7600 y de su procesadora también hay que considerarla de cara a esa decisión.

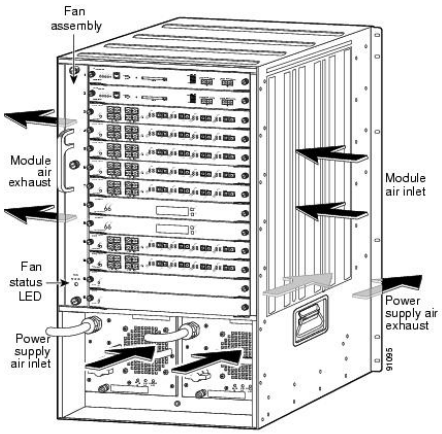

Ilustración 34- Ventilación del chasis C7613

Cada router tiene una doble procesadora, para estar protegido contra fallo de la procesadora activa. Hay varios modos de configuración de esa redundancia, que se estudiarán más adelante.

4.1.1.1.1 Supervisora SUP720-3B

Esta supervisora (incluye PFC3B y MSFC3) proporciona un throughput total de 720 Gbps (Switch Fabric de 720 Gbps Full-duplex) distribuido entre los 13 Slots mas 1 Bus compartido de 32 Gbps Full-duplex.

Slots 1 – 8 $\rightarrow$ 1 channel de 40 Gbps Full-duplex.

Slots 9 – 13 $\rightarrow$ 2 channels de 40 Gbps Full-duplex.

Ilustración 35- Supervisora SUP720-3B

El throughput total es de 720 Gbps Full-duplex:

[(8 x 1 x 40 Gbps + 5 x 2 x 40 Gbps) = 720 Gbps.

Los Slots 7 y 8 están reservados para la SUP720-3B, quedando el resto disponibles para las Line Cards.

Dentro de la propia Supervisora encontramos dos tarjeras hijas (MSFC3 y PFC3B) que proporcionan las funciones de Control de Switching y Routing por una parte y funcionalidades específicas implementadas en hardware (Routing IP, MPLS, L3VPNS, EoMPLS, etc.) sobre chips ASICs en la PFC3B.

Se diferencia entre Funciones de Control (MSFC3) y Funciones de Forwarding (PFC3B). La tarjeta MSFC3 está equipada con dos Procesadores, dos memorias SDRAM (512 MB) y otras dos BootFlash (64 MB) independientes, que son Switch Processor (SP) y Route Processor (RP), de modo que los Protocolos de Nivel 2 se ejecutan en la SP y los de Nivel 3 en la RP; así mismo la MSFC3 construye por software la tabla de forwarding (FIB) para transferírsela a los ASICs de la PFC3B quien realiza las funciones de forwarding por hardware, entre otras funciones.

La tarjeta PFC3B dispone de 512 MB de Memoria SDRAM. Soporta forwarding de paquetes IP en modo hardware hasta un máximo de 30 Mpps, y para MPLS de hasta 20 Mpps de forma centralizada.

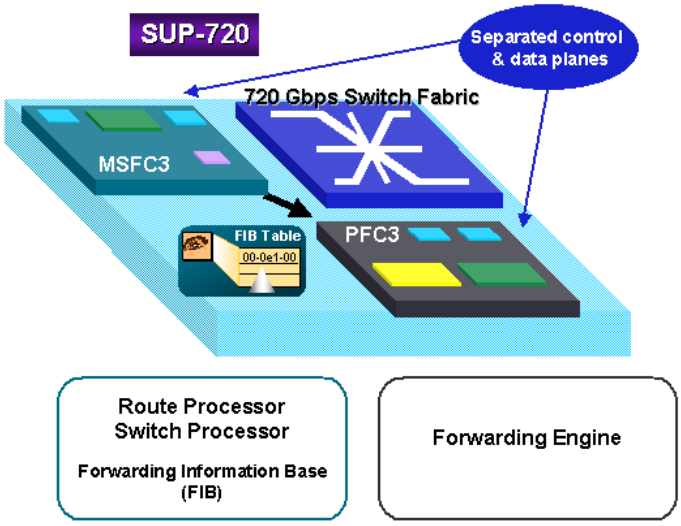

Ilustración 36- Supervisora SUP720-3B: CONTROL + DATOS

Memoria hardware de Procesadora SUP-720:

- Memoria MSFC3 SP: 512 MB SDRAM.
- Memoria MSFC3B RP: 512 MB SDRAM.
- Memoria PFC3B : 512 MB SDRAM.

Estas Memorias pueden ser actualizadas hasta un máximo de 1 GB.

En la siguiente captura observamos que tanto en el Slot 7 como en el 8 tenemos insertada la Procesadora con su MSFC y PFC:

```
MPLS-MS#sh inventory
NAME: "CISCO7613", DESCR: "Cisco Systems Cisco 7600 13-slot Chassis System"
PID: CISCO7613          , VID:    , SN: SAL094236JD

NAME: "WS-C6K-VTT 1", DESCR: "VTT FRU 1"
PID: WS-C6K-VTT         , VID:    , SN: SMT0935A499

NAME: "WS-C6K-VTT 2", DESCR: "VTT FRU 2"
PID: WS-C6K-VTT         , VID:    , SN: SMT0935A761

NAME: "WS-C6K-VTT 3", DESCR: "VTT FRU 3"
PID: WS-C6K-VTT         , VID:    , SN: SMT0935A086

NAME: "WS-C6513-CL 1", DESCR: "CXXXX Clock FRU 1"
PID: WS-C6513-CL        , VID:    , SN: SMT0934C398

NAME: "WS-C6513-CL 2", DESCR: "CXXXX Clock FRU 2"
PID: WS-C6513-CL        , VID:    , SN: SMT0934C556

NAME: "module 2", DESCR: "2 port adapter Enhanced FlexWAN Rev. 2.1"
PID: WS-X6582-2PA       , VID: V06, SN: JAE0944PJ9U

NAME: "module 2/0", DESCR: "8 Port Serial Adapter ( V35)"
PID: PA-8T-V35          , VID:    , SN: 33164219

NAME: "module 2/1", DESCR: "8 port, software configurable Multichannel T1/E1
without TDM Port Adapter"
PID: PA-MC-8TE1+        , VID:    , SN: JAE0938L6QR

NAME: "module 4", DESCR: "WS-X6148A-GE-TX 48-port 10/100/1000 RJ45 EtherModule
Rev. 1.0"
PID: WS-X6148A-GE-TX    , VID: V01, SN: SAD090705BM

NAME: "module 5", DESCR: "Enhanced OSM with 4 GE WAN ports and 2 GE LAN ports
Rev. 2.2"
PID: OSM-2+4GE-WAN+     , VID:    , SN: JAB0928018C

NAME: "module 6", DESCR: "Enhanced OSM with 4 GE WAN ports and 2 GE LAN ports
Rev. 2.2"
PID: OSM-2+4GE-WAN+     , VID:    , SN: JAE0934JLEM
```

```
NAME: "module 7", DESCR: "WS-SUP720-3B 2 ports Supervisor Engine 720 Rev. 4.4"
PID: WS-SUP720-3B      , VID:    , SN: SAL09433S5S

NAME: "msfc sub-module of 7", DESCR: "WS-SUP720 MSFC3 Daughterboard Rev. 2.3"
PID: WS-SUP720         , VID:    , SN: SAL09433RYF

NAME: "switching engine sub-module of 7", DESCR: "WS-F6K-PFC3B Policy Feature
Card 3 Rev. 2.1"
PID: WS-F6K-PFC3B      , VID:    , SN: SAL09433TJT

NAME: "module 8", DESCR: "WS-SUP720-3B 2 ports Supervisor Engine 720 Rev. 4.4"
PID: WS-SUP720-3B      , VID:    , SN: SAL09433SYB

NAME: "msfc sub-module of 8", DESCR: "WS-SUP720 MSFC3 Daughterboard Rev. 2.3"
PID: WS-SUP720         , VID:    , SN: SAL09433SQE

NAME: "switching engine sub-module of 8", DESCR: "WS-F6K-PFC3B Policy Feature
Card 3 Rev. 2.1"
PID: WS-F6K-PFC3B      , VID:    , SN: SAL09423GU9

NAME: "module 9", DESCR: "WS-X6148A-GE-TX 48-port 10/100/1000 RJ45 EtherModule
Rev. 1.7"
PID: WS-X6148A-GE-TX   , VID: V04, SN: SAL12405HT5

NAME: "module 10", DESCR: "WS-X6724-SFP CEF720 24 port 1000mb SFP Rev. 3.3"
PID: WS-X6724-SFP      , VID: V05, SN: SAL11456DYU

NAME: "switching engine sub-module of 10", DESCR: "WS-F6700-CFC Centralized
Forwarding Card Rev. 4.1"
PID: WS-F6700-CFC      , VID: V06, SN: SAL1250CW74

NAME: "PS 1 PWR-4000-DC", DESCR: "DC power supply, 4000 watt 1"
PID: PWR-4000-DC       , VID:    , SN: QCS0943203R

NAME: "PS 2 PWR-4000-DC", DESCR: "DC power supply, 4000 watt 2"
PID: PWR-4000-DC       , VID:    , SN: QCS0943203M
```

Cisco ha desarrollado determinadas funcionalidades software que con posterioridad ha implementado en modo hardware (Chips ASIC) consiguiendo así unos mejores niveles de ejecución (Performance). Básicamente existen tres tipos de ASIC cuya función varía según la funcionalidad implementada:

- Port ASIC: Funciones de Buffering, QoS.
- Forwarding ASIC: Nivel 2 y Nivel 3.
- Service ASIC: ACLS, Policing, packet replication (port-mirroring, multicast).

4.1.1.2 Tipos de accesos de la serie C7600

Los tipos de accesos que ofrecen estos equipos y que están utilizados en ambas redes son los siguientes:

- E1 canalizado y transparente.
- V.35 serial.
- STM1 canalizado.
- E3 serial.
- Ethernet/FastEthernet/GigabitEthernet.

Los distintos tipos de acceso se soportarán sobre las Line Cards y Port Adapters adquiridos, detallados a continuación:

Part Number	Product Description
OSM-4OC3-POS-SI+	4-Port STM-1 MT-RJ Single-Mode + 4 GE SC Line Card
OSM-2+4GE-WAN+	2 GE LAN + 4 GE WAN SC Line Card
WS-X6148A-GE-TX	48-Port 10/100/1000 RJ45 Classic Line Card
WS-X6582-2PA	Enhanced FlexWAN Line Card
PA-MC-8TE1+	8-Port E1 Channelized RJ48c Port Adapter
PA-MC-STM-1SMI	1-Port STM-1 Channelized SC Single-Mode Port Adapter
PA-E3	1-Port E3 Serial BNC Port Adapter
PA-8T-V35	8-Port E1 Serial V35 Port Adapter
	1-Port E3 ATM Port Adapter

Tabla 4. –Tipos de acceso en Line Cards y Port Adapters

El detalle de las tarjetas lo resumimos a continuación

4.1.1.2.1 4-Port STM-1 MT-RJ Single-Mode + 4 GE SC Line Card

Se trata de una Line Card OSM (Optical Services Module) que soporta servicios avanzados IP y MPLS basados en los “Network Processor” PXF (Parallel Express Forwarding) desarrollados por Cisco para tal fin.

Ocupa un Slot en el Chasis C7613, con unas dimensiones de 3 cm de alto x 35,6 cm de ancho x 40,6 cm de profundo y un peso de 5 Kg, soportando OIR (Online Insertion and

Removal) por lo que se puede insertar y extraer del Chasis C7613 en caliente sin que afecte al funcionamiento del mismo. La potencia requerida por esta Line Card es de 120W, con un consumo de 2,44A a 42V.

Dispone de acceso WAN y LAN simultáneo con 4 puertos STM-1 y otros tantos Gigabit Ethernet. El tipo de conector es MT-RJ para WAN y módulos GBIC tipo SC para LAN.

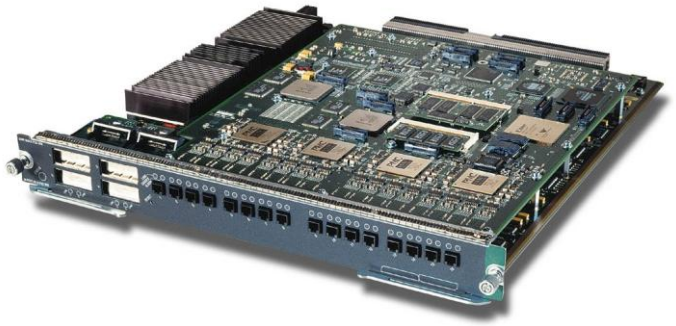

Ilustración 37- Tarjeta OSM-4OC3-POS-SI+

Esta ilustración se corresponde con una tarjeta análoga, pero con 16 puertos POS. La tarjeta real será igual, pero sólo con el grupo de cuatro puertos POS más cercano a los puertos GE, a la izquierda.

Los módulos GBIC cumplen con la norma IEEE 802.3z siendo las especificaciones ópticas las indicadas:

- 1000BASE-LX: 50 micron multimode fiber up to 550 m
- 1000BASE-LX: 9/10 micron single-mode fiber up to 5 km
- 1000BASE-LH: 62.5 micron multimode fiber up to 550 m
- 1000BASE-LH: 50 micron multimode fiber up to 550 m
- 1000BASE-LH: 9/10 micron single-mode fiber up to 10 km
- 1000BASE-ZX: 9/10 micron single-mode fiber up to 70 km
- 1000BASE-ZX: dispersion-shifted fiber up to 100 km

Ilustración 38- Módulos GBIC SC

4.1.1.2.2 2 GE LAN + 4 GE WAN SC Line Card

Esta Line Card OSM (Optical Services Module) soporta servicios avanzados IP, MPLS, AToM (EoMPLS, FRoMPLS,...) y VPLS basados en los "Network Processor" PXF (Parallel Express Forwarding) desarrollados por Cisco para tal fin.

Ocupa un Slot en el Chasis C7613, con unas dimensiones de 3 cm de alto x 35,6 cm de ancho x 40,6 cm de profundo y un peso de 4,7 Kg, soportando OIR (Online Insertion and Removal) por lo que se puede insertar y extraer del Chasis C7613 en caliente sin que afecte al funcionamiento del mismo. La potencia requerida por esta Line Card es de 359W, con un consumo de 4,35A a 42V.

Dispone de acceso WAN y LAN simultáneo con 4 puertos WAN Gigabit Ethernet y 2 LAN Gigabit Ethernet. El tipo de conector es GBIC tipo SC tanto para WAN como para LAN.

Ilustración 39- Line Card OSM 4-Port GE WAN

Los módulos GBIC utilizados son de las mismas características que en la tarjeta OSM-4OC3-POS-SI+

Soporta 802.1Q VLAN Trunking y hasta 32.000 MAC address por Puerto.

4.1.1.2.3 48-Port 10/100/1000 RJ45 Classic Line Card

Se trata de una line Card tipo Classic de 48 Puertos RJ-45 10/100/1000 Ethernet para conectar con cables UTP (Unshielded Twisted-Pair) Categoría 5 de distancia máxima 100 m.

Por ser tipo Classic la arquitectura de forwarding es CEF (Cisco Express Forwarding) Centralizado en PFC (Policy Feature Card) de Supervisor Engine 720, con conexión a Switch-Fabric a través del Bus compartido de 32 Gbps lo que supone un Throughput máximo de 32 Gbps y 15 Mpps.

Ilustración 40- Line Card 48-Port 10/100/1000

Ocupa un Slot en el Chasis C7613, con unas dimensiones de 3 cm de alto x 35,6 cm de ancho x 40,6 cm de profundo, soportando OIR (Online Insertion and Removal) por lo que se puede insertar y extraer del Chasis C7613 en caliente sin que afecte al funcionamiento del mismo. El consumo de esta Line Card es de 2,5A a 42V.

Cumple las especificaciones de los estándares Ethernet (IEEE 802.3 y 10BASE-T) Fast Ethernet (IEEE 802.3, 100BASE-TX, y 100BASE-FX), Gigabit Ethernet (IEEE 802.3z y 1000BASE-TX) y Gigabit Ethernet over copper IEEE 802.3ab. Soporta autonegociación 802.3 Triple-Speed que permite al Switch negociar velocidad (10, 100, y 1000 Mbps) y modo duplex (Half o Full) con los equipos conectados.

Soporta distintos estándares tales como IEEE 802.1q VLAN, IEEE 802.1d Spanning Tree, IEEE 802.1p LAN Layer 2 QoS/CoS Protocol for Traffic Prioritization, Cisco Enhanced Per- VLAN Spanning Tree Plus (PVST+) protocol, IEEE802.1w Rapid Spanning Tree Protocol (RSTP), IEEE 802.1s Multiple Spanning Tree (MST) protocol, Per-VLAN Rapid Spanning Tree (PVRST) protocol, Hot Standby Router Protocol (HSRP), Virtual Router Redundancy Protocol (VRRP), Cisco EtherChannel® technology, IEEE802.3ad link aggregation for fault-tolerant connectivity.

Soporta Jumbo Frames (9216 bytes por trama).

Cada uno de los 48 puertos dispone de un Buffer de 2,67 MB, estando divididos en 6 grupos de puertos (Bundles) controlado cada Bundle por un ASIC con una conexión de 1 Gigabit Ethernet:

- Grupo 1 ⏩ Puertos 1-8.
- Grupo 2 ⏩ Puertos 9-16.
- Grupo 3 ⏩ Puertos 17-24.
- Grupo 4 ⏩ Puertos 25-32.
- Grupo 5 ⏩ Puertos 33-40.
- Grupo 6 ⏩ Puertos 41-48.

Soporta Autonegociación 10/100/1000 half-full-duplex, teniendo en cuenta que los puertos Gigabit Ethernet son Full duplex.

4.1.1.2.4 Enhanced FlexWAN Line Card

El modulo FlexWAN soporta una gran variedad de Interfaces WAN canalizados y no canalizados, incluidos E1, E3, High-Speed Serial Interface (HSSI), E3 ATM, STM-1 ATM, y STM-1 packet over SONET (POS). Se pueden insertar dos Port Adapters utilizados en Routers C7200 y C7500, convirtiendo al Router C7613 en un Equipo de agregación WAN, incluyendo accesos Frame Relay, Point-to-Point Protocol (PPP), High-Level Data Link Control (HDLC), ATM, o POS.

Ilustración 41- Line Card Enhanced FlexWAN.

Está equipado con dos CPU RM7000 400MHz , una por cada Bahía (Bahía 0 y Bahía 1) donde se insertan los dos Adaptadores de Puerto (Port Adapter) (Port Adapter 0 y Port Adapter 1), dos Memorias SDRAM de 256 MB (Packet Buffer Memory), una por cada Bahía, así como dos Flash PCMCIA de 128 MB, también una por Bahía. Cada Port Adapter se gestiona independientemente por la CPU de la bahía donde se encuentra insertado. Las interfaces correspondientes de cada Port Adapter aparecen en la configuración de la siguiente forma "Serial2/0/0" y "Serial2/1/0" por ejemplo en el caso de tener un módulo FlexWAN en el Slot 2 y dos Port Adapters serial insertados en Bay 0 y Bay 1 respectivamente.

Memoria hardware de Módulo FlexWAN:

- Packet Buffer Memory: 256 MB SDRAM para Bay 0.
- Packet Buffer Memory: 256 MB SDRAM para Bay 1.

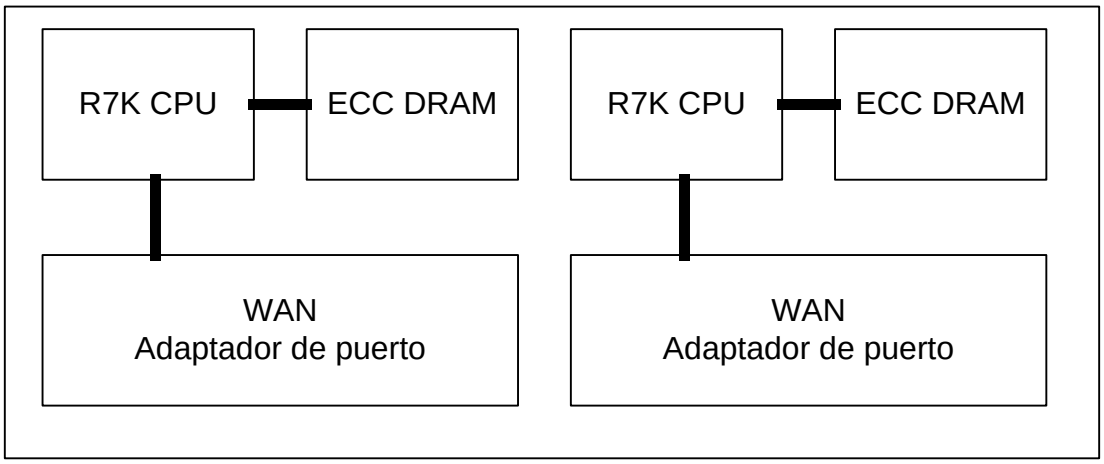

Ilustración 42- Enhanced FlexWAN + 2 Port Adapter.

Las 2 memorias por Bahía se pueden actualizar hasta un máximo de 512 MB.

Se trata de una Line Card tipo Fabric Enabled CEF256 que al utilizar una Supervisora SUP-720 requiere una versión mínima de IOS 12.2.(17a)SXA. No tiene capacidad de realizar switching local por lo que el switching se realiza de forma centralizada en la Supervisora (MSFC + PFC).

Ocupa un Slot en el Chasis C7613, con unas dimensiones de 3 cm de alto x 35,6 cm de ancho x 40,6 cm de profundo, soportando OIR (Online Insertion and Removal) por lo que se puede insertar y extraer del Chasis C7613 en caliente sin que afecte al funcionamiento del mismo, pero los Port Adapters insertados en el módulo FlexWAN no son OIR por lo que a la hora de extraerlos o insertarlos es necesario extraer antes el módulo FlexWAN.

La potencia requerida por esta Line Card es de 170W con los dos Port Adapters. Los protocolos soportados son IP, MPLS, ATOM (EoMPLS, FRoMPLS y ATMoMPLS). Las encapsulaciones son Frame Relay, Multilink Frame Relay—FRF.16, Multilink Point-to-Point Protocol (MLPPP), HDLC, ATM, HSSI, generic routing encapsulation (GRE) (supported on HSSI and Frame Relay interfaces).

4.1.1.2.5 8-Port E1 Channelized RJ48c Port Adapter

Port Adapter con 8 puertos E1 RJ48c G.703 configurables, válido para Routers 7600, 7200, 7300 y 7500. Soporta encapsulación RDSI, HDLC, X25, Frame Relay y PPP.

No es OIR (Online Insertion and Removal) en el Router 7600, por lo que no se puede insertar ni extraer del Chasis C7613 en caliente, pero el módulo FlexWAN sí que es OIR, así que a la hora de extraer el Port Adapter o insertarlo es necesario extraer antes el módulo FlexWAN.

Soporta reloj interno o de línea por cada uno de los 8 puertos.

Ilustración 43- Port Adapter 8-Port E1

4.1.1.2.6 1-Port STM-1 Channelized SC Single-Mode Port Adapter

Port Adapter con 1 puerto POS STM-1 Canalizado Fibra Monomodo, válido para Routers 7600, 7200 y 7500. Configurable 63 E1 Canalizados. Soporta encapsulación HDLC, Frame Relay y PPP.

No es OIR (Online Insertion and Removal) en el Router 7600, por lo que no se puede insertar ni extraer del Chasis C7613 en caliente, pero el módulo FlexWAN sí que es OIR, así que a la hora de extraer el Port Adapter o insertarlo es necesario extraer antes el módulo FlexWAN.

Soporta reloj interno o de línea por cada uno de los E1 Canalizados. También soporta CRC de 16 ó 32 bits.

Ilustración 44- Port Adapter 1-Port STM-1

4.1.1.2.7 1-Port E3 Serial BNC Port Adapter

Port Adapter con 1 puerto E3 Serial (Packet-over-E3) con un par de conectores BNC (cable coaxial 75 ohm), lleva integrado en la interfaz una unidad DSU (Data Service Unit) que permite conectar directamente un circuito E3 de 34.368MHz al Router, válido para Routers 7600, 7200, 7400 y 7500. Cumple con las especificaciones G.703. Soporta CRC de 16 ó 32 bits. Soporta encapsulación serie HDLC, PPP, Frame Relay.

No es OIR (Online Insertion and Removal) en el Router 7600, por lo que no se puede insertar ni extraer del Chasis C7613 en caliente, pero el módulo FlexWAN sí que es OIR, así que a la hora de extraer el Port Adapter o insertarlo es necesario extraer antes el módulo FlexWAN.

Ilustración 45- Port Adapter 1-Port E3

4.1.1.2.8 8-Port E1 Serial V35 Port Adapter

Port Adapter con 8 puertos E1 (2.048 Mbps) Serial con conector V35 Compact Serial (Necesario cable Compact Serial V35 200-pin con 8 conectores V35), válido para Routers 7600,

7200 y 7500. Se pueden configurar 8 puertos a 2 Mbps, 4 puertos a 4 Mbps, o 2 puertos a 8 Mbps. El ancho de banda agregado del PA es 16 Mbps para todos los puertos.

El tipo de cable utilizado determinará si el puerto se comporta como DTE o DCE, que en nuestro caso utilizaremos un cable pulpo V35 macho DTE.

Ilustración 46- Cable Compact Serial

No es OIR (Online Insertion and Removal) en el Router 7600, por lo que no se puede insertar ni extraer del Chasis C7613 en caliente, pero el módulo FlexWAN sí que es OIR, así que a la hora de extraer el Port Adapter o insertarlo es necesario extraer antes el módulo FlexWAN.

Ilustración 47- Port Adapter 8 puertos serie V.35.

4.1.1.2.9 1-Port E3 ATM Port Adapter.

Port Adapter con 1 puerto E3(34.368 Mbps) ATM con conector BNC. Cableado Coaxial 75 ohm. Válido para Routers C7200, C7300, C7500 y C7600, soportando las distintas clases de servicios ATM VBR-rt, CBR, VBR-nrt, ABR y UBR. Cumple con RFC2684 y RFC1483 correspondientes a encapsulación de distintos protocolos sobre ATM. RFC1577 IP y ARP sobre ATM. Configuración de PVCs y SVCs.

No es OIR (Online Insertion and Removal) en el Router 7600, por lo que no se puede insertar ni extraer del Chasis C7613 en caliente, pero el módulo FlexWAN sí que es OIR, así que a la hora de extraer el Port Adapter o insertarlo es necesario extraer antes el módulo FlexWAN.

Ilustración 48- Port Adapter 1-Port ATM E3.

4.1.1.2.10 Card WS-SVC-FWM-1-K9

La tarjeta WS-SVC-FWM-1-K9 añade las funcionalidades de firewall embebido sobre el propio Cisco 7600. Esto lo hace ideal en aplicaciones de seguridad y firewall donde la gran capacidad de conmutación y los distintos interfaces de comunicaciones son un requisito, por ejemplo como equipo frontera en redes de gestión y operación, así como en aplicaciones de firewall centralizado Esta tarjeta es capaz de proporcionar un rendimiento de hasta 5,5 Gbps, con hasta un millón de conexiones concurrentes y 100.000 intentos de conexión por segundo. Además es capaz de soportar la traslación de hasta 256.000 en modo PAT y 256.000 en modo NAT. Se puede configurar tanto en modo router como en modo transparente, lo cual permite introducirlo en la red sin suponer ningún cambio en el diseño, topología o direccionamiento de la misma.

Ilustración 49- Card WS-SVC-FWM-1-K9.

4.1.1.3 Familia C7200VXR

La familia Cisco 7200 se caracteriza por su gran versatilidad tanto en funciones como en interfaces disponibles. Además es capaz de procesar hasta 1 Mpps.

Esta familia es capaz de equiparse con distintas tarjetas procesadoras: NPE-225 (225 kpps), NPE-400 (400 kpps), NSE-1 (300 kpps) y NPE-G1 (1 Mpps). Para el caso concreto de este proyecto, la tarjeta procesadora elegida ha sido la NPE-G1.

Resumen de características:

1. Protocolos de nivel 2 y 3: ARP, IPCP, reenvío IP, IP multicast, PPPoA, VLAN, MPLS e IPv6.
2. Protocolos de routing de nivel 3: EIGRP, IGRP, ISIS, OSPF, BGP, PIM y RIP.

3. Gestión de red y seguridad: AAA, CHAP, FTP, RADIUS, SNMP, PAP y TACACS.
4. RFC 1483: Multiprotocol Encapsulation over ARM AAL 5.
5. RFC 1577: Classical IP and ARP over ATM AAL 5.
6. Cumple con NEBS level 3.
7. Dimensiones (alto x ancho x fondo): 13,34 x 42,67 x 43,18 cm.

Ilustración 50- Chasis Cisco 7200

4.1.1.4 Tipos de tarjetas de la serie 7200

4.1.1.4.1 NPE-G1

Es la tarjeta procesadora para el Cisco 7204VXR elegida para los PEs que utilizan este hardware. Esta tarjeta proporciona un rendimiento de hasta 1 Mpps utilizando CEF, incorporando 3 interfaces Gigabit Ethernet/Fast Ethernet y 1 GB de memoria DRAM.

Ilustración 51- Tarjeta procesadora NPE-G1.

4.1.1.4.2 SA-VAM2+

Esta tarjeta es la encargada de realizar todas las labores de encriptación y desencriptación sobre el router Cisco 7204VXR. La VAM2+ (VPN Acceleration Module 2+) es un port adapter que se instala en cualquier slot del Cisco 7204VXR (o Cisco 7206VXR). Este módulo proporciona aceleración hardware para AES, DES y 3DES.

Entre sus principales características destacan:

- DES en modo estándar con clave de 56 bits.
- 3DES (168 bits) con velocidades de hasta 260 Mbps.
- AES con longitudes de clave de 128, 192 o 256 bits.

- Grupos Diffie-Hellman 1,2 y 5.
- Número máximo de túneles IPsec: 5.000.

Ilustración 52- Tarjeta SA-VAM2+.

4.1.2 RED MPLS UFIN

Esta red consta de 7 nodos, todos ellos CISCO pero de diferentes categorías. Los de FII, MS, ASL, PP, CPII y VLL son Routers Cisco C7613 con doble Supervisora SUP720-3B. En el caso del nodo de RA, dado que se estimó que llevaría menos servicios es un Router C7606. En nuestro caso el 7606 va equipado con una sola supervisora.

Todos ellos comparten las funciones de P/PE puesto que los requisitos de tráfico y funcionalidad quedan totalmente cubiertos y se reducen mucho los costes.

La siguiente tabla resume el equipamiento hardware de cada equipo en la red:

Nodo	CPII	RA	ASL	MS	FII	VLL	PP
Chasis	CISCO7613	CISCO7606	CISCO7613	CISCO7613	CISCO7613	CISCO7613	CISCO7613
Supervisoras	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B
	WS-SUP720-3B		WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B
Tarjetas	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+			
	OSM-2+4GE-WAN+		OSM-2+4GE-WAN+	OSM-2+4GE-WAN+			
	WS-X6148A-GE-TX	WS-X6148A-GE-TX	WS-X6148A-GE-TX	WS-X6148A-GE-TX	WS-X6148A-GE-TX	WS-X6148A-GE-TX	WS-X6148A-GE-TX
				WS-X6148A-GE-TX			
	WS-X6724-SFP		WS-X6724-SFP	WS-X6724-SFP	WS-X6724-SFP	WS-X6724-SFP	
					WS-X6724-SFP		
	WS-X6582-2PA	WS-X6582-2PA	WS-X6582-2PA	WS-X6582-2PA	WS-X6582-2PA	WS-X6582-2PA	
	WS-X6582-2PA		WS-X6582-2PA		WS-X6582-2PA		
			WS-X6582-2PA				
	PA-MC-STM-1SMI		PA-MC-STM-1SMI		PA-MC-STM-1SMI	PA-MC-STM-1SMI	
	PA-E3		PA-E3				
	PA-E3						
	PA-MC-8TE1+	PA-MC-8TE1+	PA-MC-8TE1+	PA-MC-8TE1+	PA-MC-8TE1+	PA-MC-8TE1+	
			PA-MC-8TE1+				
			PA-A6-E3				
			PA-8T-V35	PA-8T-V35	PA-8T-V35		
	7600-SIP-400		7600-SIP-400		7600-SIP-400	7600-SIP-400	7600-SIP-400
					7600-SIP-400	7600-SIP-400	7600-SIP-400
	SPA-2XOC3-POS		SPA-2XOC3-POS		SPA-2XOC3-POS	SPA-2XOC3-POS	SPA-5X1GE-V2
					SPA-2XOC3-POS	SPA-2XOC3-POS	SPA-5X1GE-V2
Fuentes	PWR-4000-DC	PWR-2700-AC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC
	PWR-4000-DC	PWR-2700-AC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC

Tabla 5. – Tarjetería instalada en los equipos de UFIN

Geográficamente los nodos de VLL y de FII están bastante alejados del resto por lo que no tienen conexión GigabitEthernet con ellos. La unión de estos dos nodos entre es igualmente a través de un STM1-C. En el caso de las interconexiones entre los nodos de ASL, MS, RA, PP y CPII es a través de puertos GigabitEthernet de la tarjeta OSM-2+4GE-WAN+. Físicamente entre ellos hay fibra óptica monomodo directa.

Para las uniones VLL <=> ASL y FII <=> CPII se utilizan puertos POS de una tarjeta 7600-SIP-400 ya que a nivel 1 lo que hay es una red SDH que proporciona un STM1.

En la siguiente imagen se puede apreciar el esquema de interconexionado entre todos los nodos que conforman la red:

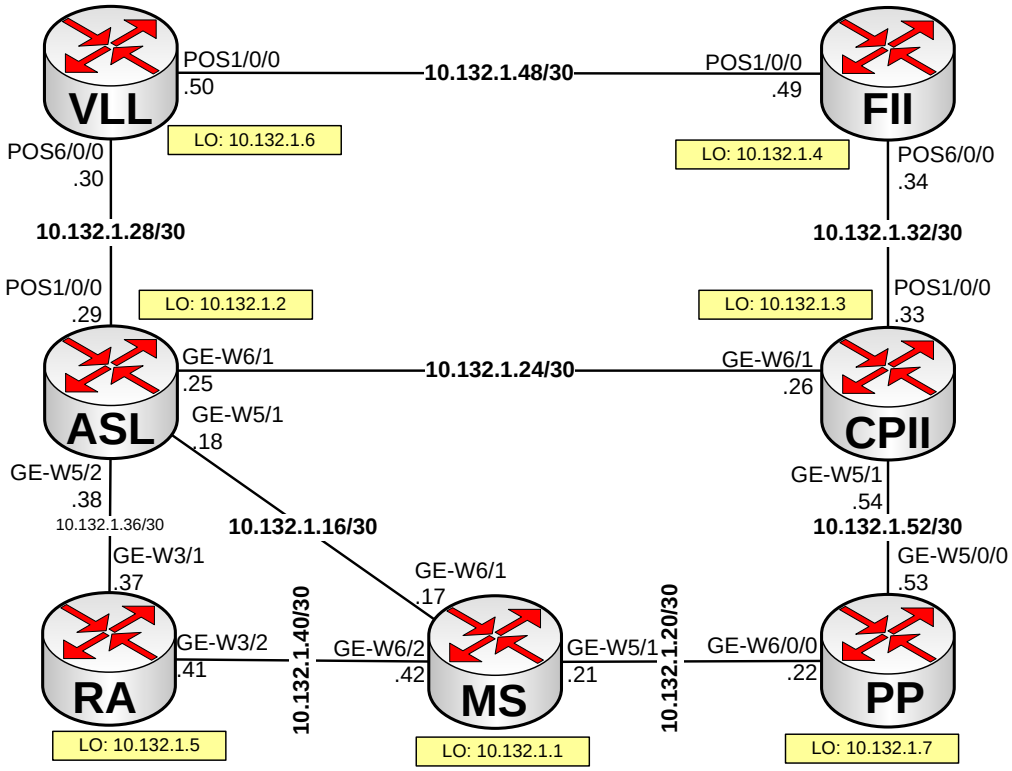

Ilustración 53- Red MPLS de UFIN

4.1.2.1 IGP

La activación de un protocolo de enrutamiento responde a la imposibilidad de tener rutas estáticas, por la no adaptación a cambios en una topología donde hay posibles varios caminos como es la red que se está diseñando. Como se definió en la parte de teoría, como IGP se prefiere un protocolo del tipo Estado de Enlace, por su rapidez de convergencia, y no sufrir los problemas de bucles de los de tipo Vector Distancia.

Está configurado OSPF por su mayor simplicidad de configuración, y dado que no se justifica la mayor escalabilidad de ISIS en una red de este tamaño, OSPF se considera suficiente.

Los router Cisco permiten la activación de varios procesos OSPF, siempre que una interfaz no pertenezca a varios procesos, incluso se puede hacer redistribuciones de rutas entre los diferentes procesos OSPF. Eso obliga a definir un número de proceso OSPF, local al router, cuando se activa. Está configurado el proceso IGP con número 1.

Está activado el protocolo de routing OSPF área 0 en los enlaces /30 de red y en una interfaz Loopback0/32 de cada router. Para activar OSPF en las interfaces anteriores se usará una wildcard exacta de 0.0.0.0 para evitar errores, que exige poner la IP de la interfaz de forma exacta.

El área 0 en OSPF está reservada para el núcleo de la red. Hay configurada un área única por una razón principal: el tamaño de la red no justifica trocearla. Por otra parte, ante el

empleo de un único área, el área 0, reservada para el backbone parece la más indicada, puesto que si la red crece, el núcleo de la red ya está definido.

Se configura el comportamiento OSPF de las interfaces /30 como tipo de red punto a punto. Se puede hacer esto puesto que la topología es conocida y es punto a punto. Si se deja que los C7613 decidan, en los enlaces GE, de tipo LAN, se forzará la elección de DR y BDR, algo no necesario para enlaces P2P.

Hay una configuración especial de costes OSPF en interfaces para evitar que el tráfico se curse por los enlaces de baja capacidad. El coste de una interfaz se puede dimensionar en el rango siguiente: 1-65535. Están penalizadas las interfaces POS para evitar que el tráfico se curse por estos enlaces sino son origen ni destino de la comunicación. Para ello se asignan los costes de la siguiente manera:

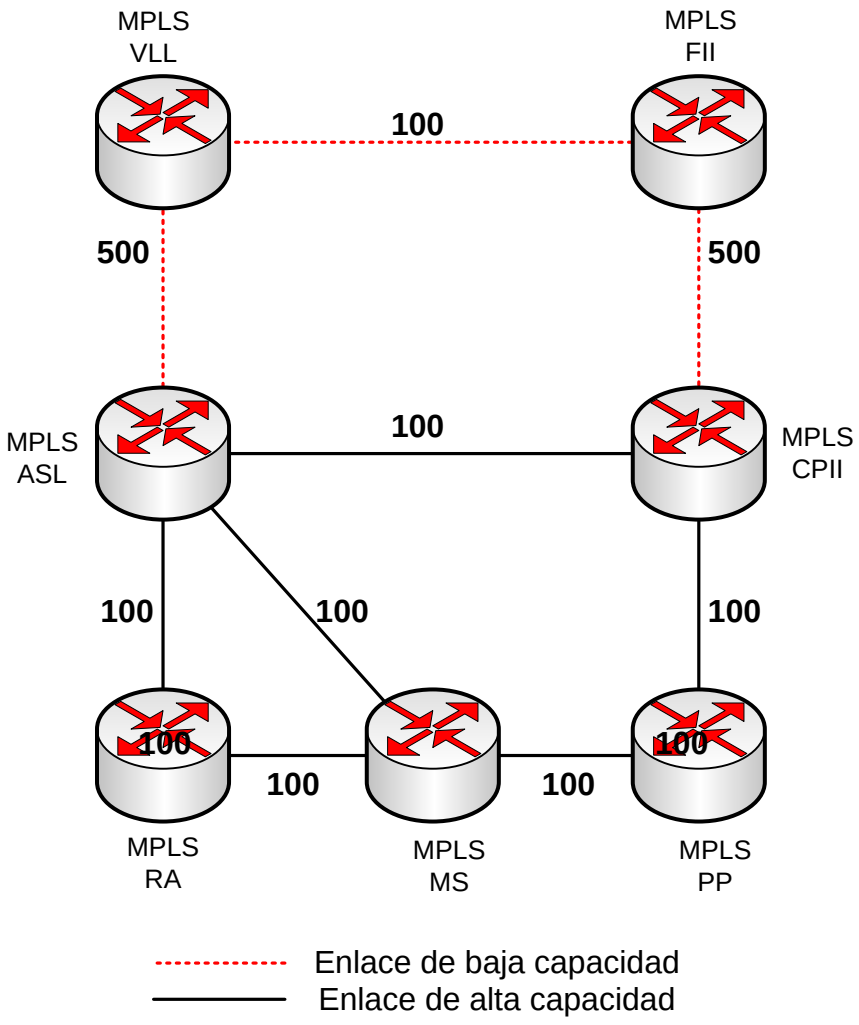

Ilustración 54- Distribución de pesos OSPF en la red de UFIN.

Con la política de costes anterior, en el caso de fallo del enlace directo entre ASL y CPII, el tráfico entre esos dos nodos prefiere el camino a través de MS y PP, y no a través de FII y VLL

Se podría haber obtenido el mismo resultado con un coste menor en las interfaces POS pero estos valores son adecuados incluso si se inserta algún nodo más en la red. Un valor 5 veces mayor en un enlace de baja capacidad supone ser preferible el paso por 5 enlaces de alta capacidad lo cual es suficiente.

El coste de las interfaces loopback está a su valor 1 por defecto.

A nivel de las interfaces del core se configurará en ambos extremos de cada interfaz la misma clave MD5, con 16 caracteres como máximo. Como índice de la clave (key ID) hay que escoger un valor entre 1 y 255. Se propone usar el valor 1 en ambos extremos.

El router-id del proceso OSPF es la interfaz loopback0 del router.

Como se pretende la mayor rapidez de convergencia posible del IGP, está configurado el uso de la funcionalidad fast-hello en todas las interfaces del core. Esta funcionalidad permite acelerar la convergencia del IGP. El parámetro "dead time" está fijado en un segundo, y mediante el parámetro "hello multiplier" se permite definir el número de paquetes OSPF hello que se envían por segundo, que configura en 3. Un valor demasiado alto puede impactar en la CPU del router, y un valor demasiado bajo puede hacer que ante un corte pequeño se pierda algún hello y se tire la adyacencia innecesariamente. Con unos valores tan pequeños de dead-interval, OSPF es demasiado sensible. Considerando que se desea estabilidad del OSPF ante inserción/extracción de tarjetas, no se usa la interfaz "OSPF fast-hello", y se usan los temporizadores siguientes de OSPF en los enlaces POS del core:

- hello-interval: 2 segundos.
- dead-interval: 6 segundos.

En los interfaces Gigabit Ethernet no se usan estos temporizadores ya que al ser fibra directa OSPF cae en el momento que el interfaz pasa al estado down.

Con estos ajustes se tiene un tiempo de convergencia de sobre 6 segundos en el IGP que sigue siendo alto. Debido a eso están ajustados una serie de temporizadores para lograr más rapidez de convergencia, puesto que se siguen unos algoritmos de backoff exponenciales que permiten conseguir ambas cosas. Un grupo de parámetros a considerar es el retardo en correr el algoritmo SPF de forma total, puesto que de forma parcial no se retrasa, y que está regulado por la funcionalidad "OSPF Shortest Path First Throttling". Se regula por tres temporizadores:

- "Initial SPF schedule delay", es el tiempo que se espera para ejecutar SPF desde que hay un cambio topológico.
- "Minimum hold time": una vez que se ha ejecutado el algoritmo, si hubiese otro cambio, la siguiente ejecución se retrasa hasta este tiempo, y si en ese intervalo hubiese otro cambio, se va duplicando automáticamente este temporizador, pero con un límite marcado por el tercer temporizador, que es "Maximum wait time".

Se da por hecho que se produce un cambio topológico, y tras 5 ms se ejecuta el SPF. Luego, se supone que en cada intervalo se produce al menos un cambio, y entonces se va retrasando el SPF hasta llegar a un límite.

Los valores por defecto de estos temporizadores son:

- Initial SPF schedule delay 5000 msecs.
- Minimum hold time between two consecutive SPFs 10000 msecs.
- Maximum wait time between two consecutive SPFs 10000 msecs.

Se cambia el primer valor a 1000 msec, y se deja por defecto los otros 2.

Otro grupo de parámetros que se puede ajustar controla el envío y la recepción de LSA en el router.

En periodos de inestabilidad de red es posible que el envío de LSA pueda saturar al router. Es por ello que originalmente, antes de que existiese la funcionalidad "LSA throttling" lo que se hacía era retrasar el envío de las LSA 5 segundos. Esto puede proteger la CPU del router a costa de la velocidad de convergencia.

Con la funcionalidad LSA throttling, habilitada por defecto, se permite más velocidad de convergencia, pero a la vez se protege la CPU del router. Para ello el funcionamiento es que la primera LSA se envía tras un temporizador "initial LSA throttle delay" tras un cambio topológico. El envío de la siguiente LSA está controlado por un temporizador: "minimum start interval". El algoritmo hace que el tiempo se vaya doblando hasta alcanzar el "maximum interval". Una vez que el maximum interval se ha alcanzado, se considera estabilidad y los temporizadores vuelven a iniciarse. Estos temporizadores se refieren en todo momento a la misma LSA. Si por ejemplo una interfaz se levanta/cae/levanta, la LSA de tipo 1 generada por un router con todas sus interfaces será la que se envíe una y otra vez actualizada con una interfaz más o menos.

Los valores por defecto son:

- Initial LSA throttle delay: 0 msecs.
- Minimum hold time for LSA throttle 5000 msecs.
- Maximum wait time for LSA throttle 5000 msecs.

Con este comportamiento ante una red inestable se consigue el comportamiento de 5 segundos anterior para proteger la CPU del router, mientras que en un comportamiento de caída/levantamiento de una interfaz, de forma no inestable, la propagación será instantánea.

Se propone la configuración de estos parámetros como:

- Initial LSA throttle delay: 0 msecs.

- Minimum hold time for LSA throttle 500 msecs.
- Maximum wait time for LSA throttle 5000 msecs.

Por otra parte, se puede controlar la recepción de LSA con el temporizador "Minimum LSA arrival".

Si una versión actualizada de la misma LSA se recibe antes de que se agote el temporizador "Minimum LSA arrival", está será descartada. Es por ello que el temporizador "Miminum LSA arrival" debe ser menor o igual que el temporizador "minimum hold time" del router vecino que envía la LSA. El temporizador "initial throttle delay" del transmisor no se tiene en cuenta puesto que "minimum LSA arrival" no actúa la primera vez. La recomendación realmente está asumiendo también que Minimum hold time <= Maximum wait time para garantizar también que min arrival <= Maximum hold time. Es por ello que el valor de "Miminum LSA arrival" = 300 ms.

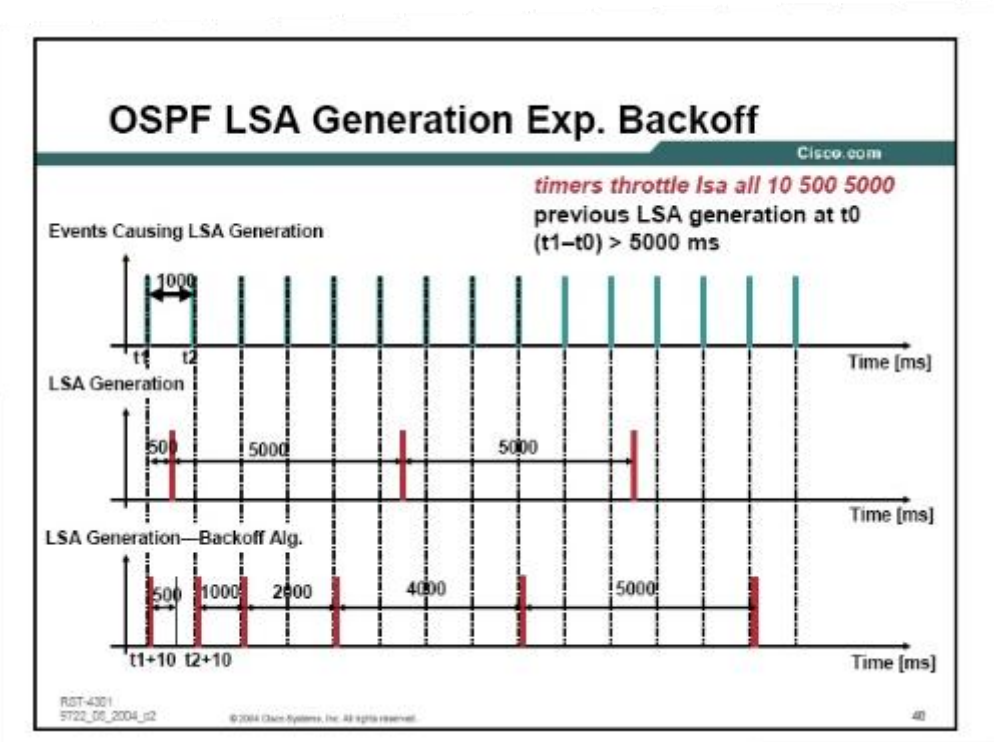

Ilustración 55- Temporizadores de algoritmo exponencial de backoff para generación de LSA.

En la ilustración anterior se muestra un ejemplo en el que tras la estabilidad (t1-t0)>max=5seg, se muestra la diferencia entre no disponer de algoritmo exponencial de backoff con el retraso de 5 segundos en la generación de LSA, y al disponer de él, donde se puede comprobar cómo la generación de la primera LSA es la los t1+initial=t1+10 ms, luego la

separación con la siguiente generación está controlada por “minimum hold-time”, y luego se va duplicando este tiempo hasta alcanzar el máximo de 5000 seg.

Las funciones de route reflector son realizadas por los nodos de ASL y CPII. Se ha decidido compartir la función de route reflector con la de PE dado a que no se prevé una alta carga de la máquina.

4.1.2.2 VTP

VTP es un protocolo que permite la asignación centralizada de VLAN en switches con un funcionamiento cliente/servidor. Con el uso de la red, siendo enlaces entre C7613 de nivel 3 de VTP, los equipos están en modo transparente, para que cada uno de ellos no dependa de la información de VTP que pudiese recibir de otros switches C7613 o de cliente.

4.1.2.3 STP

Por si se decidiese conectar algún equipo de forma redundante a nivel 2. La forma de asegurar que no haya bucles de nivel 2 es mediante el empleo del protocolo STP (spanning tree protocol). Este protocolo tiene varias variantes. STP está desactivado en los C7600. De esta forma es el cliente quien debe correr STP, y la red pasará las BPDU transparentemente, tanto en el servicio EoMPLS, VPLS y Ethernet Switching.

Por ejemplo, si conectamos un switch de acceso con dos puertos de trunk al mismo C7600, y en el C7600 no hay STP habilitado, el equipo de cliente detecta el bucle y bloquea uno de los puertos. En la siguiente ilustración se muestra el comportamiento del C7600 dejando pasar transparentemente las BPDU del cliente, como si fuese un cable. Eso provoca que el STP del switch de cliente bloquee uno de los puertos.

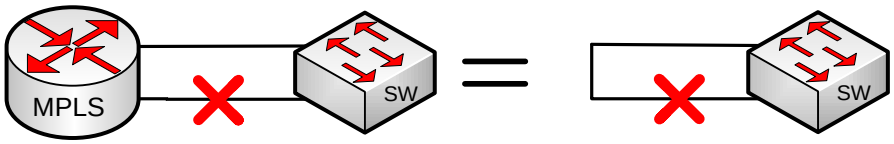

Ilustración 56- Transparencia de STP.

4.1.2.4 Direccionamiento

Respecto al direccionamiento, dado que estos enlaces corren sobre la tabla de routing global, la cual no es accesible desde ninguna VPN se utiliza direccionamiento privado utilizando la clase C 10.132.1.0/24 reservando las primeras 16 direcciones para loopbacks de los routers y el resto, subneteadas en /30 para los enlaces.

4.1.2.4.1 Interfaces de loopback.

Cada una de esas direcciones loopback se usarán como RID para OSPF, BGP, LDP, TE, para acceder por telnet, para acceder en modo lectura o escritura por SNMP a cada equipo, como dirección origen para envío de traps SNMP o mensajes de syslog.

4.1.2.5 CDP

CDP está deshabilitado por motivos de seguridad. No interesa que equipos de cliente puedan obtener información alguna de la red.

4.1.2.6 Interfaces Gigabit Ethernet de Core

Configuración por defecto. Está habilitada la negociación en ambos extremos. No se negocia velocidad, pero sí parámetros de control de flujo, información de fallos remotos, y configuración dúplex del enlace. Los enlaces GE de core estarán conectados por medio de fibras directas. Eso implica el poder confiar en la detección de caída de esos enlaces de forma instantánea. El parámetro carrier-delay, que por defecto es 2 segundos, es el tiempo que se espera para avisar a la IOS desde la detección de la caída se baja a 400 ms.

4.1.2.7 Interfaces POS STM1

Su configuración de framing SDH (Europa), frente a SONET (América), pues es lo que esperan los equipos de transmisión SDH intermedios.

Para evitar problemas de sincronismo en la transmisión de muchos 0's seguidos o 1's seguidos, se utilizará un algoritmo reversible en ambos extremos que permite que esas largas secuencias se transformen a la hora de transmitirse en otra secuencia de bits sin ese problema. La funcionalidad utiliza el mismo algoritmo que ATM, y se denomina scrambing estilo ATM.

Dado que hay una red de transmisión intermedia, está por defecto la configuración relativa al reloj, que consiste en que se tome reloj de la línea. El hecho de que exista una red de transmisión intermedia hace que aunque en la mayoría de las ocasiones las alarmas SDH se envíen de forma instantánea, ante determinadas situaciones, los equipos remotos Cisco C7600 reciben alguna alarma de forma instantánea pero que no cambia el estado UP/UP a UP/DOWN hasta 30 segundos después de haber recibido esa alarma. Eso implica el no poder confiar en los enlaces POS con transmisión por el medio para una detección de caída rápida, y por ello el IGP se ajusta en esos enlaces POS. Se ha elegido configuración de compromiso para los enlaces POS con hello-interval de 2 segundos y dead-interval de 6 segundos.

El parámetro carrier-delay, que por defecto es 2 segundos, es el tiempo que se espera para avisar a la IOS desde la detección de la caída pero está bajado a 400 ms.

4.1.3 RED MPLS DC.

Esta red consta de 14 nodos, todos ellos CISCO pero de diferentes modelos. Los de ARA, AME, COR, COR AME, FER, RIB, SNT, SEV, ESX, TEL, SAL son Routers Cisco C7609 con doble Supervisora SUP720-3B. El nodo de EDS, debido a su inicial bajo flujo de tráfico es un 7204 al igual que BCN-RRR y MAD-RRR que tienen funciones de route reflector. El resto de nodos tienen funciones de PE con la particularidad de que los nodos de TEL y ESX solo tienen un enlace con el resto de la red ya que se emplean como conexión a internet. Respecto a

supervisoras, todos los equipos tienen doble supervisora salvo los 7609 FER y SAL y todos los 7200 (EDS, BCN-RRR y MAD-RRR).

La siguiente tabla resume el equipamiento hardware de cada equipo en la red:

Nodo	COR	AME	ARA	COR AME	RIB	SNT	SEV
Chasis	CISCO7609	CISCO7609	CISCO7609	CISCO7609	CISCO7609	CISCO7609	CISCO7609
Supervisoras	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B
	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B
Tarjetas	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+
	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+		OSM-2+4GE-WAN+		
	WS-X6748-GE-TX	WS-X6748-GE-TX	WS-X6748-GE-TX	WS-X6748-GE-TX	WS-X6748-GE-TX	WS-X6748-GE-TX	WS-X6748-GE-TX
	WS-X6748-GE-TX						
	WS-X6582-2PA	WS-X6582-2PA	WS-X6582-2PA		WS-X6582-2PA	WS-X6582-2PA	WS-X6582-2PA
	WS-X6582-2PA	WS-X6582-2PA					
	PA-MC-STM-1SMI	PA-MC-STM-1SMI	PA-MC-STM-1SMI		PA-MC-STM-1SMI	PA-MC-STM-1SMI	PA-MC-STM-1SMI
	PA-MC-STM-1SMI	PA-MC-STM-1SMI					
	PA-2E3	PA-E3					
	PA-2E3						
	WS-SVC-FWM-1			WS-SVC-FWM-1			
Fuentes	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC
	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC	PWR-4000-DC

Nodo	SAL	FER	ESX	TEL	MAD RRR	BCN RRR	EDS
Chasis	CISCO7609	CISCO7609	CISCO7609	CISCO7609	CISCO7204VXR	CISCO7204VXR	CISCO7204VXR
Supervisoras	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	WS-SUP720-3B	NPE-G1	NPE-G1	NPE-G2
			WS-SUP720-3B	WS-SUP720-3B			
Tarjetas	OSM-2+4GE-WAN+	OSM-2+4GE-WAN+					
	WS-X6748-GE-TX	WS-X6748-GE-TX	WS-X6748-GE-TX	WS-X6748-GE-TX			
		PA-MC-STM-1SMI	7600-SIP-400	7600-SIP-400			
			SPA-2X1GE	SPA-2X1GE			
			SPA-2X1GE				
			SPA-2X1GE				
							PA-GE
Fuentes	PWR-4000-DC	PWR-4000-DC	WS-CAC-2500W	PWR-4000-DC			
			WS-CAC-2500W	PWR-4000-DC			

Tabla 6. – Tarjetería instalada en los equipos de DC

La unión entre todos los routers es por enlaces GE-WAN ya que todos los enlaces menos uno tienen como red de transmisión una red DWDM. El caso del enlace SNT – SEV es un circuito de operador púbico que pese a ir sobre puertos GE-WAN es un FastEthernet.

En la siguiente imagen podemos ver como están interconectados los routers:

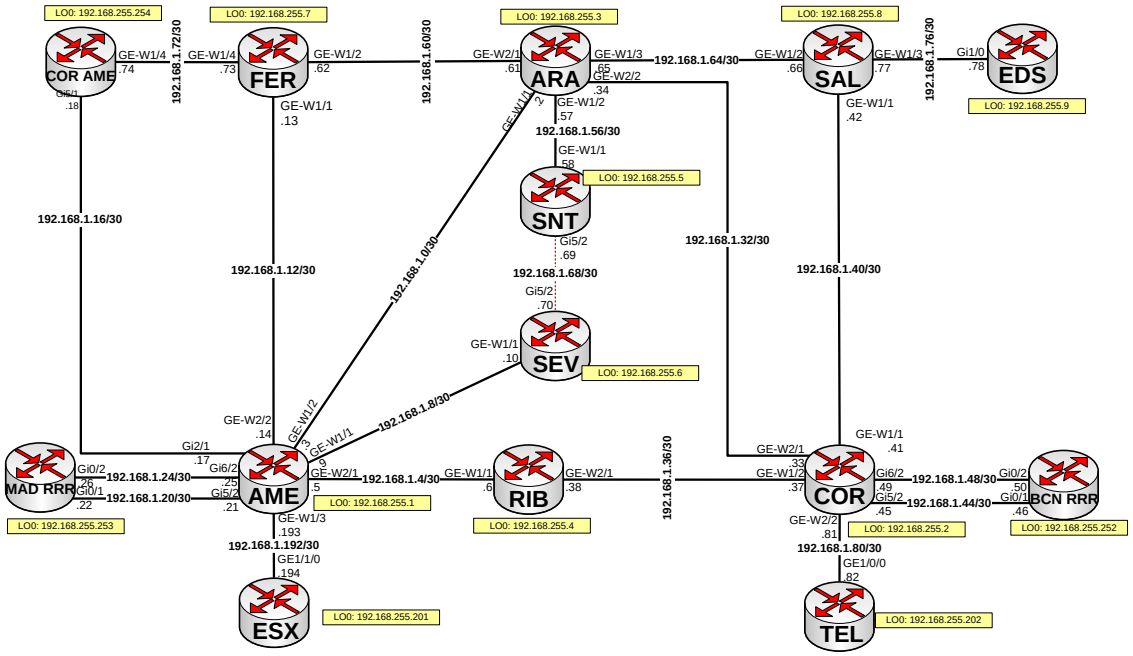

Ilustración 57- Red MPLS de DC

4.1.3.1 IGP

El protocolo que se ejecuta en la red como IGP es IS-IS. A nivel lógico, IS-IS es un protocolo de routing que se puede dividir en Áreas. Su estructura es similar a la de OSPF: Existen routers solo de Área (Nivel 1, o IntraÁrea), Router de BackBone (Nivel 2 o Área BackBone), y Routers Entre-Áreas (Nivel 1-2, o Inter-Área). Debido al número de routers de la red, no hay necesidad de dividir en áreas el backbone, y sólo hay configurado una única Área, aunque eso sí, de nivel dos (que es el área especial de backbone), por si en un futuro fuese necesario dividir en Áreas más pequeñas.

En routers Cisco, IS-IS utiliza la métrica del interfaz para el cálculo de rutas. Por defecto, asigna un valor de 10 a cada interfaz (se pueden asignar valores entre 0 y 63 con la métrica normal y entre 1 y 16.777.214 utilizando la métrica ampliada, obligatoria cuando se aplica Ingeniería de Tráfico), independientemente del ancho de banda que tenga. Si se dejan estos valores por defecto, IS-IS queda como un protocolo de cuenta de saltos, con un valor de 10 para cada salto. Por ello, y para ajustar mejor el coste de las rutas, se asigna un coste a cada interfaz, en función del ancho de banda que tiene. Basada en el cálculo 10^8/Ancho de banda del interfaz (Kbps) de la siguiente tabla:

Tipo de interfaz	Coste
Gigabit Ethernet	100
Fast Ethernet	1000

Tabla 7. – Pesos de IS-IS configurados en los enlaces de DC

Esta tabla se toma como valores de referencia, se pueden modificar por cuestiones de ingeniería de tráfico como es el caso de la red. El enlace SNT-SEV debería tener configurado un valor de 1000, pero en ese caso nunca se utilizaría ya que la suma del resto de caminos entre nodos más alejados no llega a 10 enlaces. Es por ello por lo que se configura un valor de 250, para que las comunicaciones con origen o destino en los nodos de SNT y SEV elijan siempre este enlace. Según esto el diagrama de la red en base a los pesos queda de la siguiente manera.

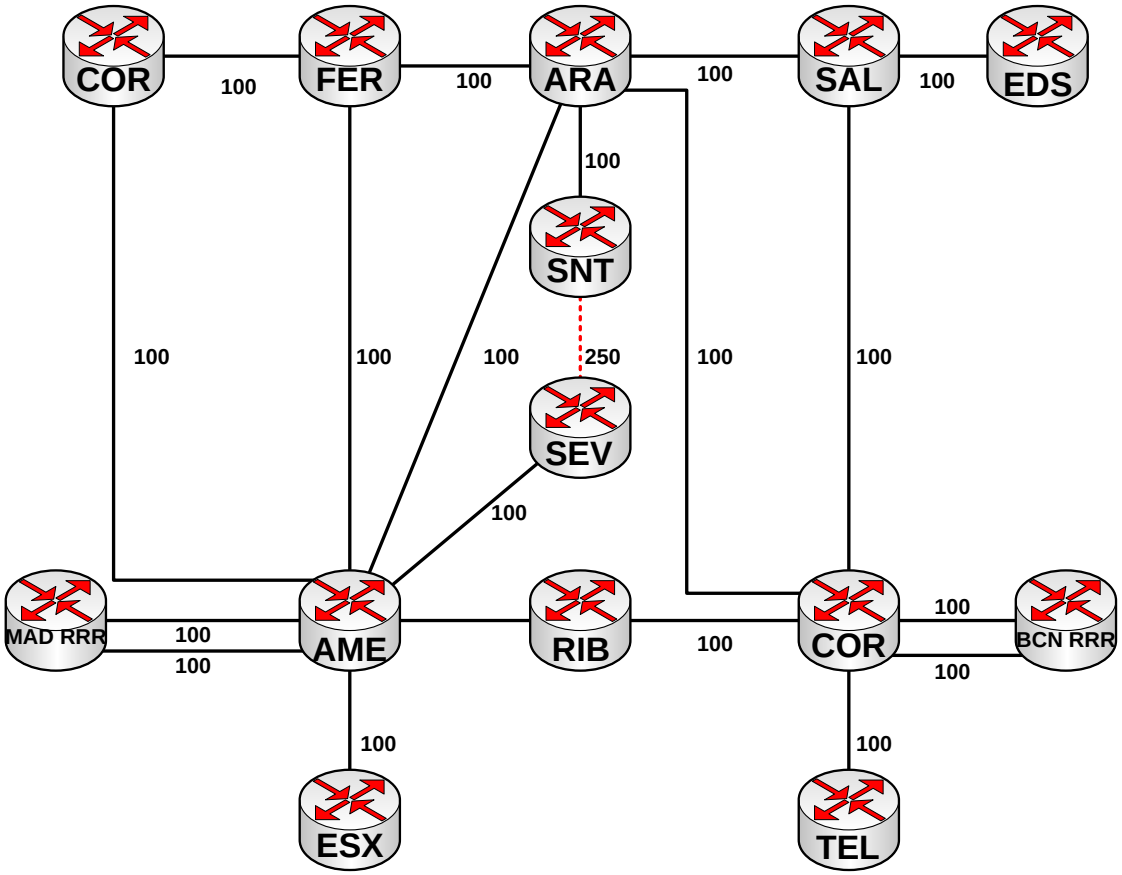

Ilustración 58- Esquema de la red en base a los pesos de IS-IS configurados

Como IS-IS es un protocolo de estado de enlace, genera una tabla de rutas en función de los datos que recibe de otros routers que hablan ISIS. Dicha tabla se ve afectada cada vez que hay un cambio (cuando se añade o quita un interfaz, cuando cae o levanta un interfaz, etc.)

Cuando un router sufre un cambio, empieza a mandar paquetes por sus interfaces IS-IS a sus vecinos, para que actualicen sus tablas de rutas (estos paquetes de llaman LSP, o Link-State PDU). Los routers IS-IS descubren a sus vecinos y forman adyacencias intercambiando IS-IS Hello Pus. Estos Hellos se envían cada 10 segundos.

Una vez que la adyacencia está establecida, los Hellos se usan como keepalive. En cada paquete Hello, va un hold-time, que indica a los vecinos lo que deben esperar para oír el

siguiente hello, antes de declarar al router como caído. En los routers Cisco, es por defecto 3 veces el valor del Hello Interval.

Al igual que se hizo en la red de UFIN con OSPF, en la red de DC en IS-IS, los enlaces Gigabit Ethernet WAN están configurados como si se trataran de enlaces punto a punto desde el punto de vista de topología de red para IS-IS, que no demoren la detección de caída del enlace, que den por caído el enlace en el en caso de flapeos en el interfaz y que se aumente el tamaño de las MTU a 4470. Esto acelerará la creación de adyacencias entre los nodos que conforman el punto a punto Gigabit Ethernet.

IS-IS soporta sumarización de Rutas. Una ruta sumarizada puede incluir múltiples grupos de direcciones para un nivel dado. Las rutas aprendidas por otros protocolos, también se pueden sumarizar. La métrica utilizada para anunciar la sumarización, es la ruta más corta de todas las rutas más específicas. En el caso concreto de esta red, en principio no es necesario realizar sumarización de rutas puesto que el dimensionado y topología de la red no lo precisa.

Por defecto, la autentificación esta desactivada pero en este caso está habilitada utilizando MD5 todos los tipos de paquete que utiliza IS-IS (LSP, LAN Hello, Serial Hello, CSNP y PSNP). Para ello, deberá aplicarse autenticación de IS-IS a nivel de interfaz (LAN Hello, Serial Hello) y a nivel de proceso

4.1.3.2 QoS

Gracias a QoS, se pueden ofrecer servicios diferenciados en una red MPLS. Se hace definiendo una clase de servicio en cada paquete IP y asignando una preferencia a cada paquete en la cabecera del mismo. MPLS QoS soporta los siguientes servicios diferenciados en una red MPLS:

- Clasificación y marcado de paquetes: Se analiza el tráfico entrante en la red y en función de unos determinados patrones se clasifica y se marcan los bits correspondientes (bits de precedencia en los paquetes IP y experimentales en las tramas MPLS).

- Detección y disminución de la congestión: Monitoriza el tráfico de red para anticiparse y prevenir la congestión y botella. Consiste en ir descartando selectivamente el tráfico de menor prioridad, cuando un interfaz comienza a estar saturado. Esta funcionalidad se realiza mediante RED (Random Early Detection) o WRED (Weighted Random Early Detection) de QoS.

- Gestión de la congestión: Asigna una asignación del ancho de banda para todo el tráfico de red. Utiliza pesos (o prioridades) para determinar cuánto ancho de banda se asigna a cada clase de tráfico y con qué prioridad. Esto lo hace WFQ (Weighted Fair Queueing) de QoS para plataformas no GSR. MDDR (Modified Deficit Round Robin) para plataformas GSR.

4.1.3.2.1 Salida de internet

Una de las diferencias notables entre la red de UFIN y la de DC está en el servicio de Internet. Actualmente en la red de DC se presta el servicio de internet de toda la empresa. Esto se hace con dos puntos de interconexión a internet con dos ISPs distintos que tienen las rutas de internet en la tabla de routing global. El resto de nodos tienen una ruta por defecto hacia ellos. Hay que distinguir entre distintos tipos de accesos debido a las características particulares de cada uno, entre los que hay que diferenciar:

- Servicio de acceso a Internet (sin VPN)
- Servicio de acceso a Internet (con VPN)
- Interconexión con otros Operadores (servicios tunelizados)
- Servicio de ISP para los clientes finales – Internet Full Routing

El hecho de ofrecer servicio de ISP con Internet Full Routing para los clientes de la red obliga a mantener la tabla completa de Internet en los equipos PE de la red MPLS.

La interconexión con Internet se realiza en ESX y en TEL. Hay 4 conexiones, dos de ellas de tránsito a Internet con ISP1 (contra ESX) y con ISP2 (contra TEL) y dos al punto neutro de interconexión de ESX. Las características de la conexión a Internet de la red son:

La solución de conexión a Internet que está implementada en la red de se conoce como dual-homing o doble proveedor con equipamiento redundado. Se dispone de dos routers Cisco 7600, localizados en TEL y ESX, que implementarán sendas conexiones con dos proveedores: ISP1 e ISP2 respectivamente. Esta solución provee la máxima redundancia en caso de fallo de equipamiento o línea.

La configuración de este dual-homing tendrá que cumplir los siguientes requisitos:

- El diseño será lo más simple posible, para facilitar su operación, gestión y mantenimiento.
- El funcionamiento elegido será Activo/Backup.
- Todo el tráfico tanto de entrada como de salida a Internet se cursará por el enlace principal cuando éste sea accesible
- En caso de caída del enlace principal o del acceso al mismo el tráfico se cursará por el enlace de backup

El enlace principal será el establecido entre TEL e ISP 2 y el enlace de respaldo será el establecido entre ESX e ISP 1.

El resto de nodos de la red no implementarán sesiones BGP con los routers de Internet. Los routers de Internet generarán la ruta por defecto para el resto de la red.

El camino entre los routers de la red MPLS e Internet es óptimo, incluyendo los casos:

- Situación normal con correcto funcionamiento de todos los elementos involucrados.
- Situación con fallo de alguno de los elementos de la solución de interconexión.

En la siguiente ilustración vemos como está configurado:

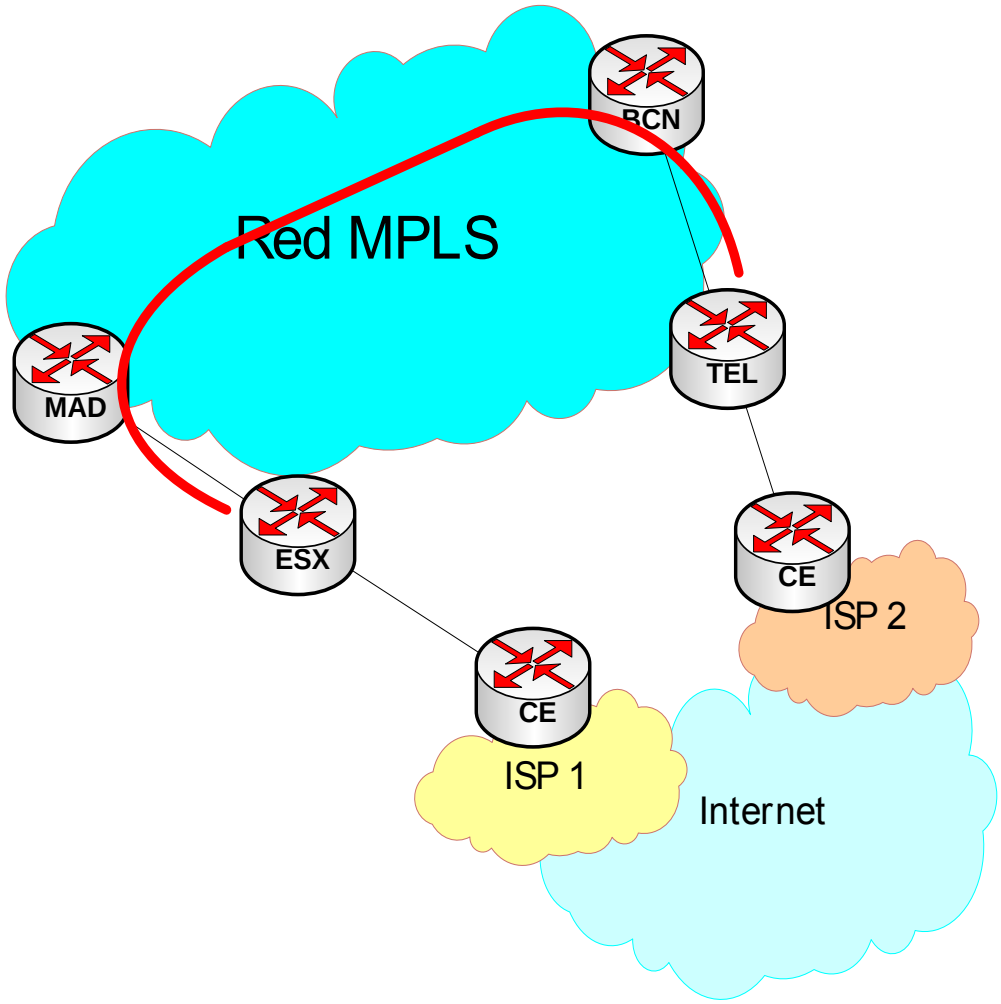

Ilustración 59- Esquema de la conexión a internet.

Antes de empezar a describir el proceso de unificación de ambas redes, es necesario añadir que hemos supuesto que debido a la imperiosa necesidad de intercambio de datos entre las dos compañías se establecieron dos enlaces entre ambas redes. Estos enlaces, aunque unen las dos redes MPLS son enlaces a nivel 3, es decir, por ellos no se habla LDP ni

ningún tipo de IGP interno. Sobre el enlace físico se establecieron subinterfaces para unir las VPNs que interesaba el intercambio de información. Los enlaces se establecieron de tal manera:

ASL ⇔ FER

CPII ⇔ AME

Se eligieron estos nodos por facilidad en la transmisión, no atiende a ningún otro diseño. Básicamente son los que están más cerca y son más fáciles de interconectar. La siguiente ilustración muestra el concepto de la interconexión.

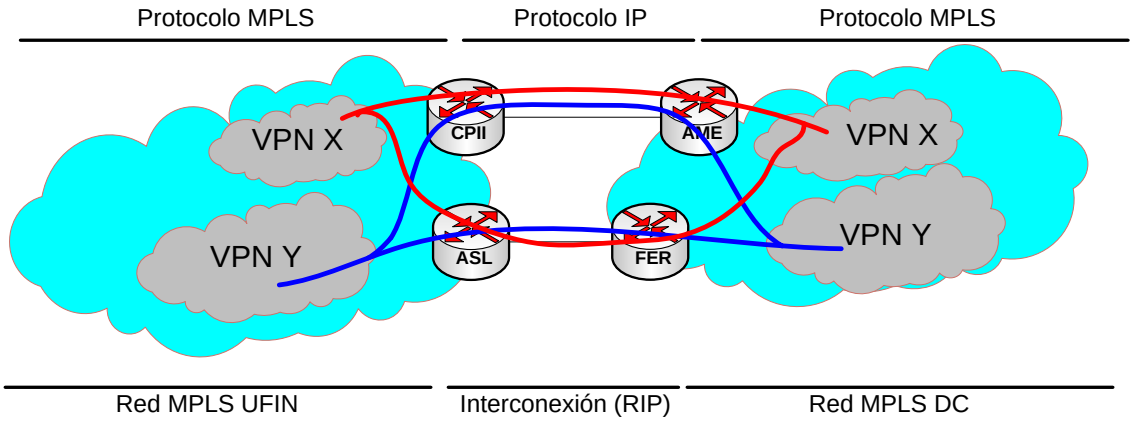

Ilustración 60- Conexión de nivel 3 de ambas redes previa a la integración

Por cuestiones de redundancia, la unión se hace por doble vía. Como protocolo de routing se utiliza RIP penalizando el peso de la ruta secundaria. Utilizamos RIP en vez de BGP (que es el protocolo más efectivo para estas conexiones) ya que, en previsión de una futura integración a nivel VPN MPLS, no se puede establecer una adyacencia eBGP si ambas redes tienen el mismo Sistema Autónomo. El tiempo de conmutación, en este esquema, es bastante alto (en torno a 3 minutos) pero debemos asumirlo para facilitar la posterior integración.

Estas conexiones aparecen en los esquemas puesto que existen e interconectan ambas redes pero al ser especiales están indicadas con la leyenda "ENLACE DE NIVEL 3"

4.2 Fase 0. Arquitectura física final

4.2.1 Topología de red

Tanto el la realización del proyecto como en esta memoria hemos separado el diseño de la implementación de la topología final de la red. Como estamos suponiendo que estamos en un caso real, no podemos disponer de los enlaces inmediatamente, se pueden producir retrasos en algunos por diversas cuestiones como pueden ser: solicitud de circuito de operador público, necesidad de hacer canalizaciones, impedimentos jurídicos, etc… Por eso, en este apartado definiremos el diseño teórico de cómo será la arquitectura final de los enlaces y la

implementación de los msimos se irá intercalando con el resto de fases, catalogados como "*Alta de enlace X (Implementación de la fase 0)*".

A la hora de elegir que nodos se unen con otros nodos, lo primero que debemos tener en cuenta es de donde a donde viajan los datos para dimensionar bien los enlaces. Al estar trabajando con tecnología MPLS el paquete de datos puede viajar por el camino que los nodos elijan según parámetros de capacidad, congestión, errores, etc. por lo que podemos hacer ingeniería de tráfico pero hasta un punto. En nuestro caso, y en general en cualquier red de comunicaciones, el segmento que más tráfico acumula son los CPDs ya que allí se encuentran los servidores de aplicaciones, tanto estándar como correo electrónico, servidores de ficheros, etc... como los propios del negocio de cada empresa.

Por ello, en nuestra red, la mayoría del tráfico se concentrará en CPII y en EDS, que son los nodos que se interconectan directamente con los CPDs. También es importante remarcar que los CPDs hablan entre sí, por cuestiones de redundancia principalmente, y que la disponibilidad entre ellos debe de ser alta también por lo que tendremos en cuenta los enlaces entre CPDs para diseñar nuestra red.

Por último analizaremos la conectividad del resto de PEs con los CPDs, más o menos todos tienen la misma carga de tráfico hacia los CPD por lo que no podemos establecer una jerarquía física. Cada servicio conectado a cada PE de la red MPLS puede acceder a los CPDs indistintamente por lo que establecer prioridades.

Otro aspecto a tener en cuenta es la red de transmisión de la que disponemos, algunos enlaces por muy operativo que fuese disponer de ellos son imposibles ya que no disponemos de transmisión y solicitar un enlace de la magnitud de Gigabit Ethernet a un operador público no es rentable.

En un escenario teórico libre, interconectaríamos los nodos de EDS y CPII y el resto de nodos conectados por doble vía a esos dos nodos. Como esto no es posible ya que estamos suponiendo un escenario real, con sus limitaciones de enlaces, adoptamos la medida de que ningún PE de la red estuviese a más de dos saltos (por enlaces de alta capacidad) de los nodos de EDS y CPII.

Otra tarea importante que hay que tener en cuenta a la hora de interconectar las redes es que se puede dar el caso de liberar recursos, dos nodos cercanos con posibilidad de liberar uno es un caso que se nos ha dado, esto es importante ya que no solo ahorras costes sino que una red es más estable cuantos menos elementos la compongan, además de que facilita mucho la operación y mantenimiento de la misma. Por cercanía y facilidad de la red de transmisión se puede migrar el acceso de los nodos de FER y de COR AME a otros nodos (ASL y AME) y eliminarlos de la red.

Así, la arquitectura final física, que se irá logrando durante todas las fases será la siguiente:

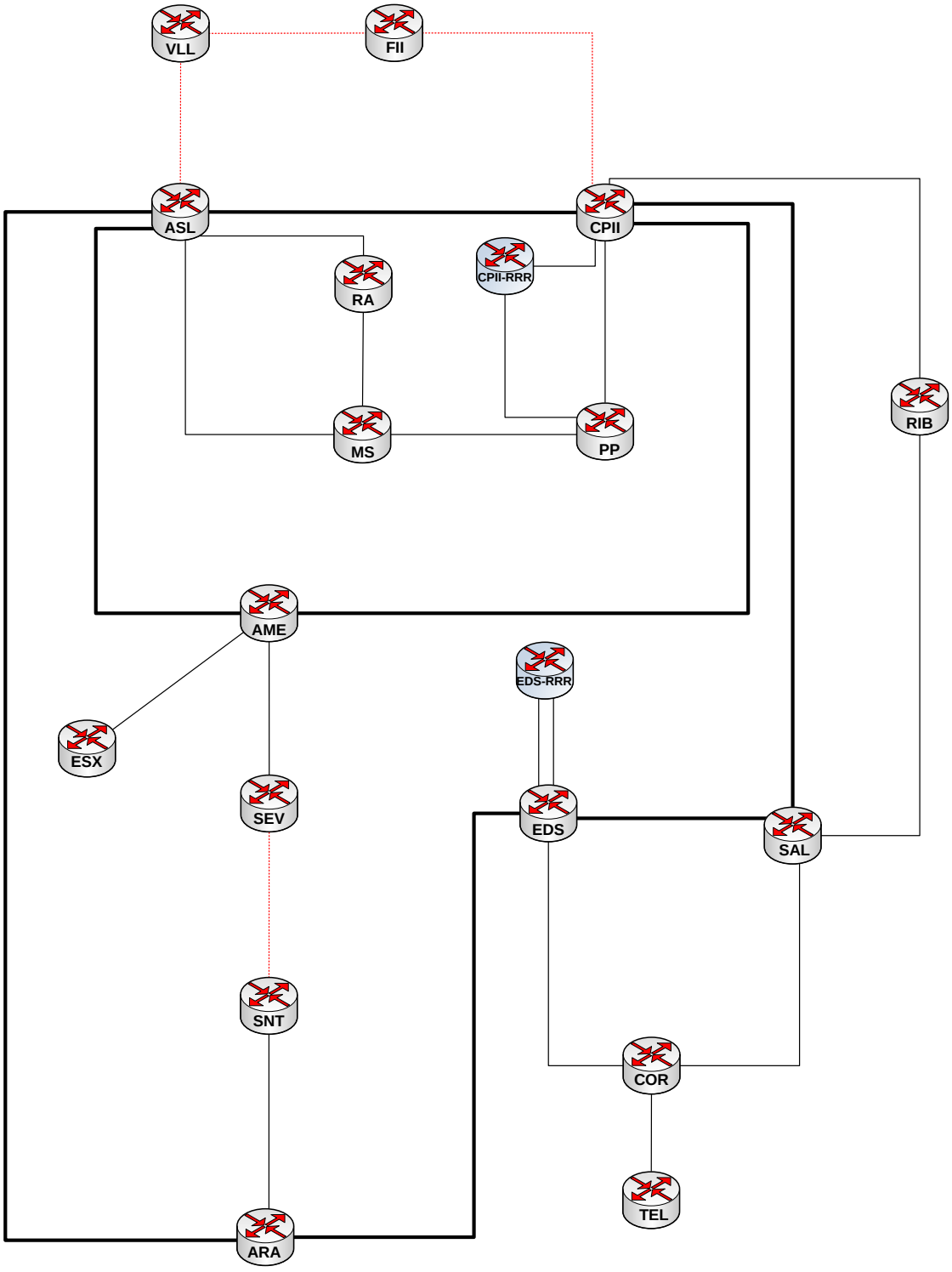

Ilustración 61- Arquitectura final planificada en un principio

Podemos apreciar las siguientes características:

- Anillos troncales formados por:

⇔ CPII ⇔ SAL ⇔ EDS ⇔ ARA ⇔ ASL ⇔ // ⇔AME ⇔ ASL ⇔ CPI ⇔

- ESX y TEL siguen cumpliendo la función que tenían cuando las redes estaban separadas que era la de tener la conexión a internet.
- MAD-LNS y BCN-LNS solo reciben los ADSLs de backup y están separados de la función de route reflector
- CLP-RRR y EDS-RRR son los route reflectors

4.2.2 PLAN DE DIRECCIONAMIENTO DE CORE

En una red MPLS, el direccionamiento de los enlaces PE – CE puede variar y dependiendo del uso, puede ser público o privado e incluso repetirse a nivel de MPLS VPN. Para la gestión de los equipos y para el funcionamiento del IGP se necesita conectividad IP entre los nodos en la tabla de routing global por robustez. Si no va a correr ningún otro servicio en la tabla de routing global, el direccionamiento puede ser privado dentro del plan de direccionamiento de la compañía. Pensando en la unificación de las redes, si el direccionamiento no se solapa entre ambos planes, como es el caso, se puede dejar el direccionamiento de enlace y de loopbacks de los nodos tal y como está, pero por cuestiones de orden y estilo se decide cambiar el direccionamiento de los enlaces y de las loopback para homogeneizarlo. Tenemos tres opciones:

1. Ajustar la red de DC al direccionamiento de UFIN.
2. Ajustar la red de UFIN al direccionamiento de DC.
3. Crear un nuevo plan de direccionamiento totalmente nuevo y ajustar ambas redes.

En este caso la elección se basó en varias cuestiones. La red de DC tiene más nodos por lo que es más laborioso cambiar su direccionamiento, además, estaba reservada una /24 para interconexiones y otra /24 para direcciones de loopback. En la red de UFIN había una /25 reservada tanto para interconexiones como para loopbacks y tiene menos nodos y enlaces. Como con una /24 para interconexiones y otra para loopbacks es suficiente se descarta crear un plan de direccionamiento nuevo por lo que se decidió que se usaría el plan de direccionamiento de la red de DC para usarlo como global.

El direccionamiento de gestión involucrado en los escenarios iniciales y final es el siguiente:

DC:

192.168.255.0/24 ➔ Loopbacks nodos MPLS.

192.168.1.0/24 ➔ Enlaces entre nodos.

UFIN:

10.132.2.0/25 ➔ Interconexión y Loopbacks nodos MPLS.

Final:

192.168.255.0/24 ➔ Loopbacks nodos MPLS.

192.168.1.0/24 ➔ Enlaces entre nodos.

El cambio de las direcciones IP de los enlaces y de las direcciones de loopback se hace junto con la actualización de la versión de IOS de cada nodo ya que, como es necesario reiniciar cada equipo, se alarga brevemente la ventana de trabajo y se aprovecha para cambiar el direccionamiento.

4.3 FASE 1. UNIFICACIÓN DE LAS REDES A NIVEL IP

En esta fase se establecen varios objetivos:

1.- Conexión a nivel IP de ambas redes

2.- Homogeneización del IGP (protocolo de routing interno).

3.- Homogeneización de la MTU

4.3.1 CONEXIÓN A NIVEL IP DE AMBAS REDES.

Para realizar esta unión hemos elegido los nodos de CPII y SAL (posteriormente formarán parte del anillo de CORE), se establece un enlace físico en fibra óptica entre ellos, el conector en ambos lados es GBIC/SFP 1000Base-Sx (hasta 70km). A este enlace le damos el direccionamiento 192.168.1.84/30 que es el siguiente libre en el plan de direccionamiento que vamos a usar, la 85 para Sal y la 86 para CPII.

Una vez conectados y configurados los puertos se comprueba la conectividad lanzando pings de diferentes tamaños. El % de éxito es del 100% en todos los casos y los tiempos de respuesta son los esperados:

```
SAL#ping ip 192.168.1.86 repeat 1000
success rate is 100 percent (1000/1000), round-trip min/avg/max = 8/10/24 ms

SAL#ping ip 192.168.1.86 repeat 1000 size 1500
Success rate is 100 percent (1000/1000), round-trip min/avg/max = 8/11/80 ms

SAL#show int ge-WAN ¼
GE-WAN1/4 is up, line protocol is up
Hardware is GigabitEthernet Interface, address is 0016.47e9.01c0 (bia
0016.47e9.01c0)
Description: ==== sal-rpe-01__GE1/4<--> MPLS-CPII__GE5/2 ====
Internet address is 192.168.1.85/30
MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec,
reliability 255/255, txload 1/255, rxload 1/255
Encapsulation ARPA, loopback not set
Keepalive not supported
Full-duplex, 1000Mb/s, link type is auto, media type is SX
output flow-control is unsupported, input flow-control is unsupported
ARP type: ARPA, ARP Timeout 04:00:00
Last input 00:00:10, output 00:00:10, output hang never
Last clearing of "show interface" counters never
Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
```

```
Queueing strategy: fifo
Output queue: 0/40 (size/max)
5 minute input rate 3000 bits/sec, 3 packets/sec
5 minute output rate 3000 bits/sec, 3 packets/sec
5071 packets input, 3376102 bytes, 0 no buffer
Received 2 broadcasts (0 IP multicasts)
0 runts, 0 giants, 0 throttles
0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
0 watchdog, 48 multicast, 0 pause input
5561 packets output, 3585140 bytes, 0 underruns
0 output errors, 0 collisions, 2 interface resets
0 babbles, 0 late collision, 0 deferred
0 lost carrier, 0 no carrier, 0 PAUSE output
0 output buffer failures, 0 output buffers swapped out
```

- Impacto:

Esta tarea no tiene impacto en el servicio ya que obviamente por ese nuevo enlace no hay otro tipo de tráfico que no sea señalización.

Con esta conexión la red total toma el siguiente aspecto:

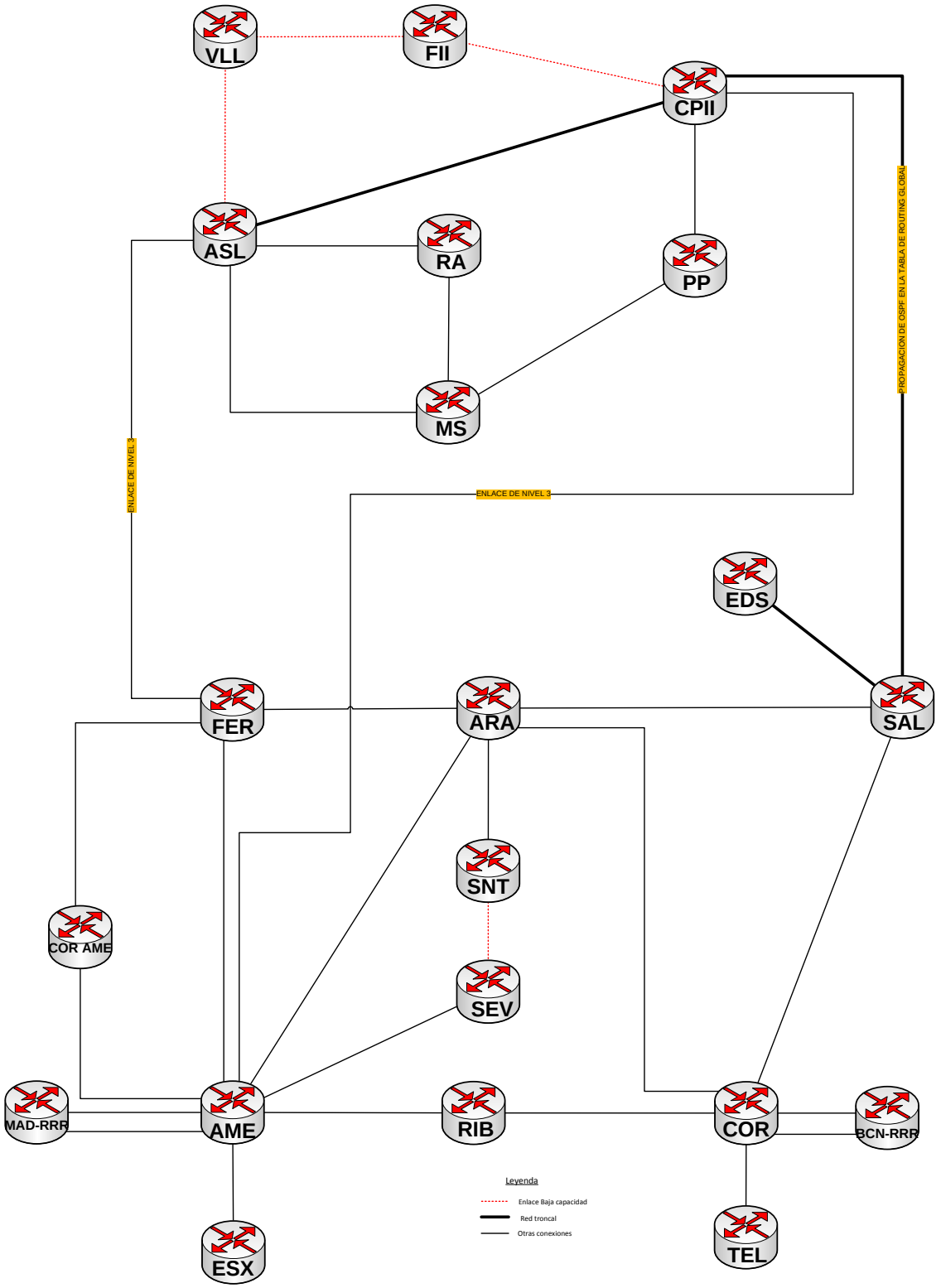

Ilustración 62- Red MPLS tras el enlace físico entre CPII y SAL

4.3.2 HOMOGENEIZACIÓN DEL IGP (PROTOCOLO DE ROUTING INTERNO)

En este momento y dado que ya hay conectividad IP se puede establecer un protocolo de routing entre ambas redes, la consecuencia de esto es que simplemente se conozcan a nivel IP en la tabla de routing global. No se pretende aún que hablen MPLS (intercambien etiquetas).

Es justo ahora cuando hay que decidir el protocolo de routing a elegir como IGP. Hay varias alternativas:

1. Dejar ISIS en el lado de DC y OSPF en el lado de UFIN:

Esto desde un punto de vista funcional es perfectamente correcto y estable pero no parece una buena solución en cuanto a estilo y mantenimiento dado que hay dos procesos de routing corriendo en la tabla de routing global y esto supone dificultad de mantenimiento.

2. Elección de cambiar los nodos de la red de UFIN a ISIS o los de DC a OSPF:

ISIS y OSPF son protocolos del estado de enlace, y ambos utilizan el mismo algoritmo de Dijkstra para calcular la mejor ruta a través de la red. Conceptualmente son similares. Ambos soportan máscaras de red de longitud variable y pueden utilizar multicast para descubrir vecinos. Usan paquetes “hello” y soportan autenticación par establecer adyacencias.

Mientras que el OSPF corre sobre IP, ISIS es un protocolo de capa de red según la torre ISO, este hecho pudo haber permitido que el OSPF sea más ampliamente utilizado. ISIS no utiliza IP para transportar mensajes.

Los routers construyen una representación topológica de la red. Este mapa indica las subredes que cada router ISIS puede alcanzar, y la trayectoria (más corta) del costo más bajo a una subred.

ISIS también se diferencia de OSPF en los métodos por los cuales intercambia mensajes para conocer la topología y cambiarla en caso de necesidad.

Puesto que OSPF es más popular, este protocolo tiene un sistema más rico de extensiones y de características agregadas. Sin embargo ISIS intercambia menos mensajes y puede escalar mejor por lo que es útil para redes más grandes. Dado el mismo sistema de recursos, ISIS puede soportar más routers en un área que OSPF. Esto hace que ISIS sea más útil en ambientes de ISP. Además, ISIS es neutral con respecto al tipo de direcciones de red que puede encaminar. OSPF en cambio, fue diseñado para IPv4. Así ISIS se adaptó más rápido a IPv6 mientras que OSPF necesitó una nueva versión OSPFv3.

A la vista de todo esto entre OSPF e ISIS se ha elegido OSPF por su mayor simplicidad de configuración, y dado el tamaño de la red (21 equipos) y que no se prevee un crecimiento significativo, se considera que OSPF es suficiente. Otro factor a tener en cuenta es que los operadores de la red, tienen más conocimiento de OSPF que de ISIS por lo que también se tuvo este punto muy en cuenta.

Una vez se ha decidido que OSPF será nuestro nuevo protocolo de routing interno hay que definirlo por completo en la red de DC. Una cosa importante a reseñar es que suponemos estar en un escenario de producción por lo que todos los trabajos se deben realizar de forma que no tengan impacto, o de ser inevitable, la ventana de trabajo debe ser lo menor posible, esto es un condicionante muy importante a la hora de realizar las tareas en red. Otro hándicap importante es que los trabajos se realizan en remoto por lo que es imprescindible hacerlos de

manera que no afecten a la gestión (p.e No se puede tocar la dirección por la que se gestiona el equipo)

En un escenario de desarrollo o maqueta, se iría nodo por nodo borrando ISIS y configurando OSPF, en un determinado momento se levantarían todas las máquinas, empezarían a intercambiar rutas y al cabo de unos instantes la red sería estable y segura.

Antes de describir como realizamos la implantación de OSPF es importante conocer el método de elección de rutas en los nodos MPLS ante dos rutas procedentes de dos protocolos de routing distintos, el router discrimina gracias al concepto de *distancia administrativa.* Este parámetro es configurable pero por defecto en los protocolos que nos interesa es de la siguiente forma:

Protocolo	Distancia administrativa
ISIS	115
OSPF	110

Tabla 8. –Distancia administrativa de ISIS y OSPF

Dado esto, vamos a ir nodo por nodo de la red de DC configurando OSPF pero con una distancia administrativa de 120, esto provoca que cada router va conociendo el resto de la red por ambos protocolos, prefiriendo siempre la ruta de ISIS. De esta forma tendríamos toda la red de DC conociendo a todos los nodos por dos protocolos de routing pero sin haberse producido impacto.

Para entender el siguiente ejemplo nos ponemos en situación. Ya se ha configurado OSPF en todos los routers y se ha cambiado la distancia administrativa de algunas redes. Esta es la tabla de rutas del nodo de SEV pero aún no se ha cambiado esa distancia.

```
SEV#show ip route
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is 192.168.1.9 to network 0.0.0.0

i L2 217.125.159.0/24 [115/200] via 192.168.1.9, GE-WAN1/1
     81.0.0.0/30 is subnetted, 1 subnets
i L2    81.46.0.40 [115/200] via 192.168.1.9, GE-WAN1/1
     80.0.0.0/23 is subnetted, 1 subnets
i L2    80.58.124.0 [115/200] via 192.168.1.9, GE-WAN1/1
     172.30.0.0/16 is variably subnetted, 5 subnets, 2 masks
```

```
i L2    172.30.254.0/23 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    172.30.252.0/23 [115/300] via 192.168.1.9, GE-WAN1/1
C       172.30.6.0/24 is directly connected, Vlan100
i L2    172.30.5.0/24 [115/250] via 192.168.1.69, GigabitEthernet5/2
i L2    172.30.8.0/24 [115/300] via 192.168.1.9, GE-WAN1/1
     192.168.200.0/32 is subnetted, 7 subnets
i L2    192.168.200.1 [115/100] via 192.168.1.9, GE-WAN1/1
i L2    192.168.200.2 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    192.168.200.3 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.200.4 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.200.5 [115/250] via 192.168.1.69, GigabitEthernet5/2
S       192.168.200.6 [1/0] via 172.30.6.1
i L2    192.168.200.8 [115/300] via 192.168.1.9, GE-WAN1/1
     77.0.0.0/8 is variably subnetted, 10 subnets, 6 masks
i L2    77.72.110.32/30 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    77.72.104.61/32 [115/400] via 192.168.1.9, GE-WAN1/1
i L2    77.72.104.60/32 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    77.72.110.11/32 [115/300] via 192.168.1.9, GE-WAN1/1
B       77.72.109.0/24 [200/0] via 10.201.5.5, 7w0d
i L2    77.72.108.0/24 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    77.72.108.0/22 [115/100] via 192.168.1.9, GE-WAN1/1
i L2    77.72.107.0/27 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    77.72.106.0/27 [115/200] via 192.168.1.9, GE-WAN1/1
B       77.72.111.128/25 [200/0] via 10.201.5.5, 7w0d
     10.0.0.0/8 is variably subnetted, 23 subnets, 4 masks
O E2    10.120.120.8/29 [120/50] via 192.168.1.9, 00:03:55, GE-WAN1/1
O E2    10.132.2.128/26 [120/50] via 192.168.1.9, 00:03:55, GE-WAN1/1
i L2    10.192.155.9/32 [115/200] via 192.168.1.9, GE-WAN1/1
O E1    10.132.1.252/30 [120/620] via 192.168.1.9, 00:03:55, GE-WAN1/1
O E1    10.132.1.248/30 [120/920] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.7/32 [120/501] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.6/32 [120/1001] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.5/32 [120/601] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.4/32 [120/901] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.3/32 [120/401] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.2/32 [120/501] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.1/32 [120/601] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.28/30 [120/1000] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.24/30 [120/500] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.20/30 [120/600] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.16/30 [120/600] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.40/30 [120/700] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.36/30 [120/600] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.32/30 [120/900] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.52/30 [120/500] via 192.168.1.9, 00:03:55, GE-WAN1/1
O       10.132.1.48/30 [120/1400] via 192.168.1.9, 00:03:55, GE-WAN1/1
i L2    10.201.5.4/30 [115/110] via 192.168.1.9, GE-WAN1/1
i L2    10.201.5.0/30 [115/210] via 192.168.1.9, GE-WAN1/1
     195.75.250.0/32 is subnetted, 1 subnets
i L2    195.75.250.38 [115/250] via 192.168.1.9, GE-WAN1/1
     192.168.255.0/32 is subnetted, 14 subnets
i L2    192.168.255.202 [115/400] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.201 [115/250] via 192.168.1.9, GE-WAN1/1
```

```
i L2    192.168.255.254 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.253 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.252 [115/400] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.7 [115/200] via 192.168.1.9, GE-WAN1/1
C       192.168.255.6 is directly connected, Loopback0
i L2    192.168.255.5 [115/250] via 192.168.1.69, GigabitEthernet5/2
i L2    192.168.255.4 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.3 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.2 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.1 [115/100] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.9 [115/400] via 192.168.1.9, GE-WAN1/1
i L2    192.168.255.8 [115/300] via 192.168.1.9, GE-WAN1/1
     192.168.0.0/24 is variably subnetted, 4 subnets, 3 masks
O       192.168.0.0/28 [120/3000] via 192.168.1.9, 00:03:58, GE-WAN1/1
O       192.168.0.16/30 [120/3000] via 192.168.1.9, 00:03:58, GE-WAN1/1
O       192.168.0.253/32 [120/3001] via 192.168.1.9, 00:03:58, GE-WAN1/1
O       192.168.0.254/32 [120/3001] via 192.168.1.9, 00:03:58, GE-WAN1/1
     192.168.1.0/30 is subnetted, 22 subnets
i L2    192.168.1.72 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.76 [115/400] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.64 [115/300] via 192.168.1.9, GE-WAN1/1
C       192.168.1.68 is directly connected, GigabitEthernet5/2
i L2    192.168.1.80 [115/400] via 192.168.1.9, GE-WAN1/1
O       192.168.1.84 [120/400] via 192.168.1.9, 00:03:58, GE-WAN1/1
i L2    192.168.1.40 [115/400] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.44 [115/400] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.32 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.36 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.56 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.60 [115/300] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.48 [115/400] via 192.168.1.9, GE-WAN1/1
C       192.168.1.8 is directly connected, GE-WAN1/1
i L2    192.168.1.12 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.0 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.4 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.24 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.16 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.20 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.252 [115/200] via 192.168.1.9, GE-WAN1/1
i L2    192.168.1.192 [115/250] via 192.168.1.9, GE-WAN1/1
     192.168.48.0/30 is subnetted, 1 subnets
i L2    192.168.48.252 [115/200] via 192.168.1.9, GE-WAN1/1
i*L2 0.0.0.0/0 [115/400] via 192.168.1.9, GE-WAN1/1
```

Resaltado en negrita y marcado con una ‘i’ están las rutas de los nodos MPLS que con los que aún habla por IS-IS ya que aún no se le ha cambiado la distancia administrativa. Marcados con una ‘O’ están las rutas de los nodos de la parte de UFIN (comienzan por 10.132.x.x) y los de la parte de DC que ya ha sido cambiada su distancia administrativa y por tanto, este nodo de SEV las prefiere ante las rutas de IS-IS.

En sombreado aparecen señaladas dos redes de ejemplo de 2 redes de interconexión que ya tiene cambiada la distancia administrativa y el nodo de SEV la ve aún por ISIS. Entonces cambiamos la distancia administrativa de IS-IS:

```
sev-rpe-01(config)#router ospf 2328
sev-rpe-01(config-router)# distance 110
sev-rpe-01(config-router)#
sev-rpe-01(config-router)#end
SEV#show ip route isis
```

Vemos que ya no se aprende nada por ISIS:

```
SEV#show ip route isis
```

Consultamos de nuevo la tabla de rutas OSPF:

```
SEV#show ip route ospf
O E1 217.125.159.0/24 [110/1200] via 192.168.1.9, 00:00:13, GE-WAN1/1
     81.0.0.0/30 is subnetted, 1 subnets
O       81.46.0.40 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
     80.0.0.0/23 is subnetted, 1 subnets
O E1    80.58.124.0 [110/1200] via 192.168.1.9, 00:00:13, GE-WAN1/1
     172.30.0.0/16 is variably subnetted, 5 subnets, 2 masks
O E1    172.30.254.0/23 [110/1201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    172.30.252.0/23 [110/1300] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       172.30.5.0/24 [110/301] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       172.30.8.0/24 [110/301] via 192.168.1.9, 00:00:13, GE-WAN1/1
     192.168.200.0/32 is subnetted, 7 subnets
O E1    192.168.200.1 [110/1100] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    192.168.200.2 [110/1300] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    192.168.200.3 [110/1200] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    192.168.200.4 [110/1200] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    192.168.200.5 [110/1300] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    192.168.200.8 [110/1300] via 192.168.1.9, 00:00:13, GE-WAN1/1
     77.0.0.0/8 is variably subnetted, 9 subnets, 6 masks
O E1    77.72.110.32/30 [110/1300] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       77.72.104.61/32 [110/401] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       77.72.104.60/32 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    77.72.110.11/32 [110/1300] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    77.72.108.0/22 [110/1100] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    77.72.107.0/27 [110/1300] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    77.72.106.0/27 [110/1201] via 192.168.1.9, 00:00:13, GE-WAN1/1
     10.0.0.0/8 is variably subnetted, 23 subnets, 4 masks
O E2    10.120.120.8/29 [110/50] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E2    10.132.2.128/26 [110/50] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    10.192.155.9/32 [110/1201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    10.132.1.252/30 [110/620] via 192.168.1.9, 00:00:13, GE-WAN1/1
O E1    10.132.1.248/30 [110/920] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.7/32 [110/501] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.6/32 [110/1001] via 192.168.1.9, 00:00:13, GE-WAN1/1
```

```
O       10.132.1.5/32 [110/601] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.4/32 [110/901] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.3/32 [110/401] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.2/32 [110/501] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.1/32 [110/601] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.28/30 [110/1000] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.24/30 [110/500] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.20/30 [110/600] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.16/30 [110/600] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.40/30 [110/700] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.36/30 [110/600] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.32/30 [110/900] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.52/30 [110/500] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.132.1.48/30 [110/1400] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.201.5.4/30 [110/101] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       10.201.5.0/30 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
     195.75.250.0/32 is subnetted, 1 subnets
O E1    195.75.250.38 [110/1200] via 192.168.1.9, 00:00:13, GE-WAN1/1
     192.168.255.0/32 is subnetted, 14 subnets
O       192.168.255.202 [110/401] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.201 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.254 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.253 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.252 [110/401] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.7 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.5 [110/301] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.4 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.3 [110/201] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.2 [110/301] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.1 [110/101] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.9 [110/401] via 192.168.1.9, 00:00:13, GE-WAN1/1
O       192.168.255.8 [110/301] via 192.168.1.9, 00:00:13, GE-WAN1/1
     192.168.0.0/24 is variably subnetted, 4 subnets, 3 masks
O       192.168.0.0/28 [110/3000] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.0.16/30 [110/3000] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.0.253/32 [110/3001] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.0.254/32 [110/3001] via 192.168.1.9, 00:00:15, GE-WAN1/1
     192.168.1.0/30 is subnetted, 22 subnets
O       192.168.1.72 [110/300] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.76 [110/400] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.64 [110/300] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.80 [110/400] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.84 [110/400] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.40 [110/400] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.44 [110/400] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.32 [110/300] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.36 [110/300] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.56 [110/300] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.60 [110/300] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.48 [110/400] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.12 [110/200] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.0 [110/200] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.4 [110/200] via 192.168.1.9, 00:00:15, GE-WAN1/1
```

```
O       192.168.1.24 [110/200] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.16 [110/200] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.20 [110/200] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.252 [110/201] via 192.168.1.9, 00:00:15, GE-WAN1/1
O       192.168.1.192 [110/200] via 192.168.1.9, 00:00:15, GE-WAN1/1
     192.168.48.0/30 is subnetted, 1 subnets
O E1    192.168.48.252 [110/1200] via 192.168.1.9, 00:00:15, GE-WAN1/1
O*E1 0.0.0.0/0 [110/500] via 192.168.1.9, 00:00:15, GE-WAN1/1
```

Aquí ya se puede ver que ha aprendido las nuevas rutas por OSPF y las deja de aprender por IS-IS.

Una vez se configuran todos los routers de la red de DC la situación es que la parte de UFIN tiene únicamente OSPF como protocolo de routing y la parte de DC tiene ISIS y OSPF prefiriendo OSPF.

Una vez hemos llegado a esta situación solo hay que borrar ISIS nodo por nodo, esto no provoca ningún tipo de corte de servicio ya que el router ya posee las rutas de sus vecinos por OSPF y en ningún momento dejan de ser accesibles.

Como curiosidad, a la hora de configurar OSPF se elige una etiqueta que identifica el proceso, esta etiqueta es local, hemos elegido 2328 por ser el número de la RFC de OSPF, en un futuro se cambiará en la parte de UFIN para homogeneizarlo. Esto ha sido así ya que en la red de DC hay varios procesos OSPF corriendo y ya está cogido el identificativo 1 ya que es típicamente el configurado. De cara a homogeneizar la red en el número de proceso de OSPF, durante la actualización de las versiones de IOS se cambiará este número ya que implica corte de servicio en todo el nodo y se aprovecha el reinicio del cambio de versión para ello.

Respecto a la política de pesos se ha utilizado la filosofía de la red de UFIN, todos los enlaces tienen coste 100 salvo los 3 de baja capacidad (VLL –ASL y FII – CPII que son POS sobre STM1 y SEV – SNT que es un FastEthernet) que tienen coste de 500. Aplicando reglas dinámicas teniendo en cuenta estos cálculos la red MPLS tomará las decisiones de enrutamiento.

Un detalle importante que se nombró en la situación inicial de las redes, es que en la parte de DC los nodos de TEL y ESX tienen la conexión a internet y sobre la tabla de routing global, esto implica que hay que configurar el OSPF de estos equipos como ruta por defecto de la red, aplicando pesos de 100 (TEL) y 1000 (ESX) para seguir enrutando internet correctamente.

4.3.3 HOMOGENEIZACIÓN DE LA MTU

Otra de las tareas que componen esta primera fase es homogeneizar la MTU en todos los enlaces. Como vimos en la parte de teoría, en una red MPLS hay dos conceptos relativos a la MTU, la MTU física del enlace y la MTU MPLS, por defecto estos valores son iguales y nosotros lo mantendremos así en las dos redes. Esta homogeneización, por tanto, corresponde a la MTU global del enlace.

Si consideramos os estándares de MTU para cada tipo de enlace, se establecerian de la siguiente manera:

- GigabitEthernet: 1500
- POS (STM1): 4470

Esta MTU se puede cambiar acorde a las necesidades de la red ya que hay aplicaciones sensibles a la fragmentación y sobre todo que no se puede fragmentar si se dan servicios de nivel 2.

En la red de UFIN la MTU está establecida en 1542 para los enlaces GE-WAN y 4470 para los enlaces POS (STM1).

En la red de DC la MTU está establecida en todos los enlaces en 4470 salvo un enlace (ARA – SNT) que está a 1900.

A la hora de tomar una decisión sobre la MTU hemos tenido en cuenta los siguientes criterios:

1.- Aplicaciones sensibles: En la red de UFIN hay establecida una VPN de nivel 3 que no permite fragmentación, esto hace imposible establecer una MTU inferior a 1542.

2.- Servicios de nivel 2 en cada red: En la red de DC hay un pseudowire que cuyo tamaño mínimo de MTU es 1600.

3.- Un enlace troncal, el que hay entre SNT y ARA a pesar de ser un interfaz GE-WAN corre sobre una red de transmisión de Alcatel que pone una limitación de MTU de 1900.

A la vista de esto, se ha tomado la decisión de establecer el tamaño de MTU en 1900 unificando todos los enlaces salvo los dos enlaces POS sobre SDH debido a la naturaleza física de los mismos. Así no habrá nunca varias fragmentaciones dentro de la red MPLS mientras no intervengan los enlaces POS

La MTU es un parámetro que se configura por interfaz, por lo que hay que trabajar en paralelo en los interfaces de nodos vecinos que comparten el enlace.

La configuración es sencilla, en el ejemplo vamos a cambiar el enlace entre CPII y ASL que sigue el siguiente esquema:

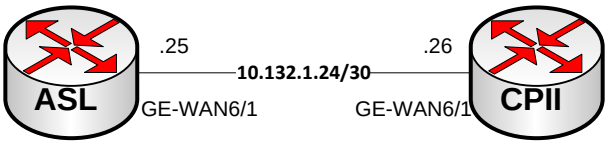

Ilustración 63- Enlace Gigabit Ethernet ASL-CPII

Basta con introducir el siguiente comando en los puertos que unen ambos nodos:

```
MPLS-CPII#configure terminal
MPLS-CPII(config)#interface GE-WAN6/1
MPLS-CPII (config-if)#mtu 1900
```

```
MPLS-ASL#configure terminal
MPLS-ASL (config)#interface GE-WAN6/1
MPLS-ASL (config-if)#mtu 1900
```

Comprobamos en CPII que el enlace sigue levantado:

```
MPLS-CPII#sh ip inter brief | i 6/1
GE-WAN6/1              10.132.1.26     YES NVRAM  up                    up
```

Y entonces hacemos un ping con la opción "df-bit" (don't fragment) que quiere decir que envíe el ping pero que si se ha de fragmentar el paquete en alguno de los enlaces que lo descarte. La opción "size" indica el tamaño del paquete ICMP que queremos enviar.

Comprobamos que un paquete de 1900 bytes llega sin problemas:

```
MPLS-CPII#ping 10.132.1.25 df-bit size 1900
Type escape sequence to abort.
Sending 5, 1900-byte ICMP Echos to 10.132.1.25, timeout is 2 seconds:
Packet sent with the DF bit set
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/3/4 ms
```

Pero en el momento que aumentamos en un byte el paquete, este no llega al destino pese a estar directamente conectado.

```
MPLS-CPII#ping 10.132.1.25 df-bit size 1901
Type escape sequence to abort.
Sending 5, 1901-byte ICMP Echos to 10.132.1.25, timeout is 2 seconds:
Packet sent with the DF bit set
.....
Success rate is 0 percent (0/5)
```

- Impacto:

De nuevo a la hora de realizar estos trabajos hubo que planificarlos de modo que no hubiese corte de servicio. Esta vez no fue difícil ya que si se cambia la MTU a la vez en los dos puertos que comparten enlace, las adyacencias de OSPF no caen ya que el intercambio de paquetes es pequeño y no se produce impacto en el servicio.

En este punto, la foto de las redes a estas alturas es que están unidas por un enlace solo a nivel IP (no MPLS) y se habla el mismo protocolo de routing IGP y los enlaces tienen la misma MTU

4.3.4 SUSTITUCIÓN DEL NODO DE EDS (IMPLEMENTACIÓN FASE 0)

Previamente a la fase 2 se va a insertar el nuevo nodo de EDS. El nodo de EDS es de la serie 7200 en vez de la 7600 como en el resto de PEs de la red. Según la arquitectura establecida, el nodo EDS va ser uno de los puntos centrales de comunicación con el CPD. Si bien un 7200 puede cumplir perfectamente las funciones de PE decidimos sustituirlo ya que, según veremos en el punto 4.5.1, utilizamos esta máquina para hacer de route reflector y en

su lugar ponemos un 7600 del que se disponía para una ampliación de la red existente. Además con este cambio, todos los nodos que hacen de PE son 7600 y queda la red más homogénea en lo que a hardware se refiere.

La sustitución de este nodo es relativamente sencilla ya que la funcionalidad de PE de este nodo se supone mínima, es decir, tiene pocos accesos, apenas la conexión con el CPD en varias VPNs.

Antes de nada hay que instalar, alimentar y verificar a nivel de hardware este equipo antes de integrarlo en la red. La verificación del hardware la realizamos utilizando el test GOLD (Generic Online Diagnostic)

La configuración del nuevo nodo se basa en la que tenía el antiguo (7200), cambiando algunos comandos básicos que al cambiar de plataforma y versión de IOS son diferentes.

La sustitución se realizará en cinco pasos:

4.3.4.1 Conexión física con ARA

Inicialmente se pone en servicio el enlace con ARA mientras que el enlace con SAL se cambiará una vez se haya realizado la integración en red y la migración de los servicios.

El escenario local es el siguiente para entender mejor los ejemplos:

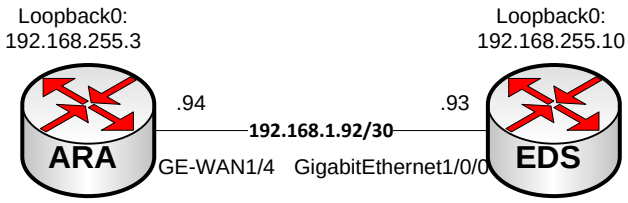

Ilustración 64- Conexión transitoria ARA – EDS

4.3.4.2 Comprobación del enlace físico y conectividad

Una vez se ha establecido el enlace con ARA lo comprobamos mediante un "show interfaces" y un ping.

```
EDS#show interfaces gi1/0/0
GigabitEthernet1/0/0 is up, line protocol is up
Hardware is GigEther SPA, address is 001e.f741.e000 (bia 001e.f741.e000)
Description: ==== eds-rpe-01__Giga1/0/0 <--> ara-rpe-01__GE1/4 ====
Internet address is 192.168.1.93/30
MTU 1900 bytes, BW 1000000 Kbit, DLY 10 usec,
reliability 255/255, txload 1/255, rxload 1/255
Encapsulation ARPA, loopback not set
Keepalive not supported
Full Duplex, 1000Mbps, link type is auto, media type is SX
output flow-control is unsupported, input flow-control is unsupported
Carrier delay is 0 msec
ARP type: ARPA, ARP Timeout 04:00:00
Last input 00:00:04, output 00:00:00, output hang never
```

```
Last clearing of "show interface" counters never
Input queue: 0/75/4/0 (size/max/drops/flushes); Total output drops: 0
Queueing strategy: Class-based queueing
Output queue: 0/40 (size/max)
5 minute input rate 5000 bits/sec, 4 packets/sec
5 minute output rate 5000 bits/sec, 5 packets/sec
7153916 packets input, 792682899 bytes, 0 no buffer
Received 338 broadcasts (0 IP multicasts)
0 runts, 0 giants, 0 throttles
4 input errors, 0 CRC, 4 frame, 0 overrun, 0 ignored
0 watchdog, 369 multicast, 0 pause input
840 packets output, 589 bytes, 0 underruns
0 output errors, 0 collisions, 3 interface resets
0 babbles, 0 late collision, 0 deferred
0 lost carrier, 0 no carrier, 0 pause output
0 output buffer failures, 0 output buffers swapped out
```

Comprobamos que está levantado y que negocia a la velocidad máxima del enlace 1000 Mbps y full-dupplex. Con un ping vemos que el enlace no pierde tráfico:

```
EDS#ping 192.168.1.94  size 1500 repeat 1000
Type escape sequence to abort.
Sending 1000, 1500-byte ICMP Echos to 192.168.1.94, timeout is 2 seconds:
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!
Success rate is 100 percent (1000/1000), round-trip min/avg/max = 8/9/112 ms
```

4.3.4.3 Comprobación de los procesos OSPF

Comprobamos que se han establecido las adyacencias de OSPF correctamente:

```
EDS#sh ip ospf neighbor
Neighbor ID     Pri   State           Dead Time   Address         Interface
192.168.255.3     0   FULL/  -00:00:38    192.168.1.94    GigabitEthernet1/0/0
192.168.255.8     0   FULL/  -00:00:38    192.168.1.77    GigabitEthernet2/0/0
192.168.43.254    0   FULL/  -            -       192.168.43.254  OSPF_SL19
192.168.43.200    0   FULL/  -            -       192.168.43.200  OSPF_SL18
192.168.43.199    0   FULL/  -            -       192.168.43.199  OSPF_SL17
192.168.43.198    0   FULL/  -            -       192.168.43.198  OSPF_SL16
```

```
192.168.43.197    0   FULL/  -           -        192.168.43.197  OSPF_SL15
192.168.43.196    0   FULL/  -           -        192.168.43.196  OSPF_SL14
192.168.43.195    0   FULL/  -           -        192.168.43.195  OSPF_SL13
192.168.43.194    0   FULL/  -           -        192.168.43.194  OSPF_SL12
192.168.43.193    0   FULL/  -           -        192.168.43.193  OSPF_SL11
10.2.254.85       0   FULL/  -00:00:35    192.168.45.42   GigabitEthernet3/2
192.168.43.252    0   FULL/  -           -        192.168.43.252  OSPF_SL21
192.168.43.253    0   FULL/  -           -        192.168.43.253  OSPF_SL20
```

4.3.4.4 Migración de las conexiones con los CPEs

Una vez integrado el nodo de EDS, ponemos fuera de servicio los interfaces del nodo antiguo (el 7200) y nos proponemos a migrar las conexiones de los CPEs al nuevo nodo (7600). Los nuevos interfaces físicos del 7600 están configurados de manera idéntica que los del 7200. Una vez migradas las conexiones comprobamos las adyacencias de EIGRP (es el protocolo utilizado en esta conexión PE-CPE) en la tabla de rutas, tanto del CPE como del nodo que se están intercambiando rutas adecuadamente.

4.3.4.5 Comprobación física del enlace con SAL

Y por último, establecemos el enlace que faltaba EDS – SAL. Localmente las conexiones quedan de la siguiente manera:

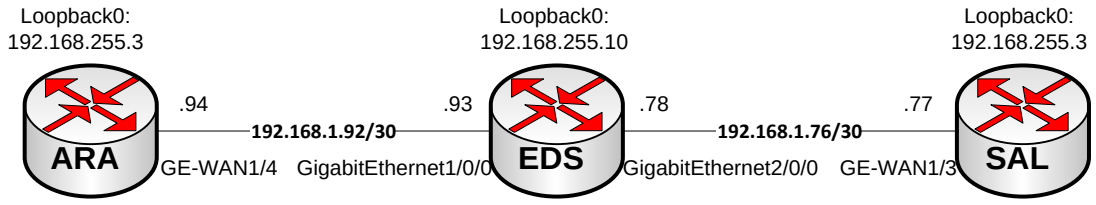

Ilustración 65- Conexión definitiva EDS

Tras esta sustitución, el esquema de toda la red es el siguiente:

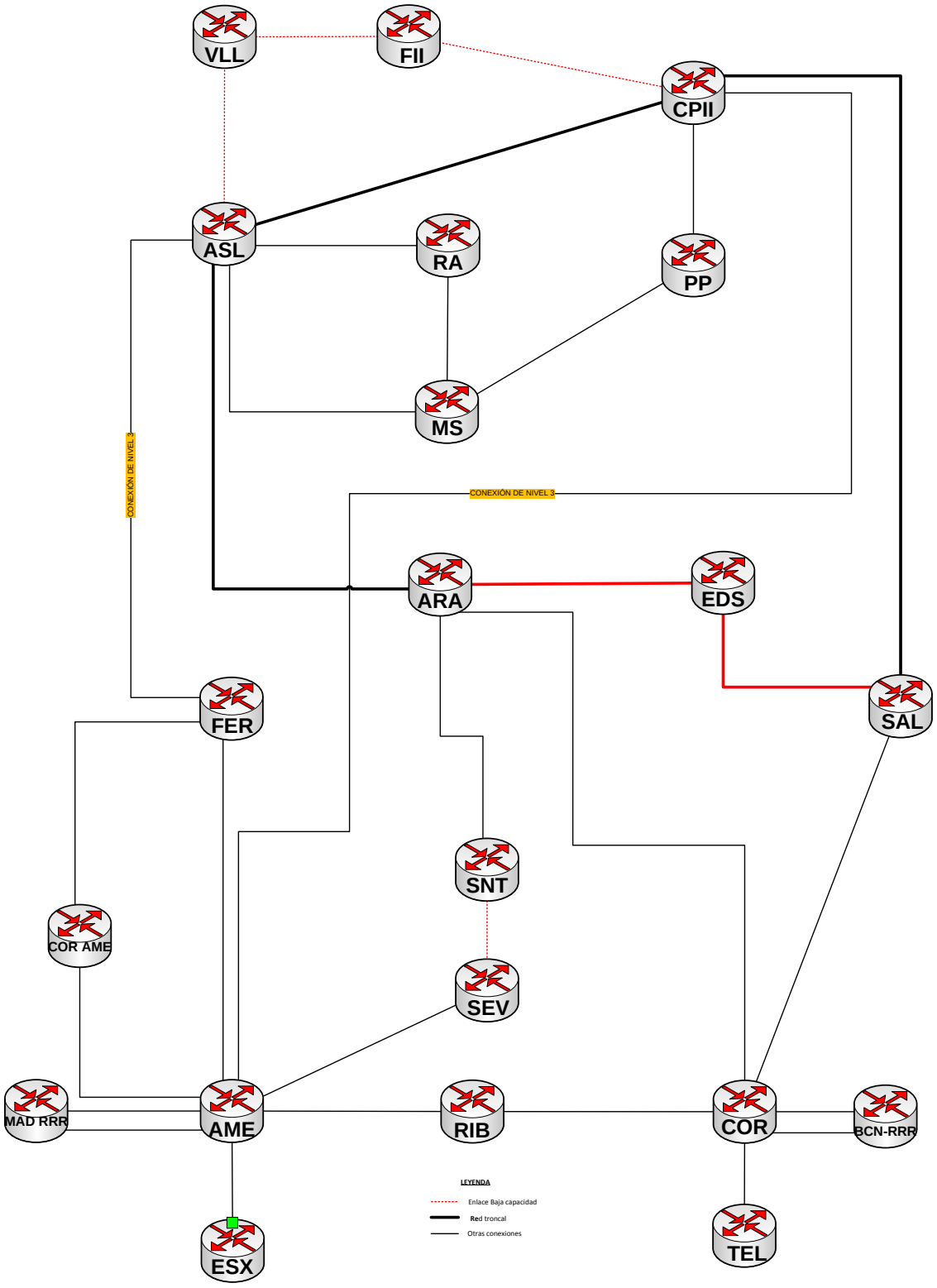

Ilustración 66- Red MPLS tras la inserción del nodo de EDS

- Impacto:

El impacto sobre la red en este cambio es un corte de servicio de unos 20 minutos en los CPEs conectados a este PE. Es el tiempo que transcurre en la reconexión física de los CPEs y el tiempo de actualización de las rutas de los protocolos dinámicos que se establecen entre el nodo y los CPEs.

4.4 FASE 2. UNIFICACIÓN DE LAS REDES A NIVEL MPLS

Para tener una red unificada a nivel MPLS es necesario propagar las etiquetas entre las dos redes existentes. Para ello, es necesaria la propagación de etiquetas.

La propagación de etiquetas entre las dos zonas de red requiere la configuración de LDP entre los nodos conectados. Esta tarea debe realizarse en el enlace entre los nodos de CPII y SAL. Igualmente deberá ser realizada en los nuevos enlaces que se vayan estableciendo según la arquitectura diseñada.

El protocolo de distribución de etiquetas es parte fundamental de la red MPLS por lo que debe ser unificado en toda la red. Existen dos diferencias de configuración entre ambas redes. En primer lugar en la parte de red de UFIN se está empleando sesiones targeted para los servicios de nivel dos por lo que será necesario añadir dicha configuración en el lado de DC. Una sesión targeted es una adyacencia de LDP entre nodos no directamente conectados. Si bien no es necesario establecer esta sesión para la distribución de etiquetas, si es necesario si se pretenden dar servicios de AToM, servicios de nivel 2. Dado que en la red de UFIN si se daban estos servicios se extiende esta capacidad a la nueva parte de DC. Se podría dejar sin configurar en la parte de DC pero no podríamos hablar de una red homogénea si no se pueden establecer los mismos servicios entre cualquier origen y destino de la red.

Por otro lado en la parte de red de DC se ha configurado la opción “mpls ldp explicit-null”. Como vimos en la parte de teoría, el uso de explicit-null implica que aunque sea el último salto, se ha de añadir una etiqueta con el campo de etiqueta a 0 pero respetando el resto de campos, en particular el de EXP que es donde residen los parámetros de QoS. El uso de esta función está relacionado con la forma de implementar la calidad de servicio en la red, lo que hace es que mantiene la etiqueta MPLS para respetar la calidad de servicio en función del campo EXP hasta el último salto. En función de si se desea aplicar la calidad de servicio en base a la precedencia contenida en el paquete IP o en los bits experimentales de MPLS se usará implicit-null o explicit null label.

Con esta situación de partida, en esta tarea:

1. Configuraremos LDP con soporte para sesiones targeted en los nodos de DC para soporte de los servicios de nivel 2.

Al configurar LDP en las dos partes del enlace, como medida de seguridad pondremos una clave

```
mpls ldp neighbor <peer IP add> password xxxxxx
```

En realidad esto no es necesario ya que estamos en un entorno privado y totalmente controlado, es muy remota la posibilidad de que se inserte un nuevo router en la red y se conecte a otro nodo de la red de core pero dado que la versión de IOS de los equipos nos lo permite lo configuraremos igualmente. La clave es local a cada enlace por lo que no es necesario tener un cuidado especial con su elección.

Para añadir más seguridad y de nuevo no es necesario dado el entorno en el que nos manejamos, limitaremos la aceptación de paquetes "hello" de otros nodos mediante una lista de acceso.

```
mpls ldp discovery targeted-hello accept from ACL_THELLO_LDP
ip access-list standard ACL_THELLO_LDP
permit 10.132.1.7
:
permit 10.132.1.1
```

Es importante reseñar que cada vez que se añada un nuevo nodo es necesario incluirlo en la lista de acceso del resto de nodos.

2. Forzaremos el router-id en los equipos de la red de DC en los que no está explícitamente configurado como loopback0.

Esto se hace para tener más controlado la identificación del nodo ante sus vecinos. De no fijarlo, el propio nodo sigue los criterios por defecto para su asignación.

3. Uniformizaremos la opción de "mpls ldp explicit-null" en todos los nodos conforme los requerimientos del QoS definidos por los servicios prestados

Esta tarea es incluir únicamente un comando a nivel global.

- Impacto:

Estas tareas no suponen impacto ninguno ya que en el caso de la tarea 1, establecer una nueva sesión LDP entre nodos no supone corte ya que no hay ningún servicio por ese enlace.

En la tarea 2 el impacto es una posible caída de los enlaces sobre los cuales hay establecidas sesiones targeted con el nodo que se fija, en caso de que el router-id sea distinto del que tenía establecido. Al ser la red MPLS dinámica en cuanto al routing del tráfico no supone impacto sobre ningún servicio ofrecido.

La tarea 3 tampoco supone corte de servicio ya que el campo IP precedence y el campo EXP de la etiqueta MPLS tiene el mismo valor por tanto no supone diferencia el estar o no estar.

4.5 FASE 3. UNIFICACIÓN DE LAS REDES A NIVEL MPLS VPN

Dentro de esta fase se contemplan las siguientes tareas:

1.- Establecer la arquitectura definitiva en lo que implica a los route reflectors
2.- Propagación del MP-BPG

4.5.1 ESTABLECER LA ARQUITECTURA DEFINITIVA EN LO QUE IMPLICA A LOS ROUTE REFLECTORS

Los route reflectors son los encargados de establecer el routing de la red a nivel de VPNs y como tal son parte vital de una red MPLS. Con las redes separadas teníamos un par de RRs por red, en el caso de la red de UFIN estos además tenían función de RR, P y PE, en el caso de la red de DC hacen de RR y de LNS.

4.5.1.1 Elección de la arquitectura

En la red de UFIN eran los nodos de ASL y CPII los encargados de esta misión, por el lado de DC son los BCN-RRR y MAD-RRR.

Teníamos varias alternativas:

1. Dejar todos los nodos que actúan como RR con su función: Esto implicará redundancia cuádruple en la red, no lo estimamos necesario aparte de que se complica la arquitectura y el mantenimiento.
2. Dejar los RR de la red de UFIN como RR: Esto supondría pasar toda la red de DC a depender de los dos RR de la red de UFIN. Dado que no se pueden liberar de su función de PE, los RR de la red de UFIN, ASL y CPII no pueden absorber la misión de ser a la vez PE y RR de toda la red.
3. Dejar los RR de la red de DC como RR: Esto supondría pasar toda la red de UFIN a depender de los RR de la red de DC. Al igual que en el caso de la red de UFIN estos equipos tienen otra función, hacen de LNS para los backup por ADSL.
4. La última opción es establecer dos equipos dedicados separados de otras funciones. Esto es importante ya que se separa la funcionalidad de los equipos y si los RR están bien configurados no es necesario tocar su configuración salvo que se añadan nuevos nodos a la red. Cualquier nodo PE está cambiando la configuración continuamente ya que siempre se están dando de alta y de baja servicios, esto implica cierto peligro ya que cualquier error en la configuración puede provocar la caída de los nodos más importantes de la red. Su ubicación es importante ya que tienen que tener buena conectividad con todos los nodos de la red (la ubicación la detallaremos a continuación)

4.5.1.2 Elección de la plataforma

Una vez nos decidimos por la cuarta opción, detallamos los requisitos necesarios que deben cumplir los equipos a utilizar como Route Reflector en el core de la nueva red integrada

y las recomendaciones resultantes de contrastar dichas especificaciones con la gama de equipos disponibles por Cisco.

Los equipos que vayan a realizar estas funciones deben soportar como mínimo las siguientes funciones:

- Soporte BGP Mulitprotocolo
- Soporte MPLS
- Soporte de funcionamiento Route-Reflector sobre MPBGP y con varios VRF
- Soporte LDP
- Soporte de un número elevado de rutas por si se desea usarlos para gestionar el routing de Internet.

Además de estas funciones es conveniente que el equipo sea compatible al máximo con la base instalada para favorecer las labores de operación y mantenimiento de la red.

La gama de equipos instalados que podrían adaptarse a estas funciones incluye los Cisco 7600, Cisco 7200 y Cisco 3800, por tanto el equipo seleccionado debe pertenecer a una de estas gamas.

De las tres familias de routers, los Cisco 7600 tienen una diferencia importante de funcionamiento respecto a los otros dos: el routing y otros procesos de nivel tres se hacen por hardware. Esto tiene consecuencias favorables y desfavorables.

- Favorables: Todos esos procesos se realizan generalmente más rápido que cuando se hacen por hardware.
- Desfavorables: Las funcionalidades y capacidades están limitadas por la implementación hardware y ninguna actualización de software puede proporcionar funciones no soportadas por hardware.

En este último caso hay que considerar el tamaño de la tabla de rutas. El uso de un equipo 7600 incluso si acepta el tamaño de la tabla de rutas actual puede plantear limitaciones en un futuro no muy lejano.

Por esta razón descartamos el uso de la familia 7600 como route reflector de la nueva red.

Un route reflector no necesita un número elevado de interfaces. Por ese motivo de cada gama se ha elegido un equipo que manteniendo toda la potencia de proceso no tenga un elevado número de interfaces lo que conlleva un ahorro importante.

Entonces hemos seleccionado un equipo básico de ambas gamas: Cisco 7201 y Cisco 3825.

La serie Cisco 7200 soporta una CPU modular y actualizable y un motor de conmutación de paquetes basado en hardware que ofrece la máxima flexibilidad y capacidad de escalado en el rendimiento. El módulo procesador NPE-G2 ofrece una capacidad de

conmutación de dos millones de paquetes por segundo (2 mpps). La serie 7200 admite numerosas opciones de conectividad, muchas de las cuales no están soportadas en la serie Cisco 3800, incluyendo E3 canalizado, POS, SS7, emulación de circuito OC-3 y FDDI (si bien no utilizaremos este tipo de conexiones en los RR). La serie 3800 está orientada a las oficinas de mediana y gran empresa que requieren un alto grado de servicios y aplicaciones. Los equipos 7200 son dispositivos de agregación para instalar en nodos que requieran una alta densidad de conectividad de nivel medio o varias líneas de conectividad de alta velocidad, como nodos de agregación regional o y nodos centrales corporativos.

La siguiente tabla compara las características de ambos equipos.

	CISCO 3825	CISCO 7201
Velocidad de conmutación	0.35 mpps	2 mpps
Fuente redundante	Opcional Externa	Opcional interna
Interfaces GigaEthernet incorporados	1 RJ45, 1 RJ45+SFP	2 SFP, 2 RJ45+SFP
Interfaces externos		
POS	No	Si
E1 Canalizado	Si	Si
E3 Canalizado	No	Si
E3 Sin canalizar	Si	Si
ATM IMA	No	Si
E3 ATM	Si	Si
STM-1 ATM	Si	Si
Conectividad directa a mainframe IBM	No	Si
Concentración WAN de alta velocidad	No	Si
Conmutación LAN integradada	Si	No
Módem digital y analógico	Si	No
Primario RDSI	Si	Si
Líneas asíncronas	Si	Si

Tabla 9. –Comparativa de cisco 3825 y cisco 7201

En base a la potencia de ambos equipos y no a las interfaces disponibles en ambos elegimos para los route reflectors el modelo 7201 con un mínimo de un Gigabyte de memoria.

Esta elección tiene además la ventaja de que ya existen otros equipos de core (los dedicados a internet en ESX y TEL, y los LNS de AME y COR) de la familia 72xx y por tanto las versiones de IOS estarán unificadas y será más sencillo plantear su actualización.

4.5.1.3 Conexión de los nuevos RR a la red MPLS

Dada la relevancia de la función de los route reflector es recomendable que estén conectados al núcleo de la red, con una buena conectividad con todos los nodos de la red.

La ubicación física que se decidió fue utilizar los centros de CLP/PP (en conexión dos a dos (doble vía a dos nodos distintos) dada la facilidad de redundancia en cuanto a transmisión entre CLP y PP que explicaremos un poco más adelante) y EDS (en conexión de doble vía pero a un solo nodo MPLS, el de EDS). El volumen de tráfico de las sesiones de BGP no es excesivamente alto por lo que no es necesario agregar los dos enlaces disponibles en los equipos. No son equipos que deban cursar tráfico de red por lo que no requieren enlaces de alta capacidad. Los enlaces GigaEthernet son recomendables ya que aportan una baja latencia pero no son necesarios por volumen de tráfico. Por ello es preferible conectar el route reflector a dos nodos para mayor redundancia. En EDS no teníamos esta facilidad de redundancia de transmisión y por eso se conecta por doble vía a un solo nodo.

4.5.1.3.1 Conexión del RR de CLP

La conexión por doble vía del route reflector de CLP es de la siguiente manera: La vía principal es mediante un latiguillo directo entre puertos GigabitEthernet de ambos equipos y la secundaria que es con PP es un tanto especial. No se trata de un enlace directo de fibra ni por SDH/DWDM sino que es a través de una cadena de switches que disponemos entre ambas ubicaciones. Son dos Cisco 3750 interconectados entre ellos por fibra óptica en Gigabit Ethernet que separan los diferentes tráficos por Vlans. Lo utilizamos habitualmente como vía secundaria entre servicios que se ofrecen en CLP y PP.

El esquema a nivel global de cómo es esta redundancia incluida la de RR es:

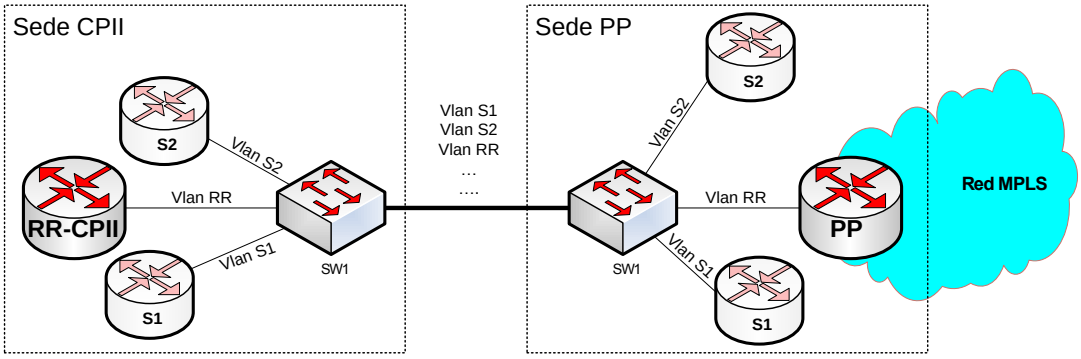

Ilustración 67- Redundancia CP-PP a nivel 2.

Como vemos, cada servicio se conecta a un puerto físico distinto tanto en CLP como en PP y encapsulado a su Vlan correspondiente. Luego la comunicación entre CLP y PP es por una

vía de fibra sobre un puerto físico que está configurado como un trunk y transporta todas las Vlans. Ya en PP de nuevo cada servicio tiene su puerto físico asignado y que es donde se conectan los equipos.

En lo que refiere a los route reflector en particular (y en general para cualquier servicio) es una red /30 de interconexión normal. Vemos el esquema a nivel 3:

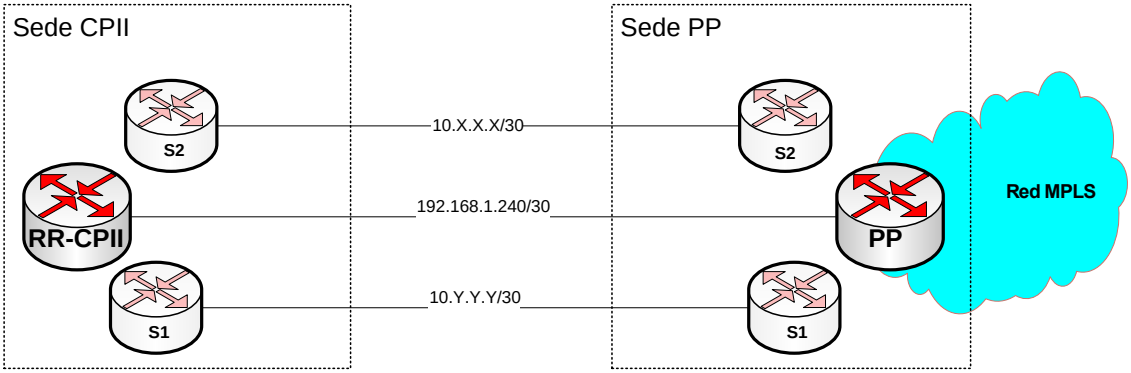

Ilustración 68- Redundancia CP-PP a nivel 3.

Dado este punto, contamos con la ventaja de que en la parte de UFIN estaba proyectado la instalación de un nuevo nodo (CISCO 7600) por lo que, dado que la plataforma para los route reflector elegida es 7200 y en EDS el nodo en servicio era de la serie 7200, se sustituyó en el punto anterior por un 7600 y solo es necesario comprar un 7200 que tendrá funcionalidad de route reflector.

4.5.1.3.2 Conexión del RR de EDS

Este equipo es el que antes prestaba funciones de PE en EDS y se sustituyó por un 7600. Al encontrarse en la misma ubicación que el PE actual de EDS simplemente hay que conectarlos por doble vía (dos vías independientes sin agregar tráfico). No realizamos la arquitectura llevada a cabo en CPII (doble vía a doble nodo) dado que no tenemos la facilidad de transmisión para la redundancia que si tuvimos en el escenario de CPII. En un esquema de detalle, la conexión sería la siguiente a nivel 2:

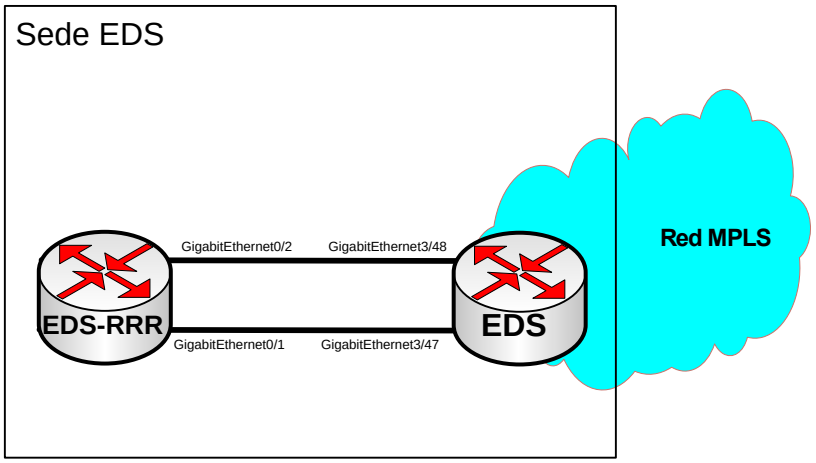

Ilustración 69- Redundancia EDS- EDS-RRR a nivel 2.

Como hemos visto, no es necesario agregar tráfico por lo que la conexión a nivel 3 sigue el siguiente esquema local:

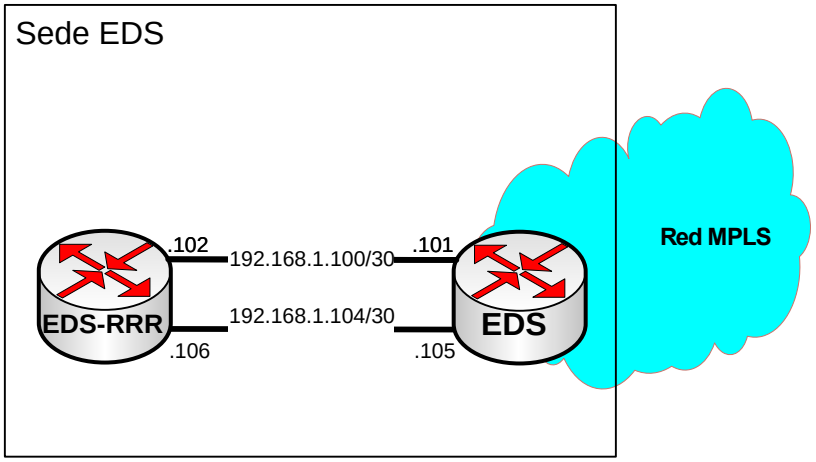

Ilustración 70- Redundancia EDS- EDS-RRR a nivel 3.

Comprobamos que se han configurado una /30 de enlace para cada línea y como es habitual en conexiones de doble vía, las ruta principal se establece mediante pesos de BGP.

Con esta situación, los dos nuevos route reflector conectados a la red pero sin funcionalidad, el esquema de la red, en cuanto a route reflector se refiere es la siguiente:

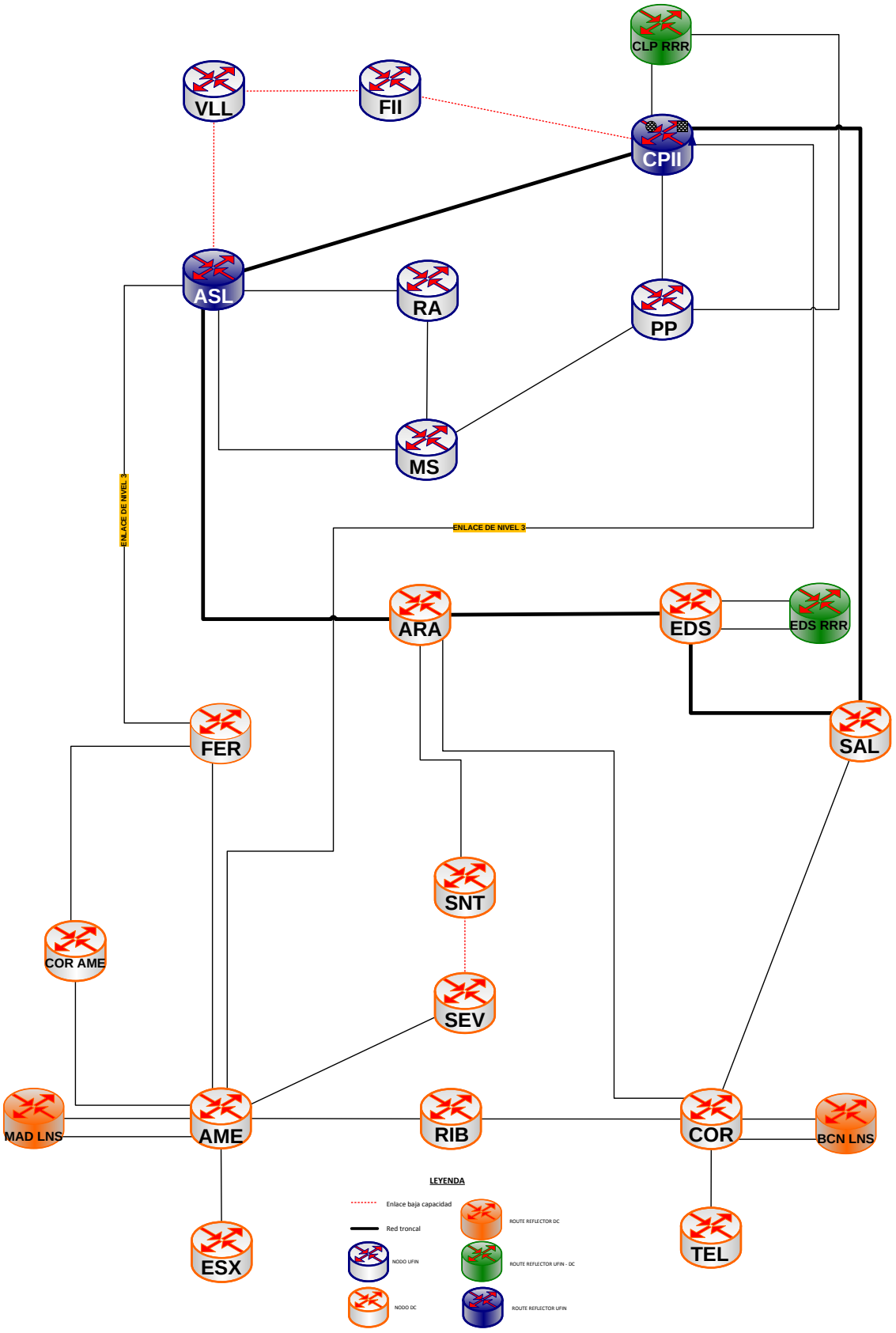

Ilustración 71- Red MPLS según los route reflector de la red (fase 1)

La antigua parte de UFIN está en azul, con los route reflector de esta parte (ASL y CPII) resaltados también en azul. Por otro lado tenemos la parte de DC en naranja, con sus route reflector (BCN-LNS y MAD-LNS) resaltados también en naranja. A esto añadimos los que serán nuevos route reflector de la red en verde (CLP-RRR y EDS-RRR).

El procedimiento a la hora de establecer la nueva arquitectura es el siguiente. Primero se establecerán los nuevos route reflector como route reflector de la parte de UFIN, se ha decidido hacerlo así ya que la migración a los nuevos RR se hará en dos pasos, primero la parte de UFIN y luego la parte de DC. La parte de DC es más complicada ya que como base se queda la parte de UFIN y es necesario migrar el MP-BGP. Dada esta complejidad separamos las tareas en lo que a RR se refiere en las partes de UFIN y DC.

4.5.1.4 Migración de la parte de UFIN a la nueva arquitectura de RR

El objetivo de esta tarea es sustituir los actuales RR de la parte de UFIN. Actualmente esta función la realizan los nodos de ASL y CPII además de realizar funciones de PE. Los nuevos RR de esta parte serán los 7200 CLP-RRR y EDS-RRR que actualmente solo están conectados a la red pero sin funcionalidad. Los nodos de ASL y CPII pasarán a tener funcionalidad exclusiva de PE tal y como están el resto de nodos.

Los pasos conceptualmente serían:

1.- Configuración de los nuevos RR y establecimiento de una adyacencia BGP entre vecinos-rr de los nodos CLP-RRR y EDS-RRR.

2.- Establecer una adyacencia BGP que denominaremos vecinos-tmp-rr de los nuevos RR (CLP-RRR y EDS-RRR) con los antiguos (CPII y ASL).

3.- Mover las adyacencias BGP que los PEs tienen contra ASL y CPII a CLP-RRR y EDS-RRR. Esta adyacencia se denomina parent-rr en la dirección PE ➔ RR y cliente-rr en la dirección RR ➔ PE.

4.- Sustituir las adyacencias vecino-tmp-rr entre los nodos MPLS de ASL y CPII con los de CLP-RRR y EDS-RRR por unas adyacencias del tipo parent-rr y cliente-rr en ambas direcciones.

4.5.1.4.1 Configuración de los nuevos RR y sus adyacencias.

En el nodo de EDS-RRR primero hay que comprobar que la versión de IOS es la correcta y eliminar la antigua de la memoria flash. Una vez hecho eso hay que configurar los 3 tipos de adyacencias que tendrá este nodo y son:

- Vecino-rr: es la adyacencia BGP que se establece entre los RR (CLP-RRR y EDS-RRR)
- Cliente-rr: es la adyacencia BGP que se establecerá entre los nuevos RR (CLP-RRR y EDS-RRR) y cada PE (ASL, CPII, RA, MS, VLL, FII, PP)

- Vecino-tmp-rr: es la adyacencia que se establece entre los nuevos RR (CLP-RRR y EDS-RRR) y los antiguos (ASL y CPII)

Una vez configurado el primer punto, en CLP-RRR se deben configurar las mismas adyacencias y comprobamos que se ha establecido esa sesión en CLP-RRR:

```
CLP-RRR#show ip bgp VPNv4 all summary

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.255.9 4 64987      53      53     3006    0    0 00:04:18        0
10.132.1.1      4 64987       0       0        0    0    0 never    Idle
10.132.1.2      4 64987       0       0        0    0    0 never    Idle
10.132.1.3      4 64987       0       0        0    0    0 never    Idle
10.132.1.4      4 64987       0       0        0    0    0 never    Idle
10.132.1.5      4 64987       0       0        0    0    0 never    Idle
10.132.1.6      4 64987       0       0        0    0    0 never    Idle
10.132.1.7      4 64987       0       0        0    0    0 never    Idle
```

Vemos que se ha establecido y que lleva 4:18 segundos activa sin ningún cambio de estado. Nótese que ya figuran el resto de nodos con función PE pero en estado Idle. La sesión está establecida pero aún no se reciben prefijos Esto es así ya que en los nuevos RR ya están configuradas las sesiones BGP contra el resto de PEs pero estos no están configurados aún. Comprobamos lo mismo en el nodo de EDS-RRR:

```
eds-rrr# show ip bgp VPNv4 all summary

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.132.1.1      4 64987       0       0        0    0    0 never    Idle
10.132.1.2      4 64987       0       0        0    0    0 never    Idle
10.132.1.3      4 64987       0       0        0    0    0 never    Idle
10.132.1.4      4 64987       0       0        0    0    0 never    Idle
10.132.1.5      4 64987       0       0        0    0    0 never    Idle
10.132.1.6      4 64987       0       0        0    0    0 never    Idle
10.132.1.7      4 64987       0       0        0    0    0 never    Idle
192.168.255.18 4 64987     102     102     5772    0    0 00:08:27        0
```

Vemos que se ha establecido perfectamente la adyacencia entre ambos route reflector. Conceptualmente la situación actual sería la siguiente:

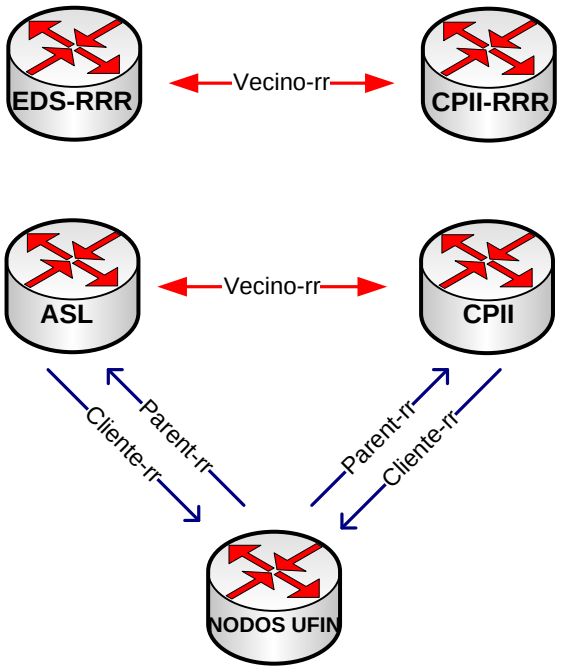

Ilustración 72- Relación temporal de la red de UFIN entre nodos (fase 1).

Una vez hecho esto, el siguiente paso es establecer las adyacencias entre los nuevos RR y los viejos para que los cuatro sean RR de toda la red.

- Impacto:

Estos trabajos no suponen impacto en la red ya que aún no están funcionando como RRs, no se han modificado las adyacencias en los PEs de la red. Tienen capacidad de routing total pero aún están aislados del resto de nodos y de los RR antiguos.

4.5.1.4.2 Establecimiento de adyacencias entre los nuevos RR y los antiguos.

En este paso se establecen las adyacencias BGP entre los RR nuevos y los antiguos. Estas adyacencias son sesiones BGP estándar sin hacer clientes los nodos de ASL y CPII de los de CLP-RRR y EDS-RRR por el momento. Serán unas adyacencias temporales que desaparecerán al final de estos trabajos.

Los nuevos RR de CLP-RRR y EDS-RRR ya están configurados de la tarea anterior por lo que solo hay que crearla en los nodos de ASL y CPII y comprobar que se ha establecido correctamente.

La configuración es la misma y procedemos a realizar las mismas comprobaciones que se hicieron en el punto anterior. En el nodo de EDS-RRR comprobamos lo siguiente:

```
eds-rrr#show ip bgp VPNv4 all summary
Neighbor        V  AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down  State/PfxRcd
10.132.1.1      4 64987       0       0        0    0    0 never    Idle
10.132.1.2      4 64987     433     116     5828    0    0 00:08:23     2768
10.132.1.3      4 64987     534     163     5828    0    0 00:12:22     3005
10.132.1.4      4 64987       0       0        0    0    0 never    Idle
```

```
10.132.1.5       4 64987       0       0        0    0    0 never    Idle
10.132.1.6       4 64987       0       0        0    0    0 never    Idle
10.132.1.7       4 64987       0       0        0    0    0 never    Idle
192.168.255.18   4 64987     210     209     5828    0    0 00:16:19        0
```

Vemos que han levantado las sesiones contra los antiguos RR y que además está aprendiendo prefijos de cada uno de ellos (sombreado en oscuro). La sesión contra el otro RR nuevo (CLP-RRR) está levantada pero sigue sin recibir prefijos ya que no se ha configurado aún ASL y CPII contra el RR nuevo de CPII (CLP-RRR).

Las direcciones corresponden a los siguientes equipos: 10.132.1.2 es ASL, 10.132.1.3 es CPII y 192.168.255.18 es CLP-RRR. El resto son los PEs de la red.

Ahora mismo lo que tenemos son 4 nodos MPLS configurados para hacer funciones de route-reflector de la misma red por lo que cualquier otro PE que establezca una relación de parent-rr con cualquiera de los 4 puede tener VPNs configuradas. Realmente solo los nodos de ASL y CPII tienen relaciones de cliente-rr con el resto de PEs y estos, una relación de parent-rr con ellos.

Visto de una manera conceptual, la situación es la siguiente:

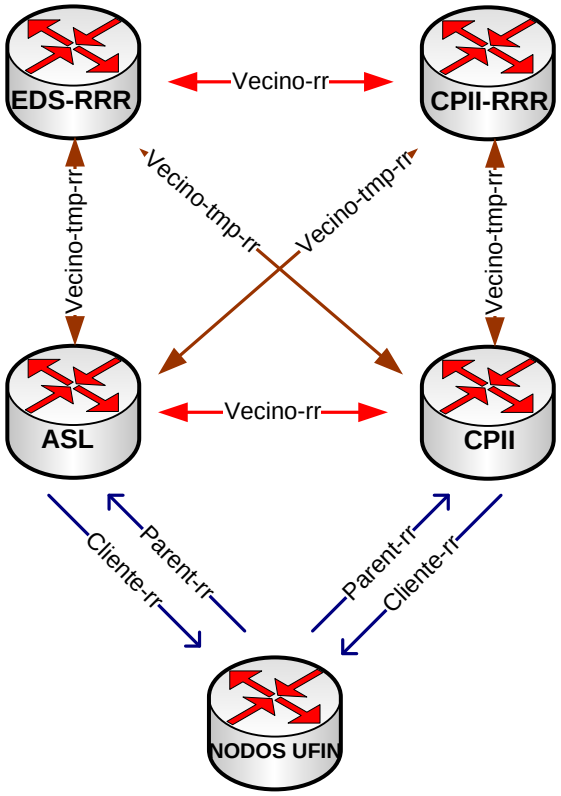

Ilustración 73- Relación temporal de la red de UFIN entre nodos (fase 2).

- Impacto:

Estos trabajos no suponen impacto en la red ya que igual que en el paso anterior solo estamos añadiendo capacidades a los nuevos RR sin migrar servicios aún.

4.5.1.4.3 Mover las adyacencias de los PEs desde los antiguos RR a los nuevos

Llegados a este punto, el siguiente paso es mover las relaciones de parent-rr que tienen el resto de nodos de la red de UFIN con ASL y CPII a los nodos de CLP-RRR y EDS-RRR que están adecuadamente configurados para realizar la función de RR a partir de ahora.

Configuramos las nuevas adyacencias parent-rr en los PEs. Las de cliente-rr no son necesarias ya que se configuraron en el primer punto.

Vemos las comprobaciones que se realizaron para el caso concreto del nodo de VLL que fue el siguiente en ser migrado pero en todos son las mismas, se comprueban las sesiones BGP tanto en el PE (en este caso VLL) como en cualquiera de los nuevos RR (CLP-RRR o EDS-RRR) una vez configurados las nuevas adyacencias de parent-rr:

```
MPLS-VLL#show ip bgp VPNv4 all summary

Neighbor             V      AS  MsgRcvd  MsgSent     TblVer    InQ  OutQ  Up/Down
State/PfxRcd
10.132.1.2       4 64987  217645  209182   354296    0     0 1w5d          1346
10.132.1.3       4 64987  218573  209394   354296    0     0 1w5d          1348
192.168.255.9    4 64987     536      56   354296    0     0 00:00:27      1348
192.168.255.18   4 64987     535      55   354296    0     0 00:00:26      1348
```

Comprobamos que se ven las 4 adyacencias de parent-rr, dos contra los antiguos nodos RR (CPII es 10.132.1.2 y ASL es 10.132.1.2) y contra los nuevos CLP-RRR es 192.168.255.18 y EDS-RRR es 192.168.255.9. Además se puede comprobar que reciben el mismo número de prefijos tanto de los RR antiguos como de los nuevos.

Justo en este momento el esquema conceptual es el siguiente:

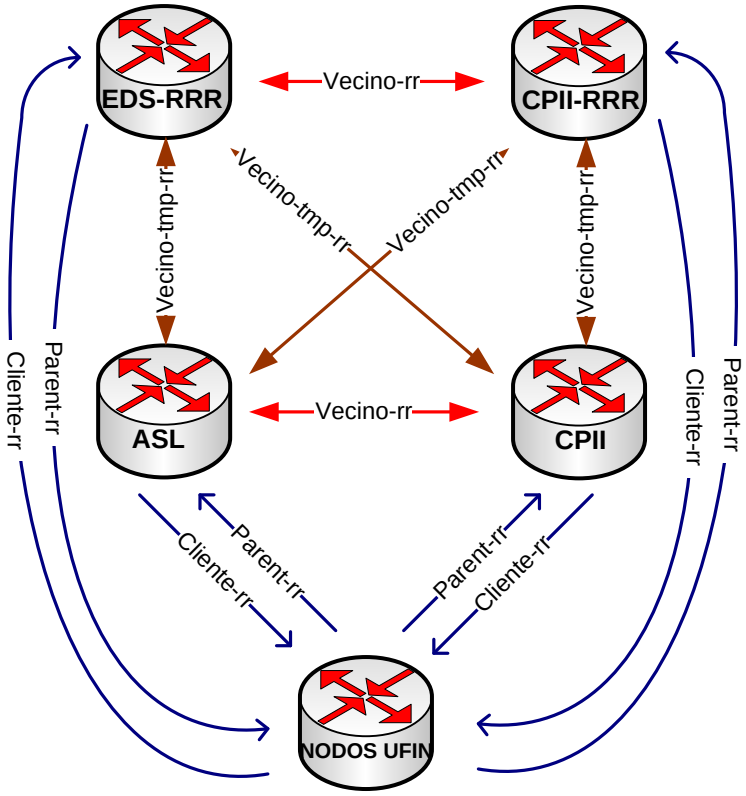

Ilustración 74- Relación temporal de la red de UFIN entre nodos (fase 3.1).

Tras ello borramos las adyacencias de parent-rr existentes contra los nodos de ASL y CPII y lo comprobamos de la siguiente manera:

```
MPLS-VLL#show ip bgp VPNv4 all summary

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd

192.168.255.9   4 64987     722     160   355037    0    0 00:09:20      1348
192.168.255.18  4 64987     723     159   355037    0    0 00:09:18      1348
```

Nótese que en el caso de los PEs no aparecen en la tabla de vecinos BGP el resto de nodos. Esto es totalmente normal ya que los PEs solo establecen adyacencias con los RR y es en los RR donde si figuran todos los PEs como vecinos BGP.

Una vez finalizada esta tarea con todos los nodos PE de la red y en una situación estable, comprobamos en el RR de CLP-RRR como queda su tabla de vecinos BGP:

```
CLP-RRR#show ip bgp VPNv4 all summary

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.132.1.1      4 64987  111791  115887    28304    0    0 6d11h         498
10.132.1.2      4 64987  111347  115282    28304    0    0 6d11h        1820
10.132.1.3      4 64987  111296  115208    28304    0    0 6d11h        3370
10.132.1.4      4 64987  109665  115628    28304    0    0 6d11h         240
10.132.1.5      4 64987  109509  115761    28304    0    0 6d11h          29
```

```
10.132.1.6      4 64987  109561  115717    28304    0    0 6d11h          176
10.132.1.7      4 64987  109963  116013    28304    0    0 6d11h           83
192.168.255.9   4 64987  115367  115597    28304    0    0 6d12h            0
```

Comprobamos que están todos los nodos con sus adyacencias BGP establecidas tras 6 días y 11 horas de funcionamiento. Los nodos con direcciones 10.132.1.1-7 son los PEs y el 192.168.255.9 es el RR de EDS-RRR.

Conceptualmente la red queda de la siguiente manera:

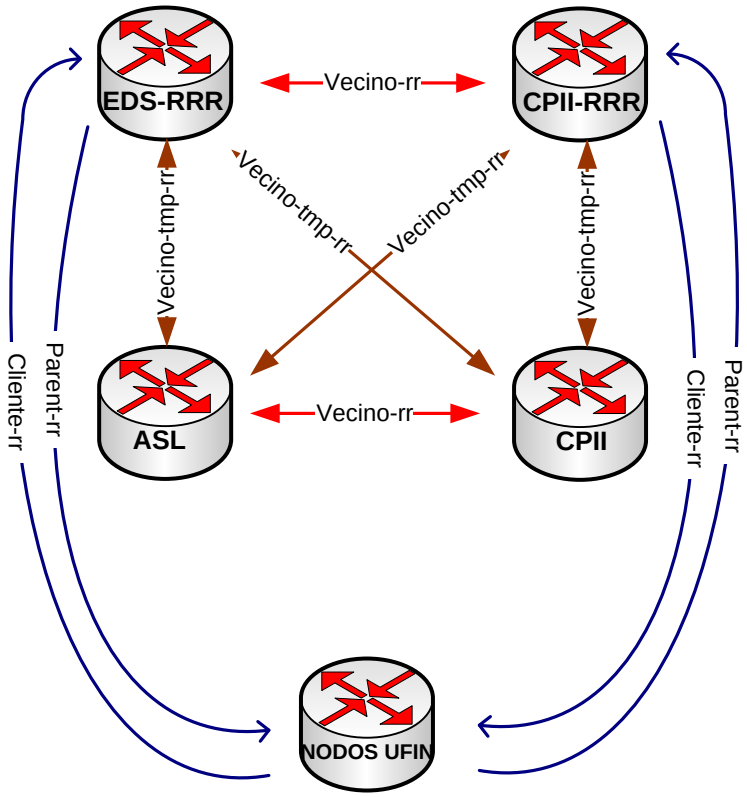

Ilustración 75- Relación temporal de la red de UFIN entre nodos (fase 3.2).

- Impacto:

Estos trabajos no suponen impacto en la red ya que, primero, se han creado las nuevas adyacencias contra los nuevos RR y después se han eliminado las adyacencias que cada PE tenía con los antiguos RR (ASL y CPII). Gracias a la adyacencia vecino-tmp-rr no se produce corte de servicio en este paso.

4.5.1.4.4 Sustitución de adyacencias de los antiguos RR

Ya solo queda sustituir las adyacencias de vecino-tmp-rr que existen entre los nuevos RR (CLP-RRR y EDS-RRR) y los antiguos (CPII y ASL). En realidad la adyacencia ya está establecida, solo hay que cambiar el tipo. Pasará de ser vecino-tmp-rr bidireccional a cliente-rr en los nuevos RR y de parent-rr en los antiguos RR para que los nodos de ASL y CPII sean una función equivalente a cualquier otro PE de la red.

La tabla de adyacencias BGP no varía ya que siguen establecidas, solo cambia el tipo y eso no se refleja en esta tabla. Solo cambia de manera conceptual y finalmente es así:

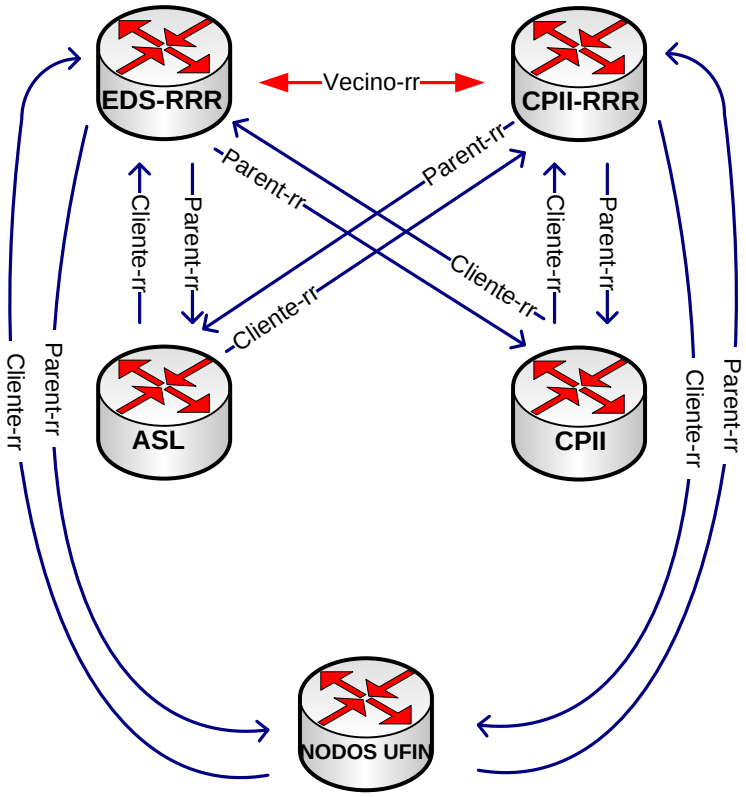

Ilustración 76- Relación temporal de la red de UFIN entre nodos (fase 4 / vista 1).

Que es lo mismo que el siguiente esquema simplificado.

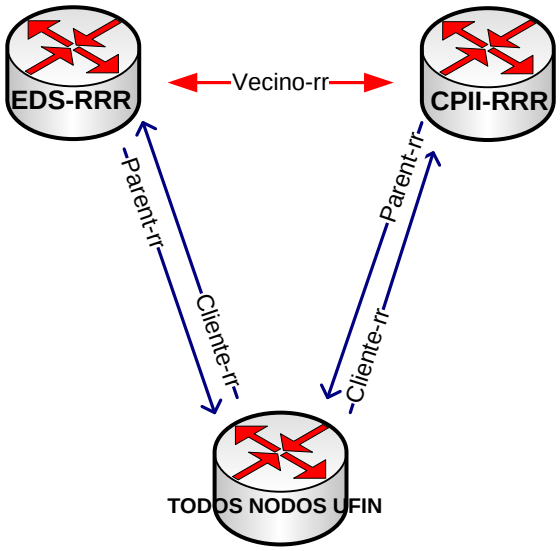

Ilustración 77- Relación temporal de la red de UFIN entre nodos (fase 4 / vista 2).

Aquí ya se han eliminado las adyacencias de vecino-rr entre los nodos de ASL y CPII y han pasado a ser otros PEs estándar de la red. También se ha eliminado la adyacencia vecino-tmp-rr que nos sirvió de paso intermedio.

- Impacto

En este paso si se produce corte de servicio en la red ya que los nodos de ASL y CPII han pasado de tener adyacencias de parent-rr contra el resto de PEs a tener adyacencias de cliente-rr contra los nuevos RR (CLP-RRR y EDS-RRR). El corte de servicio es de menos de 10 minutos por cada VPN que tenía configurado el nodo como PE (recordemos los nodos de ASL y CPII hacían función de RR a la vez que de PE). El resto de servicios AToM no se ven afectados.

Una vez concluidos los trabajos de sustitución de los RR en la red de UFIN la situación total de la red es la siguiente:

Los nodos MPLS de la parte de UFIN tienen como RR los equipos CLP-RRR y EDS-RRR que son equipos exclusivos para esta función.

Aún en la parte de DC son los BCN-LNS y MAD-LNS los encargados de la función de RR.

En la siguiente ilustración vemos en verde todos los nodos de la parte de UFIN, con ASL y CPII realizando las mismas funciones de PE que el resto de nodos. Aquí los nodos de CLP-RRR y EDS-RRR ya si hacen funciones de RR de la parte de UFIN y también la harán de la parte de DC en el siguiente paso. En este momento la situación es la siguiente:

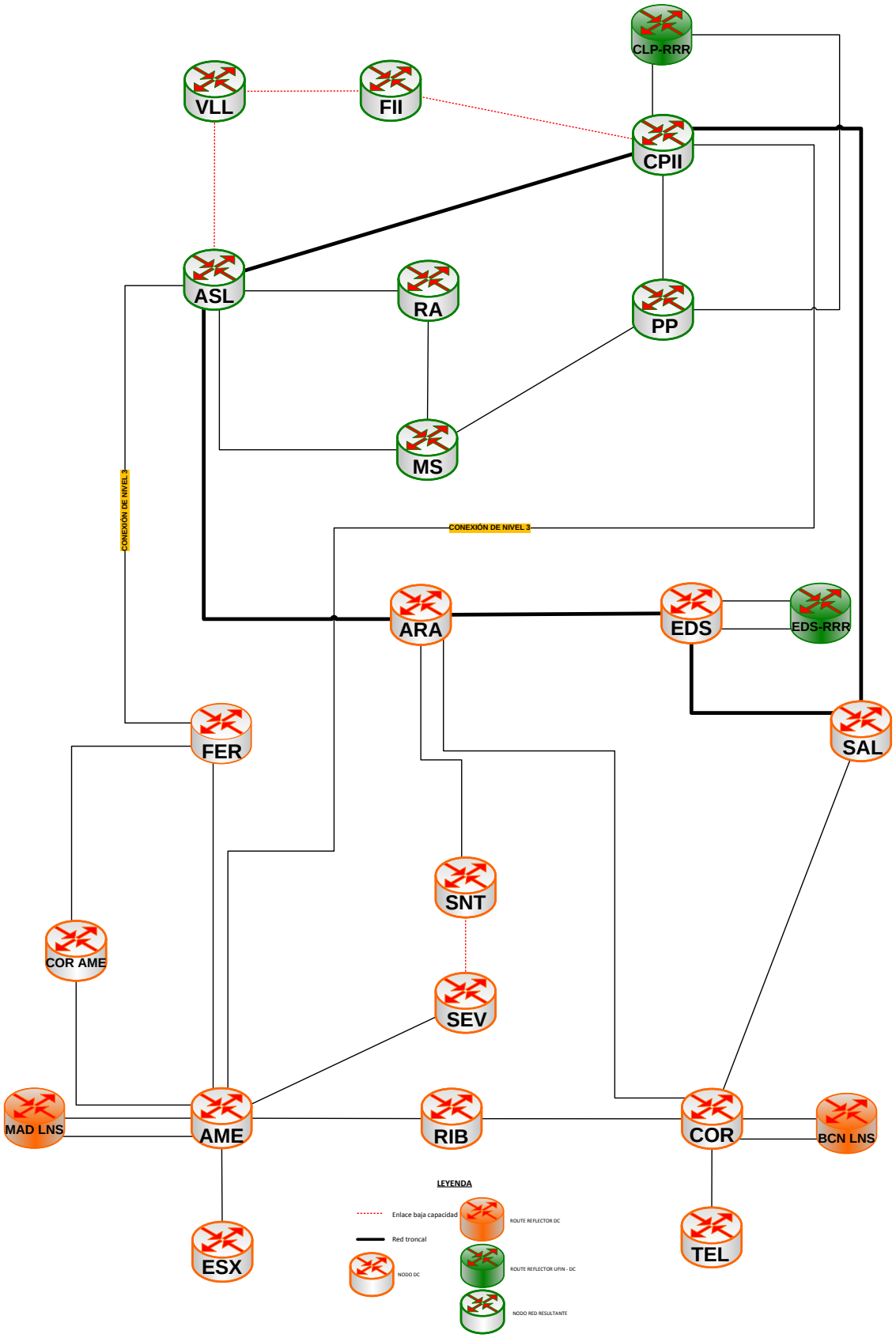

Ilustración 78- Red MPLS según los route reflector de la red (fase 2)

4.5.2 PROPAGACIÓN DEL MP-BPG

A la finalización de este paso ya tendremos la arquitectura definitiva en cuando a route reflector se refiere y, por tanto, dispondremos de una red homogénea a nivel de VPN. Se podrán propagar VPNs por todos los nodos que conforman la totalidad de la red.

El objetivo de esta tarea no es otro que unificar el número de Sistema Autónomo (AS). En la parte de UFIN es 64987 y en la parte de DC es 65550, ámbos son privados. Se ha tomado la determinación de unificarlos con el AS 64987 ya que en la red de UFIN, cuando hay una configuración de sede 2CE – 2PE (dos equipos de sede conectados a dos PEs distintos para ofrecer redundancia) es BGP el protocolo de routing que se establece. Dado que en la configuración de BGP de los CPEs se especifica el número de AS contra el que se establece la sesión BGP, el cambio de AS supondría la reconfiguración de absolutamente todos los CPEs configurados con la arquitectura 2CE – 2PE de la red. En la parte de DC se hace con OSPF por lo que no tenemos esta restricción y se puede cambiar el número de AS con un impacto menor y menos costoso.

La situación de la red a nivel de configuración es que tenemos, todos los nodos de la parte de UFIN configurados de modo que sus route reflector son los definitivos, CLP-RRR y EDS-RRR.

Configuración relativa a BGP en el route reflector de CLP (192.168.255.18):

```
router bgp 64987
 bgp router-id 192.168.255.18
 no bgp default ipv4-unicast
 no bgp default route-target filter
 bgp cluster-id 640
 bgp log-neighbor-changes
 neighbor CLIENTE-RR peer-group
 neighbor CLIENTE-RR remote-as 64987
 neighbor CLIENTE-RR update-source Loopback0
 neighbor CLIENTE-RR timers 5 15
 neighbor VECINO-RR peer-group
 neighbor VECINO-RR remote-as 64987
 neighbor VECINO-RR update-source Loopback0
 neighbor VECINO-RR timers 5 15
 neighbor 10.132.1.1 peer-group CLIENTE-RR
 neighbor 10.132.1.2 peer-group CLIENTE-RR
 neighbor 10.132.1.3 peer-group CLIENTE-RR
 neighbor 10.132.1.4 peer-group CLIENTE-RR
 neighbor 10.132.1.5 peer-group CLIENTE-RR
 neighbor 10.132.1.6 peer-group CLIENTE-RR
 neighbor 10.132.1.7 peer-group CLIENTE-RR
 neighbor 192.168.255.9 peer-group VECINO-RR

address-family vpnv4
  neighbor CLIENTE-RR send-community both
  neighbor CLIENTE-RR route-reflector-client
```

```
  neighbor CLIENTE-RR next-hop-self
  neighbor CLIENTE-RR advertisement-interval 1
  neighbor VECINO-RR send-community both
  neighbor VECINO-RR next-hop-self
  neighbor VECINO-RR advertisement-interval 1
  neighbor 10.132.1.1 activate
  neighbor 10.132.1.2 activate
  neighbor 10.132.1.3 activate
  neighbor 10.132.1.4 activate
  neighbor 10.132.1.5 activate
  neighbor 10.132.1.6 activate
  neighbor 10.132.1.7 activate
  neighbor 192.168.255.9 activate
exit-address-family
```

En negrita vemos las relaciones de adyacencia con todos los nodos de la red de UFIN como CLIENTE-RR y de VECINO-RR con el route reflector de EDS.

La configuración relativa a BGP en el route reflector de EDS (192.168.255.18) es exactamente igual que en el de CPII solo que su relación de VECINO-RR es con CPII:

```
router bgp 64987
 bgp router-id 192.168.255.9
 no bgp default ipv4-unicast
 no bgp default route-target filter
 bgp cluster-id 640
 bgp log-neighbor-changes
 neighbor CLIENTE-RR peer-group
 neighbor CLIENTE-RR remote-as 64987
 neighbor CLIENTE-RR update-source Loopback0
 neighbor CLIENTE-RR timers 5 15
 neighbor VECINO-RR peer-group
 neighbor VECINO-RR remote-as 64987
 neighbor VECINO-RR update-source Loopback0
 neighbor VECINO-RR timers 5 15
 neighbor 10.132.1.1 peer-group CLIENTE-RR
 neighbor 10.132.1.2 peer-group CLIENTE-RR
 neighbor 10.132.1.3 peer-group CLIENTE-RR
 neighbor 10.132.1.4 peer-group CLIENTE-RR
 neighbor 10.132.1.5 peer-group CLIENTE-RR
 neighbor 10.132.1.6 peer-group CLIENTE-RR
 neighbor 10.132.1.7 peer-group CLIENTE-RR
 neighbor 192.168.255.18 peer-group VECINO-RR

address-family vpnv4
  neighbor CLIENTE-RR send-community both
  neighbor CLIENTE-RR route-reflector-client
  neighbor CLIENTE-RR next-hop-self
  neighbor CLIENTE-RR advertisement-interval 1
  neighbor VECINO-RR send-community both
  neighbor VECINO-RR next-hop-self
  neighbor VECINO-RR advertisement-interval 1
```

```
  neighbor 10.132.1.1 activate
  neighbor 10.132.1.2 activate
  neighbor 10.132.1.3 activate
  neighbor 10.132.1.4 activate
  neighbor 10.132.1.5 activate
  neighbor 10.132.1.6 activate
  neighbor 10.132.1.7 activate
  neighbor 192.168.255.18 activate
exit-address-family
```

Por su parte, cada nodo de la red UFIN tiene una configuración de BGP con relación de adyacencia exclusivamente de PARENT-RR contra los dos route reflector de la red, el ejemplo es del nodo de RA (10.132.1.5):

```
router bgp 64987
 bgp router-id 10.132.1.5
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor PARENT-RR peer-group
 neighbor PARENT-RR remote-as 64987
 neighbor PARENT-RR update-source Loopback0
 neighbor PARENT-RR timers 5 15
 neighbor 192.168.255.9 peer-group PARENT-RR
 neighbor 192.168.255.18 peer-group PARENT-RR

address-family vpnv4
  neighbor PARENT-RR send-community both
  neighbor PARENT-RR next-hop-self
  neighbor PARENT-RR advertisement-interval 1
  neighbor 192.168.255.9 activate
  neighbor 192.168.255.18 activate
```

Esta es la configuración que al término de esta fase deben tener absolutamente todos los nodos de la red salvo los dos RR.

Las denominaciones de PARENT-RR, CLIENT-RR y VECINO-RR solo son etiquetas que ayudan a la operación de los equipos.

Por otro lado, en la parte de DC, los nodos de bcn-rrr y mad-rrr de la red de DC tienen la siguiente configuración relativa a BGP ya que son los actuales route reflector de la red. El nodo de bcn-rrr tiene la configuración:

```
router bgp 65500
 bgp log-neighbor-changes
 bgp graceful-restart restart-time 120
 bgp graceful-restart stalepath-time 360
 bgp graceful-restart
 timers bgp 20 60
 neighbor peer_mpbgp peer-group
 neighbor peer_mpbgp remote-as 65500
```

```
 neighbor peer_mpbgp password 7 105D0C1A1712065B
 neighbor peer_mpbgp update-source Loopback0
 neighbor peer_rr peer-group
 neighbor peer_rr remote-as 65500
 neighbor peer_rr password 7 0518030C33495A59
 neighbor peer_rr update-source Loopback0
 neighbor 192.168.255.1 peer-group peer_mpbgp
 neighbor 192.168.255.1 description IBGP -> MAD-RPE-01
 neighbor 192.168.255.2 peer-group peer_mpbgp
 neighbor 192.168.255.2 description IBGP -> BCN-RPE-01
 neighbor 192.168.255.3 peer-group peer_mpbgp
 neighbor 192.168.255.3 description IBGP -> ARA-RPE-01
 neighbor 192.168.255.4 peer-group peer_mpbgp
 neighbor 192.168.255.4 description IBGP -> RIB-RPE-01
 neighbor 192.168.255.5 peer-group peer_mpbgp
 neighbor 192.168.255.5 description IBGP -> SNT-RPE-01
 neighbor 192.168.255.6 peer-group peer_mpbgp
 neighbor 192.168.255.6 description IBGP -> SEV-RPE-01
 neighbor 192.168.255.7 peer-group peer_mpbgp
 neighbor 192.168.255.7 description IBGP -> FER-RPE-01
 neighbor 192.168.255.8 peer-group peer_mpbgp
 neighbor 192.168.255.8 description IBGP -> SAL-RPE-01
 neighbor 192.168.255.10 peer-group peer_mpbgp
 neighbor 192.168.255.10 description IBGP -> EDS-RPE-01
 neighbor 192.168.255.201 peer-group peer_mpbgp
 neighbor 192.168.255.201 description eBGP -> ESX-RII-01
 neighbor 192.168.255.202 peer-group peer_mpbgp
 neighbor 192.168.255.202 description IBGP -> ESX-RII-02
 neighbor 192.168.255.253 peer-group peer_rr
 neighbor 192.168.255.253 description IBGP -> MAD-RRR-01
 neighbor 192.168.255.254 peer-group peer_mpbgp
 neighbor 192.168.255.254 description IBGP -> COR-RPE-01
 !
 address-family ipv4
  neighbor peer_mpbgp send-community both
  neighbor peer_mpbgp route-reflector-client
  neighbor 192.168.255.1 activate
  neighbor 192.168.255.2 activate
  neighbor 192.168.255.3 activate
  neighbor 192.168.255.4 activate
  neighbor 192.168.255.5 activate
  neighbor 192.168.255.6 activate
  neighbor 192.168.255.7 activate
  neighbor 192.168.255.8 activate
  neighbor 192.168.255.10 activate
  no neighbor 192.168.255.201 activate
  no neighbor 192.168.255.202 activate
  neighbor 192.168.255.253 activate
  neighbor 192.168.255.254 activate
  no auto-summary
  no synchronization
 exit-address-family
```

Y naturalmente, el router de mad-rrr tiene la configuración calcada salvo que tiene la adyacencia cruzada con el route reflector:

```
router bgp 65500
 bgp log-neighbor-changes
 bgp graceful-restart restart-time 120
 bgp graceful-restart stalepath-time 360
 bgp graceful-restart
 timers bgp 20 60
 neighbor peer_mpbgp peer-group
 neighbor peer_mpbgp remote-as 65500
 neighbor peer_mpbgp password 7 0832494D1B1C1147
 neighbor peer_mpbgp update-source Loopback0
 neighbor peer_rr peer-group
 neighbor peer_rr remote-as 65500
 neighbor peer_rr password 7 111A1C0605171F5C
 neighbor peer_rr update-source Loopback0
 neighbor 192.168.255.1 peer-group peer_mpbgp
 neighbor 192.168.255.1 description IBGP -> MAD-RPE-01
 neighbor 192.168.255.2 peer-group peer_mpbgp
 neighbor 192.168.255.2 description IBGP -> BCN-RPE-01
 neighbor 192.168.255.3 peer-group peer_mpbgp
 neighbor 192.168.255.3 description IBGP -> ARA-RPE-01
 neighbor 192.168.255.4 peer-group peer_mpbgp
 neighbor 192.168.255.4 description IBGP -> RIB-RPE-01
 neighbor 192.168.255.5 peer-group peer_mpbgp
 neighbor 192.168.255.5 description IBGP -> SNT-RPE-01
 neighbor 192.168.255.6 peer-group peer_mpbgp
 neighbor 192.168.255.6 description IBGP -> SEV-RPE-01
 neighbor 192.168.255.7 peer-group peer_mpbgp
 neighbor 192.168.255.7 description IBGP -> FER-RPE-01
 neighbor 192.168.255.8 peer-group peer_mpbgp
 neighbor 192.168.255.8 description IBGP -> SAL-RPE-01
 neighbor 192.168.255.10 peer-group peer_mpbgp
 neighbor 192.168.255.10 description IBGP -> EDS-RPE-01
 neighbor 192.168.255.201 peer-group peer_mpbgp
 neighbor 192.168.255.201 description eBGP -> ESX-RII-01
 neighbor 192.168.255.202 peer-group peer_mpbgp
 neighbor 192.168.255.202 description IBGP -> ESX-RII-02
 neighbor 192.168.255.252 peer-group peer_rr
 neighbor 192.168.255.252 description IBGP -> BCN-RRR-01
 neighbor 192.168.255.254 peer-group peer_mpbgp
 neighbor 192.168.255.254 description IBGP -> COR-RPE-01
 !
  address-family vpnv4
  neighbor peer_mpbgp send-community both
  neighbor peer_mpbgp route-reflector-client
  neighbor peer_rr send-community both
  neighbor 192.168.255.1 activate
  neighbor 192.168.255.1 weight 40960
  neighbor 192.168.255.2 activate
```

```
  neighbor 192.168.255.2 weight 40960
  neighbor 192.168.255.3 activate
  neighbor 192.168.255.3 weight 40960
  neighbor 192.168.255.4 activate
  neighbor 192.168.255.4 weight 40960
  neighbor 192.168.255.5 activate
  neighbor 192.168.255.5 weight 40960
  neighbor 192.168.255.6 activate
  neighbor 192.168.255.6 weight 40960
  neighbor 192.168.255.7 activate
  neighbor 192.168.255.7 weight 40960
  neighbor 192.168.255.8 activate
  neighbor 192.168.255.8 weight 40960
  neighbor 192.168.255.10 activate
  neighbor 192.168.255.10 weight 40960
  neighbor 192.168.255.252 activate
  neighbor 192.168.255.252 weight 40960
  neighbor 192.168.255.254 activate
  neighbor 192.168.255.254 weight 40960
  bgp scan-time import 5
 exit-address-family
```

Por su lado, el resto de nodos de la red de DC están configurados exactamente igual. Solo tienen adyacencias con los dos route reflector, en este caso vemos la configuración del nodo de ARA:

```
router bgp 65500
 bgp log-neighbor-changes
 bgp graceful-restart restart-time 120
 bgp graceful-restart stalepath-time 360
 bgp graceful-restart
 timers bgp 20 60
 neighbor peer_mpbgp peer-group
 neighbor peer_mpbgp remote-as 65500
 neighbor peer_mpbgp password 7 06150A225E4B1D49
 neighbor peer_mpbgp update-source Loopback0
 neighbor 192.168.255.252 peer-group peer_mpbgp
 neighbor 192.168.255.252 description IBGP -> BCN-RRR-01
 neighbor 192.168.255.253 peer-group peer_mpbgp
 neighbor 192.168.255.253 description IBGP -> MAD-RRR-01
 !
address-family vpnv4
 neighbor peer_mpbgp activate
 neighbor peer_mpbgp send-community both
 neighbor 192.168.255.252 peer-group peer_mpbgp
 neighbor 192.168.255.253 peer-group peer_mpbgp
 bgp scan-time import 5
 exit-address-family
```

Una vez hemos visto las situaciones iniciales previas a esta tarea, definimos los pasos que se van a dar:

1. Cambio de AS en los nodos de DC
2. Mallado de los Route Reflector
3. Migración conexiones Route Reflector

4.5.2.1 Cambio de AS en los nodos de DC

En la primera fase se reconfigurarán los nodos de DC modificando únicamente el número de AS de BGP. Esto supone que al terminar esta fase la red estará en una situación idéntica que al inicio de la operación con la salvedad de que el AS de BGP de los equipos será el mismo en toda la red. De esta forma se establece un punto estable y aunque no se pudieran ejecutar las fases posteriores no sería necesario dar marcha atrás en estos cambios.

Dado que actualmente no existen conexiones BGP desde los equipos afectados con otros equipos externos la red debe continuar operando exactamente igual que antes.

En esta tarea, no se tocará configuración de ningún nodo de la red de UFIN salvo los dos route reflector, ya que serán los encargados de reflejar rutas de la red resultante.

Previamente al cambio del AS de BGP se pondrán fuera de servicio las líneas de conexión con el CPD, la interconexión de nivel 3 existente entre las redes de UFIN y DC. Conforme se migren los nodos se restablecerán las conexiones de forma que los procesos de routing levanten de forma ordenada.

El diagrama conceptual es el siguiente al inicio:

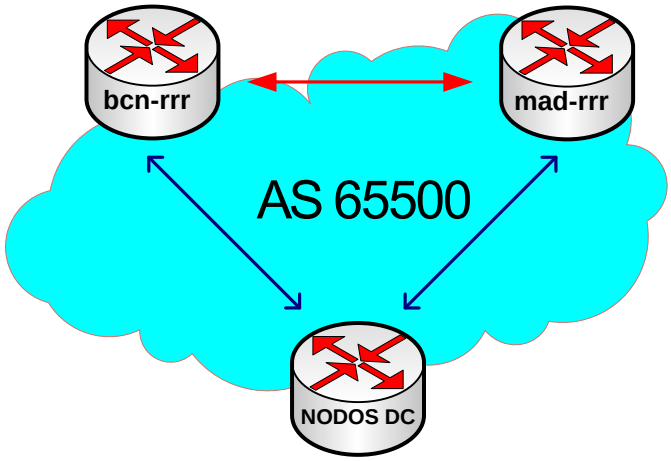

Ilustración 79- Esquema de la red de DC previa en lo relativo a los route reflector

En un primer paso, le cambiamos el número de AS a los route reflector por lo que en ese momento quedan de tal manera:

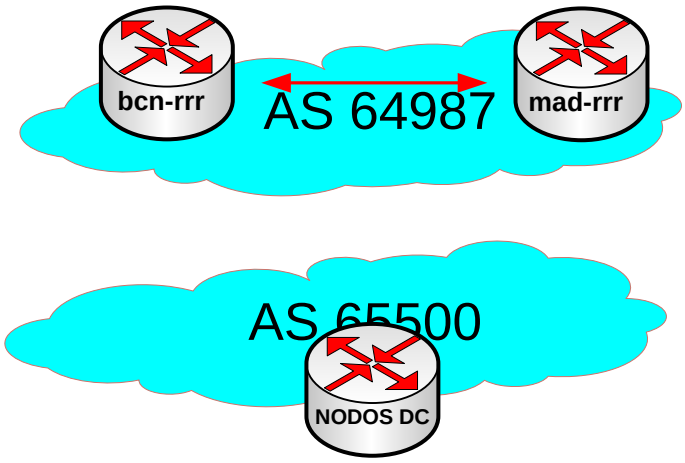

Ilustración 80- Paso 1 del cambio de número de AS en la red de DC

Quedan aún las adyacencias entre los route reflector con el nuevo AS configurado pero se pierden las adyacencias de cada PE con los route reflector ya que los PEs aún tienen configurado el AS antiguo.

El siguiente esquema se refiere a una VPN particular, la VPN 1. Se indica como es el flujo de comunicaciones entre sedes de esa VPN 1 en distintos PEs.

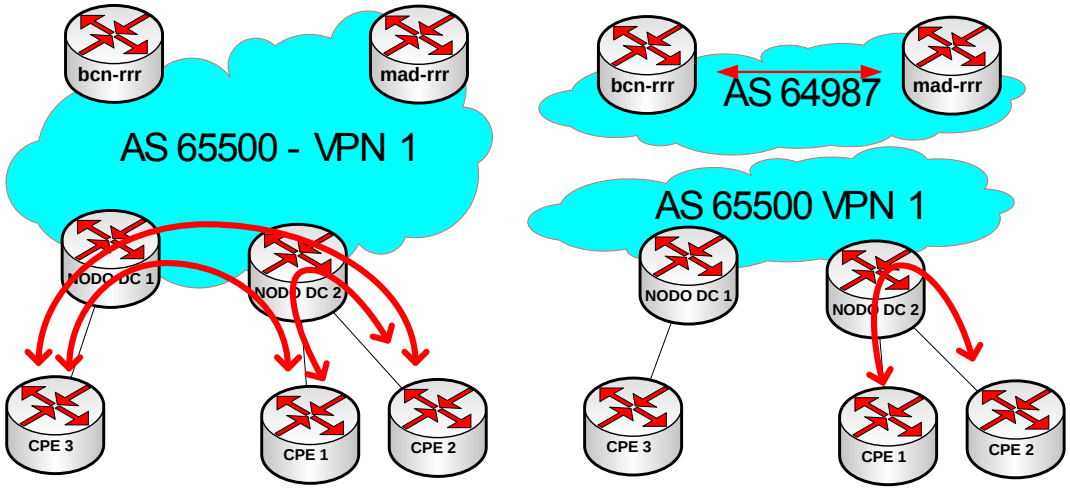

Ilustración 81- Flujo de comunicación durante el paso 1 del cambio de número de AS en la red de DC

En la situación inicial es posible la comunicación entre CPEs conectados a cualquier PE ya que los RR y los PE están dentro del mismo AS e intercambian rutas. Al cambiar el AS de los RR y no el de los PEs se pierde la adyaciencias entre PEs y los RR por lo que los RR ya no pueden reflejar las rutas aprendidas de un PE al resto de PEs de la red. Esto produce que CPEs que cuelgan del mismo PE en la misma VPN mantengan la comunicación pero entre dos CPEs conectados a distintos PEs no existirá comunicación.

El segundo paso es ir migrando uno a uno los PEs de la red de DC al nuevo sistema autónomo, según se van añadiendo nodos al nuevo AS, se van recuperando los servicios. El siguiente esquema muestra conceptualmente como se va recuperando la comunicación en la VPN 1 según se van reconfigurando los nodos:

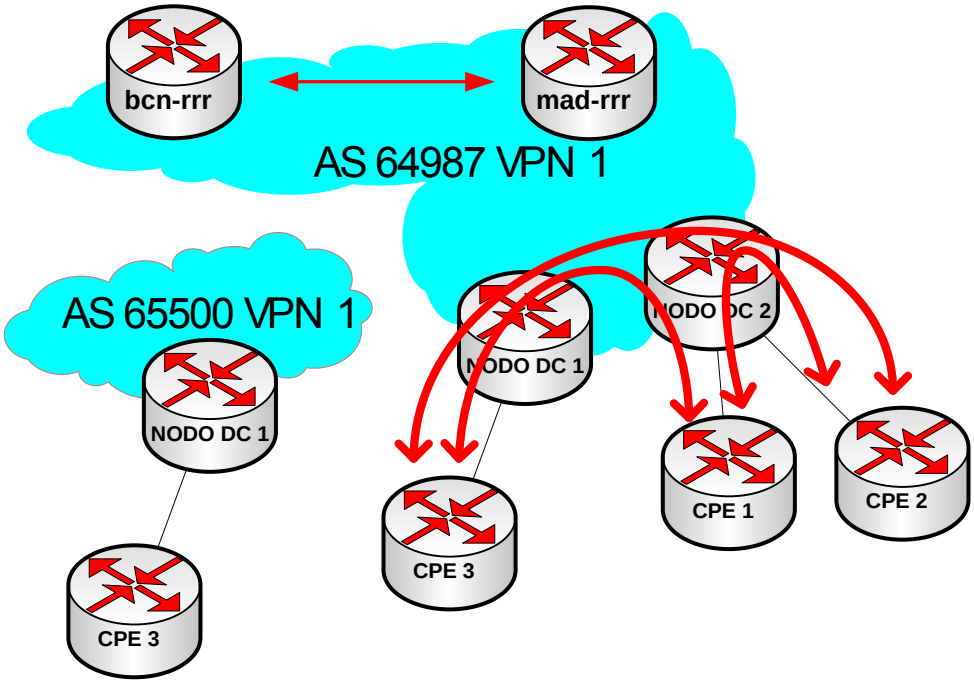

Ilustración 82- Flujo de comunicación durante el paso 1 del cambio de número de AS en la red de DC durante la migración de los nodos.

Según se añaden PEs al nuevo sistema autónomo se van reflejando las rutas aprendidas de cada PE al resto y se va recuperando la conexión gradualmente.

- Impacto:

En este momento, absolutamente todas las sedes conectadas a los diferentes PEs de la red dejan de poder comunicar con el resto de sedes en el resto de PEs ya que al no establecerse las adyacencias de BGP entre nodos no se produce intercambio de rutas. Los servicios de nivel 2 no se ven afectados ya que los nodos no han perdido conectividad en la tabla de routing global. Se van recuperando según se van migrando los PEs al nuevo AS. Se estima un corte de servicio de 45 minutos, el tiempo necesario en reescribir la configuración de BGP en los nodos de la parte de DC.

Una vez cambiado el número de sistema autónomo en la parte de DC, la situación que tenemos es la de dos redes, con dos pares de route reflector cada red y un mismo sistema autónomo pero entre ambas redes no se puede intercambiar datos a nivel de VPN ya que cada parte de nodos establece sus sesiones de BGP con su par de route reflector, eso sí, a nivel de BGP las redes son idénticas. La siguiente ilustración esquematiza la situación actual:

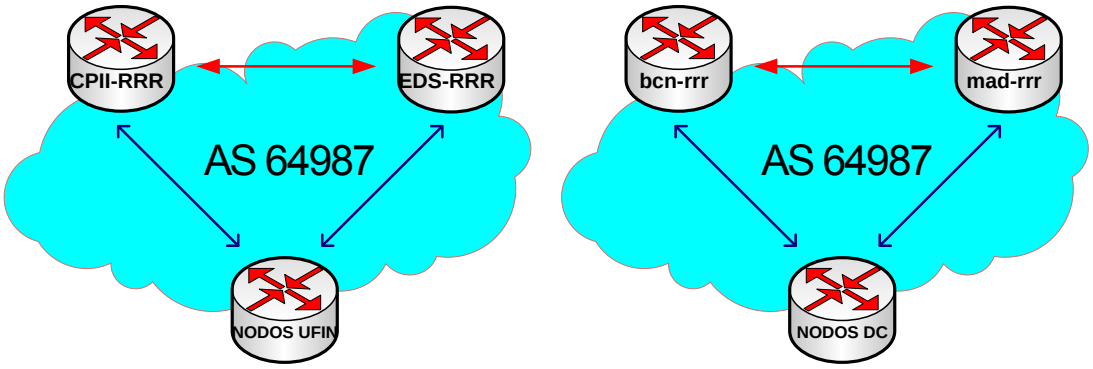

Ilustración 83- Situación a nivel BGP de ambas redes previo al mallado de route reflector

4.5.2.2 Mallado de los route reflector y migración de los nodos de DC

Este paso consiste en configurar en los route reflector (los dos pares) las adyacencias de BGP contra el resto de RR, con esto se logra la integración efectiva a nivel de MP-BGP, pasando la red a comportarse como una única red. A nivel de configuración se trata de configurar y activar los route reflector de UFIN en la parte de BGP y activarlos. Esto se consigue ya que, para cualquier nodo, es suficiente estar conectado a un route reflector ya que estos se comunican entre sí. El siguiente ejemplo, el EDS-RRR es válido para todos los route reflector cambiando sus vecinos:

```
router bgp 64987
 bgp router-id 192.168.255.9
 no bgp default ipv4-unicast
 no bgp default route-target filter
 bgp cluster-id 640
 bgp log-neighbor-changes
 neighbor VECINO-RR peer-group
 neighbor VECINO-RR remote-as 64987
 neighbor VECINO-RR update-source Loopback0
 neighbor VECINO-RR timers 5 15
 neighbor 192.168.255.18 peer-group VECINO-RR
 neighbor 192.168.255.252 peer-group CLIENTE-RR
 neighbor 192.168.255.253 peer-group CLIENTE-RR
!
 address-family ipv4
  no synchronization
  no auto-summary
 exit-address-family
 !
 address-family vpnv4
  neighbor VECINO-RR send-community both
  neighbor VECINO-RR next-hop-self
  neighbor VECINO-RR advertisement-interval 1
  neighbor 192.168.255.18 activate
  neighbor 192.168.255.252 activate
  neighbor 192.168.255.253 activate
```

Comprobamos que las adyacencias son efectivas:

```
eds-rrr-01#show ip bgp all summary
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.10, local AS number 64987
BGP table version is 64396, main routing table version 64396
6067 network entries using 855447 bytes of memory
15322 path entries using 1041896 bytes of memory
375/372 BGP path/bestpath attribute entries using 28500 bytes of memory
26 BGP rrinfo entries using 624 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
140 BGP extended community entries using 13908 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1940423 total bytes of memory
BGP activity 10588/4521 prefixes, 33602/18280 paths, scan interval 5 secs

Neighbor        V  AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.255.18  4 64987    1056      64    64396    0    0 00:04:23     3028
192.168.255.252 4 64987    2415     849    64396    0    0 01:10:24     3029
192.168.255.253 4 64987    5091    2850    64396    0    0 02:59:21     3029
...
```

En este momento la situación es de toda una red a nivel de VPN en el mismo AS, el 64987 con cuatro route reflector mallados a nivel BGP entre ellos y los nodos de la parte de DC apuntando a mad-rrr y bcn-rrr y los de la parte de UFIN que apuntan a CLP-RRR y EDS-RRR. La conectividad a nivel de VPN es total ya que los route reflector intercambian rutas entre sí. El esquema conceptual sería el siguiente:

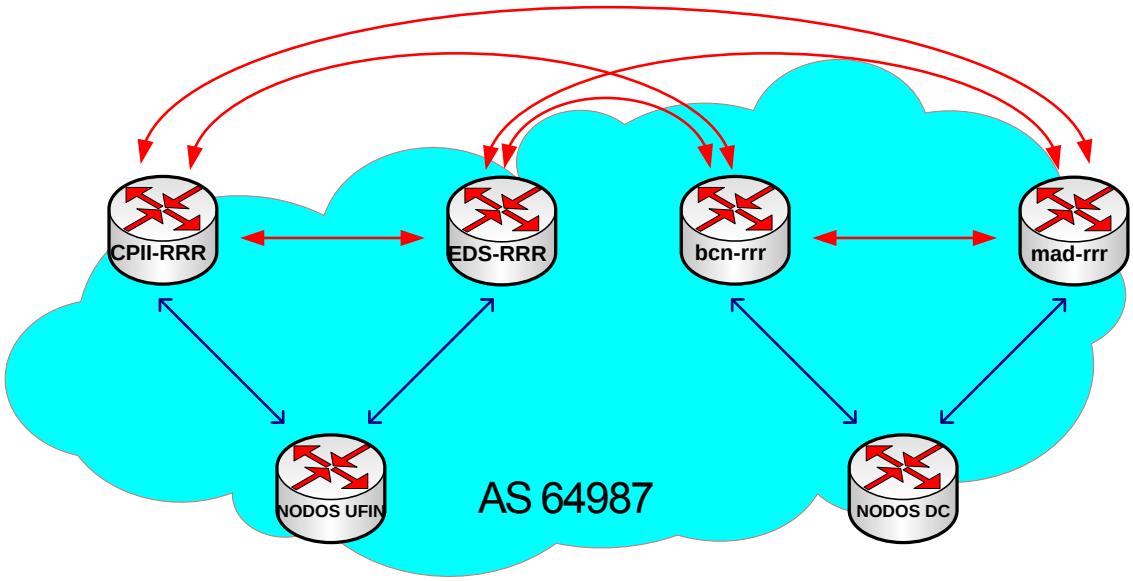

Ilustración 84- Situación a nivel BGP de ambas tras al mallado de route reflector

Tras este paso, ya solo falta migrar las conexiones de los nodos cliente de DC de route reflector dentro de la red. También los nodos mad-rrr y bcn-rrr pasarán de ser route reflector de la red de DC a ser clientes de route reflector de la red global. Se modifica la configuración de BGP de todos los nodos de la red de DC quedando exactamente igual que el resto de nodos de UFIN teniendo como vecinos de BGP los nuevos route reflector:

```
router bgp 64987
 bgp router-id 192.168.255.253
 no bgp default ipv4-unicast
 no bgp default route-target filter
 bgp log-neighbor-changes
 neighbor PARENT-RR peer-group
 neighbor PARENT-RR remote-as 64987
 neighbor PARENT-RR update-source Loopback0
 neighbor PARENT-RR timers 5 15
 neighbor 192.168.255.9 peer-group PARENT-RR
 neighbor 192.168.255.18 peer-group PARENT-RR
 !
 address-family vpnv4
  neighbor PARENT-RR send-community both
  neighbor PARENT-RR next-hop-self
  neighbor PARENT-RR advertisement-interval 1
  neighbor 192.168.255.9 activate
  neighbor 192.168.255.18 activate
 exit-address-family
```

Una vez hecho, comprobamos en los antiguos route reflector que ya solo comparten adyacencias de BGP con los nuevos route reflector en ambos nodos consultando sus vecinos de BGP:

```
bcn-rrr-01#sh ip bgp all summary
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.252, local AS number 64987
BGP table version is 1433556, main routing table version 1433556
7897 network entries using 1081889 bytes of memory
13097 path entries using 890596 bytes of memory
975/971 BGP path/bestpath attribute entries using 120900 bytes of memory
20 BGP rrinfo entries using 480 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
2 BGP community entries using 48 bytes of memory
600 BGP extended community entries using 35376 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 2129337 total bytes of memory
BGP activity 123863/115954 prefixes, 834282/821185 paths, scan interval 5 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.255.9   4 64987     421      15    64396    0    0 00:01:01        4909
192.168.255.18  4 64987     525      12    64396    0    0 00:01:51        4909
```

Y en el mad-rrr:

```
mad-rrr-01#sh ip bgp all summary
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.253, local AS number 64987
BGP table version is 1223488, main routing table version 1223488
7640 network entries using 1046680 bytes of memory
12580 path entries using 855440 bytes of memory
976/974 BGP path/bestpath attribute entries using 121024 bytes of memory
20 BGP rrinfo entries using 480 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
2 BGP community entries using 48 bytes of memory
600 BGP extended community entries using 35376 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 2059096 total bytes of memory
BGP activity 119177/111537 prefixes, 711973/699393 paths, scan interval 15
secs

Neighbor      V    AS MsgRcvd MsgSent TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.255.9  4 64987  692    284    1223488  0    0   00:03:12    4645
192.168.255.18 4 64987  803    296    1223488  0    0   00:04:00    4645
```

Hacemos lo propio con todos los nodos de la red de DC, donde hay que configurarlos para que establezcan vecindades BGP con los nuevos route reflector en vez de con los antiguos, en el ejemplo lo comprobamos con el de COR antes del mallado:

```
COR#sh ip bgp all summary
For address family: IPv4 Unicast
BGP router identifier 192.168.255.2, local AS number 64987
BGP table version is 1, main routing table version 1

Neighbor      V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.255.252 4 64987     203     401        1    0    0 00:00:21    532
192.168.255.253 4 64987     244     549        1    0    0 00:00:40    532
```

Y después de configurar los route reflector definitivos:

```
COR#sh ip bgp all summary
For address family: IPv4 Unicast
BGP router identifier 192.168.255.2, local AS number 64987
BGP table version is 1, main routing table version 1

Neighbor        V  AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.255.9   4 64987    3041     407    16    0    0 00:01:12     860
192.168.255.18  4 64987    3368     407    16    0    0 00:03:27     860
```

Durante todo este proceso, se ha modificado varias veces la tabla de adyacencias de BGP de los nuevos route reflector ya que se han ido añadiendo nuevos vecinos, aquí

mostramos la evolución de la tabla de adyacencias con los eventos que muestra el log del equipo. Se muestra comentado y en cursiva las acciones que se han ido realizando:

```
eds-rrr-01#term mon
eds-rrr-01#show ip bgp all summary
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.9, local AS number 64987
BGP table version is 431679, main routing table version 431679
3794 network entries using 534954 bytes of memory
5095 path entries using 346460 bytes of memory
375/371 BGP path/bestpath attribute entries using 28500 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
1 BGP community entries using 24 bytes of memory
302 BGP extended community entries using 11964 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 921926 total bytes of memory
BGP activity 70767/66973 prefixes, 249092/243997 paths, scan interval 15 secs

Neighbor             V      AS  MsgRcvd  MsgSent     TblVer   InQ  OutQ  Up/Down
State/PfxRcd
10.132.1.1      4 64987  196498  204625   431679    0    0 1w4d          523
10.132.1.2      4 64987  194589  206914   431679    0    0 1w4d         1832
10.132.1.3      4 64987  194101  204905   431679    0    0 1w4d         2200
10.132.1.4      4 64987  189460  206433   431679    0    0 1w4d          241
10.132.1.5      4 64987  189308  206434   431679    0    0 1w4d           27
10.132.1.6      4 64987  189564  206433   431679    0    0 1w4d          180
10.132.1.7      4 64987  189707  206433   431679    0    0 1w4d           86
192.168.255.18  4 64987  206131  206438   431679    0    0 1w4d            0
eds-rrr-01#
eds-rrr-01# !aquí se configura el bcn-rrr-01
eds-rrr-01#sh ip bgp all sum
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.9, local AS number 64987
BGP table version is 431679, main routing table version 431679
3794 network entries using 534954 bytes of memory
5095 path entries using 346460 bytes of memory
375/371 BGP path/bestpath attribute entries using 28500 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
1 BGP community entries using 24 bytes of memory
302 BGP extended community entries using 11964 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 921926 total bytes of memory
BGP activity 70767/66973 prefixes, 249092/243997 paths, scan interval 15 secs

Neighbor             V      AS  MsgRcvd  MsgSent     TblVer   InQ  OutQ  Up/Down
State/PfxRcd
10.132.1.1      4 64987  196505  204631   431679    0    0 1w4d          523
```

```
10.132.1.2      4 64987  194595  206920   431679    0    0 1w4d          1832
10.132.1.3      4 64987  194108  204911   431679    0    0 1w4d          2200
10.132.1.4      4 64987  189466  206439   431679    0    0 1w4d           241
10.132.1.5      4 64987  189314  206440   431679    0    0 1w4d            27
10.132.1.6      4 64987  189570  206439   431679    0    0 1w4d           180
10.132.1.7      4 64987  189714  206439   431679    0    0 1w4d            86
192.168.255.18  4 64987  206137  206444   431679    0    0 1w4d             0
192.168.255.252 4 64987       0       0        0    0    0 never    Idle

000337: Nov 23 01:28:03.271: %BGP-5-ADJCHANGE: neighbor 192.168.255.252 Up
eds-rrr-01#sh ip bgp all sum
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.9, local AS number 64987
BGP table version is 431840, main routing table version 431840
3955 network entries using 557655 bytes of memory
5256 path entries using 357408 bytes of memory
395/391 BGP path/bestpath attribute entries using 30020 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
1 BGP community entries using 24 bytes of memory
316 BGP extended community entries using 12464 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 957595 total bytes of memory
BGP activity 70928/66973 prefixes, 249253/243997 paths, scan interval 15 secs

Neighbor             V      AS  MsgRcvd  MsgSent     TblVer    InQ  OutQ  Up/Down
State/PfxRcd
10.132.1.1      4 64987  196514  204661   431840    0    0 1w4d           523
10.132.1.2      4 64987  194605  206950   431840    0    0 1w4d          1832
10.132.1.3      4 64987  194117  204941   431840    0    0 1w4d          2200
10.132.1.4      4 64987  189475  206469   431840    0    0 1w4d           241
10.132.1.5      4 64987  189324  206470   431840    0    0 1w4d            27
10.132.1.6      4 64987  189580  206469   431840    0    0 1w4d           180
10.132.1.7      4 64987  189723  206469   431840    0    0 1w4d            86
192.168.255.18  4 64987  206147  206454   431840    0    0 1w4d             0
192.168.255.252 4 64987      24     507   431840    0    0 00:00:02      161
eds-rrr-01#
eds-rrr-01#
eds-rrr-01#
000338: Nov 23 01:41:43.679: %SYS-5-CONFIG_I: Configured from console by
agonzalezca on vty0 (10.132.2.144)
eds-rrr-01#!aquí se configura el mad-rrr-01
000339: Nov 23 01:41:56.456: %BGP-5-ADJCHANGE: neighbor 192.168.255.253 Up
eds-rrr-01#
eds-rrr-01#sh ip bgp all sum
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.9, local AS number 64987
BGP table version is 432277, main routing table version 432277
4364 network entries using 615324 bytes of memory
5665 path entries using 385220 bytes of memory
413/409 BGP path/bestpath attribute entries using 31388 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
1 BGP community entries using 24 bytes of memory
```

```
328 BGP extended community entries using 13368 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1045348 total bytes of memory
BGP activity 71351/66987 prefixes, 249676/244011 paths, scan interval 15 secs

Neighbor             V      AS  MsgRcvd  MsgSent     TblVer   InQ  OutQ  Up/Down
State/PfxRcd
10.132.1.1      4 64987  196712  204885   432277    0    0 1w4d          523
10.132.1.2      4 64987  194797  207174   432277    0    0 1w4d         1832
10.132.1.3      4 64987  194310  205165   432277    0    0 1w4d         2200
10.132.1.4      4 64987  189668  206693   432277    0    0 1w4d          241
10.132.1.5      4 64987  189516  206694   432277    0    0 1w4d           27
10.132.1.6      4 64987  189772  206693   432277    0    0 1w4d          180
10.132.1.7      4 64987  189915  206693   432277    0    0 1w4d           86
192.168.255.18  4 64987  206342  206649   432277    0    0 1w4d            0
192.168.255.252 4 64987     661    1205   432277    0    0 00:16:27      161
192.168.255.253 4 64987     483     537   432277    0    0 00:02:34      409
eds-rrr-01#
eds-rrr-01#!se van configurando el resto de nodos
eds-rrr-01#
000348: Nov 23 02:55:59.824: %BGP-5-ADJCHANGE: neighbor 192.168.255.8 Up
000349: Nov 23 02:56:00.680: %BGP-5-ADJCHANGE: neighbor 192.168.255.7 Up
000350: Nov 23 02:56:38.759: %BGP-5-ADJCHANGE: neighbor 192.168.255.254 Up
eds-rrr-01#
eds-rrr-01#sh ip bgp all sum
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.9, local AS number 64987
BGP table version is 438147, main routing table version 438147
6907 network entries using 973887 bytes of memory
13359 path entries using 908412 bytes of memory
791/787 BGP path/bestpath attribute entries using 60116 bytes of memory
10 BGP rrinfo entries using 240 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
1 BGP community entries using 24 bytes of memory
498 BGP extended community entries using 30152 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1972879 total bytes of memory
BGP activity 74096/67189 prefixes, 257885/244526 paths, scan interval 15 secs

Neighbor             V      AS  MsgRcvd  MsgSent     TblVer   InQ  OutQ  Up/Down
State/PfxRcd
10.132.1.1      4 64987  197592  206732   438147    0    0 1w4d          519
10.132.1.2      4 64987  195653  209021   438147    0    0 1w4d         1825
10.132.1.3      4 64987  195168  207012   438147    0    0 1w4d         2178
10.132.1.4      4 64987  190520  208540   438147    0    0 1w4d          241
10.132.1.5      4 64987  190369  208541   438147    0    0 1w4d           27
10.132.1.6      4 64987  190625  208540   438147    0    0 1w4d          180
10.132.1.7      4 64987  190768  208540   438147    0    0 1w4d           86
192.168.255.1   4 64987     193    1367   438147    0    0 00:03:56     1248
192.168.255.2   4 64987     146    1457   438147    0    0 00:04:17      234
192.168.255.3   4 64987      66    1204   438147    0    0 00:02:56      102
```

```
192.168.255.4   4 64987      74    1239   438147    0    0 00:03:01       76
192.168.255.5   4 64987      80    1172   438147    0    0 00:02:26       39
192.168.255.6   4 64987      76    1112   438147    0    0 00:02:11      102
192.168.255.7   4 64987      60    1008   438147    0    0 00:01:15      120
192.168.255.8   4 64987      96    1102   438147    0    0 00:01:16      606
192.168.255.10  4 64987     208    1625   438147    0    0 00:16:41       40
192.168.255.18  4 64987  207650  207954   438147    0    0 1w4d            0
192.168.255.201 4 64987   72168   79932   438147    0    0 01:12:11      525
192.168.255.202 4 64987   71594   79349   438147    0    0 01:14:19      128
192.168.255.252 4 64987    2051    3013   438147    0    0 01:29:12     2733
192.168.255.253 4 64987    1897    2345   438147    0    0 01:15:19     2988
192.168.255.254 4 64987      20    1001   438147    0    0 00:00:37        9
```

- Impacto:

Esta tarea no tiene impacto ninguno sobre la red ya que mientras los PEs tengan conexión con un route reflector, y este esté mallado con el resto siempre va a conocer las rutas de cada VPN.

4.5.2.3 Migración de conexiones de los route reflector

Este paso consiste únicamente en cambiar los antiguos route reflector de ser route reflector a ser cliente de route reflector. En este caso, basta con calcar la configuración de cualquier PE de ambas redes, de hecho ya ni deberíamos considerar que hay dos redes.

La configuración es la siguiente, en este caso del antiguo bcn-rrr pero es exacta en todos los nodos de la red salvo los route reflector:

```
router bgp 64987
 bgp router-id 192.168.255.253
 no bgp default ipv4-unicast
 no bgp default route-target filter
 bgp log-neighbor-changes
 neighbor PARENT-RR peer-group
 neighbor PARENT-RR remote-as 64987
 neighbor PARENT-RR update-source Loopback0
 neighbor PARENT-RR timers 5 15
 neighbor 192.168.255.9 peer-group PARENT-RR
 neighbor 192.168.255.18 peer-group PARENT-RR
 !
 address-family vpnv4
  neighbor PARENT-RR send-community both
  neighbor PARENT-RR next-hop-self
  neighbor PARENT-RR advertisement-interval 1
  neighbor 192.168.255.9 activate
  neighbor 192.168.255.18 activate
 exit-address-family
```

Y el siguiente ejemplo muestra la tabla de adyacencias en un route reflector varias semanas después en una situación absolutamente estable. Podemos comprobar un dato

curioso, no aprende ninguna ruta del otro route reflector ya que él ya conoce todas las rutas directamente de los PEs:

```
eds-rrr-01#show ip bgp all summary
For address family: VPNv4 Unicast
BGP router identifier 192.168.255.9, local AS number 64987
BGP table version is 1035140, main routing table version 1035140
5133 network entries using 723753 bytes of memory
5142 path entries using 349656 bytes of memory
878/876 BGP path/bestpath attribute entries using 66728 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
1 BGP community entries using 24 bytes of memory
545 BGP extended community entries using 32076 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1172285 total bytes of memory
BGP activity 161198/156065 prefixes, 555386/550244 paths, scan interval 15
secs

Neighbor             V       AS  MsgRcvd  MsgSent     TblVer    InQ  OutQ  Up/Down
State/PfxRcd
10.132.1.2       4 64987  807000 1125347  1035140    0    0 6w4d          584
10.132.1.3       4 64987  806142 1122269  1035140    0    0 6w4d          953
10.132.1.4       4 64987  793074 1119330  1035140    0    0 6w4d          237
10.132.1.5       4 64987  138910  154378  1035140    0    0 1w1d           22
10.132.1.6       4 64987  792098 1118223  1035140    0    0 6w4d          178
10.132.1.7       4 64987  791908 1118224  1035140    0    0 6w4d           86
192.168.255.1    4 64987  622862  918735  1035140    0    0 5w0d          420
192.168.255.2    4 64987  629592  923219  1035140    0    0 5w0d          273
192.168.255.3    4 64987  661812  919524  1035140    0    0 5w0d          112
192.168.255.4    4 64987  620909  921867  1035140    0    0 5w0d          286
192.168.255.5    4 64987  615845  918505  1035140    0    0 5w0d           39
192.168.255.6    4 64987  113918  122343  1035140    0    0 6d13h          91
192.168.255.7    4 64987  619326  919427  1035140    0    0 5w0d           69
192.168.255.8    4 64987 1031924  918438  1035140    0    0 5w0d          551
192.168.255.10   4 64987  601374  913518  1035140    0    0 5w0d           42
192.168.255.11   4 64987  604526  878380  1035140    0    0 4w5d          478
192.168.255.18   4 64987 1116471 1117497  1035140    0    0 6w4d            0
192.168.255.201 4 64987  673282  988528  1035140    0    0 5w4d            3
192.168.255.202 4 64987  672708  987945  1035140    0    0 5w4d            3
192.168.255.252 4 64987  623160  918871  1035140    0    0 5w0d          217
192.168.255.253 4 64987  661116  918656  1035140    0    0 5w0d          496
192.168.255.254 4 64987  615630  917288  1035140    0    0 5w0d            2
```

El siguiente esquema nos muestra la topología física de la red con los nodos coloreados según su función en la red:

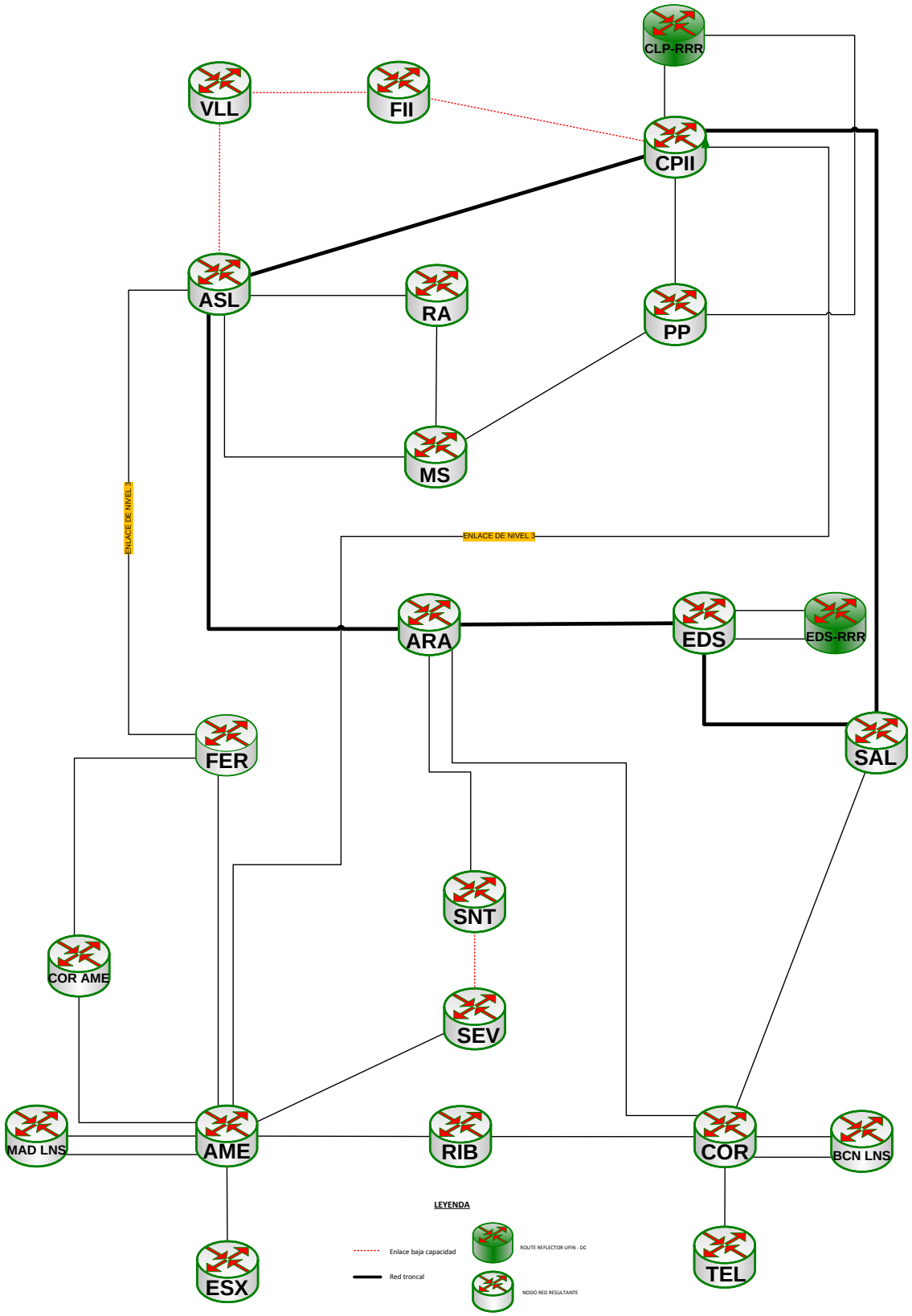

Ilustración 85- Esquema final según los route reflector de la red.

4.5.3 ELIMINACIÓN DEL NODO DE FER Y ALTA DE ENLACE ASL-AME (IMPLEMENTACIÓN FASE 0)

Estos trabajos constan de las siguientes tareas:

1. Movimiento a otros nodos de las conexiones del equipo de FER
2. Desconexión de la red troncal del nodo de FER
3. Establecimiento del nuevo enlace ASL-AME a partir de las conexiones del nodo de FER.

Por cuestiones de diseño, en la red de DC, este nodo se utiliza para recoger las conexiones de respaldo de las sedes importantes. Una vez unificadas ambas redes, la filosofía es que los respaldos van al nodo MPLS que tenga más facilidad de transmisión y no se centralizan solo en un nodo. Dado esto y debido a la situación de la red de transmisión resultante de su integración, este nodo se puede suprimir.

4.5.3.1 Movimiento a otros nodos de las conexiones del equipo de FER

Antes de llegar a este punto, se han ido moviendo a otros nodos, dependiendo de la red de transmisión, todos los backups que conectaban a este nodo.

Los trabajos consisten en:

- Configurar la VPN MPLS a la que pertenece el puerto de no estar creada, los protocolos de routing adecuados,
- Configurar el proceso de routing CE-PE que corresponda de la misma manera que estuviese en FER.
- Configurar el puerto de entrada al nuevo nodo con los mismos parámetros que tenía en FER (dirección IP, VRF, etc…) Si es un puerto Ethernet hay que configurar las colas de QoS de igual forma. Si el puerto de conexión es un tributario de una tarjeta STM1-c hay que configurar el KLM con la misma parametrización. Si es un E1 o E3 canalizado también habría que configurar los time-slots.
- Movimiento del circuito físico. Dependiendo de la tecnología y las posibilidades o se ha creado un circuito nuevo o se ha movido el existente en la red de transmisión. Por ejemplo, si el circuito acaba en un tributario del un STM1-c se hacen las crosconexiones necesarias en la red SDH.

  - <u>Impacto:</u>

No supone impacto en la red puesto que se trata de vías secundarias. En el momento del cambio se pierde la redundancia pero no hay afectación en el servicio.

4.5.3.2 Desconexión de la red troncal del nodo de FER

Esta parte simplemente supone la desinstalación del equipo. En local a través del cable de consola se borra toda la configuración dejando su configuración de fábrica y tras ello se desconectan las tres vías troncales que tenía contra los nodos de COR AME, AME, y ASL (esta última es una conexión a nivel 3 que se puso inicialmente para extender la funcionalidad a toda la red). Una vez se ha hecho esto, el nodo está aislado de la red. Se puede quitar la alimentación y desmontarlo físicamente. Su uso se limita a tenerlo en cuenta para futuras ampliaciones de la red. La tarjeta de puertos Ethernet WS-X6748-GE-TX se reutiliza y se pasará al nodo de CPII puesto que era necesaria la ampliación. En los equipos Cisco de la familia 760x la inserción de tarjetas se puede hacer en caliente, se reconoce automáticamente por lo que no es necesario programar ventana de trabajo.

Es importante reseñar, que este equipo tiene otra vía con el nodo de COR AME. Si eliminamos el nodo de FER de la red, el de COR AME queda con una única conexión al resto de la red lo cual es bastante peligroso por temas de saturación de enlace y redundancia. Se ha hecho así ya que el nodo de COR AME tiene como única función en la red la de proporcionar una vía de respaldo para la gestión de la red MPLS. Visto esto, y dado que ya hay gestión por la parte de la red de UFIN se puede asumir este riesgo y dejar el nodo de COR AME con una única vía a la red. En un escenario real si habría tenido más sentido desmontar primero el nodo de COR y después el de FER, pero dado el bajo riesgo de dejar COR AME por una sola vía, se estima más necesario disponer del enlace de ASL-AME, que además forma parte de la nueva troncal, que disponer de una cuarta vía para la gestión de la red.

- Impacto:

Estos trabajos no suponen impacto en la red. Al eliminarse las conexiones de este equipo, puede que se estuviera utilizando esta vía si el LDP calcula basándose en el IGP que es el mejor camino para un FEC dado. Esto, dada la complejidad de los cálculos es muy complicado de ver. Si algún paquete estaba siendo encaminado por este enlace, al eliminarlo se va a recalcular el LDP y tomará otro camino casi automáticamente por lo que las comunicaciones finales no lo notarán.

4.5.3.3 Establecimiento del nuevo enlace ASL-AME a partir de las conexiones del nodo de FER

El nodo de FER tenía dos conexiones, una contra ASL y otra contra AME, cuando se decidió la arquitectura final, se estableció que AME y ASL formarían parte del anillo troncal por lo que es necesario unirlos de manera directa. Dada la red de transmisión, geográficamente FER es un punto por el que pasa el camino de fibra óptica resultante por lo que es necesario realizar un tránsito en un repartidor óptico de FER para dar continuidad al circuito.

Previo a levantar el enlace hay que configurar, en el lado de ASL el puerto como puerto de troncal y no de VPN MPLS que era el servicio que se estaba prestando por ahí. Esa vía ya no se estaba usando puesto que en la fase anterior se integraron a nivel de VPN MPLS

En la siguiente ilustración vemos el detalle (a nivel 3) de lo que suponen los cambios en la red.

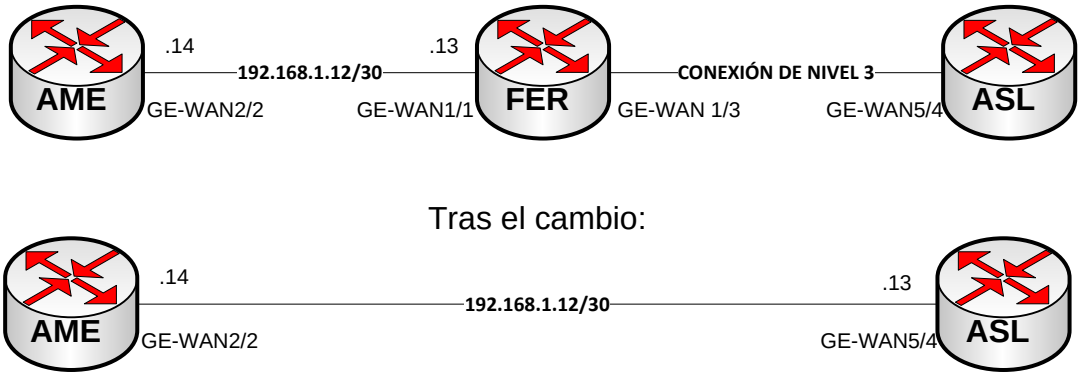

Ilustración 86- Detalle de la eliminación del nodo de FER.

- Impacto:

El alta de este enlace no supone impacto en la red. Se trata de ofrecer otro camino más al LDP para que pueda encaminar tráfico dado un inicio y fin de las comunicaciones determinado.

Tras el alta de este nuevo enlace, el esquema físico de la red queda como sigue a continuación:

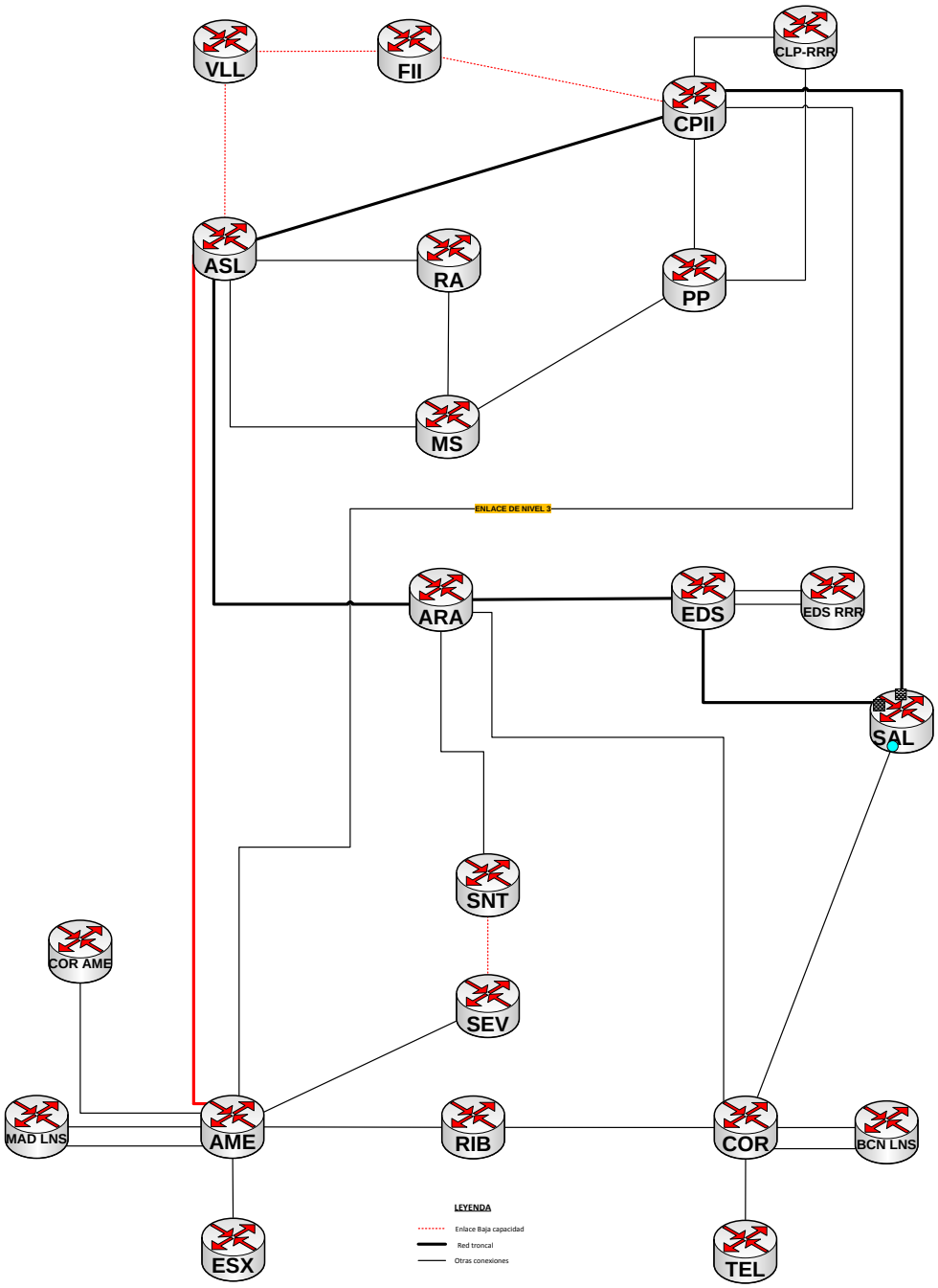

Ilustración 87- Red MPLS tras la eliminación del nodo de FER.

4.5.4 Eliminación del nodo de COR (implementación Fase 0)

Inmediatamente después de eliminar el nodo de FER se puede eliminar el de COR AME. Como se avanzó en el punto anterior, la única función del nodo de COR AME es la de ser vía secundaria para la gestión de la red MPLS. Una vez se han integrado las redes, se disponen de cuatro vías para la gestión, con dos es más que de sobra para este cometido y se ha

decidido que se quedarán la de MS y la de FII. Estas son las dos ubicaciones donde se gestionaba la red exUFIN y lo seguirá siendo para toda la red puesto que es en estas dos ubicaciones donde se encuentra el centro de operación de red.

La eliminación de este equipo es extremadamente sencilla puesto que su función se considera innecesaria. Los trabajos consisten en eliminar las conexiones del equipo, configurarlo con sus valores de fábrica, desalimentarlo y eliminarlo del rack. En el extremo de AME se borra la configuración del interfaz GigabitEthernet 2/1 y se deja desconectado para otros usos.

- Impacto:

El impacto de estos trabajos de nuevo es nulo puesto que no hay servicio en ese equipo.

Tras ello, el esquema de la red es como se muestra en la siguiente figura:

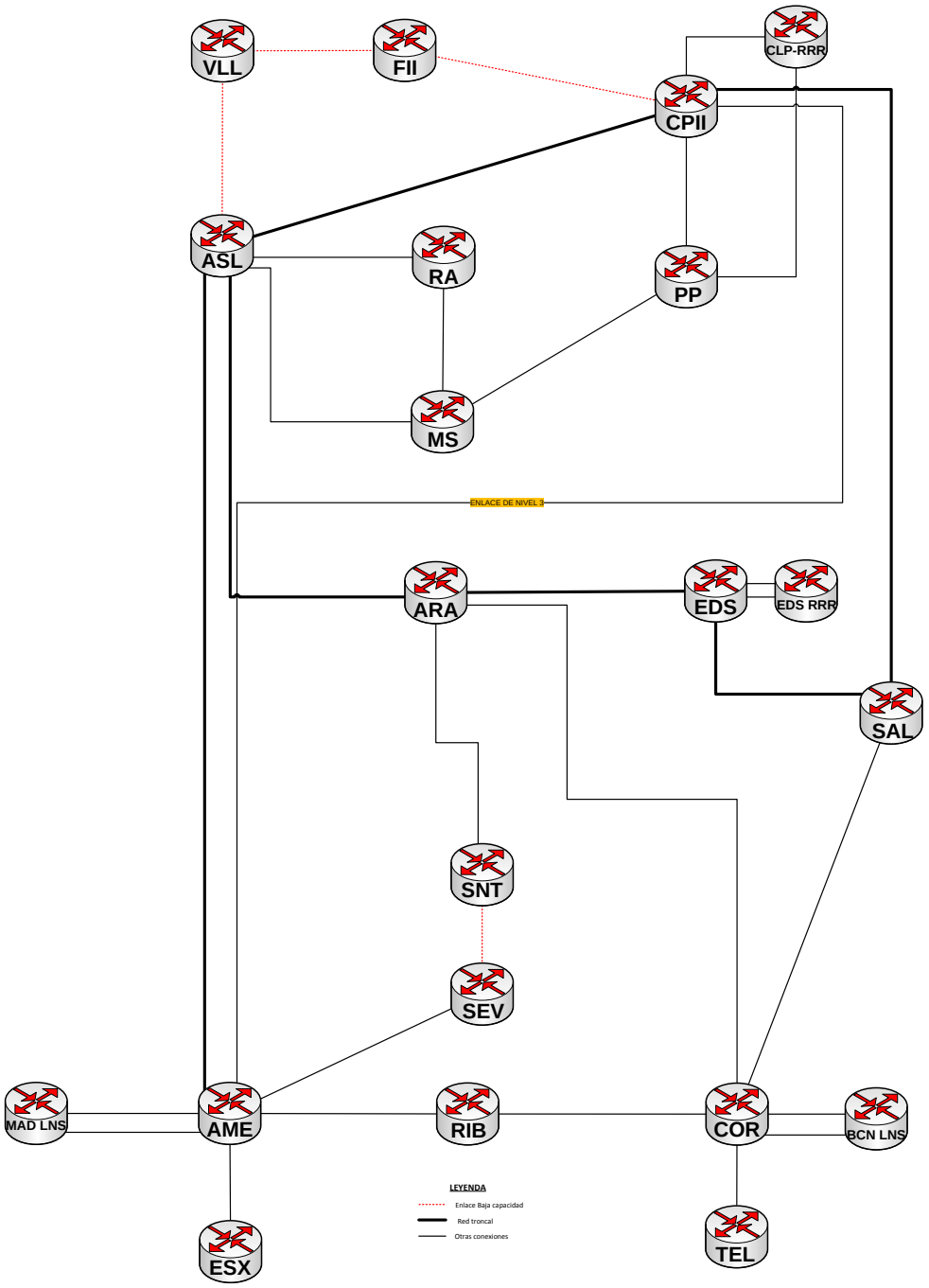

Ilustración 88- Red MPLS tras la eliminación del nodo de COR AME.

4.5.5 Alta del enlace RIB-CPII y baja del enlace RIB-AME (implementación Fase 0)

El siguiente enlace que estuvo disponible es el que unirá el nodo de RIB con el de CPII y se aprovecha y se da de baja el RIB-AME. Se podrían dejar los dos, pero puesto que por las decisiones de diseño expuestas en el apartado “Arquitectura final” el nodo de RIB iría unido al

de CPII, se elimina la de AME-CPII ya que AME y CPII cubren la misma área geográfica y no es necesario este enlace.

Para este caso, dado que no podemos dejar por una única vía el nodo, no vamos a reutilizar el puerto que conecta RIB-AME, primero estableceremos el nuevo enlace y luego daremos de baja el antiguo.

Por tanto las dos partes de estos trabajos son:

1. Alta del enlace RIB-CPII
2. Baja del enlace RIB-AME

4.5.5.1 Alta del enlace RIB-CPII

Como en el alta de todos los enlaces, las tareas a realizar son, configuración del enlace en cuanto se refiere a direccionamiento IP, LDP e IGP (OSPF). La MTU tal y como se homologó será de 1900 bytes.

Una vez están configurados los nuevos puertos en ámbos extremos, levantamos los interfaces y comprobamos la conectividad:

```
rib-rpe-01#ping 192.168.1.161 source 192.168.1.162 repeat 1000 size 1900 df-
bit

Type escape sequence to abort.
Sending 1000, 1900-byte ICMP Echos to 192.168.1.161, timeout is 2 seconds:
Packet sent with a source address of 192.168.1.162
Packet sent with the DF bit set
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!
Success rate is 100 percent (1000/1000), round-trip min/avg/max = 4/6/24 ms
```

Vemos que no se pierde ningún paquete y los tiempos son lo esperado dada la distancia del enlace.

Comprobamos que se establecen las adyacencias de OSPF con el nuevo enlace (192.168.255.13) además de mantener los que ya tenía con AME y COR

```
rib-rpe-01#sh ip ospf 2328 neigb

Neighbor ID     Pri   State           Dead Time   Address         Interface
192.168.255.13    0   FULL/  -        00:00:36    192.168.1.161   GE-WAN2/2
192.168.255.2     0   FULL/  -        00:00:38    192.168.1.37    GE-WAN2/1
192.168.255.1     0   FULL/  -        00:00:32    192.168.1.5     GE-WAN1/1
```

También comprobamos las adyacencias de LDP con sus nodos vecinos:

```
rib-rpe-01#sh mpls ldp nei
    Peer LDP Ident: 192.168.255.2:0; Local LDP Ident 192.168.255.4:0
        TCP connection: 192.168.255.2.646 - 192.168.255.4.25574
        State: Oper; Msgs sent/rcvd: 20968/20957; Downstream
        Up time: 1w5d
        LDP discovery sources:
          GE-WAN2/1, Src IP addr: 192.168.1.37
        Addresses bound to peer LDP Ident:
          192.168.255.2   192.168.1.45    192.168.1.41    192.168.1.37
          192.168.1.33    192.168.1.81    192.168.1.49    192.168.1.250
          172.30.2.254
    Peer LDP Ident: 192.168.255.1:0; Local LDP Ident 192.168.255.4:0
        TCP connection: 192.168.255.1.646 - 192.168.255.4.54740
        State: Oper; Msgs sent/rcvd: 19173/19169; Downstream
        Up time: 1w4d
        LDP discovery sources:
          GE-WAN1/1, Src IP addr: 192.168.1.5
        Addresses bound to peer LDP Ident:
          192.168.255.1   192.168.1.21    192.168.1.9     192.168.1.1
          192.168.1.193   192.168.1.5     192.168.1.14    192.168.1.157
          192.168.1.17    1.1.1.1         192.168.1.25
     Peer LDP Ident: 192.168.255.13:0; Local LDP Ident 192.168.255.4:0
        TCP connection: 192.168.255.13.32336 - 192.168.255.4.646
        State: Oper; Msgs sent/rcvd: 96/97; Downstream
        Up time: 00:03:02
        LDP discovery sources:
          GE-WAN2/2, Src IP addr: 192.168.1.161
        Addresses bound to peer LDP Ident:
          192.168.1.154   192.168.1.126   192.168.1.109   192.168.48.9
          192.168.255.13  10.132.1.3      192.168.1.133   192.168.1.158
          192.168.1.86    192.168.1.161
```

- Impacto:

No hay impacto en la red, añadir un enlace supone ofrecer otro camino más al tráfico. Si se reencamina tráfico por este enlace no supone afectación a ningún servicio.

4.5.5.2 Baja del enlace RIB-AME

Una vez se ha establecido el enlace RIB-CPII, ya podemos eliminar el de RIB-AME. En este caso es bastante sencillo porque basta por poner en "shutdown" el interfaz GE-WAN 1/1 y limpiar la configuración del mismo.

- Impacto:

En el caso de dar de baja un enlace, supone que el tráfico que se está cursando por aquí debe ser repartido por el resto de enlaces. La conmutación es muy rápida y no supone corte de servicio, además que, en caso de producirse, es muy complicado determinar qué tráfico es el afectado.

Tras estos trabajos el esquema de la red queda como se muestra en la siguiente ilustración:

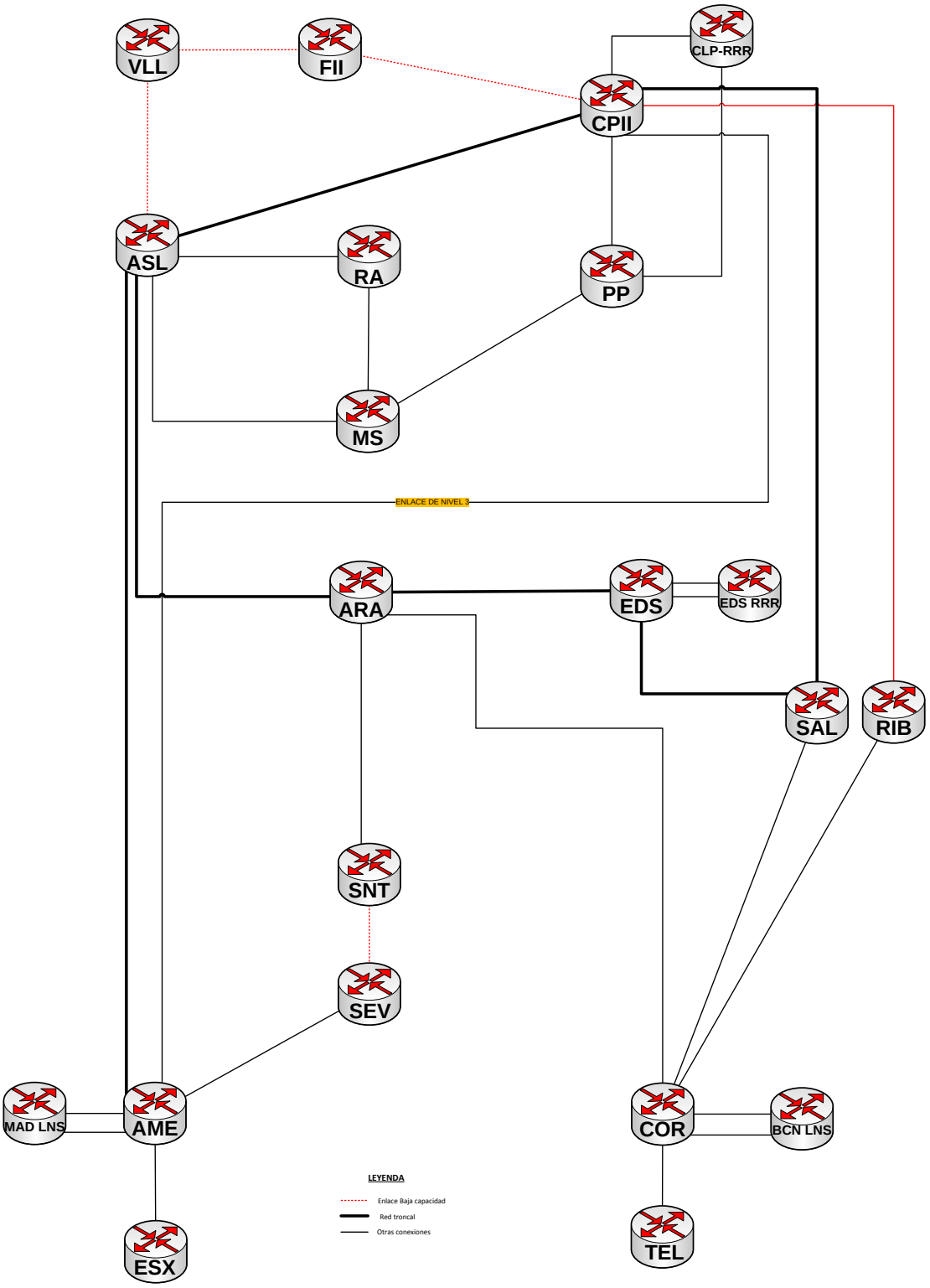

Ilustración 89- Red MPLS tras el cambio de conexiones de RIB-AME por RIB-CPII

4.5.6 ALTA DEL ENLACE RIB-SAL Y BAJA DEL ENLACE RIB-COR (IMPLEMENTACIÓN FASE 0)

Para acabar la arquitectura física en lo que al nodo de RIB se refiere, hay que modificar un enlace, según la arquitectura diseñada, el nodo de RIB conectará con EDS en vez de con

COR. COR y SAL cubren la misma área geográfica por lo que no es necesario mantener dos enlaces.

Hay que realizar dos tareas:

1. Alta del enlace SAL-RIB
2. Baja del enlace COR-RIB.

4.5.6.1 Alta del enlace SAL-RIB

Los pasos son exactamente los mismos que en el caso anterior, se trata de configurar los puertos que se van a enfrentar (direccionamiento IP) y también el IGP (OSPF) y LDP.

Una vez configurados ambos extremos se activan los puertos y se realizan las comprobaciones:

```
rib-rpe-01#ping 192.168.255.8 source 192.168.255.4 repeat 1000 size 1896 df-
bit

Type escape sequence to abort.
Sending 1000, 1896-byte ICMP Echos to 192.168.255.8, timeout is 2 seconds:
Packet sent with a source address of 192.168.255.4
Packet sent with the DF bit set
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!
Success rate is 100 percent (1000/1000), round-trip min/avg/max = 4/6/104 ms
```

Y comprobamos que se ha establecido correctamente las adyacencias de OSPF y LDP:

```
rib-rpe-01#sh ip ospf 2328 nei

Neighbor ID     Pri   State           Dead Time   Address         Interface
192.168.255.13    0   FULL/  -        00:00:12    192.168.1.161   GE-WAN2/2
192.168.255.8     0   FULL/  -        00:00:11    192.168.1.98    GE-WAN1/2
192.168.255.2     0   FULL/  -        00:00:13    192.168.1.37    GE-WAN2/1
192.168.255.1     0   FULL/  -        00:00:17    192.168.1.5     GE-WAN1/1
rib-rpe-01#sh mpls ldp nei
    Peer LDP Ident: 192.168.255.2:0; Local LDP Ident 192.168.255.4:0
```

```
        TCP connection: 192.168.255.2.646 - 192.168.255.4.25574
        State: Oper; Msgs sent/rcvd: 20968/20957; Downstream
        Up time: 1w5d
        LDP discovery sources:
          GE-WAN2/1, Src IP addr: 192.168.1.37
        Addresses bound to peer LDP Ident:
          192.168.255.2   192.168.1.45    192.168.1.41    192.168.1.37
          192.168.1.33    192.168.1.81    192.168.1.49    192.168.1.250
          172.30.2.254
    Peer LDP Ident: 192.168.255.1:0; Local LDP Ident 192.168.255.4:0
        TCP connection: 192.168.255.1.646 - 192.168.255.4.54740
        State: Oper; Msgs sent/rcvd: 19173/19169; Downstream
        Up time: 1w4d
        LDP discovery sources:
          GE-WAN1/1, Src IP addr: 192.168.1.5
        Addresses bound to peer LDP Ident:
          192.168.255.1   192.168.1.21    192.168.1.9     192.168.1.1
          192.168.1.193   192.168.1.5     192.168.1.14    192.168.1.157
          192.168.1.17    1.1.1.1         192.168.1.25
    Peer LDP Ident: 192.168.255.8:0; Local LDP Ident 192.168.255.4:0
        TCP connection: 192.168.255.8.26946 - 192.168.255.4.646
        State: Oper; Msgs sent/rcvd: 98/97; Downstream
        Up time: 00:04:20
        LDP discovery sources:
          GE-WAN1/2, Src IP addr: 192.168.1.98
        Addresses bound to peer LDP Ident:
          192.168.255.8   192.168.1.77    192.168.1.42    192.168.1.85
          192.168.1.98
    Peer LDP Ident: 192.168.255.13:0; Local LDP Ident 192.168.255.4:0
        TCP connection: 192.168.255.13.32336 - 192.168.255.4.646
        State: Oper; Msgs sent/rcvd: 96/97; Downstream
        Up time: 00:03:02
        LDP discovery sources:
          GE-WAN2/2, Src IP addr: 192.168.1.161
        Addresses bound to peer LDP Ident:
          192.168.1.154   192.168.1.126   192.168.1.109   192.168.48.9
          192.168.255.13  10.132.1.3      192.168.1.133   192.168.1.158
          192.168.1.86    192.168.1.161
```

Todos los datos son correctos por lo que ya hemos concluido la arquitectura física del nodo de RIB. En la siguiente ilustración vemos como queda el esquema total de la red MPLS tras este cambio:

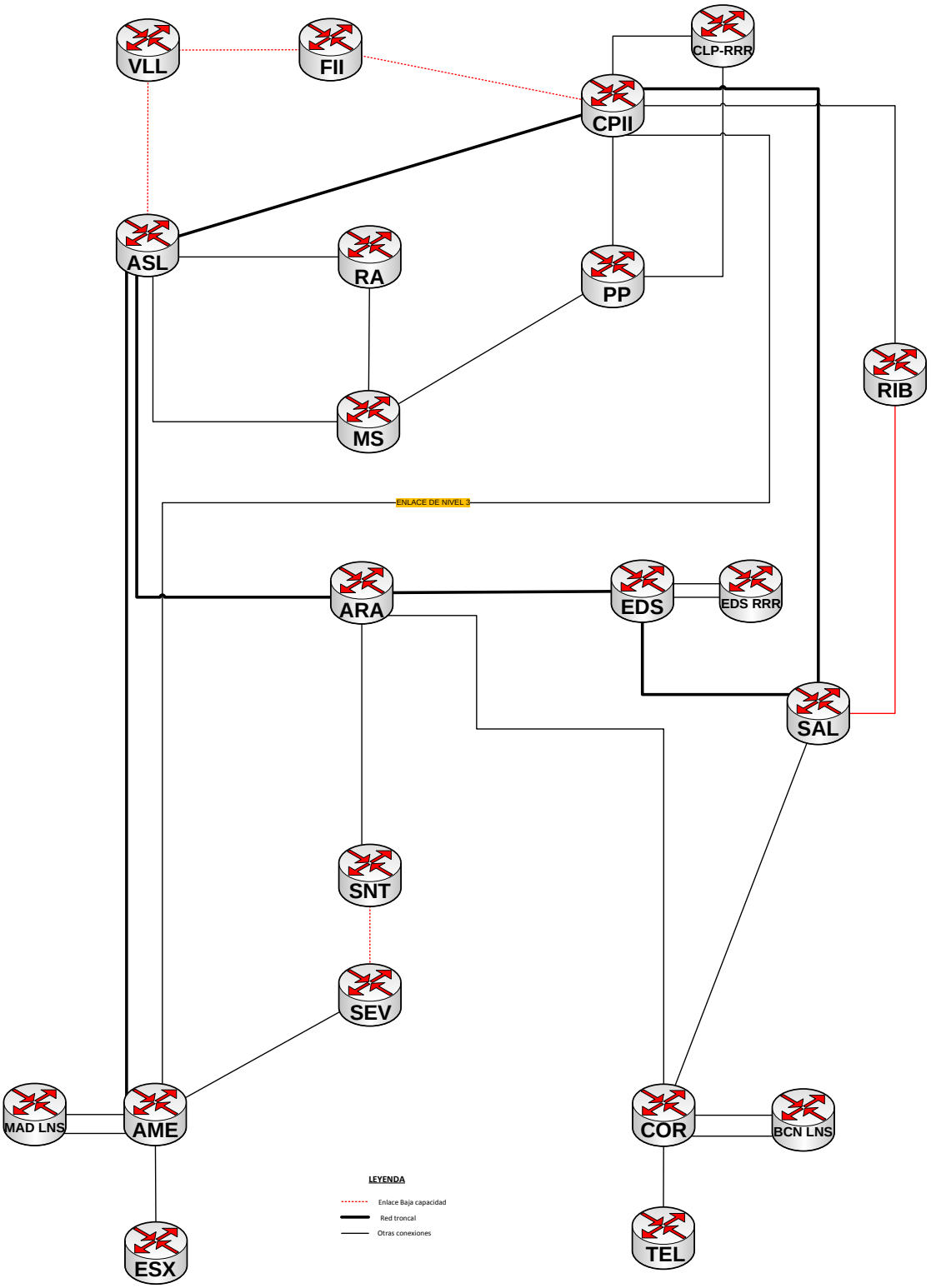

Ilustración 90- Red MPLS tras el cambio de conexiones de RIB-COR por RIB-SAL

4.6 FASE 4. HOMOGENEIZACIÓN DE VERSIONES DE IOS

Durante esta fase, lo que se pretende es la homogeneización de las versiones de software de los nodos MPLS. El software de los nodos se denomina IOS (Internetworking Operating System) y es propietario de Cisco.

Se realizará en 3 pasos:

1. Elección de la versión de IOS a instalar
2. Pruebas de homologación en el hardware de la red
3. Instalación de la nueva versión en los nodos MPLS

4.6.1 ELECCIÓN DE LA VERSIÓN DE IOS A INSTALAR

Antes de la migración, en la parte de DC existían las siguientes versiones:

- Advanced IP Services 12.2(18)SXF6
- Advanced IP Services 12.2(18)SXF7
- Advanced IP Services 12.2(18)SXF8

Los requisitos para soportar estas versiones son:

- ADVANCED IP SERVICES SSH
- s72033-advipservicesk9_wan-mz.122-18.SXF17a.bin
- Fecha de lanzamiento: 19/Mar/2010
- Tamaño: 79052.04 KB (80949284 bytes)
- Requisito mínimo de memoria: DRAM:512 MB Flash:128 MB

La parte de UFIN está homogénea en lo que a versión de IOS se refiere y tienen cargada la versión:

- Advanced Enterprise 12.2(33)SRC

Los requisitos para soportarla son:

- ADVANCED ENTERPRISE SERVICES SSH
- c7600s72033-adventerprisek9-mz.122-33.SRC.bin
- Fecha de lanzamiento: 14/Jan/2008
- Tamaño: 121563.04 KB (124480548 bytes)

- Requisito mínimo de memoria: DRAM:512 MB Flash:256 MB

Por último, cuando se sustituyó el nodo de EDS, referido en esta memoria en el primer punto de la fase 2, no estaba hecho el estudio de la versión a instalar y se instaló la que se creyó que sería la definitiva, cosa que no fue así tras el estudio. Esta es:

- Advanced Enterprise 12.2(33)SRC6

Los requisitos para soportar esta versión son los siguientes:

- ADVANCED ENTERPRISE SERVICES SSH
- c7600s72033-adventerprisek9-mz.122-33.SRC6.bin
- Fecha de lanzamiento: 11/Mar/2010
- Tamaño: 122076.07 KB (125005892 bytes)
- Requisito mínimo de memoria: DRAM:512 MB Flash:256 MB

En el momento de la realización del proyecto, existían cuatro versiones alternativas de software disponibles para la actualización, que son:

- IOS 122-33.SRC6
- IOS 122-33.SRD5
- IOS 122-33.SRE2
- IOS 150-1.S

Pasamos a analizar cada una de ellas:

4.6.1.1 IOS 122-33.SRC6

Los requisitos son:

- ADVANCED ENTERPRISE SERVICES SSH
- c7600s72033-adventerprisek9-mz.122-33.SRC6.bin
- Fecha de lanzamiento: 11/Mar/2010
- Tamaño: 122076.07 KB (125005892 bytes)
- Requisito mínimo de memoria: DRAM:512 MB Flash:256 MB

Esta versión está en su ciclo final de vida. Aún esta soportada por Cisco pero no se desarrollan nuevas revisiones y la corrección de errores se realizará mediante migración a versiones superiores.

Salvo ocho equipos en la red, dos de los cuales se dan de baja, podría implantarse en todos los nodos de la red.

Corrige algunos bugs existentes en las versiones actuales aunque introduce alguna problemática nueva.

Consideramos que esta versión si bien permite la unificación de la red aporta poco mas, siendo de preveer una nueva actualización en un plazo de tiempo breve.

4.6.1.2 IOS 122-33.SRD5

Los requisitos son:

- ADVANCED ENTERPRISE SERVICES SSH
- c7600s72033-adventerprisek9-mz.122-33.SRD5.bin
- Fecha de lanzamiento: 03/Oct/2010
- Tamaño: 132167.82 KB (135339844 bytes)
- Requisito mínimo de memoria: DRAM:512 MB Flash:256 MB

Esta versión esta soportada por Cisco y se desarrollaran nuevas revisiones hasta mayo de 2012.

Salvo ocho equipos en la red, dos de los cuales se darán de baja, podría implantarse en todos los nodos de la red. El coste orientativo de la actualización es de 1050€ por nodo.

Corrige determinados bugs existentes en las versiones actuales y es susceptible de ir corrigiendo los restantes u otros que se detecten en sucesivas revisiones.

Consideramos que esta versión permite la unificación de la versión de IOS en la red con unas garantías aceptables. No corrige todos los bugs existentes pero si avanza en este sentido respecto a la versión anterior. El coste requerido es similar al caso anterior por lo que esto no debe ser una traba.

4.6.1.3 IOS 122-33.SRE2

Los requisitos son:

- ADVANCED ENTERPRISE SERVICES SSH
- c7600s72033-adventerprisek9-mz.122-33.SRE2.bin
- Fecha de lanzamiento: 10/Aug/2010
- Tamaño: 153825.61 KB (157517420 bytes)
- Requisito mínimo de memoria: DRAM:1024 MB Flash:256 MB

Esta versión esta soportada por Cisco. No existe anuncio de fin de soporte.

Para poder realizar la actualización a esta versión sería necesario actualizar todas las tarjetas supervisoras. El coste mínimo orientativo de la actualización es de 5050€ por nodo, salvo los que requieren actualización de Flash en cuyo caso serían 6100€ por nodo.

Es de esperar que corrija la mayoría de los bug existentes en las versiones actuales y es susceptible de ir corrigiendo los restantes u otros que se detecten en sucesivas revisiones.

Consideramos que esta versión ofrece mayores ventajas a nivel técnico ya que el número de bug corregidos es mayor. No obstante la inversión requerida es considerable y a la fecha no corrige algunos bugs que afectan a la red e igualmente pueden ser corregidos en la línea de desarrollo anterior.

4.6.1.4 IOS 150-1.S

Los requisitos son:

- ADVANCED ENTERPRISE SERVICES SSH
- c7600s72033-adventerprisek9-mz.150-1.S.bin
- Fecha de lanzamiento: 30/Jul/2010
- Tamaño: 162365.02 KB (166261772 bytes)
- Requisito mínimo de memoria: DRAM:1024 MB Flash:256 MB

Para poder realizar la actualización a esta versión sería necesario actualizar todas las tarjetas supervisoras. El coste mínimo orientativo de la actualización es de 5050€ por nodo, salvo los que requieren actualización de Flash en cuyo caso serían 6100€ por nodo.

Es de esperar que corrija la mayoría de los bug existentes en las versiones actuales y es susceptible de ir corrigiendo los restantes u otros que se detecten en sucesivas revisiones.

Consideramos que es una línea de software muy reciente y dado que no hay ninguna funcionalidad en la misma que aconseje su implantación inmediata se recomienda esperar a que se alcance una situación de madurez en el desarrollo de esta versión, máxime teniendo en cuenta que el coste de esta versión es considerable, como en el caso anterior.

Como resultado del análisis realizado elegimos **migrar a la versión de IOS 12.2(33)SRD5**, recientemente publicada por Cisco. Lógicamente ha pesado el presupuesto en nuestra decisión ya que técnicamente no es la más avanzada pero si es la que cumple con las expectativas con un coste más bajo.

4.6.2 PRUEBA DE HOMOLOGACIÓN EN EL HARDWARE DE LA RED

La versión a la que se ha sido probada en laboratorio es la c7600s72033-adventerprisek9- mz.122-33.SRD5. Es decir, la Advanced Enterprise 12.2(33) SRD5.

Para estas pruebas de homologación se han tomado dos configuraciones, una de cada una de las redes que se están unificando: CPII de la red de UFIN y MAD de la red de DC.

Como complemento para algunas de las funciones y hardware no incluidos en estos equipos (específicamente la tarjeta de Firewall WS-SVC-FWM-1) se ha usado parte de la configuración del equipo COR de la red de DC.

Todas las configuraciones existentes se han cargado correctamente en estos equipos y han pasado pruebas básicas de funcionamiento.

Sólo ha existido un comando utilizado en la red de DC que no está soportado en esta nueva versión. Se trata del uso de comunidades extendidas en la redistribución de rutas EIGRP en BGP y su funcionalidad es evitar la creación de bucles de enrutado. Esta misma funcionalidad se puede obtener mediante el uso de la funcionalidad SoO (Site of Origin). Además en el proceso de integración de redes está planificada la eliminación del protocolo de routing EIGRP de la red, por lo que este problema no afectará a la configuración una vez eliminado este protocolo.

Respecto a los bugs que se han detectado en las versiones ejecutándose actualmente en la red, esta versión corrige la mayor parte de ellos. Hay un bug que sólo ha aparecido una vez en un equipo (CPII) y no va a ser corregido en esta rama. Sería necesario migrar a la rama SRE.

El segundo bug no está documentado ni afirmativa ni negativamente en la rama SRD ni SRE.

El hardware evaluado se muestra en la siguiente tabla junto con el resultado de las pruebas.

Hardware	Resultado
CISCO7609	OK
WS-SUP720-3B	OK
WS-X6582-2PA	OK (versión 1.8 de firmware mínimo)
PA-E3	OK
WS-X6148A-GE-TX	OK
WS-SVC-FWM-1	OK
WS-X6748-GE-TX	OK
PA-MC-8TE	OK
7600-SIP-400	OK
SPA-2X1GE	OK
OSM-2+4GE-WAN+	OK

Tabla 10. –Hardware homologado en maqueta

Las siguientes tarjetas no se han podido probar en laboratorio pero Cisco afirma en su herramienta de configuración que está soportada:

Hardware	Resultado
SPA-2XOC3-POS	OK
SPA-5X1GE-V2	OK
PA-MC-STM-1SMI	OK
PA-8T-V35	OK
PA-MC-8TE1	OK
PA-2E3	OK
PA-A6-E3	OK
WS-X6724-SFP	OK

Tabla 11. –Hardware homologado por indicaciones de la página de Cisco

A continuación se describe la lista de bugs detectados en las versiones antiguas de las redes de DC y UFIN junto con el estado en la versión elegida:

CSCsm06762

Produce la pérdida de ciertas rutas aprendidas por RIP

Estado en la versión homologada: **Resuelto**

CSCsw36285

El comando "show policy-map interface" da información incorrecta (los contadores de cada clase no muestran los valores reales)

Estado en la versión homologada: **Resuelto**

CSCso65821

Aplicarlo a una interfaz se propaga la configuración a todos los puertos de la tarjeta

Estado en la versión homologada: **Resuelto**

CSCtc09913

Bloqueo de sesiones Telnet

Estado en la versión homologada: **Presente**

Este bug no se va a resolver en la rama SRD. La primera versión de la gama SR en la que está solucionado es en la 12(33)SRE1.

CSCsq76103

Crash del router al realizar un comando “show” en paralelo con otro comando.

Estado en la versión homologada: **Indeterminado.**

La web de Cisco no proporciona información fiable sobre la presencia o no de este bug en esta versión ni en las anteriores de esta rama.

CSCsd55059

Lentitud en la respuesta a una petición de lectura de la flash.

Estado en la versión homologada: **Resuelto.**

CSCti77937

Inestabilidad de rutas en la interconexión de dos VPNs

Este bug es privado de Cisco y no se puede verificar normalmente en la web de Cisco. La única referencia es la información escrita por el ingeniero de Cisco en el caso correspondiente (SR 615533891) que indica textualmente:

This seems indeed be related to CSCti77937 as previously informed. I also tested SRD release, but the issue is still there. The behavior is not seen in SRE releases.

Por tanto el bug sigue presente.

Estado en la versión homologada: **Presente**

Los siguientes comandos no son aceptados por la IOS SRD5 pero su ausencia no resulta crítica. Las razones se explican en cada comando:

```
tag-switching tdp router-id Loopback0
```

Este comando sólo se aplica con el protocolo de etiquetas TDP y la red utiliza el protocolo LDP. Por tanto este comando es innecesario y no tiene ninguna función.

```
fair-queue (en la interfaz)
```

Este comando es obsoleto y ha sido eliminado de la versión de IOS evaluada. La recomendación de Cisco es aplicarlo a través de la class-default del policy-map aplicado a la interfaz:

```
Router(config)# policy-map policy-map-name
Router(config-pmap)# class class-default
Router(config-pmap-c)# fair-queue
Router(config-pmap-c)# fair-queue dynamic-queues
Router(config-pmap-c)# fair-queue queue-limit packets
```

Sin embargo al aplicarlo da el siguiente error:

```
fair-queue command is not supported in output direction for this interface
Configuration failed!
```

La ausencia de este comando en la configuración no es crítica ya que se aplica únicamente a la clase por defecto que no estamos usando en la red.

Otro tema que hay que evaluar dada su importancia, es el de los comandos por defecto, dependiendo de la versión los comandos adoptan por defecto unos valores u otros.

C7600S72033-ADVENTERPRISEK9-MZ.122-33.SRC

Los siguientes comandos no existen en la versión SRC y si en la versión SRD5. Se trata de comandos que configuran la configuración entre procesos del router. Decidimos dejarlos por defecto:

```
ipc holdq threshold upper 0
ipc holdq threshold lower 0
ipc ru-action on
ipc header-cache permanent 6000 500
ipc buffers min-free 112
ipc buffers max-free 1120
```

```
ipc buffers permanent 504
```

El siguiente comando no está documentado ni aparece en la configuración de la versión SRC. Según la documentación de cisco el valor por defecto existe y es el mismo que en la versión SRD5, por tanto lo debe modificarse la configuración al portarla:

```
ip icmp redirect subnet
```

El siguiente comando no está documentado ni aparece en la configuración de la versión SRC. Dado que son comandos de depuración que no afectan al funcionamiento del equipo hasta que no se activa el snooping en un puerto, no es necesaria ninguna configuración especial:

```
ip dhcp snooping database write-delay 300
ip dhcp snooping database timeout 300
```

El siguiente comando no existe en la versión SRD5 y afecta únicamente a la configuración del servicio CNS (Cisco Networking Services) que no está activo en nuestra red

```
cns message format notification version 1
```

S72033-ADVIPSERVICESK9_WAN-MZ.122-18.SXF6

Los siguientes comandos que existen en la configuración por defecto de la versión SRD5, no están soportados en la SXF6:

```
parser config partition no existe
no service scripting
no service call-home
no logging count
logging buginf
ipc holdq threshold upper 0
ipc holdq threshold lower 0
ipc ru-action on
ipc header-cache permanent 6000 500
ipc buffers min-free 112
ipc buffers max-free 1120
ipc buffers permanent 504
call-home
alert-group configuration
alert-group diagnostic
alert-group environment
alert-group inventory
alert-group syslog
rate-limit 20
profile "CiscoTAC-1"
no active
```

```
destination preferred-msg-format xml
destination message-size-limit 3145728
no destination transport-method http
destination transport-method email
destination address email callhome@cisco.com
destination address http
https://tools.cisco.com/its/service/oddce/services/DDCEService
subscribe-to-alert-group diagnostic severity minor
subscribe-to-alert-group environment severity minor
subscribe-to-alert-group syslog severity major pattern ".*"
subscribe-to-alert-group configuration periodic monthly 18 10:44
subscribe-to-alert-group inventory periodic monthly 18 10:29
ip http secure-port 443
ip http secure-active-session-modules all
ip http max-connections 5
ip http timeout-policy idle 180 life 180 requests 1
ip http active-session-modules all
ip http client cache memory pool 100
ip http client cache memory file 2
ip http client cache ager interval 5
ip http client connection timeout 10
ip http client connection retry 1
ip http client connection idle timeout 30
ip http client response timeout 30
logging history Unknown
cns image retry 60
netconf max-sessions 4
netconf lock-time 10
```

El resto de los comandos por defecto adoptan el mismo valor por defecto en ambos entornos por lo que no plantean ningún problema.

4.6.3 Instalación de la nueva versión en los nodos MPLS

Una vez que se ha elegido la versión de IOS a instalar y se ha probado en el hardware disponible en la red solo falta cargarla en los nodos.

Antes del momento de la instalación hay que cargar la nueva versión de IOS en la memoria del nodo, esto se hace mediante FTP desde el segmento de gestión. Se podría hacer en local a través de la conexión de la consola pero a 9600 bps se tardaría demasiado. El proceso es exactamente igual en todos los equipos, por ello, solo describiremos un caso.

El nodo elegido es el de ESX y el proceso es el siguiente:

Se cargan los ficheros en la flash (sup-bootdisk:) de ambas procesadoras:

- c7600s72033-advipservicesk9-mz.122-33.SRD5.BIN
- c7600-fpd-pkg.122-33.SRD5.pkg

Se comprueba la versión cargada que hay actualmente:

```
ESX#show version
Cisco Internetwork Operating System Software
IOS (tm) s72033_rp Software (s72033_rp-ADVIPSERVICESK9_WAN-M), Version
12.2(18)SXF7, RELEASE SOFTWARE (fc1)
Technical Support: http://www.cisco.com/techsupport
Copyright (c) 1986-2006 by cisco Systems, Inc.
Compiled Thu 23-Nov-06 06:00 by kellythw
Image text-base: 0x40101040, data-base: 0x42D98000

ROM: System Bootstrap, Version 12.2(17r)S4, RELEASE SOFTWARE (fc1)
BOOTLDR: s72033_rp Software (s72033_rp-ADVIPSERVICESK9_WAN-M), Version
12.2(18)SXF7, RELEASE SOFTWARE (fc1)

ESX uptime is 1 year, 40 weeks, 2 days, 7 hours, 6 minutes
Time since ESX switched to active is 1 year, 40 weeks, 2 days, 7 hours, 8
minutes
System returned to ROM by  power cycle (SP by power on)
System restarted at 12:57:34 MET Mon Feb 9 2009
System image file is "sup-bootdisk:s72033-advipservicesk9_wan-mz.122-
18.SXF7.bin"


This product contains cryptographic features and is subject to United
States and local country laws governing import, export, transfer and
use. Delivery of Cisco cryptographic products does not imply
third-party authority to import, export, distribute or use encryption.
Importers, exporters, distributors and users are responsible for
compliance with U.S. and local country laws. By using this product you
agree to comply with applicable laws and regulations. If you are unable
to comply with U.S. and local laws, return this product immediately.

A summary of U.S. laws governing Cisco cryptographic products may be found at:
http://www.cisco.com/wwl/export/crypto/tool/stqrg.html

If you require further assistance please contact us by sending email to
export@cisco.com.

cisco CISCO7609 (R7000) processor (revision 1.2) with 458720K/65536K bytes of
memory.
Processor board ID FOX103405RJ
SR71000 CPU at 600Mhz, Implementation 0x504, Rev 1.2, 512KB L2 Cache
Last reset from s/w reset
SuperLAT software (copyright 1990 by Meridian Technology Corp).
X.25 software, Version 3.0.0.
Bridging software.
TN3270 Emulation software.
1 SIP-400 controller (6 GigabitEthernet).
2 Virtual Ethernet/IEEE 802.3 interfaces
58 Gigabit Ethernet/IEEE 802.3 interfaces
1917K bytes of non-volatile configuration memory.
8192K bytes of packet buffer memory.
```

```
65536K bytes of Flash internal SIMM (Sector size 512K).
Configuration register is 0x2102
```

Se verifica que están los archivos en memoria antes de proceder a la actualización.

```
ESX#dir
Directory of sup-bootdisk:/
1  -rw-  124480548  Apr  1  2008  07:21:52  +02:00  c7600s72033-adventerprisek9-
mz.122-33.SRC.bin
2 -rw- 33554432 Jan 5 2008 17:39:22 +01:00 sea_log.dat
3 -rw- 27872768 Apr 1 2008 07:52:48 +02:00 c7600-fpd-pkg.122-33.SRC.pkg
4 -rw- 109464 Aug 6 2010 13:01:12 +02:00 ETL3_Av_out_36000_33_3_61_0
5 -rw- 39106560 Nov 19 2010 12:08:40 +01:00 c7600-fpd-pkg.122-33.SRD5.pkg
6  -rw-  133391108  Nov  19  2010  12:13:40  +01:00  c7600s72033-advipservicesk9-
mz.122-33.SRD5.bin
512114688 bytes total (153575424 bytes free)
ESX#
ESX#dir slavesup-bootdisk:
Directory of slavesup-bootdisk:/
1 -rw- 27872768 Apr 1 2008 07:56:06 +02:00 c7600-fpd-pkg.122-33.SRC.pkg
2 -rw- 33554432 Jan 5 2008 17:40:52 +01:00 sea_log.dat
3  -rw-  124480548  Apr  1  2008  08:04:36  +02:00  c7600s72033-adventerprisek9-
mz.122-33.SRC.bin
4 -rw- 39106560 Nov 19 2010 12:17:44 +01:00 c7600-fpd-pkg.122-33.SRD5.pkg
5  -rw-  133391108  Nov  19  2010  12:29:02  +01:00  c7600s72033-advipservicesk9-
mz.122-33.SRD5.bin
512114688 bytes total (153690112 bytes free)
```

El borrado de la actual imagen se realizará varias semanas después por si ocurriese cualquier imprevisto que implicase volver a la versión anterior.

El siguiente paso es configurar el arranque del equipo para que lo haga con la nueva versión:

```
boot-start-marker
boot system flash sup-bootdisk:c7600s72033-advipservicesk9-mz.122-
33.SRD5.bin
boot-end-marker
```

Ya por último solo queda reiniciar el nodo para que vuelva a arrancar con la nueva configuración. Una vez ha arrancado hacemos comprobaciones de los servicios, básicamente comprobar las adyacencias de BGP, OSPF y comprobar los servicios de nivel 2. Una vez hecho esto ya disponemos de la nueva versión, lo volvemos a comprobar:

```
ESX#sh ver
Cisco IOS Software, c7600s72033_rp Software (c7600s72033_rp-ADVIPSERVICESK9-
M), Version 12.2(33)SRD5, RELEASE SOFTWARE (fc2)
Technical Support: http://www.cisco.com/techsupport
Copyright (c) 1986-2010 by Cisco Systems, Inc.
```

```
Compiled Sat 02-Oct-10 21:54 by prod_rel_team

ROM: System Bootstrap, Version 12.2(17r)S4, RELEASE SOFTWARE (fc1)
BOOTLDR: Cisco IOS Software, c7600s72033_rp Software (c7600s72033_rp-
ADVIPSERVICESK9-M), Version 12.2(33)SRD5, RELEASE SOFTWARE (fc2)

 ESX uptime is 0 weeks, 0 days, 0 hours, 19 minutes
Uptime for this control processor is 0 weeks, 0 days, 0 hours, 19 minutes
System returned to ROM by reload (SP by reload)
System restarted at 20:17:48 MET Thu Nov 18 2010
System image file is "sup-bootdisk:c7600s72033-advipservicesk9-mz.122-
33.SRD5.bin"
Last reload type: Normal Reload


This product contains cryptographic features and is subject to United
States and local country laws governing import, export, transfer and
use. Delivery of Cisco cryptographic products does not imply
third-party authority to import, export, distribute or use encryption.
Importers, exporters, distributors and users are responsible for
compliance with U.S. and local country laws. By using this product you
agree to comply with applicable laws and regulations. If you are unable
to comply with U.S. and local laws, return this product immediately.

A summary of U.S. laws governing Cisco cryptographic products may be found at:
http://www.cisco.com/wwl/export/crypto/tool/stqrg.html

If you require further assistance please contact us by sending email to
export@cisco.com.

cisco CISCO7609 (R7000) processor (revision 1.2) with 458720K/65536K bytes of
memory.
Processor board ID FOX103405RJ
SR71000 CPU at 600Mhz, Implementation 0x504, Rev 1.2, 512KB L2 Cache
Last reset from s/w reset
1 SIP-400 controller (6 GigabitEthernet).
2 Virtual Ethernet interfaces
10 Gigabit Ethernet interfaces
1917K bytes of non-volatile configuration memory.
8192K bytes of packet buffer memory.

65536K bytes of Flash internal SIMM (Sector size 512K).
Configuration register is 0x2102
```

Tal y como referenciamos en la fase 1 en el paso de Homogeneización del IGP, aprovechamos el reinicio que se produce durante la actualización para homogeneizar tanto el número de proceso de OSPF como el direccionamiento de loopback y los enlaces. Para describirlo, tomamos el ejemplo del nodo de VLL ya que el de ESX tiene el direccionamiento definitivo.

Lo primero de todo es cambiar la dirección de la interfaz de loopback, en nuestra red, la loopback del nodo tiene el identificador 0 por lo que realizamos la siguiente sustitución:

```
interface Loopback0
ip address 192.168.255.16 255.255.255.255
!
interface Loopback255
ip address 10.132.1.6 255.255.255.255
!
ip radius source-interface Loopback255
!
snmp-server trap-source Loopback255
!
logging source-interface Loopback255
!
mpls ldp router-id Loopback0 force
!
ntp source Loopback0
```

Hay que configurar otra dirección de loopback, en este caso con el identificador 255 ya que esta dirección, no solo es de gestión, sino que es necesaria para referenciar todos los servicios de nivel 2 que ofrece el router, cuando se configura una crosconexión a nivel 2 entre nodos, se referencia el nodo destino mediante una dirección alcanzable a nivel IP. Se podrían migrar todos los servicios, pero es algo que preferimos separarlo en tareas en distintos días para una mejor identificación de causa frente a incidencias que pudiera surgir. La tarea de cambiar los enlaces de nivel 2 se describe en la sección de trabajos futuros.

El siguiente parámetro a cambiar es el identificador de OSPF, tan solo hay que reescribir el proceso, con los mismos parámetros. Hay que reseñar que este cambio no se puede realizar mediante telnet con origen el segmento de gestión ya que al borrar el proceso de OSPF el nodo pierde conectividad a nivel 3 con el resto de la red. Hay que conectarse mediante telnet desde un nodo adyacente. A pesar de tener listas de acceso configuradas en el line vty, esto es posible ya que en esa lista de acceso, están incluidas las direcciones de enlace de los nodos adyacentes precisamente para poder conectarse desde un nodo adyacente ante un posible fallo del OSPF.

4.6.4 Alta del enlace EDS-COR (implementación Fase 0).

Según se definió en el apartado de “Arquitectura final”, el nodo de COR irá conectado al de EDS en vez de al de ARA. Esto supone, al igual que en el resto de enlaces que se han dado de alta, la configuración de los puertos en ambos equipos en lo que se refiere a direccionamiento IP, IGP (OSPF) y LDP. Tras la configuración y activación de la transmisión del enlace, se levantan los puertos y se hacen las comprobaciones pertinentes que son las mismas que en el resto de establecimiento de enlaces.

Dado que ya hemos descrito el alta de varios enlaces, en el caso de este activamos el modo de monitorización de eventos y vemos lo que se refleja durante las fases. Los eventos los hemos recogido en ambos nodos.

En este caso vemos los de COR, el puerto de conexión es el GE-WAN 3/2.

```
COR# !configuramos el interfaz
000466: Mar 23 22:20:09.672: %SYS-5-CONFIG_I: Configured from console by
agonzalezca on vty2 (10.132.2.144)
COR# !lo levantamos
000468: Mar 23 22:20:38.503: %LINK-3-UPDOWN: Interface GE-WAN2/3, changed
state to up
COR# !el interfaz entra en el proceso de routing
000469: Mar 23 22:20:38.511: %IFDAMP-5-UPDOWN: interface GE-WAN2/3 update IP
Routing state to UP, interface is not suppressed
000470: Mar 23 22:20:39.379: %LDP-5-NBRCHG: LDP Neighbor 192.168.255.10:0 (6)
is UP !establece las adyacencias de LDP
COR#
000472: Mar 23 22:20:53.207: %OSPF-5-ADJCHG: Process 2328, Nbr 192.168.255.10
on GE-WAN2/3 from LOADING to FULL, Loading Done !establece la adyacencia de
OSPF
COR#
COR#
```

Comprobamos las adyacencias de OSPF contra los equipos de EDS, TEL, SAL y ARA.

```
COR#   show ip ospf neighbor

Neighbor ID     Pri   State           Dead Time   Address         Interface
192.168.255.10    0   FULL/  -        00:00:39    192.168.1.38    GE-WAN2/3
192.168.255.202   0   FULL/  -        00:00:34    192.168.1.82    GE-WAN2/2
192.168.255.8     0   FULL/  -        00:00:37    192.168.1.42    GE-WAN1/1
192.168.255.3     0   FULL/  -        00:00:35    192.168.1.34    GE-WAN2/1
```

En este momento, las conexiones de la red MPLS resultante son:

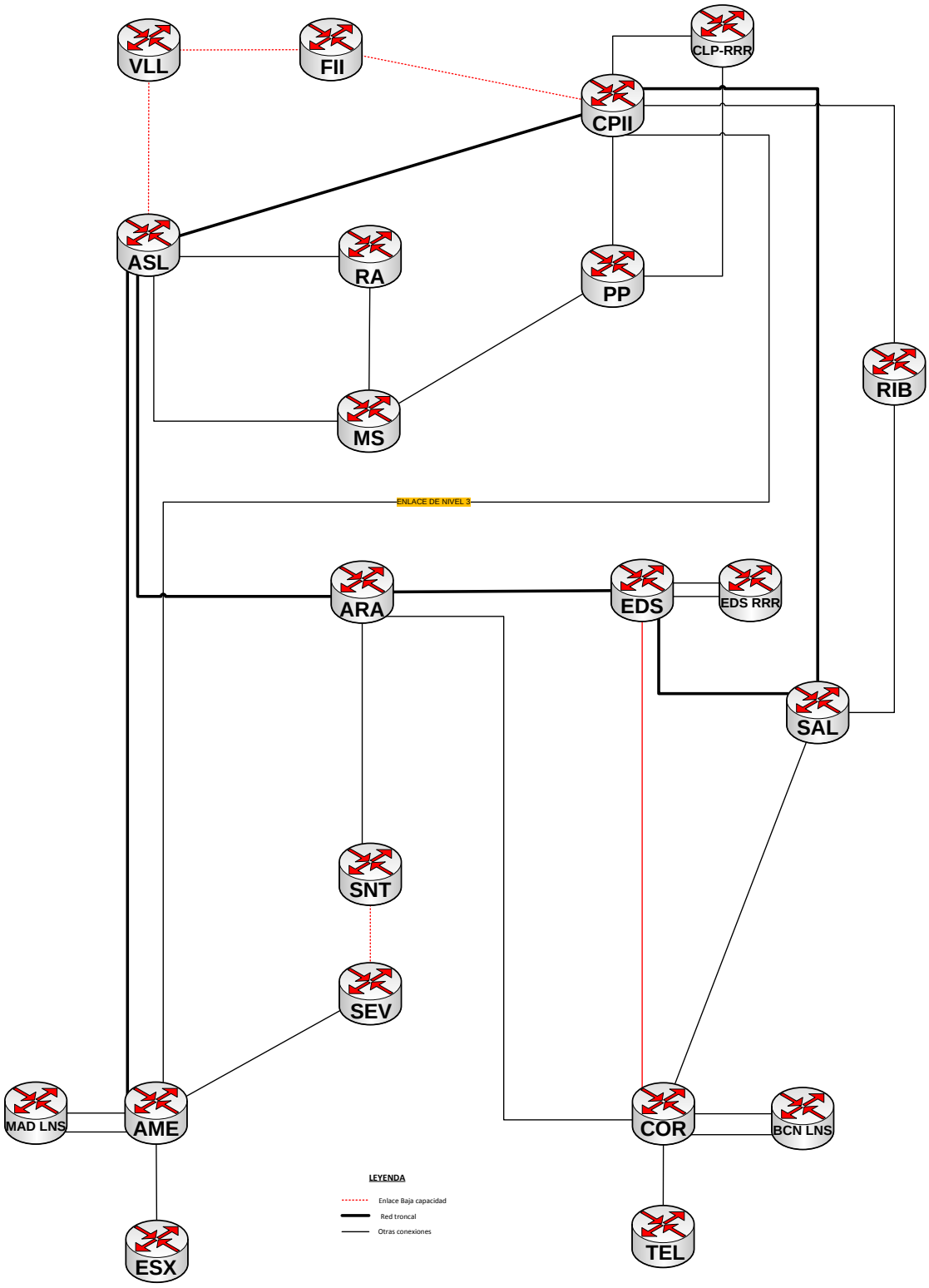

Ilustración 91- Red MPLS tras el alta del enlace EDS-COR

4.6.5 BAJA DEL ENLACE COR-ARA. (IMPLEMENTACIÓN FASE 0)

Para concluir la arquitectura final definida, tan solo falta eliminar el enlace que une los equipos de ARA y COR. Como siempre, dejamos los enlaces deshabilitados en ARA y COR y borramos la configuración asociada.

La red queda configurada de la siguiente manera:

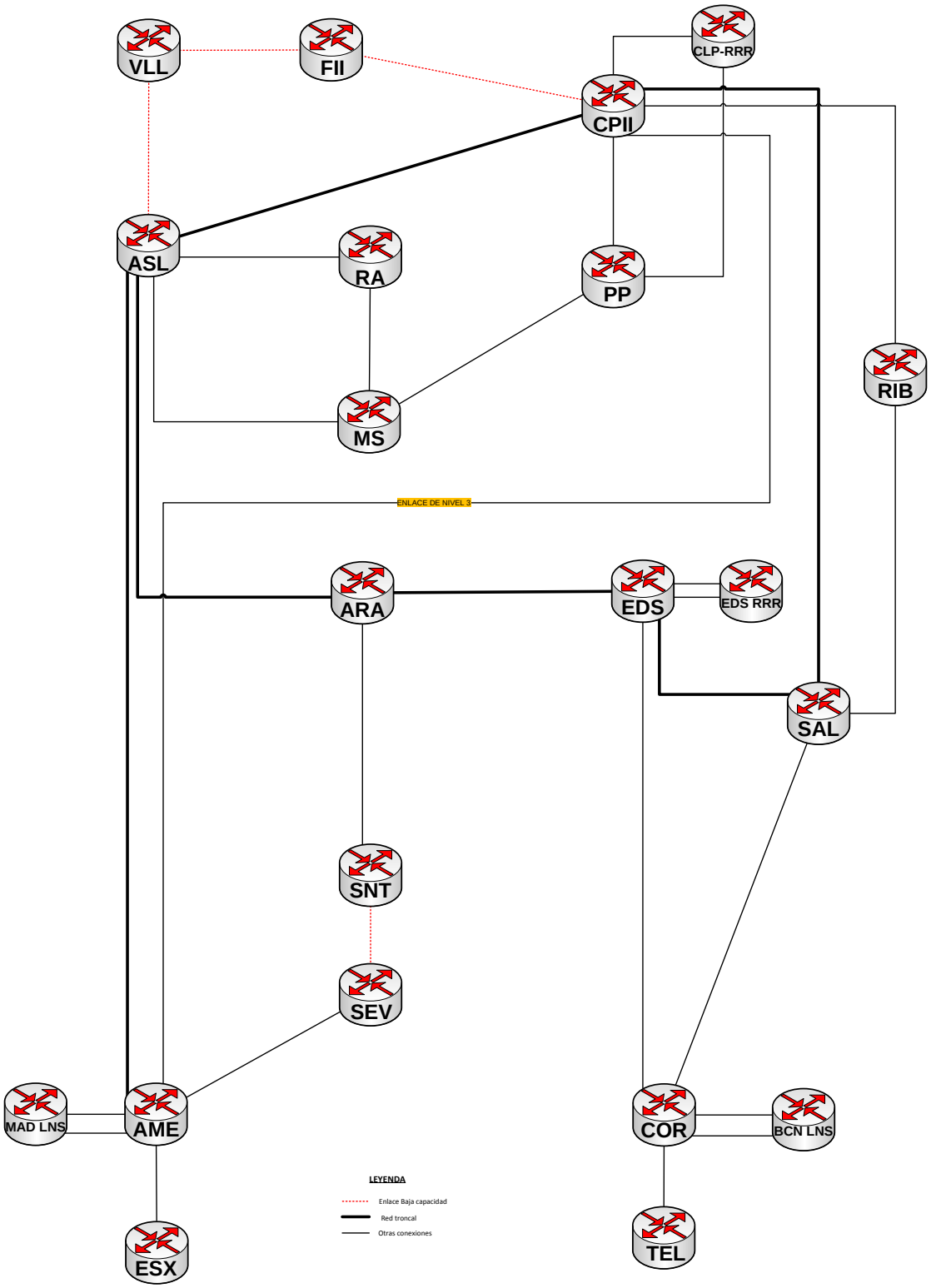

Ilustración 92- Red MPLS tras la baja del enlace ARA-COR

4.6.6 CONVERSIÓN ENLACE AME-CPII DE ENLACE DE NIVEL 3 A ENLACE TRONCAL. (IMPLEMENTACIÓN FASE 0)

Por último y para finalizar la arquitectura física de la red, solo falta por convertir, el enlace que existía de nivel 3 entre ambas redes a enlace troncal MPLS. La configuración del enlace, cuando tenía funciones de nivel 3, era en subinterfaces, cada uno de ellos establecía comunicación entre las VPNs que debían ser propagadas en ambas redes. Puesto que ya tenemos una única red MPLS, no tiene sentido que este enlace sea de nivel 3 ya que solo añade complejidad al routing. Tan solo tenemos que cambiar la configuración de ámbos enlaces y no es necesario realizar pruebas de comunicación puesto que el enlace está en servicio. Con este cambio además se conforma el segundo anillo de conexión troncal sobre el cual se ha diseñado el resto de los enlaces.

Por tanto, tras este cambio, no varía la arquitectura física, si no que es un cambio lógico quedando de la siguiente manera.

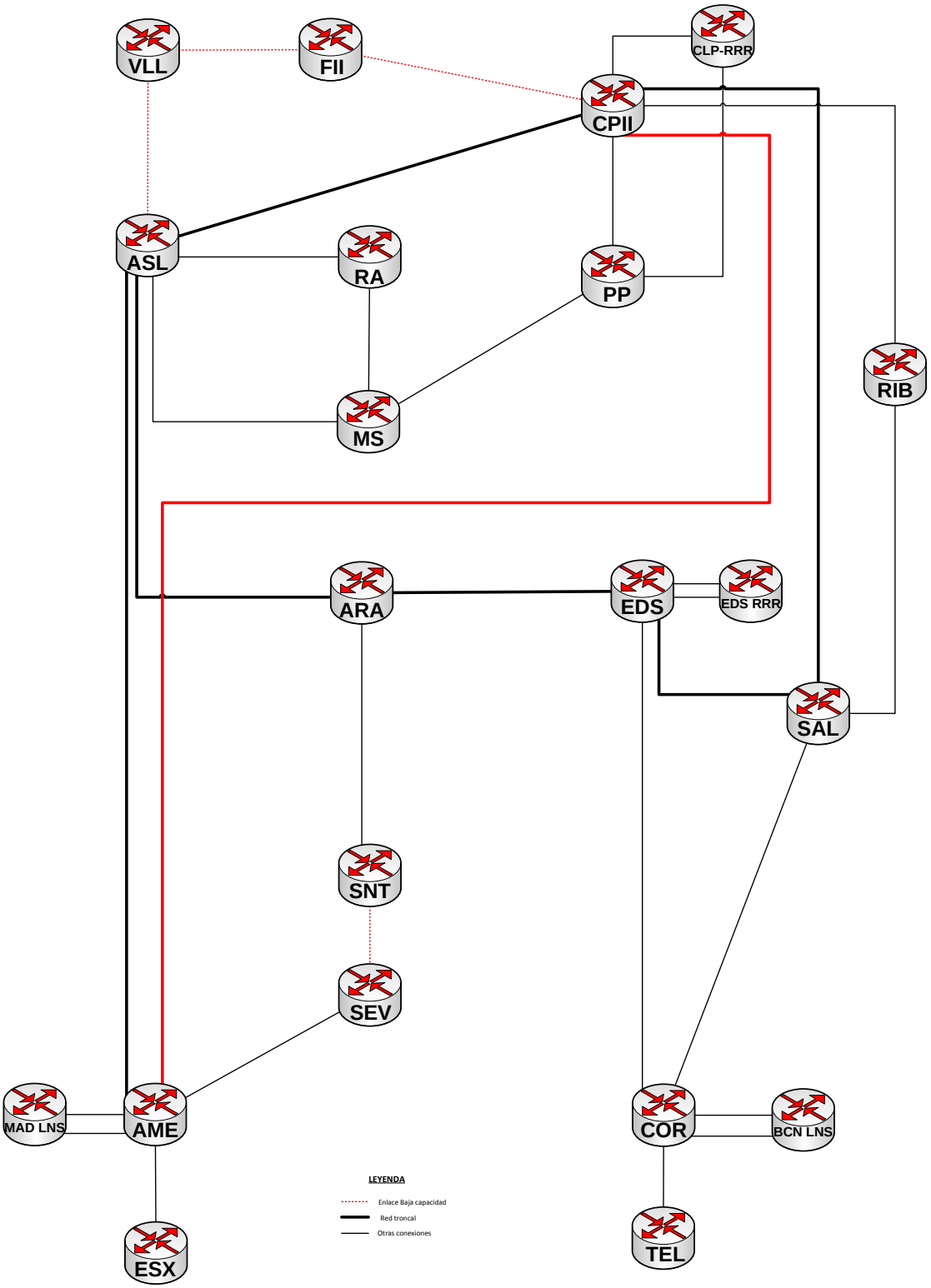

Ilustración 93- Red MPLS tras el cambio lógico del enlace CPII-AME. Es el esquema de la arquitectura física final

En este último esquema no se aprecian bien los dos anillos troncales y el resto de conexiones, es así puesto que se ha pretendido desde el primer cambio mostrar las redes

separadas y como se han ido interconectando. En el apartado de arquitectura final se aprecia el dibujo de la red con los nodos totalmente integrados de manera gráfica.

4.7 FASE 5. HOMOGENEIZACIÓN DE LA CALIDAD DE SERVICIO EN LOS ENLACES TRONCALES.

En caso de producirse congestión en la red, uno de los parámetros más importantes es disponer de un plan potente de calidad de servicio. Así cuando haya saturación en los enlaces troncales y se produzcan descarte de paquetes, se priorizará más unos datos que otros en base a coste, criticidad de los datos o cualquier otro criterio que se considere oportuno.

En la siguiente tabla se muestra como estaban catalogados los valores de los campos EXP de la trama MPLS y el campo IPP del paquete IP en la parte de UFIN:

Nombre de la cola	Marcación	% BW asignado	Observaciones
Core_Control	EXP 6 y 7 IPP 6 y 7	5	Control.
Core_Platino	EXP5 IPP5	10	Multimedia (voz, vídeo y telepresencia).
Core_Oro	EXP4 IPP4	10	Uso exclusivo tráfico crítico.
Core_Plata_FR	EXP3 IPP3	20	Tráfico no crítico. Accesos N-3 de tipo FR y emulación de FR sobre MPLS.
Core_Plata_VPN	EXP2 IPP2	20	Tráfico no crítico. Accesos N-3 de tipo Ethernet.
Core_Plata_Eth	EXP1 IPP1	20	Tráfico no crítico. Servicios de N-2 (EoMPLS y pseudowires).
class-default		15	Resto de tráfico. Internet y copias de seguridad.

Tabla 12. –Utilización de las colas en la red de UFIN

En el caso de la red de DC, los valores configurados en los enlaces troncales son:

Nombre de la cola	Marcación	% BW asignado	Observaciones
MPLS_Routing	EXP 6 y 7 IPP 6 y 7	5	Control
MPLS_Voz	EXP5	25	Multimedia (voz, vídeo y telepresencia)
N/A	EXP4		
MPLS_DatosCriticos	EXP3	45	Datos no críticos prioritarios
N/A	EXP2		
N/A	EXP1		
MPLS_BestEffort	EXP0	25	Datos no críticos no prioritarios

Tabla 13. –Utilización de las colas en la red de DC

Para la unificación de las redes, puesto que las clases de servicio empleadas en ambas redes son compatibles, todas ellas pueden convivir y no es necesaria una redefinición del plan de calidad de servicio. Para ello se mantienen todas las colas empleadas en la red de UFIN excepto la cola Core_Plata_FR que suponemos que se habían migrado todos los servicios anes de la integración y la clase de servicio no se empleaba. Suponemos también que, a día de hoy se produce marcación de tráfico asignado a todas las clases restantes. Las colas empleadas en la parte de DC se integran con las existentes en la parte de UFIN y solo es necesario redimensionar los anchos de banda asignados, de modo que las troncales puedan albergar todo el tráfico de la red.

Por otro lado, al incluir en toda la red el comando “mpls ldp explicit-null” se fuerza a que la etiqueta MPLS no sea eliminada en el penúltimo salto y se mantenga hasta el nodo MPLS que finalmente realizará la entrega de tráfico al CE. Bajo esta configuración ya no es necesario que en la red procedente de UFIN el tráfico en la troncal se diferencie a través del campo EXP propio de la etiqueta MPLS y del campo IPP propio de la cabecera IP, sino que sólo es necesario mirar el campo EXP de la etiqueta MPLS, ya que el paquete IP permanecerá “enmascarado” hasta su entrega al último nodo de la red.

El tráfico de control tiene un comportamiento especial, es marcado en los equipos con IPP6 ó 7. Para este tipo de tráfico hay que tener en cuenta tanto el valor EXP de la etiqueta MPLS como el valor IPP del paquete IP que puede generarse en cualquier nodo MPLS.

Teniendo en cuentas todas estas consideraciones, la marcación de tráfico y asignación de ancho de banda en los enlaces troncales quedará de la siguiente manera:

Nombre de la cola	Marcación	% BW asignado	Observaciones
Core_Control	EXP 6 y 7 IPP 6 y 7	5	Control
Core_Platino_Multimedia	EXP5	20	Multimedia (voz, vídeo y telepresencia)
Core_Oro_Critico	EXP4	10	Uso exclusivo de tráfico crítico.
Core_Plata_P3	EXP3	20	Tráfico no crítico prioritario
Core_Plata_P2	EXP2	20	Tráfico no crítico
Core_Servicios_N1	EXP1	5	Servicios de N-2 (EoMPLS y pseudowires)
Core_BestEffort	EXP0	20	Resto de tráfico

Tabla 14. –Utilización de las colas en la red MPLS final

Ésta es la asignación de colas de partida que facilitará la integración de las redes MPLS. Con esta configuración se garantiza la calidad de servicio deseada para cada tipo de tráfico existente incluso en casos de saturación de los enlaces y no es necesario modificar la marcación en los CPEs procedentes de ninguna de las dos redes.

El porcentaje de ancho de banda de cada cola se redefinirá tras la unificación de las redes MPLS, una vez esté consolidada la arquitectura final, entonces será posible determinar los anchos de banda reales consumidos por cada tipo de tráfico en cada enlace y determinar el porcentaje de ancho de banda necesario por cola.

Los tráficos puramente de “nivel 2” (EoMPLS y emulación de FR sobre MPLS) tendrán un ancho de banda independiente del resto de servicios y garantizado en función de los servicios contratados, ya que estos servicios son de capacidad y no deben incurrir en la misma sobresuscripción del resto. Aún más, se debería tender a no tener sobresuscripción para los mismos y que su cola correspondiente sea capaz de soportar la suma de las capacidades de acceso.

Esta configuración no es una configuración optimizada sino que será necesaria una redefinición y unificación de las clases de servicio ofrecidas en la red, para lo cual se propondrá como trabajo futuro el estudio de las colas actuales y la eliminación de las redundantes.

5 PRESUPUESTO

Aquí definimos el valor monetario que ha tenido el proyecto en lo que se refiere tanto a horas de estudio de la situación como horas de ejecución diferenciando entre jornada de ingeniero en horario diurno o nocturno. Una jornada, tanto nocturna como diurna son 8 horas.

5.1 FASE 0. ARQUITECTURA FINAL

El alcance de las tareas relacionadas con la definición de la arquitectura física final de la red será el siguiente:

5.1.1 TOPOLOGÍA DE RED

Tarea	Objetivo	Duración
Documento de trabajo.	Redactar apartado de arquitectura física de la red indicando los puntos de interconexión.	0.5 días.
Actualizar e integrar normas de ingeniería.	Actualización e integración de las normas de configuración de ambas redes.	2 días
Alta/baja de enlace	Trabajo de establecimiento o baja de un enlace de la red	1 día
Soporte alta/baja de enlace	Soporte de persona onsite por si hay problemas o es necesario actuación en local en la alta y baja de enlaces	1 día

Tabla 15. –Estimación en jornadas de la topología de la red

5.1.2 DIRECCIONAMIENTO IP

Tarea	Objetivo	Duración
Estudio de situación actual.	Estudiar el direccionamiento (servicios) actualmente propagado en el backbone y posibles solapamientos.	1 día.
Documento de trabajo	Actualizar documento de trabajo indicando los resultados del estudio del direccionamiento.	0,5 días.

Tabla 16. –Estimación en jornadas de análisis del direccionamiento IP

5.2 FASE 1. INTEGRACIÓN A NIVEL IP DE AMBAS REDES

El alcance de las tareas en jornadas de la unificación de las redes a nivel IP es:

5.2.1 HOMOGENEIZACION DEL IGP

Tarea	Objetivo	Duración
Comparativa ISIS vs OSPF	Recopilación de información de funcionalidades especificas OSPF e ISIS	0.5 días.
Documento de trabajo y normas de ingeniería.	Redactar apartado de comparativa de funcionalidades entre ISIS y OSPF. Se incluye en este apartado las conclusiones consensuadas con el cliente.	2 días
Elección de protocolo IGP.	Toma de decisión sobre protocolo a implantar.	0,5 días
Plan de migración del IGP.	Elaboración y redacción del plan de migración detallado para la implantación del IGP.	1 día
Ejecución de la migración.	Ejecución nocturna de la migración del IGP.	1 día

Tabla 17. –Estimación en jornadas de la homogeneización del IGP

5.2.2 ANÁLISIS DE MTU

Tarea	Objetivo	Duración
Estudio de servicios y recomendación	Estudio de servicios y recomendación de MTU a configurar	0,5 día.

MTU.	en las interfaces del backbone.	
Implantación MTU.	Implantación del valor de MTU homogéneo en las interfaces del backbone.	1 día. (Noche)

5.3 Fase 3. Integración a nivel VPN MPLS de ambas redes

Tarea	Objetivo	Duración
Estudio de la situación actual.	Estudio de configuración específica de BGP y MP-BGP en las redes de DC y UFIN.	3 días.
Propuesta de configuración.	Propuesta de configuración común de BGP/MP-BGP para ambas redes.	1,5 días
Elaboración del plan de migración.	Elaboración de documento de migración en el que se indican los pasos a seguir, equipos implicados, servicios afectados, etc.	2 días
Ejecución del plan de migración.	Ejecución del plan de migración.	1 día (noche)
Actualización de documentos	Actualización de los documentos de ingeniería de la red integrada	3 días

Tabla 18. –Estimación en jornadas de la integración de las redes a nivel VPN MPLS

5.4 Fase 4. Actualización de la versión de IOS

De manera previa se elabora un documento de migración general que aplique a la totalidad de los nodos a migrar. La elaboración de este documento requerirá de los siguientes recursos:

Tarea	Objetivo	Duración
Elaboración de documento de migración.	Elaboración de documento de migración con el procedimiento de carga de IOS y las pruebas genéricas de servicio a realizar (desde el punto de vista del equipo P/PE – Internet, VPNL3, VPNL2).	2 días.

A continuación se planifica y lleva a cabo la migración de los distintos nodos, definiéndose a continuación las necesidades de recursos para cada una de las migraciones a ejecutar:

Tarea	Objetivo	Duración
Elaboración de documento de migración.	Particularización de documento de migración para el nodo en particular. En función de la configuración del equipo se indican pruebas específicas de servicio para este nodo y se anexan las capturas tras la migración.	2 días.
Ejecución de migración de IOS	Trabajo nocturno de actualización de versión de IOS en el nodo.	1 días. (noche)
Soporte migración de IOS.	Soporte de persona onsite por si hay problemas durante la carga de la versión de IOS y es necesario conectarse por consola.	1 días. (noche)

Tabla 19. –Estimación en jornadas de la homogeneización de la versión de IOS

5.5 Fase 5. Homogeneización de la calidad de servicio en los enlaces troncales

Tarea	Objetivo	Duración
Validación de QoS	Validar la configuración de QoS válida para los servicios de la red homogénea	3 días.

Tabla 20. –Estimación en jornadas de la homogeneización de la QoS

5.6 Total servicios

Concepto	Precio unitario (€)	Unidades	Precio (€)
Jornada de ingeniero diurna (8 horas)	700	25	17.500
Jornada de ingeniero nocturna (8 horas)	1225	42	51.450
Total			68.950

5.7 EQUIPAMIENTO

Tal y como describimos en el capítulo de desarrollo, concretamente en la fase 0, ha sido necesaria la compra de un equipo Cisco 7201, con funciones de route reflector, además hemos adquirido la versión de IOS correspondiente y un conector SFP ZX. La cuantía se describe en la siguiente tabla:

Descripción	Unidades	Precio (€)
Cisco 7201 Chassis, 1GB Memory, dual P/S, 256MB Flash	1	20.338,98
Cisco 7201 AC Power Supply option System	2	0
AC Power Cord (Europe), C13, CEE 7, 1.5M	2	0
Cisco 7200 NPE G2 IOS ADVANCED IP SERVICES	1	3.813,56
Cisco 7201 Series 1GB Memory System	1	0
Cisco 7201 Compact Flash Disk, 256 MB System	1	0
1000BASE-ZX SFP	1	3.385,59
Total equipamiento		**27.538,13**

Tabla 21. –Coste de equipamiento

5.8 COSTE TOTAL.

Una vez tenemos los costes de equipamiento y de servicios, los sumamos y obtenemos coste total de la realización de este proyecto. Los vemos en la siguiente tabla:

Concepto	Precio (€)
Servicios	68.950
Equipamiento	27.538,13
Total Proyecto	**96.488,13**

Tabla 22. –Coste total del proyecto

6 CONCLUSIONES

Una vez se ha finalizado el proyecto de unificación de ambas redes, teniendo en cuenta que no pretende ser teórico sino un supuesto aproximado a la realidad del mundo empresarial, se han obtenido las conclusiones que se detallan a continuación.

Primero, como comentamos, este proyecto no es una guía de integración de redes MPLS, es la aplicación de los pasos básicos que hay que dar, particularizados a las dos redes iniciales propuestas. Puesto que se parte de dos redes en servicio, algunos procedimientos no han sido los óptimos desde el punto de vista teórico, pero hay que tener en cuenta que se han tenido que adaptar a las necesidades existentes previamente y a las características de las redes originales. También se ha tenido en cuenta sobremanera minimizar el impacto sobre los servicios que prestan ambas redes, por eso algunos pasos se han visto incrementados en su complejidad.

Dado el coste, tanto monetario como de recursos en la empresa, lo primero que hay que evaluar detalladamente es la conveniencia de la unificación de ambas redes MPLS. Gracias a la interconexión de nivel 3 que existía en un primer momento, se podían cubrir las necesidades de ofrecer servicios de nivel 3 entre ambas redes. Pero no se pueden prestar servicios de nivel 2 aunque estos no son muy frecuentes en la red. Otro inconveniente es que no se puede ofrecer una calidad de servicio extremo a extremo. También esta conexión de nivel 3 genera gran cantidad de problemas y supone bastantes limitaciones. La primera es que ante la imposibilidad de configurar eBGP, se debe configurar otro protocolo como RIP, que es el más sencillo. Si se usara RIP el tiempo de convergencia de RIP es de tres minutos, por lo que ante caída de la vía principal se plantea una indisponibilidad inaceptable en servicios críticos. También supone un cuello de botella ya que tan sólo existe un camino para ir de una red a otra y por último tenemos la problemática de la duplicidad de redes, ya que los segmentos del CPD se empiezan a ver por EDS y CPII, y con este tipo de conexión no hay mucha versatilidad para

evitar bucles de nivel 3. Otro de los motivos y quizás el más importante, es la sencillez de operación y mantenimiento que supone una única red. Por todo esto, se justifica el tremendo esfuerzo que supone unificar las redes.

Debido a diversos factores como el plazo en el que se dispondría de los nuevos enlaces para la realización de toda la unificación de las redes y dada la necesidad que se tenía para poder ofrecer servicios globales, era importante conocer en qué momento se podría disponer de la red con una funcionalidad aceptable. Dadas las fases, se ha descrito que en la fase 1 se establece comunicación en la tabla de routing global entre ambas redes, pero no se puede dar ningún tipo de servicio puesto que LDP no está establecido, aún no se puede prestar ningún servicio sobre MPLS. En la segunda fase, se establece la comunicación a nivel MPLS pero tan sólo se pueden ofrecer servicios de nivel 2, que suponen la minoría de los servicios prestados por lo que se tenía la red al 5% de su capacidad. Es en la fase 3, la unión de las redes a nivel VPN MPLS, con la homogeneización del MP-BGP cuando se dispone del 100% de las capacidades que ofrece una red MPLS. La cuarta fase, la homogeneización de la versión de IOS en sí, no es imprescindible, pero supone eliminar una importante fuente de incompatibilidades. La quinta fase, un potencial problema en cualquier unificación de redes no lo fue tanto en este supuesto ya que el mapeo del tráfico en colas de cada red era bastante similar, sólo fue necesario realizar unos pequeños ajustes. Si por ejemplo, el tráfico de nivel 2 de la red de DC y el crítico de la red de UFIN hubiesen estado mapeados en la misma cola habría supuesto un grave problema ya que reciben tratamientos distintos en la red. Habría sido necesario cambiar todo el plan de QoS en los CPEs de alguna de las dos redes.

7 TRABAJOS FUTUROS

7.1 UNIFICACIÓN DE QOS

En este caso, se trata de estudiar cómo están definidas las colas y que servicio dan. Al haber homogenizado ambas redes, dado que el plan de colas lo permitía, lo que se ha hecho ha sido que convivan ambos planes, en este apartado es el momento de evaluar las necesidades reales de la red y replanificar todo el plan de calidad de servicio, creando colas nuevas si fuera necesario (improbable), eliminando las que tienen funciones duplicadas (muy probable) y redefiniendo los anchos de banda asignados a cada clase de servicio en los enlaces troncales.

Asociado a la redefinición de las clases de servicio de la red MPLS, será necesaria la modificación de la configuración de los CPEs para el marcado de tráfico acorde a las clases definidas.

Tal y como se comentó, antes de unir las redes existía una filosofía distinta a lo que se refiere a remarcado de entrada en la red MPLS. En la parte de UFIN, el CPE es gestión también de UFIN por lo que el tráfico se marcaba en el CPE y en la entrada a la red MPLS se confiaba en ese marcado de tráfico y se podían hacer políticas de tráfico multicolor (varias colas). En la parte de DC se desconfiaba del tráfico entregado por el CPE por lo que era necesario este remarcado. Es también necesidad establecer las nuevas políticas y adaptar la configuración a estos nuevos requerimientos.

7.2 INTEGRACIÓN DE VPNS COMUNES

En la integración de las redes MPLS de dos empresas, habrá un tráfico que sea necesario compartir, básicamente lo que se refiere a tráfico corporativo como puede ser el correo electrónico, acceso a bases de datos de nóminas, datos de empleados, etc.…

Típicamente, como es el caso, en ambas redes existe una VPN MPLS que denominaremos de tráfico de intranet que cursa este tipo de datos. En esta VPN el direccionamiento es corporativo y no se solapa en ambos planes de direccionamiento por lo que está listo para ser unificado. Esto supone rehacer la configuración de todos los enlaces que pertenezca a sedes donde se curse tráfico de intranet.

Por supuesto que habrá VPNs en las que no se deba mezclar el tráfico, bien porque son de cliente externo bien porque son de proyectos especiales, en este caso no será necesario realizar ningún cambio.

7.3 INTEGRACIÓN DE LA GESTIÓN

Cuando las redes están separadas, ambas tienen un mecanismo de gestión de los equipos. Típicamente se gestionan en la tabla de routing global con sus direcciones de loopback0. En un punto de la red, un interfaz configurado en la tabla de routing global y no en ninguna VPN MPLS se interconectará con el segmento de gestión. Este es el caso de ambas redes por lo que hay que unificar los segmentos de gestión (no es integración de redes MPLS) pero sí que hay que unificar como se va a realizar la gestión de los nodos. Estos trabajos se realizan en esta tarea.

7.4 OPTIMIZACIÓN DE LA CONFIGURACIÓN

En estos trabajos se trata de realizar los cambios denominados estéticos. Un ejemplo de ellos es los servicios de nivel 2. Los servicios de nivel 2 se establecen entre 2 nodos de la red y se identifica origen y destino mediante la dirección de loopback del nodo. En el momento que se cambió el direccionamiento, en la fase 4 del desarrollo junto con la elevación de la versión de IOS de los equipos no se hizo un cambio de loopback puro. Lo que se hizo fue sustituir el valor de la loopback0 por el nuevo, pero la antigua dirección IP de loopback0 se utiliza como loopback255. Esto se hizo así para diferenciar los trabajos. De cara a un fallo en la red tras los trabajos es preferible que no se realicen muchas acciones a la vez para saber qué está fallando. Además de los servicios de nivel 2, esta dirección se usa para el tráfico SNMP, NTP, autenticación de radius, etc…

Vemos el ejemplo del nodo de VLL:

Antes del cambio:

```
interface Loopback0
ip address 10.132.1.6 255.255.255.255
```

Tras el cambio:

```
interface Loopback0
ip address 192.168.255.16 255.255.255.255
!
interface Loopback255
ip address 10.132.1.6 255.255.255.255
ip radius source-interface Loopback255
```

```
!
snmp-server trap-source Loopback255
!
logging source-interface Loopback255
!
mpls ldp router-id Loopback0 force
!
ntp source Loopback0
```

De lo que se trata en este trabajo posterior es ir cambiando todos los servicios que apuntan a la dirección de loopback255, para que al final, poder borrar ese interfaz.

En el caso de la gestión SNMP, habría que cambiar en los gestores de red, en nuestro caso HPOpenView y Ciscoworks la dirección a la que apuntan.

8 BIBLIOGRAFÍA

8.1 REFERENCIAS IMPRESAS

- **MPLS Fundamentals**
  Luc De Ghein.
  Cisco Press. 2007

- **Advanced MPLS VPN Solutions, Revision 1.0: Student Guide**
  Cisco Press. 2000

- **Understanding MPLS/VPN. Security Issues**
  Michael Behringer.
  Cisco Press. 2003.

- **Troubleshooting MPLS Networks**
  Yussuf Hassan, Rajiv Asati.
  Cisco Press. 2004.

- **Introduction to Intermediate System-to-Intermediate System Protocol**
  Cisco Press.
  Abril 2010.

8.2 REFERENCIAS EN INTERNET

- Glosario de terminus:
  http://ldc.usb.ve/~poc/RedesII/Grupos/G5/glosario.htm

- Funcionamiento de ISIS:
  http://www.nada.kth.se/kurser/kth/2D1490/06/hemuppgifter/bhatia-manral-diff-isis-ospf-01.txt.html

- Herramienta de configuración de Cisco:
  https://apps.cisco.com/qtc/config/html/configureHomeGuest.html

- Comandos de IOS SRD5:
  http://www.cisco.com/en/US/docs/ios/ios_xe/qos/configuration/guide/legacy_qos_cli_deprecation_xe.html

- RFC 3031
  Multiprotocol Label Switching Architecture
  http://www.ietf.org/rfc/rfc3031.txt
  E. Rosen, A. Viswanathan, R. Callon
  Enero 2001

- Tutoriales acerca de MPLS
  http://www.convergedigest.com/tutorials/#MPLS

- RFC 1771
  A Border Gateway Protocol 4 (BGP-4)
  http://www.ietf.org/rfc/rfc1771.txt
  Y. Rekhter, T.J. Watson, T. Li
  Marzo 1995

- RFC 2858
  Multiprotocol Extensions for BGP-4
  http://www.ietf.org/rfc/rfc2858.txt
  T. Bates, Y. Rekhter, R. Chadra, D. Katz
  Junio 2000

- RFC 2328
  OSPF Version 2
  http://www.ietf.org/rfc/rfc2328.txt
  J. Moy
  Abril 1998

- RFC 1195
  Use of OSI IS-IS for Routing in TCP/IP and Dual Environments
  http://www.ietf.org/rfc/rfc1195.txt
  R. Callon
  Diciembre 1990

9 GLOSARIO.

- ***ADSL***: Asymmetric Digital Subscriber Line (Linea de abonado digital simétrica) es un tipo de DSL que consiste en una transmisión analógica de datos digitales apoyada en el par simétrico de cobre que lleva la línea telefónica convencional
- ***AS:*** Autonomous System (Sistema autónomo). Una colección de redes bajo un único dominio administrativo.
- ***ATM:*** Asynchronous Transfer Mode (Modo de Transferencia Asíncrono). Técnica de comunicación basada en la conmutación de paquetes
- ***Backbone:*** Un sistema de transmisión utilizado para interconectar redes de distribución de menor velocidad.
- ***BGP:*** Border Gateway Protocol. Protocolo de enrutamiento diseñado para su uso en redes controladas por distintas organizaciones.
- ***CBR:*** Constant Bit Rate. Una categoría de servicio ATM que le caracteriza por un ancho de banda máximo constante, a menudo empleado para aplicaciones de tiempo real.
- ***CE:*** Customer Edge. Es un router que se ubica en el extremo de la red del cliente que permite el acceso al núcleo MPLS.
- ***CPD:*** Protocolo de descubrimiento de Cisco (Cisco Discovery Protocol), se utiliza para obtener información acerca de los dispositivos vecinos, como los tipos de dispositivos conectados, las interfaces de router conectadas, las interfaces utilizadas para crear las conexiones y los números de modelos de los dispositivos.
- ***CoS:*** Class of Service. (Clase de Servicio)

- ***Coste:*** Término aplicado para evaluar las características de una ruta dentro de una red de comunicaciones tales como: ancho de banda, carga, número de saltos, confiablidad, retardo, etc.

- ***CPE:*** Customer Premies Equipment. Equipo ubicado en la propiedad del usuario.

- ***DVB:*** Digital Video Broadcasting (Broadcast de video digital) es una organización que promueve estándares aceptados internacionalmente de televisión digital, en especial para televisión vía satélite

- ***DWDM*** Dense Wavelength Division Multiplexing (Multiplexación por división en longitudes de onda densas). DWDM es una técnica de transmisión de señales a través de fibra óptica usando la banda C (1550 nm).

- ***EGP:*** Exterior Gateway Protocol. (Protocolo de Gateway Exterior). Protocolo de enrutamiento diseñado para su uso entre redes controladas por distintas organizaciones.

- ***E-LSR:*** Un LSR que se encuentra en el borde de un dominio MPLS.

- ***Ethernet:*** Un diseño para red de área local muy popular, patentado por la XEROX Corp., caracterizado por transmisión en banda base que emplea el protocolo CSMA/CD como mecanismo para el control de acceso; modificado por el IEEE para constituir el estándar IEEE-802.3. Puede utilizar diferentes tipos de cables, entre los cuales están el par trenzado y la fibra óptica.

- ***E1:*** Correspondiente europeo del T1. Se refiere a una tasa de transmisión de 2048Mbit/s, capaz de acomodar 32 canales PCM de los cuales 30 son para voz y 2 para señalización.

- ***GPRS:*** General Packet Radio Service (servicio general de paquetes vía radio). Es una extensión de GSM para dispositivos móviles.

- ***FEC:*** Forwarding Equivalence Class. (Clase Equivalente de Envío). Los paquetes que son enviados sobre el mismo punto de salida en la red a lo largo del camino y que reciben el mismo tratamiento de envío a lo largo de la ruta se dice que pertenecen a una misma FEC.

- ***FIB:*** Forwarding Information Base. (Base de Información de Envío). Una tabla de envío de una arquitectura IP tradicional.

- ***Forwarding:*** Procedimiento o mecanismo de envío de paquetes de información sobre una red de comunicaciones.

- ***Frame Relay:*** Protocolo para la transmisión de paquetes a mayor velocidad que el X.25 y con más eficiencia, gracias a la eliminación del chequeo de errores en cada tramo. Permite una utilización más flexible del ancho de banda.

- ***HDLC:*** High-level Data Link Control.
- ***ICMP:*** Internet Control Message Protocol (Protocolo de Control de Internet)
- ***IETF:*** Internet Engineering Task Force. (Grupo de Trabajo para la Ingeniería de Internet)
- Ingeniería de Tráfico: TE es el proceso de direccionamiento de tráfico a través del Backbone para facilitar el uso eficiente del ancho de banda disponible entre un par de routers.
- ***IGP:*** Internal Gateway Protocol (Protocolo Gateway Interior). Protocolo de enrutamiento diseñado para su uso en redes controladas o administradas por una sola organización.
- ***IP:*** Internet Protocol. (Protocolo de Internet). Protocolo enrutado, la versión más empleada en ipv4 mientras que la versión IPv6 está en desarrollo y figura como una alternativa para optimizar el mecanismo de direccionamiento actual.
- ***LFIB:*** Lable Forwarding Information Base. (Base de Información de Envío de Etiquetas). Una tabla que forma parte de la estructura de datos en una arquitectura MPLS.
- ***LIB:*** Label Information Base. (Base de Información de Etiquetas). Una tabla que forma parte de la estructura de control en la arquitectura MPLS.
- ***LSP:*** Label Switched Path. (Ruta o camino de Conmutación de Etiquetas).
- ***LSR:*** Label-Switching Router (Router de Conmutación de Etiquetas). Un router que puede enviar paquetes basado en el valor de una etiqueta adherida al paquete.
- ***Métrica:*** Valor que asigna un dispositivo de red como un router para evaluar el coste de una ruta.
- ***MPLS:*** MultiProtocol Label Switching (Protocolo Múltiple por Conmutación de Etiquetas). Un conjunto de normas IETF para permitir que el tráfico de comunicaciones se envíe basado en etiquetas.
- ***MPLS-TE:*** Ingeniería de Tráfico aplicada a MPLS.
- ***OSPF:*** Open Shortest Path First (Primero la Ruta Libre más Corta). Protocolo de enrutamiento de estado de enlace.
- ***PDH:*** Plesyochronous Digital Hierarchy (Jerarquía Plesiocrónica Digital). Jerarquía Digital utilizada anteriormente en sistemas de telecomunicaciones, sustituida hoy en día por SONET/SDH, Jerarquía Digital Sincrónica.
- ***PE:*** Provider Edge. Son los routers que se colocan en los extremos del núcleo MPLS. Los routers PE clasifican paquetes de ingreso desde los routers CE, según los valores de prioridad IP asociados.

- ***Protocolos Enrutados:*** Es cualquier protocolo de red que ofrezca suficiente información en su dirección de capa de red, como para permitir que un paquete sea enviado de un host a otro. Por ejemplo: IP.
- ***Protocolos de Enrutamiento:*** Es cualquier protocolo que soporte un protocolo enrutado y quye suministre los mecanismos necesarions para compartir información de enrutamiento. Por ejemplo: RIP, IGRP, EIGRP, OSPF.
- ***PVC:*** Permanent Virtual Channel (Circuito Virtual Permanente).
- ***QoS:*** Quality of Service (Calidad de Servicio). Un dimensionamiento del rendimiento que refleja la calidad y el servicio y su disponibilidad.
- ***RDSI:*** Red Digital de Servicios Integrados. Es una red que procede por evolución de la red telefónica existente, que al ofrecer conexiones digitales de extremo a extremo permite la integración de multitud de servicios en un único acceso
- ***RFC:*** Request For Comments. Un documento de la IETF.
- ***Route Reflector:*** (Reflector de rutas). Es el equipo encargado en una red MPLS de establecer sesiones MP-BGP con todos los nodos para el intercambio de rutas.
- ***RSVP:*** Resource reServeVation Protocol. Reserva ancho de banda a lo largo de una ruta de un origen a un destino específico en aplicaciones de Ingeniería de Tráfico.
- ***SDH:*** Syncronous Digital Hierarchy (Jerarquía Digital Sincrónica). Nombre con el que la ITU-T adaptó SONET. Método de multiplexaje a velocidades desfe 155.250 (STM-1) hasta 10Gb/s (STM-48).
- ***SP:*** Service Provider (Proveedor de servicios).
- ***TCP:*** *Transmission control protocol.* (Protocolo de control de transmisión). es un protocolo de comunicación orientado a conexión y fiable. Es un protocolo de capa 4 según el modelo OSI.
- ***VPN:*** Virtual prívate network (Red privada virtual) es una tecnología de red que permite una extensión de la red local sobre una red pública.
- ***VRF:*** Virtual Routing and Forwarding (Enrutamiento y reenvio virtual). Es una técnica que permite a varias instancias de routing coexistir en el mismo router a la misma vez.