DISEÑO DE UNA RED VPLS JERÁRQUICA

Nombre del Proyecto	**Diseño de una red VPLS jerárquica**
Autor	Pablo Sesmero Orihuela
Tutor	Ricardo Romeral Ortega
Universidad	Carlos III de Madrid
Titulación	Ingeniería Técnica de Telecomunicación

ÍNDICE GENERAL

LISTADO DE IMÁGENES

ESPECIFICACIONES DEL DISEÑO

RESULTADOS PRUEBAS

1. Objetivo del Proyecto

El objetivo principal de este Proyecto Fin de Carrera sobre el Diseño de una Red VPLS (Virtual Private LAN Service) jerárquica es proporcionar una solución completa a la problemática de las Redes Privadas Virtuales, que utilizan la mayoría de las grandes empresas, a través de redes públicas o compartidas.

La demanda de ancho de banda y capacidad por parte de las grandes compañías requiere un uso más eficiente de la tecnología. Los enlaces de gran capacidad hoy día han evolucionado hacia circuitos de 1, 10, 40 y hasta 100 Gigabits por segundo, basándose en tecnología Ethernet, más económica y con mucha más capacidad. El creciente despliegue de plataformas de transporte basadas en Ethernet requiere un diseño y adecuación por parte de las redes de datos que se apoyan en dicha infraestructura. Las redes de nivel 2, basadas en VPLS, aprovechan al máximo dicha tecnología dado que se trata de Ethernet encapsulado en otros protocolos de transporte.

Además del aumento de ancho de banda debido a la tecnología Ethernet, las grandes empresas demandan una reducción considerable del tiempo y latencias que se introduce en las redes de dimensiones considerables, nivel metropolitano, nacional e internacional. Las empresas han empezado a desplegar masivamente dispositivos de uso final basados en IP, por ejemplo, telefonía IP, videoconferencias, aplicaciones críticas en tiempo real, etc., por lo tanto los equipos de datos que manejan el tráfico de las VPNs, deben reducir los tiempos de tránsito sobre su red. La latencia y retardo introducido por una red puede afectar muy negativamente el resultado final de una determinada aplicación. La voz, por ejemplo, transportada y paquetizada en IP, es crítica en cuanto a tiempo de retardo, ya que puede resultar imposible una comunicación si el tráfico supera ciertos umbrales de calidad en cuanto a latencias. El tiempo de retardo que introduce la red se debe a la latencia de los propios circuitos físicos, y a la conmutación realizada en los equipos de datos. El procesamiento y enrutamiento del tráfico requiere recursos del equipo y tiempo. Analizar la información de cada paquete de datos, origen y destino, requiere un mínimo período de tiempo, que dependiendo de la tecnología escogida, y la implementación de los protocolos de control puede ser mayor o menor. En el caso del enrutamiento clásico por IP, por ejemplo, la tabla de rutas se debe consultar completamente para determinar la mejor salida en cada equipo. Este proceso es relativamente rápido, pero se optimiza apoyándose en otros protocolos de transporte como MPLS. La conmutación de tráfico basada en MPLS es sensiblemente más rápida. Las redes

VPLS implementan su plano de control sobre MPLS, de forma que la conmutación en cada uno de los equipos de la red es considerablemente más rápida.

Por otra parte, el tráfico prioritario no puede sufrir excesivos retardos por lo que se trata de priorizar a lo largo de toda la red para que en situaciones de congestión no se vea afectado y se priorice frente a otros equipos. Tráfico que transporte video o voz debe protegerse por todo el camino que siga en la red. Todos los protocolos de control y transporte tienen mecanismos de calidad de servicio que marcan el tráfico prioritario para poder identificarlo y procesarlo con garantías ante situaciones de congestión en la red. De esta forma, además de reducir la latencia en ese tipo de tráfico, ya que es tratado por los equipos antes que ningún otro tipo de tráfico, se mejora el efecto y la sensación en usuarios y aplicaciones finales.

Otro argumento a favor de la elección de una red basada en VPLS, es la abstracción del direccionamiento IP de los clientes. El operador de red, que proporciona el transporte y la conectividad entre sedes de un mismo cliente, no tiene que preocuparse de la conectividad IP entre él mismo y el cliente. La tecnología VPLS proporciona conectividad directa entre los equipos de una misma compañía. Por decirlo de otra forma, los equipos de las delegaciones de los clientes, se puede abstraer del nivel de transporte y pasar a interpretar la red VPLS como un simple cable Ethernet, de forma que sus equipos se verán unos a otros como si estuvieran directamente conectados entre sí. A nivel físico, el tráfico atravesará números equipos de transmisión y conmutación, pero a un nivel lógico, los equipos de los usuarios se verían como en una gran red local, aunque estén a cientos de kilómetros de distancia. El direccionamiento IP y su administración serían independientes completamente de la red del operador. La res es completamente transparente en ese sentido.

Un punto adicional a tener en cuenta a la hora de escoger equipamiento diseñado para VPLS es el económico, ya que la conmutación MPLS y de nivel 2 supone una inversión menor frente al equipamiento de nivel 3 habitual. Se requieren un conocimiento más avanzado en cuanto a tecnologías de conmutación con grandes anchos de bandas y comportamientos de nivel 2, pero la inversión económica es menor. Es decir, los riesgos que conlleva una inversión más económica sobre equipos VPLS, supone un incremento en el conocimiento y diseño de la propia red VPLS.

Circuitos de transmisión Ethernet y equipos de conmutación más económicos, la creciente demanda de ancho de banda tanto para grandes redes de datos de empresas como para el consumo de Internet, reducción en los tiempos de latencia para aplicaciones críticas, y simplicidad en la interacción con los equipos de cliente, son motivos conducen al responsable del diseño de una red de datos de un operador a considerar y elegir una red VPLS.

2. Introducción

Una plataforma basada en tecnología VPLS ofrece hoy día numerosas ventajas de cara al operador de red que decida explotar este tipo de infraestructura para sus servicios, como para el cliente o usuario final que decida contratar o disfrutar de las soluciones que se pueden mantener sobre estas redes.

El objetivo final que se pretenderá conseguir de estas plataformas será mayoritariamente servicios de Redes Privadas Virtuales (VPNs) sobre una plataforma de red compartida, con accesos a la misma dedicados. Esto es lo que las grandes operadoras de telecomunicaciones han ofrecido y ofrecerán a las grandes compañías que precisen de redes de comunicaciones, tanto de voz como de datos, seguras, potentes y eficaces.

Una Red Privada Virtual es una tecnología de red que permite una extensión de la red local sobre una red pública o no controlada, como por ejemplo Internet. Ejemplos comunes son, la posibilidad de conectar dos o más sucursales de una empresa utilizando como vínculo Internet, permitir a los miembros del equipo de soporte técnico la conexión desde su casa al centro de red, o que un usuario pueda acceder a su equipo doméstico desde un sitio remoto, como por ejemplo un hotel.

Estas redes se extienden sobre un área geográfica amplia, a veces un país o un continente, y contiene una colección de elementos de red dedicados a ejecutar programas de usuario (aplicaciones). En los últimos años las redes se han convertido en un factor crítico para cualquier organización. Cada vez en mayor medida, las redes transmiten información vital, por tanto dichas redes cumplen con atributos tales como seguridad, fiabilidad, alcance geográfico y efectividad en costes.

La tecnología que subyace a este objetivo es lo que marca la diferencia entre una plataforma de red y otra, y lo que decantará al cliente o usuario final a la hora de seleccionar la que mejores prestaciones ofrezca en función de las aplicaciones que necesite para su propio negocio. De este modo, dependiendo de cómo trabaje el usuario final, que requerimientos exija para sus comunicaciones tendrá que optar por una solución u otra.

Es fundamental conocer de antemano las necesidades de un usuario o cliente porque de ellas dependerá el despliegue que se tenga que hacer para satisfacer sus requisitos. No todas las soluciones desarrolladas para Redes Privadas Virtuales consiguen el mismo resultado.

Influirán por tanto los detalles propios de cada usuario y la red que necesite construir en la decisión sobre un diseño óptimo, eficiente y que cumpla las expectativas iniciales. Desde qué tipo de aplicaciones usará el sistema completo, a si necesita soluciones de voz implementadas sobre algún tipo de protocolo específico, o si la conectividad entre todos los elementos de su red debe ser total, parcial, o concentrada en algún punto, hasta el tráfico total que desee cursar deben analizarse muy en detalle en la primera fase del diseño para poder decidir correctamente qué tecnología, o cuáles, son las adecuadas para cumplir el objetivo

A menudo, la opción de múltiples tecnologías interconectadas entre sí para simular una única Red Privada Virtual resulta la solución más óptima para un determinado usuario o cliente, y desde luego esa opción debe ser contemplada desde un inicio. Obviamente, realizar inversiones en varias plataformas o despliegues de red supondrá unos costes superiores a la solución simple, pero también es obvio que el objetivo se cumple con creces, dotando además a la implementación de mayor robustez puesto que la solución compuesta ofrece vías alternativas de conectividad ya que no dependería únicamente de una plataforma.

Dichas plataformas suponen una inversión para el operador de telecomunicaciones que pretende ofrecer servicios de comunicación de voz y datos. Sobre cada plataforma se implementarán distintas soluciones para distintos clientes, de modo que es muy importante conseguir el aislamiento lógico del tráfico de cada uno de los clientes. Esto se consigue mediante protocolos de control, supervisados por el operador de red, que a base de añadir información adicional al tráfico del usuario consiguen gestionar y organizar la plataforma según sus propios criterios.

Este tipo de tráfico controlado y generado por la operadora de telecomunicaciones hay que tenerlo en cuenta a la hora de dimensionar la capacidad de las plataformas. Será un volumen muy pequeño en comparación con el propio del usuario, pero imprescindible para conseguir aislar las Redes Privadas Virtuales de los clientes, y para optimizar el uso de la propia plataforma de red de la operadora.

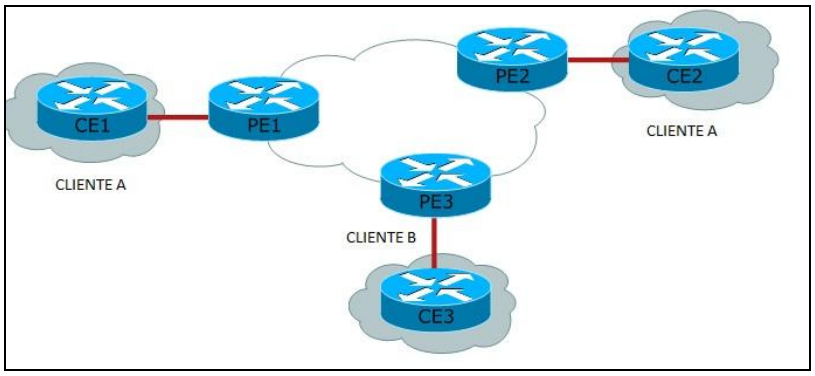

Imagen 1 - Redes Privadas Virtuales

Las tecnologías que han ido ofreciendo servicios basados en Redes Privadas Virtuales evolucionan continuamente, para ofrecer mejores a cambio de una mayor complejidad en la implementación.

Asimismo, la complejidad de la plataforma de red supondrá un mejor servicio final, entendiendo como mejor, más flexibilidad, más rapidez, más eficiencia y más seguridad. De modo que dependiendo de los protocolos utilizados por una tecnología u otra ofrecerán conectividad entre los elementos del cliente a un nivel o a otro. Redes Privadas Virtuales que sólo quieran conectividad física pueden requerir únicamente dicho caudal sobre un enlace virtual, e implementar ellos mismos sus soluciones de enrutamiento y gestión para sus aplicaciones, o bien, pueden demandar más complejidad a la solución aportada por la operadora para sus servicios finales, y requerir conectividad a nivel de routing necesitando así más inteligencia en los protocolos de control para manejar el tráfico del cliente.

Además del tipo de enlace que necesite un usuario dependiendo de los elementos de red de que disponga, bien nivel de enlace, o bien nivel de red, también es importante conocer el tipo de conectividad que necesita para sus usuarios. No todas las infraestructuras de red soportan la mima conectividad, siendo la conectividad uno a uno la más sencilla, y la conectividad denominada todos con todos la más compleja.

Es frecuente encontrar los escenarios en los que el propio cliente centraliza sus aplicaciones, servidores y otros elementos de servicios de datos o voz en uno o dos puntos, seguramente para

economizar su inversión y explotar los recursos más costosos, y sean el resto de delegaciones o sucursales las que tengan que acceder a esos servidores centrales para obtener la información deseada. Estas soluciones pueden plantear serias dudas a la hora de elegir una determinada tecnología para implementar su Red Privada Virtual. Si el objetivo del cliente es por ejemplo, sólo el acceso a esos servidores, quizá pueda optar por una solución más simple, en la que cada uno de sus usuarios tenga conectividad punto a punto con esas centrales. Al elegir una solución más sencilla, evidentemente será más económica.

Por el contrario, el cliente puede necesitar una solución algo más compleja e inteligente, a pesar de centralizar sus recursos en varios puntos localizados. Las posibilidades de crecimiento y escalabilidad no se ven reducidas o condenadas a una implementación menos eficiente desde el punto de vista de la plataforma de red ya que no necesitaría recursos exponencialmente a sus nuevas necesidades.

La conectividad todos con todos es posible sólo con determinadas tecnologías de red de una manera eficiente. Incluso, el cliente puede retocar esta solución de todos con todos de forma que sea él en su aplicación final quien evite que realmente el tráfico sea compartido por todos sus usuarios.

3. Tecnología

3.1.1. Redes IP

En una primera división del mundo de las redes de telecomunicaciones se podría fijar el criterio sobre el tipo de arquitectura básica que siguen, entendiendo como arquitectura los protocolos involucrados en cada uno de los niveles de comunicación que se establecen para conseguir la conectividad a distancia.

Las redes fundamentan su estructura en pilas de protocolos que cubren cada una de las particularidades esenciales para una comunicación final extremo a extremo de dos usuarios, aplicaciones o máquinas.

Las pilas de protocolos en esencia son el orden fijado de antemano por los dos elementos que pretenden establecer comunicación. De este modo, pueden esperar el tipo y el formato del tráfico o señal que van tanto a recibir como a transmitir. Si el orden fijado se respeta y se conserva, ambos extremos recibirán información y la transmitirán de acuerdo con lo establecido en cada uno de los protocolos y en el orden que esperan.

Siendo así la arquitectura del paradigma de las redes de comunicaciones, las pilas de protocolos tienen que ser un concepto estándar a nivel global, y deben ser respetados por todos los participantes en la red, es decir, fabricantes, ingenieros, usuarios, aplicaciones, etc.… De otro modo sería imposible la comunicación general entre elementos de red desarrollados conforma a otros paradigmas, pilas de protocolos o estándares.

Existen varias arquitecturas consolidadas y definidas que consiguen una estandarización del modelo adecuadamente, respetando y conservando sus propios principios de modo que consiguen el éxito de su cometido, que no es otro que comunicar dos extremos finales.

Sin embargo, a nivel mundial se puede afirmar que el modelo más asentado e implementado es el basado en la arquitectura TCP/IP. Estos dos protocolos cubren varias capas de una arquitectura

de pila de protocolos y dan nombre a la propia definición del conjunto. Su desarrollo gracias a Internet ha propiciado que la mayoría de los fabricantes de equipos de telecomunicaciones basen su electrónica en dichos protocolos.

El modelo TCP/IP se asemeja al estándar inicial de la OSI, el cual, fue la base o el origen de los paradigmas basados en pilas de protocolos. El modelo definido originalmente por la OSI define siete niveles en dicha pila de protocolos, los cuales, uno a uno respetados extremo a extremo proporcionan conectividad.

El objetivo inicial de esta organización de estándares fue evitar que múltiples fabricantes desarrollaran sus recursos software y hardware sin tener en cuenta la interoperabilidad entre otros proveedores, evitando así el fácil crecimiento en las redes de empresas, clientes y organizaciones.

Para enfrentar el problema de incompatibilidad de redes, la Organización Internacional para la Estandarización (OSI) investigó modelos de conexión, arquitectura de sistemas de red y modelos de pilas de protocolos, a fin de encontrar un conjunto de reglas aplicables de forma general a todas las redes. Con base en esta investigación, la OSI desarrolló un modelo de red que ayuda a los fabricantes a crear redes que sean compatibles con otras redes.

Siguiendo el esquema de este modelo se crearon numerosos protocolos, por ejemplo X.25, que durante muchos años ocuparon el centro de la escena de las comunicaciones informáticas. El advenimiento de protocolos más flexibles donde las capas no están tan demarcadas y la correspondencia con los niveles no era tan clara puso a este esquema en un segundo plano. Sin embargo sigue siendo muy usado en la enseñanza como una manera de mostrar cómo puede estructurarse una "pila" de protocolos de comunicaciones (sin importar su poca correspondencia con la realidad).

El modelo en sí mismo no puede ser considerado una arquitectura, ya que no especifica el protocolo que debe ser usado en cada capa, sino que suele hablarse de modelo de referencia. Este modelo está dividido en siete capas, que definen las funciones de los protocolos de comunicaciones. Cada capa del modelo representa una función realizada cuando los datos son transferidos entre aplicaciones cooperativas a través de una red intermedia.

En una capa no se define un único protocolo sino una función de comunicación de datos que puede ser realizada por varios protocolos. Cada protocolo se comunica con su igual en la capa equivalente de un sistema remoto. Cada protocolo solo ha de ocuparse de la comunicación con su gemelo, sin preocuparse de las capas superior o inferior. Sin embargo, también debe haber acuerdo en cómo pasan los datos de capa en capa dentro de un mismo sistema, pues cada capa está implicada en el envío de datos.

Las capas superiores delegan en las inferiores para la transmisión de los datos a través de la red subyacente. Los datos descienden por la pila, de capa en capa, hasta que son transmitidos a través de la red por los protocolos de la capa física. En el sistema remoto, irán ascendiendo por la pila hasta la aplicación correspondiente.

La ventaja de esta arquitectura es que, al aislar las funciones de comunicación de la red en capas, minimizamos el impacto de cambios tecnológicos en el juego de protocolos, es decir, podemos añadir nuevas aplicaciones sin cambios en la red física y también podemos añadir nuevo hardware a la red sin tener que reescribir el software de aplicación.

El modelo de arquitectura TCP/IP es más simple que el modelo OSI, como resultado de la agrupación de diversas capas en una sola o bien por no usar alguna de las capas propuestas en dicho modelo de referencia.

Así, por ejemplo, la capa de presentación desaparece pues las funciones a definir en ellas se incluyen en las propias aplicaciones. Lo mismo sucede con la capa de sesión, cuyas funciones son incorporadas a la capa de transporte en los protocolos TCP/IP. Finalmente la capa de enlace de datos no suele usarse en dicho paquete de protocolos

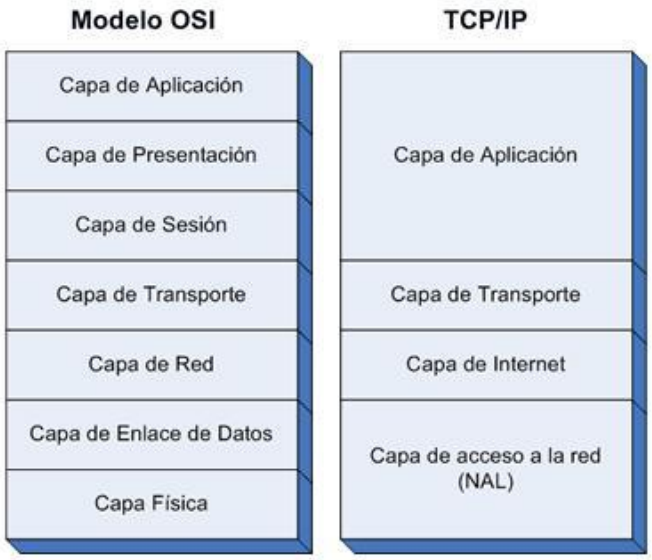

Imagen 2 - Arquitectura de Protocolos

Al igual que en el modelo OSI, los datos descienden por la pila de protocolos en el sistema emisor y la escalan en el extremo receptor. Cada capa de la pila añade a los datos a enviar a la capa inferior, información de control para que el envío sea correcto. Esta información de control se denomina cabecera, pues se coloca precediendo a los datos. A la adición de esta información en cada capa se le denomina encapsulación. Cuando los datos se reciben tiene lugar el proceso inverso, es decir, según los datos ascienden por la pila, se van eliminando las cabeceras correspondientes.

Cada capa de la pila tiene su propia forma de entender los datos y, normalmente, una denominación específica que podemos ver en la tabla siguiente. Sin embargo, todos son datos a transmitir, y los términos solo nos indican la interpretación que cada capa hace de los datos.

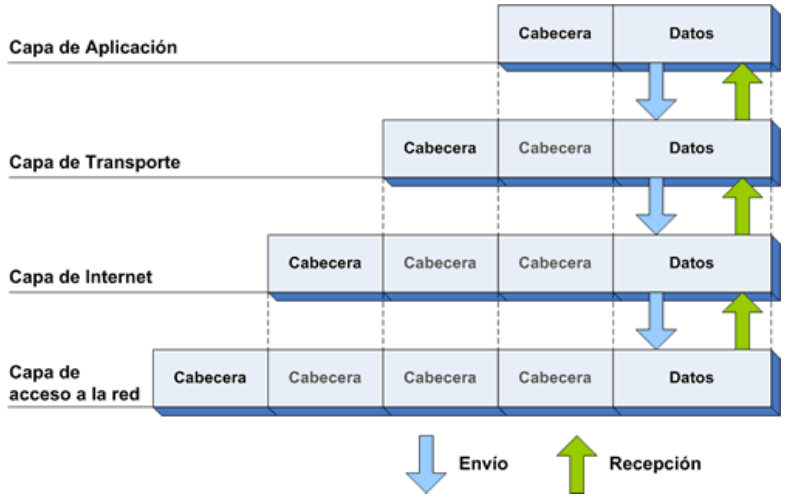

Imagen 3 - Arquitectura de Protocolos (2)

3.1.1.1. Protocolos dinámicos

Una vez se han conectado físicamente todos los elementos que formarán la red, y se les ha configurado direccionamiento sobre sus interfaces/adaptadores de red, se llega al paso de la conectividad. Los elementos de la red han de ser capaces de llegar unos a otros independientemente de si están directamente conectados, o de si tienen direccionamiento diferente.

Todo elemento de una red, que tenga capacidad de entender direcciones IP y por lo tanto sepa enrutar el tráfico en función de las direcciones IP destino, debe conocer por qué interfaz sacar el tráfico que recibe y no está destinado a él.

Un nodo de red, tiene diferentes interfaces con dirección IP, las cuales pertenecen a una red o subred diferente. Por el hecho de formar parte de una subred, ese nodo automáticamente aprende que esos interfaces serán los que utilizará para sacar el tráfico que reciba cuya dirección IP de destino pertenezca a una de esas subredes.

Parece lógico que el propio elemento de la red determine de inmediato la salida del tráfico destinado a subredes de las cuales él forma parte directamente

En una red con cierta dimensión, un elemento no tiene normalmente una interfaz en cada uno de los segmentos que forman esa red de tamaño medio o grande. Esto supondría muchos recursos hardware en el equipo, excesivos recursos a nivel de enlaces físicos o cableados, y poca eficiencia en el diseño de la misma.

Por esto, se necesita introducir manualmente rutas en los equipos para dirigir el tráfico destinado a subredes que el nodo no tiene directamente conectado. De esta forma, un elemento de red puede recibir tráfico destinado a una subred que no conoce implícitamente por no tenerla directamente conectada, pero que sí sabe enlutar porque tiene una salida configurada a mano que le indica por qué interfaz debe sacar el tráfico, y será el próximo o los próximos elementos de red que lo reciban los que ya lo conocerán por tener dicha subred directamente conectada.

Esta solución solventa el problema de conectividad entre subredes que no comparten un nodo de red, pero no soluciona la cuestión de la escalabilidad y el crecimiento de una red mayor. Introducir manualmente rutas estáticas para una red de dimensión grande puede ser una odisea, por el volumen de subredes posibles, por el posible número de elementos a configurar manualmente uno por uno, y por el continuo trabajo de mantenimiento que implicaría.

Para evitar esta afanosa tarea de configuración y mantenimiento manual, existen diferentes protocolos de enrutamiento que se habilitan en los equipos de red, y una vez sincronizados y establecidos, generan la información necesaria para enlutar tráfico a cualquier subred de forma dinámica y automática.

Estos protocolos de enrutamiento dinámico tienen que configurarse inicialmente por ingenieros de redes, y una vez activados, el intercambio de información ante posibles cambios de redes, cortes de enlaces, o desconexiones de equipos se realiza de forma automática.

Los elementos de red normalmente corren el mismo protocolo dinámico, pero también son posibles las soluciones mixtas, en las que algunos elementos tienen configurado uno o varios protocolos dinámicos y mediante políticas y filtros creados por el ingeniero, permite el intercambio de información de un protocolo a otro.

Los protocolos de enrutamiento dinámico más conocidos y utilizados son RIP, OSPF, EIGRP, BGP… siendo la mayoría definidos por estándares oficiales, y algunos como por ejemplo EIGRP, propietarios del fabricante que lo patentó, obligando así a utilizar sus propios equipos en la red para utilizarlo.

Los protocolos de enrutamiento dinámico pueden utilizarse como protocolos internos a una red o como protocolos que interconectan dos redes

3.1.1.2. IGP – Interior Gateway Protocol

En un entrono particular de redes, en el cual una serie de elementos de red, routers y switches, tienen conectividad física entre sí, existe la necesidad de que la información de enrutamiento sea estable, coherente y eficiente.

Asumiendo la ineficacia de enrutamiento estático en redes con cierto tamaño, inclusive de pocos elementos, es necesario elegir un protocolo de enrutamiento dinámico que gobierne dicha red y sea capaz de establecer la señalización del tráfico entre todos los nodos.

La elección dependerá de varios factores, recursos disponibles, requerimientos de funcionalidad, limitaciones técnicas, limitaciones a nivel político, etc. Una vez analizada la situación y el escenario sobre el que se quiere desplegar la solución de routing, se debe planificar y preparar la implementación del protocolo en base a sus cualidades y propiedades.

Los protocolos de enrutamiento presentan características diferentes, y en base a ellas se específica el diseño de la solución en detalle.

Los protocolos de pasarela internos se pueden dividir en dos grandes categorías, en función del algoritmo de cálculo que utilizan para decidir la mejor ruta. Existen varios, pero los más extendidos son el algoritmo de Bellman-Ford y el de Dijsktra

3.1.1.2.1. Protocolos de enrutamiento por Vector-Distancia

Se conoce como Vector-Distancia a la familia de protocolos que utilizan este método de encaminamiento que se basa en calcular la dirección y la distancia hasta cualquier enlace en la red. El coste de alcanzar un destino se lleva a cabo usando cálculos matemáticos como la métrica del camino.

En los protocolos de este tipo, ningún elemento tiene información completa sobre la topología de la red. En lugar de ello, se comunica con los demás equipos, enviando y recibiendo información sobre las distancias entre ellos. Así, cada router genera una tabla de enrutamiento que usará en el siguiente ciclo de comunicación, en el que los demás nodos intercambiarán los datos de las tablas. El proceso continuará hasta que todas las tablas alcancen unos valores estables. Este conjunto de protocolos tienen el inconveniente de ser algo lentos, si bien es cierto que son sencillos de manejar y muy adecuados para redes compuestas por pocas máquinas.

Los cambios son detectados periódicamente ya que la tabla de encaminamiento de cada router se envía a todos los vecinos que usan en mismo protocolo. Una vez que el router tiene toda la información, actualiza su propia tabla reflejando los cambios y luego informa a sus vecinos de los mismos. Este proceso se conoce también como “encaminamiento por rumor” ya que los nodos utilizan la información de sus vecinos y no pueden comprobar a ciencia cierta si ésta es verdadera o no

Uno de los problemas más graves de éstos protocolos es la creación de bucles, o agujeros negros para el tráfico, ya que la desaparición de un nodo en la red puede provocar que el resto siga mintiendo rutas hacia él sin conocer su verdadero estado.

Como soluciones, se suele definir un límite de saltos máximo para el cual se puede dar por caído a un equipo. Una de las modificaciones aplicadas sobre el paradigma Vector-Distancia es el conocido como **split-horizon,** y que básicamente es una norma por la cual, un nodo no anunciará una ruta a su vecino, si la ruta pasa por el propio vecino. Esto no es efectivo en todos los escenarios y topologías por lo que sólo suaviza el problema.

El ejemplo más común de este tipo de protocolos es **RIP**, (Routing Information Protocol), que tiene la ventaja de ser muy fácil de poner en marcha, aunque para calcular una ruta sólo tiene en cuenta por cuántas equipos pasará, y no otros aspectos más importantes como puede ser el ancho de banda

3.1.1.2.2. Protocolos de enrutamiento por Enlace-Estado

En este caso, cada nodo posee información acerca de la totalidad de la topología de la red. De esta manera, cada uno puede calcular el siguiente salto a cada posible nodo destino de acuerdo a su conocimiento sobre cómo está compuesta la red. La ruta final será entonces una colección de los mejores saltos posibles entre nodos.

Esto contrasta con los basados en el cálculo de la distancia entre redes, en el que cada nodo tiene que compartir su tabla de enrutamiento con sus vecinos. En los protocolos Enlace-Estado, la única información compartida es aquella concerniente a la construcción de los mapas de conectividad. Se basa en que un router comunica a los restantes nodos de la red cuáles son sus vecinos y a qué distancias está de ellos. Con la información que un nodo de la red recibe de todos los demás, puede construir un "mapa" de la red y sobre él calcular los caminos óptimos

Una vez que el router ha completado la recopilación de información, utiliza el algoritmo de Dijkstra para calcular el camino más corto a todos los nodos.

El problema fundamental de los protocolos basados en Vector-Distancia es que aantes de 1979 el máximo ancho de banda era de 56Kb posteriormente se modernizaron las líneas a 230Kbps o incluso a 1,5Mbps lo que hizo necesario el uso de mejores técnicas, que sí lo considera el Estado del enlace

Los protocolos basados en el estado del enlace, convergen más rápido, consiguen que cada nodo pueda construir una topología de la red que será coherente en todos, y se adapta mejor a cambios en la red, aunque para ello necesita siempre más recursos de CPU y memoria de los equipos

El máximo exponente de estos protocolos es **OSPF,** y además es el más utilizado en grandes redes ya que facilita el crecimiento de las mismas sin afectar lo existente. Además, este protocolo perite dividir el escenario en áreas, pequeñas o grandes, que resumen la información de routing de las mismas y establece ciertas reglas de comunicación entre ellas.

Otro de los puntos fuertes para que OSPF sea el protocolo de IGP más extendido es que se basa en las normas de código abierto, lo que significa que muchos fabricantes lo pueden desarrollar y mejorar

Por el tamaño de redes que se pueden llegar a administrar con **OSPF**, su apertura de código implementado por la mayoría de fabricantes y por las cualidades propias del protocolo, le elegimos como **IGP** para nuestra red VPLS jerárquica.

Especificación de diseño 1

3.1.2. OSPF

Este protocolo ha sido desarrollado por un grupo de trabajo del Internet Engineering task Force, cuya especificación viene recogida en el RFC 2328, ante la necesidad de crear un protocolo de routing interno que cubriera las necesidades en Internet de que protocolos como RIP ponían de manifiesto

OSPF ha sido pensado para el entorno de Internet y su pila de protocolos TCP/IP, como un protocolo interior, es decir, que distribuye información entre routers que pertenecen al mismo Sistema Autónomo, entendiendo como Sistema Autónomo un conjunto de redes bajo la misma administración y la misma política de encaminamiento

El fundamento principal en el cual se basa un protocolo de estado de enlace como OSPF es en la existencia de un mapa de la red el cual es construido por todos los nodos y regularmente actualizado. Para llevar a cabo este propósito, la red debe de ser capaz de:

- Almacenar en cada nodo el mapa de la red.
- Ante cualquier cambio en la estructura de la red actuar rápidamente, con seguridad, sin crear bucles y teniendo en cuenta posibles particiones o uniones de la red.

Además, OSPF puede estructurarse en áreas con distintos niveles de organización para funcionar como un protocolo jerárquico, de forma que resulte más fácil administrar redes de gran tamaño. Equipos con múltiples interfaces pueden formar parte de múltiples áreas a modo de nodos frontera. La idea es que cada área agrupe una serie de nodos o dispositivos finales para conseguir estructurar una red.

3.1.2.1. Tipos de áreas y routers en OSPF

Por definición del protocolo, existe un área central o backbone que formará la parte central de la red y será el área con más peso en la estructura jerárquica. Todas las áreas deben tener un router con conectividad directa al backbone, también denominado área 0 para que todas las rutas anunciadas circulen por el mismo.

En el caso de no ser posible la conexión directa de un área con el área 0, se debe definir un enlace virtual, que pase transparentemente a través de algún área intermedia o de tránsito y proporcione conectividad al backbone con dicha área remota.

Cada área mantiene una tabla de estado de enlace cuya información es sumarizada y traspasada al resto de áreas a través de los routers frontera que las interconectan. De este modo, la topología de la propia área no es conocida fuera del mismo, reduciendo así la cantidad de información que se intercambia dentro del Sistema Autónomo.

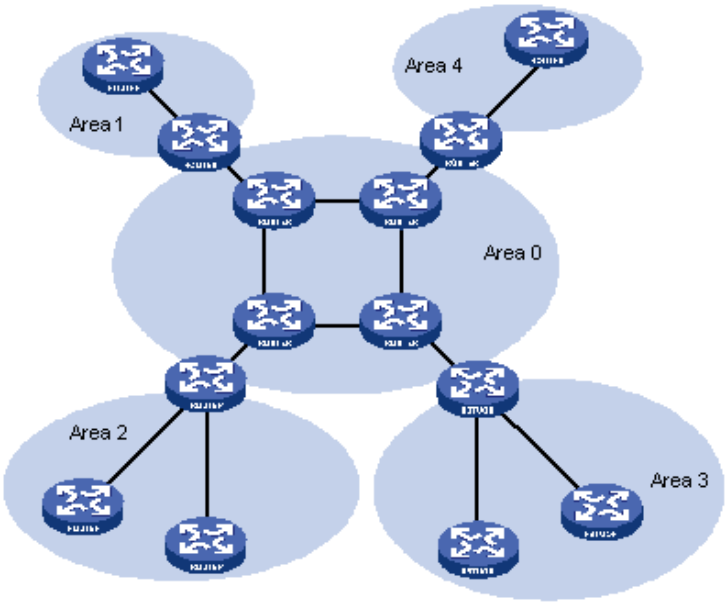

Imagen 4 - OSPF

Dependiendo de las conexiones que tengan los routers de cualquiera de las áreas, asumirán unas funciones u otras. Se definen los siguientes roles para los routers de un Sistema Autónomo gobernado por OSPF:

- **ASBR** (Autonomus System Border Router): Éste equipo, como su nombre indica, tendrá conexiones fuera del Sistema Autónomo con otras redes cuya política de routing es muy posible que desconozcamos o por lo menos no podamos administrar. Estos equipos se encargarán de introducir rutas hacia otras redes ajenas al Sistema Autónomo. Lo normal es hacerlo de forma controlada redistribuyendo sólo la información que nos interese, o bien anunciando una salida por defecto si por ejemplo se trata de una conexión a un proveedor de acceso a Internet.
- **ABR** (Area Border Router): Estos equipos interconectan áreas actuando como equipos frontera. Por definición del protocolo, tendrán conexión al área cero, y a otra.
- **IR** (Internal Router): Este equipo sólo tendrá relación de vecindad establecidas con equipos dentro de su misma área.
- **BR** (Backbone Router): Estos equipos pertenecen al área cero, y pueden ser simplemente BR o también ABR si además tienen conexión con otra área.

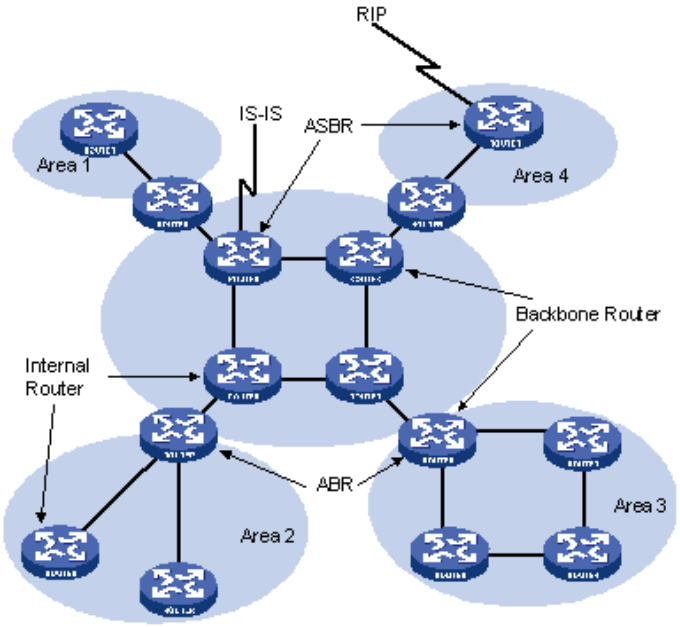

Imagen 5 - OSPF (2)

OSPF contempla la posibilidad de utilizar medios de acceso compartidos tales como segmentos Ethernet. En estos casos de acceso múltiple, la negociación entre equipos de una misma área puede suponer un problema en cuanto a carga de tráfico de routing ya que tienen que intercambiar información todos con todos. Para solucionar este inconveniente, se han definido dos tipos de roles que organizarán en cierto modo el segmento de red compartido. Se trata de:

- **DR** (Designated Router): Este role asumido por un equipo sobre uno de sus interfaces conectados al segmento compartido, establece una relación de dependencia con todos los demás equipos del segmento. Cada equipo envía su información al DR, y éste la replica a los demás evitando de este modo que cada uno tenga que enviarla individualmente a cada elemento
- **BDR** (Backup Designated Router): Este equipo contiene la misma información de la red que el DR, pero sólo replica la información cuando el principal tiene problemas de conectividad o se ha caído.

La elección de DR y BDR depende de ciertos parámetros intercambiados cuando los equipos se están descubriendo unos a otros. No se debe confundir el concepto de DR y BDR como tipos de routers OSPF. Cualquier equipo en un entrono OSPF puede tener interfaces DRs, BDRs, u otros como no DRs.

Así pues, dependiendo del tipo de conexiones que tenga un router o sus forma de acceder físicamente a la red, desempeñará unas funciones u otros. Dichas funciones pueden ser portavoz de redes externas al sistema, o actuar de frontera entre áreas internas o simplemente informar sobre sus propias redes o equipos conectados.

Las áreas también pueden definirse más estrictamente para permitir la sólo la información que se crea más conveniente. En OSPF existen dos tipos de áreas especiales que se comportan de la siguiente manera:

- **Área Stub:** En estas áreas no se reciben anuncios de redes externas al sistema Autónomo. En lugar de anunciar muchas redes ajenas, se anuncia una ruta por defecto como salida del tráfico para dicha área. Para ello, estas áreas sólo tienen un router corriendo OSPF, y por supuesto, éste no puede ser del tipo ASBR
- **Área Not-so-stubby:** Estas áreas tampoco permiten que se les anuncien redes externas al Sistema Autónomo desde otras áreas, pero si lo permiten si es directamente de otros Sistemas Autónomos. También pueden anunciar dichas redes al resto de áreas.

3.1.2.2. Tipos de adyacencias y de mensajes

Los nodos de una red basada en OSPF se conectan a ella a través de una o varias interfaces con las que se conectan a otros nodos de la red. El tipo de enlace define la configuración que asume la interfaz correspondiente. OSPF soporta tanto enlace punto a punto como enlaces sobre accesos múltiples.

Una vez detectados entre sí a través de paquetes *Hello*, los routers establecen una relación de adyacencia tras el intercambio de una serie de mensajes. En un entrono de acceso compartido, como se explica anteriormente, se designará un router DR para que establezca adyacencias con el resto de routers y replique las rutas de cada uno a los demás.

Los routers que establezcan relación de vecindad lo hará a través de interfaces que pertenezcan a la misma área, por lo tanto, un interfaz sólo puede asociarse a una única área. Lo harán utilizando los distintos tipos de mensajes definidos en el protocolo. Se tratan de:

- Mensajes **HELLO:** utilizados para descubrir y establecer relación con otros routers OSPF.
- Mensajes **DATABASE DESCRIPTION:** sumarizan contenidos de la base de datos
- Mensajes **LINK STATE REQUEST:** utilizados para solicitar la información de adyacencias de los vecinos en la fase de establecimiento e intercambio
- Mensajes **LINK STATE UPDATE:** utilizados para enviar información a los vecinos
- Mensajes **LINK STATE ACKNOWLEDGMENT:** confirmación a mensajes enviados.

Dependiendo del tipo de información incluía en los mensajes de LINK STATE se clasificarán en LSAs de diferentes tipos, por ejemplo, sumarizaciones de áreas, información externa al Sistema Autónomo, listado de routers dentro un área compartida con DR y BDR, mensajes de áreas stub, etc..

+	Bits 0–7	8–15	16–18	19–31
0	Version	Type	Packet Length	
32	Router ID			
64	Area ID			
96	Checksum		Authentication Type	
128	Authentication			
160	Authentication			
192	Network Mask			
224	Hello Interval		Options	Router Priority
256	Router Dead Interval			
288	Designated Router			
320	Backup Designated Router			
352	Neighbor ID			
384	...			

Imagen 6 - OSPF (3)

Una vez se hayan establecido todas las relaciones de vecindad entre todos los equipos correctamente, existirá una consistencia entre todos ellos a través de una tabla topológica de la red coherente, ya que cada nodo conocerá la situación de sus vecinos, y el estado de los enlaces de los mismos.

Dado que todos los routers han compuesto un mapa de la red idéntico y han ejecutado el algoritmo SPF para calcular la ruta más óptima a cada destino, OSPF es un protocolo libre de bucles, y gracias al intercambio de LSAs, rápido en cuanto a convergencia si se producen cambios inesperados en la red.

Para el diseño de nuestra red VPLS jerárquica definiremos un único un área OSPF, que por consiguiente será el área cero. Todos los nodos de la red serán **Backbone Routers,** no habiendo ASBR o ABR en ningún momento. Los enlaces utilizados para conectarse entre sí los equipos serán sobre Ethernet, simulando un punto a punto, por lo que en el establecimiento de la adyacencia se designarán un DR entre los dos vecinos.

Especificación de diseño 2

3.1.3. VPLS (Virtual Private LAN Service)

3.1.3.1. Descripción

Un servicio basado en tecnología VPLS es un servicio que emula la funcionalidad completa de una Red de Área Local tradicional independientemente de su distribución geográfica. Dicho de otro modo, una red VPLS convierte los entornos tradicionalmente considerados WAN en entornos LAN.

Un entorno WAN se considera habitualmente con elementos de un cliente o servicio que se conectan a la red individualmente y por norma general situados a gran distancia del primer nodo de red. Así, se crea un enlace punto a punto entre dicho elemento y el primer punto de la red, proporcionando la conectividad al equipo remoto con la Red Privada Virtual.

Con los protocolos tradicionales sobre los que se constituyen Redes Privadas Virtuales suele ser necesario configurar conexiones entre todos y cada uno de los equipos remotos conectados individualmente, o haciéndose valer de un equipo central que haga de repetidor del tráfico entre el resto de sedes o equipos remotos.

Aun así, es bastante costoso a nivel de recursos lógicos de la red proporcionar conectividad todos con todos, con conexiones punto a punto entre todos los elementos de cada uno de los usuarios o clientes.

A través de una red basada en tecnología VPLS convertimos el entorno WAN, en un entorno LAN, entendiendo este entorno LAN conseguido con VPLS como la simulación de un segmento Ethernet en el que se considera que todos los elementos están directamente conectados entre sí.

Una red VPLS reproduce el funcionamiento de un switch tradicional, elemento de red con el que se separan dominios de colisión, y a través del cual se consigue ampliar un simple segmento Ethernet de red.

VPLS en sí mismo está basado en los protocolos que forman la tecnología MPLS. De este modo, la red se construye sobre la pila de protocolos definida para MPLS, consiguiendo una red IP/MPLS

que proporcionará conectividad a los clientes de forma independiente y conectando sus segmentos de red de forma transparente como si estuvieran ubicados en el mismo emplazamiento físico.

Las Redes de Área Local de los clientes se extienden hasta el proveedor de servicios, y se conectan a la red que está simulando el comportamiento de un switch tradicional.

VPLS es un servicio que proporciona la posibilidad de crear Redes Privadas Virtuales sobre estructuras basadas en Ethernet. Consigue que varios segmentos de LAN dispersos geográficamente se comuniquen entre sí compartiendo el mismo dominio de difusión Ethernet, es decir, como si estuvieran conectados al mismo segmento de LAN.

Con la implementación de una red basada en VPLS, las Redes de Área Local de los usuarios se interconectan entre sí como si la red completa fuera un simple switch. Al estar simulando una LAN, se consigue el modelo de conectividad punto a multipunto, en el que todos los segmentos de LAN se comportan como uno único.

Básicamente se trata de construir VPNs para usuarios sobre una red IP compartida, tal y como muestra el siguiente esquema

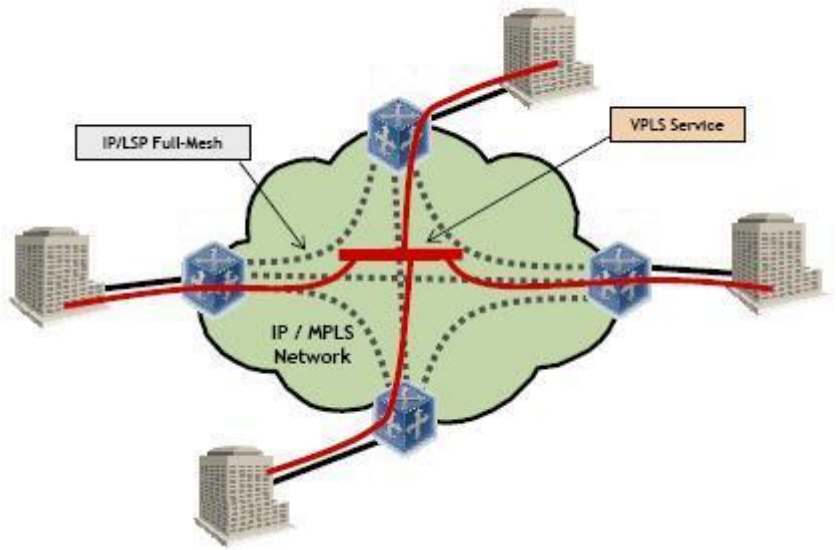

Imagen 7 - VPLS

Al implementar VPLS sobre redes IP/MPLS, se consigue en cierto modo que este tipo de redes sean capaces de proporcionar conectividad punto a multipunto y no sólo punto a punto. Aunque de cara al segmento de LAN del usuario al que se le quiere dar servicio, sólo se implementa hasta nivel 2 de la pila de protocolos TCP/IP, es decir, el equipo remoto del usuario sólo tiene que preocuparse por el nivel físico y la trama Ethernet con el equipo de red al que se conecta, estos últimos implementan la pila completa TCP/IP de cara a los demás equipos de red.

3.1.3.2. Elementos de red

Los elementos de una red VPLS o MPLS se diferencian fundamentalmente por el tipo de conexiones que tienen y su localización lógica en la topología de la red. La distinción se hace radicalmente en si el equipo tiene conexiones contra equipos de usuarios finales, o si solamente tiene conexiones contra equipos dentro de la red.

A los primeros se les denomina **PE router** (Provider Edge router), y son los nodos que ofrecen interfaces de cara al usuario remoto para que éstos puedan acceder a la red. En cierto modo son la frontera de la red MPLS, normalmente la frontera entre un operador de red y sus clientes. Estos equipos desempeñan un papel clave ya que, aun teniendo muy diferenciadas las funciones de red y las funciones de acceso, deben relacionarlas correctamente y crear el servicio de red privada virtual para el usuario basándose en la tecnología aplicada sobre ellos.

Los equipos en el interior de la red, con sólo conexiones troncales entre otros nodos de red, son denominados **P router** (Provider), y mantienen la señalización para el tráfico de todos los usuarios, utilizando la pila de protocolos correspondiente para diferenciarlo y aislarlo adecuadamente.

Finalmente, para completar la nomenclatura de los equipos involucrados en redes que soportan VPNs sobre IP/MPLS, denominamos a los equipos de los usuarios o clientes como **CE router** (Customer Edge). Éstos pueden implementar distintas arquitecturas, switches, routers, hosts con capacidad para enrutamiento…en el presente diseño se especificarán recomendaciones para utilizar fundamentalmente routers, y en su caso swtiches.

3.1.3.2.1. Equipos de acceso (PE)

Los equipos que proporcionan acceso a los enlaces de usuarios tienen claramente diferenciado el plano de acceso y el plano de red, dentro de los procesos propios del tráfico de control. Implementan las dos pilas de protocolos de forma paralela para que por un lado, el usuario pueda establecer comunicación con el equipo de acceso, de forma que éste primero se crea que lo está haciendo con el equipo remoto al cual quiere conectarse y que estará físicamente conectado al otro lado de la red IP/MPLS.

Para que esto ocurra, se implementan sobre los equipos de acceso los mismos protocolos que utiliza el usuario para establecer comunicación remota. En este caso, se pretende ofrecer una conexión Ethernet (nivel 2) de forma transparente al usuario, por lo que de cara a éste, sólo se implementa nivel físico y nivel de enlace según el paradigma TCP/IP.

Al tratarse de Ethernet en el nivel de acceso, el cliente apenas necesita desarrollar o implementar tecnología o equipamiento sofisticados, una tarjeta de red Ethernet es suficiente. De momento no se especifica tráfico de control entre los dos extremos de la conexión Ethernet. Si físicamente la conexión está arriba, será suficiente para que los interfaces Ethernet en los dos extremos también lo estén.

Como se muestra en el siguiente gráfico, el equipo de acceso presenta dos pilas de protocolos dependiendo del equipo conectado

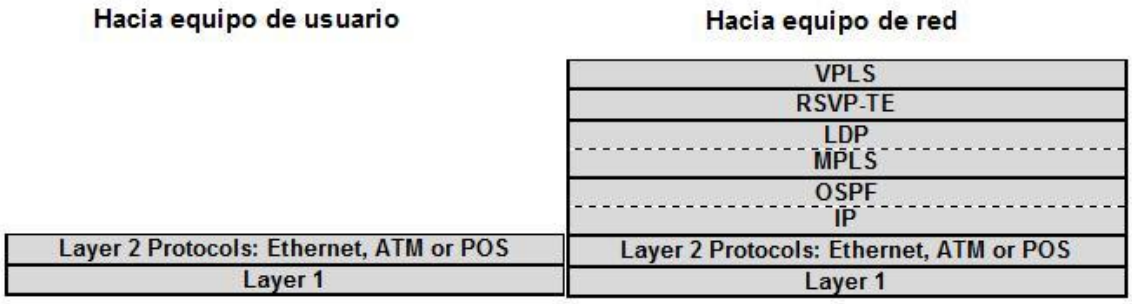

Imagen 8 - VPLS (2)

En estos equipos, en el plano del usuario o cliente, se guarda una relación entre el interfaz físico y la dirección MAC Ethernet del equipo conectado. De esta forma se puede aprender la localización del equipo como si de un puerto de un conmutador de capa 2 se tratase.

Para el usuario el tratamiento que recibe su tráfico en la red es transparente en lo que a conectividad se refiere. El equipo de usuario conectado a la red cuya VPN se implementa sobre VPLS no necesita levantar nivel IP con la red, lo hará directamente con el equipo remoto con el que quiera establecer comunicación. De este modo, el problema de direccionamiento IP que pudiera surgir entre usuario y red desaparece, asumiendo el usuario la responsabilidad de asignar direccionamiento IP de forma consistente.

En el caso de servicios de VPN sobre MPLS, la pila a implementar de cara al usuario es más compleja. Requiere levantar hasta IP, es decir, nivel 3, y seguramente algún protocolo de routing sobre el mismo para asegurar la disponibilidad del enlace o la redundancia de accesos sobre todo si el nivel de enlace está desplegado sobre Ethernet, donde ya hemos visto que no hay mecanismo de control extremo a extremo, y posibles fallos físicos intermedios no se detectarían.

Los PEs de redes MPLS asocian direcciones IP y rutas a los interfaces físicos por donde acceden los usuarios, no las direcciones físicas.

3.1.3.2.2. Equipos de red (P)

Los equipos internos a la red, sin conexiones de acceso de clientes, deben implementar la pila de protocolos necesarios para proporcionar redes privadas virtuales sobre VPLS en todos los interfaces de red que tengan conexiones hacia otros equipos. Genéricamente, estos equipos son denominados **nodos de red,** pero específicamente para redes IP/MPLS, P routers.

El plano de control en estos equipos soporta mucha señalización, ya que al ser equipos internos, tendrán numerosas conexiones con otros nodos. El plano de reenvío en estos equipos queda al margen y dicha señalización es la que consigue que se reduzca el tiempo de conmutación del tráfico al tener señalizado el camino a seguir para cada paquete de antemano.

Durante cambios topológicos en la red, caídas de otros nodos, cortes en los enlaces, se producen tormentas de tráfico de señalización entre los nodos de red, (incluyendo también a los PEs ya que terminan los túneles de tráfico de usuarios). Dichas cantidades de tráfico de control producen normalmente picos en la utilización de los recursos hardware y software de los equipos, ya que ejecutan algoritmos de protocolos, actualizan las tablas de rutas, renegocian caminos y etiquetas para encapsular y aislar los paquetes de los usuarios, intentado además en todo momento no interrumpir el reenvío de tráfico de servicio para los usuarios. Es normal en estas situaciones experimentar subidas de CPU en los equipos de forma temporal.

Debido a la carga de control y la cantidad de señalización que deben mantener estos nodos, es recomendable instalar los equipos de gama más alta dentro de los que se vayan a desplegar en la red. Sus funciones resultan clave para el éxito en la señalización de la red, y para evitar cortes de servicio a los usuarios ante fallos relativamente normales en la red.

Para el diseño de nuestra red VPLS jerárquica definiremos como **Core** de la red **cuatro nodos** que actúen como **P routers.** Estos cuatro elementos no tendrán conexiones de clientes o usuarios, estarán conectados entre sí formando un cuadrado y además, desempeñarán funciones jerárquicas que se explicarán en capítulos posteriores

Especificación de diseño 3

3.1.3.2.3. Equipos de usuario o cliente (CE)

El usuario final o cliente no debería supeditarse en exceso a especificaciones por parte del operador o proveedor de red, pero tampoco puede elegir de forma independiente los equipos a utilizar. Debe seguir ciertas recomendaciones para que el servicio que vaya a utilizar resulte lo más transparentemente posible a sus necesidades.

En primer lugar, se ha de hacer un estudio concienzudo y detallado de las necesidades que se necesitan cubrir, antes de elegir un tipo de tecnología u otro a implementar en la red. Básicamente tiene que estimar las necesidades de ancho de banda, el tipo de tráfico que cursará y las características del mismo, y la dimensión que quiere cubrir en cuanto a número de equipos y dispersión geográfica.

Una vez concluido, que el usuario necesita anchos de banda grandes, conmutación del tráfico rápida, con aplicaciones sensibles al retardo, y conectividad entre todos sus equipos como si estuviesen directamente conectados entre sí en una red local, la decisión de implementar un servicio VPLS es inmediata.

Dependiendo del proveedor u operador de red, se recomendará la instalación de unos equipos u otros, fundamentalmente razonado en base a las prestaciones de la red y los compromisos que asegure la misma.

Para el tipo de red sobre el que estamos haciendo el estudio, serían recomendables routers, por el simple hecho de que utilizarán sólo una dirección física por cada sede de cliente para conectarse a la red. Esto simplifica mucho el acceso a una red VPLS y reduce el tamaño de MACs en la VPN del usuario. Al menos, deberían ser así la mayoría de la sedes.

En el caso de necesitar un switch que conmute el tráfico de varios segmentos de red del cliente, éste no tratará el tráfico a nivel 3 por lo que no alterará la dirección física del tráfico entrante a la red, recibiéndose así multitud de direcciones MACs por el mismo acceso. Esta situación es menos deseable para la red, ya que se requiere mayor control de las direcciones MAC para evitar situaciones complejas y perjudiciales al servicio como se explica a continuación. No obstante, la utilización de algunos switches en la topología final del usuario pueden admitirse en estas redes.

Si el usuario conectase directamente equipos con la aplicación final, es decir, host o servidores, éstos deberían ser consistentes con el diseño completo del usuario, e interactuar con el resto de elementos bajo su supervisión. De cara a la red, sería un elemento conectado capaz de levantar nivel de enlace con el PE.

3.1.3.3. Aprendizaje de direcciones MAC

Es fundamental en una red VPLS que los equipos de red implementen correctamente esta funcionalidad. De hecho, es la clave del éxito del servicio, ya que para conseguir tiempos de conmutación en los nodos tan bajos resulta imprescindible.

La idea básicamente reside en aprender dónde está conectado cada equipo remoto del usuario, y por lo tanto aprender la dirección MAC de dichos equipos y asociarla al interfaz por el cual se conectan a la red. De esta forma, se puede conseguir reenviar tráfico de un modo directo.

El aprendizaje de MACs se habilita en el servicio, y cuando los equipos remotos generan tráfico entra en funcionamiento. Irá completando la tabla de MACs para conmutar tráfico rápidamente a medida que los equipos vayan introduciendo tráfico en la red. Las tablas de MACs de un equipo de red tienen temporizadores que hacen expirar las entradas de la misma cuando pasa determinado tiempo desde la última vez que transmitiera un equipo.

La primera vez que llega un paquete de un equipo que hasta ese momento no estuviera transmitiendo o no tuviera una entrada en la tabla de MACs, el equipo de acceso de la red lo primero que hace es introducir en la tabla la relación dirección MAC de origen con el interfaz por donde recibió el paquete. De esta forma el equipo de usuario que ha generado el tráfico queda asociado a un interfaz físico, y a menos que se desinstale el equipo para moverlo a otra ubicación con otro interfaz para el acceso a la red, o se sustituya por otro, esa dirección de origen generará tráfico siempre por el mismo interfaz.

El equipo de acceso para reenviar la trama hacia la red, debería conocer porque enlace de red puede localizar correctamente al destino. Si es la primera vez que esa dirección MAC es examinada por el equipo de red, lo que hará será comportarse como un switch de capa 2, y reenviará la trama por todos los interfaces que pertenecen a la VPN el usuario excepto por el que lo recibió. La red así se inundará de tramas idénticas esperando que un equipo remoto reciba una correctamente. A su paso por los demás nodos de red, la trama irá dejando un registro en la tabla de MACs de cada uno, aprendiendo que la MAC originaria de la trama está accesible por el interfaz por donde se haya recibido.

Llegado al punto en que un equipo de red reenvía la trama y el equipo remoto es el destinatario correcto, éste devolverá tráfico hacia el origen para establecer comunicación o simplemente confirmar que lo ha recibido. Ahora, el equipo que antes era destino se convierte en origen y coloca así su dirección MAC como dirección origen de la trama. Al llegar la trama a la red, el primer equipo que la reciba determinará unívocamente que la MAC de este segundo equipo se debe asociar al interfaz por el que la reciba, cerrándose así en el primer nodo el aprendizaje de MACs para esta comunicación entre dos equipos de usuarios.

En este caso, para que la respuesta llegue correctamente, los nodos no inundarán la red replicando la trama por todos los interfaces que tengan asociados ya que anteriormente aprendieron la MAC a la que ahora tienen que reenviar el tráfico. A cada paso por los nodos de red, éstos cerrarán el aprendizaje para estos dos equipos de usuario.

Finalmente, la trama la recibirá el equipo que tiene directamente conectado al destino y la comunicación se habrá completado. El paquete inicial se fue replicando por cada nodo que pasó pero la respuesta siguió un camino directo y sin replicaciones de tramas.

3.1.3.4. Tablas de MACs

Cada VPN de usuario basada en VPLS conformará una tabla de direcciones MACs donde se asociará cada dirección aprendida por el interfaz por donde se recibió el tráfico con esa MAC origen. Una VPLS estable deberá tener la misma tabla de MACs en todos los nodos de la red. Esta consistencia refleja el número de dispositivos del usuario, dando por hecho que transmitirán un mínimo de tráfico en todo momento.

El tamaño de la tabla de MACs puede elevarse considerablemente si los dispositivos de usuario son switches que sólo realicen funciones de capa 2. Siendo así, cada dispositivo remoto puede conmutar tráfico de múltiples equipos del usuario generando entradas de múltiples MACs sobre el mismo interfaz físico.

Es conveniente conectar a la red en la medida de lo posible elementos que levanten hasta nivel IP en la pila de protocolos TCP/IP en la ubicación remota del usuario. De esta forma, estos equipos al hacer routing, cambiarán siempre la dirección física

Para el diseño de nuestra red VPLS jerárquica habilitaremos la funcionalidad de **aprendizaje de MACs**, contemplando la posibilidad de definir tamaños de tablas de MACs por VPLS en función del número de dispositivos de cada usuario

Especificación de diseño 4

3.1.3.5. Detección de bucles físicos

3.1.3.5.1. Movimiento de direcciones MAC

Es posible implementar mecanismos de protección para la red. Al ser un servicio basado en nivel 2 y emulando un dominio de difusión Ethernet, estas redes son muy sensibles a bucles físicos hacia ellas. Es habitual que los usuarios monten sus equipos en sus propias sedes, de manera que haya que transportar el circuito de acceso hasta las instalaciones del operador de red. Dichos circuitos seguramente se implementen sobre soluciones de nivel 1, transmisión pura. Si en algún punto de esos circuitos se coloca un bucle físico hacia la red, esto provoca un bucle sobre el dominio de difusión, y el tráfico que fuera dirigido al equipo con el bucle en su línea de acceso o el tráfico broadcast dirigido a todos los miembros de la VPN, es devuelto a la red generando así una tormenta de tráfico que saturará rápidamente el ancho de banda disponible, afectando gravemente al servicio, ya que al ser tramas de nivel 2, no tienen contadores de saltos que puedan hacer que expiren y sean eliminados por el nodo donde se termine el contador.

Resulta muy difícil detectar estas situaciones ya que el interfaz físico con el bucle en la línea de acceso no dejará de reenviar tráfico hacia el usuario remoto, y así, la misma cantidad de tráfico volverá automáticamente a la red.

Existen algunos proveedores de estos equipos que implementan un mecanismo de control basándose en la funcionalidad de aprendizajes de MACs. Cuando se produce un bucle físico, el tráfico que es devuelto a la red conserva la dirección MAC origen del dispositivo de usuario que generó el tráfico. Esta MAC origen ahora es vista por el equipo de red que tiene el bucle en la línea como si estuviera directamente conectada a él y la asociará al interfaz que tiene el bucle. Pero no sólo esta dirección MAC. Todas aquellas que estén generando tráfico broadcast o unicast dirigido al dispositivo con el bucle en la línea, si su trama llega a este nodo de acceso, serán vistas como directamente conectadas por dicho interfaz. Esta situación, en la que las MACs aparecen sobre el mismo interfaz, y en ocasiones posteriores como MACs remotas, puede ser supervisada por un mecanismo de control que señalice este movimiento de MACs y al llegar a un determinado umbral bloquee lógicamente el interfaz con el bucle.

El umbral del mecanismo de control se basaría en número de MACs aprendidas por un interfaz y que a su vez se empiezan a ver por otro, o en número de veces que se aprende y se modifica una MAC en un intervalo de tiempo.

No es siempre precisa esta manera de detectar bucles ya que puede existir otro interfaz en la misma VPN con mucha más cantidad de tráfico, y sea éste quien reciba las replicaciones de tramas de otro que realmente tenga puesto un bucle físico. Los fabricantes suelen garantizar en porcentaje de éxito la detección del interfaz con bucle, pero nunca al cien por cien.

La sensibilidad ante bucles físicos en estas redes es enorme. Aunque la clave del éxito resida en el aprendizaje de MACs sobre un dominio de difusión único y la transparencia a nivel 3, también supone un riesgo muy elevado la posibilidad de afectación al servicio ante bucles físicos. Una sede de un usuario puedo suponer la inoperatividad del resto.

Para el diseño de nuestra red VPLS jerárquica habilitaremos la funcionalidad de **detección de bucles,** tratando de ajustar los umbrales de movimientos de MACs al máximo

Especificación de diseño 5

3.1.4. MPLS (Multi Protocol Label Switching)

Una vez tenemos definido el acceso a la red, y descrita la finalidad del servicio, esto es, mantener aislado el tráfico de cada usuario o cliente en una red compartida y administrada de forma común, analizaremos el conjunto de protocolos y tecnología que hacen que esto sea posible, realizando además el tratamiento adecuado a cada tipo de tráfico, aplicando calidad de servicio, y consiguiendo tiempos de conmutación en cada salto de nodo prácticamente inmejorables.

Para esto, implementamos en el núcleo de la red MPLS.

MPLS (Multi Protocol Label Switching) es un grupo de trabajo específico del IETF (Internet Engineering Task Force) que trata sobre el encaminamiento, envío y conmutación de los flujos de tráficos a través de la red. Este protocolo se basa en la conmutación de paquetes en base a etiquetas de señalización previamente establecidas, que encapsulan el tráfico y evitan la consulta en la tabla de enrutamiento del router a la hora de reenviar un paquete recibido.

Los routers, para encaminar el tráfico por un interfaz u otro hacia su destino, examinan la dirección IP de destino, y consultan su tabla de rutas, alimentada por el protocolo interno que utilice. En esta búsqueda, se trata de localizar aquella entrada más específica para la dirección de destino, entendiendo por más específica aquella cuya máscara de red sea mayor y coincida con la subred de destino. En caso de no tener ninguna entrada que encaje con la dirección IP destino, de utilizará la ruta por defecto. La búsqueda, al basarse en la máscara de red más coincidente, independientemente del protocolo que la señalice, no finaliza hasta que no trata la tabla entera.

Existe un potencial factor de retardo en la conmutación de redes basadas exclusivamente en IP. El tamaño de la tabla de rutas puede ser muy grande, sobre todo si la red no se administra óptimamente, y la búsqueda del interfaz de salida para el paquete puede introducir mucho retardo en la comunicación final.

MPLS sustituye la tabla de rutas basada en IP por una tabla de reenvío basada en etiquetas. Un nodo corriendo MPLS intercambia etiquetas de señalización con los vecinos de forma que identifique el origen del tráfico y el destino en base a la etiqueta que encapsule el tráfico. Así, la

conmutación es extremadamente más rápida, ya que la búsqueda es precisa, sobre un valor de longitud determinado y entrada única en la tabla.

El intercambio de etiquetas se hace entre nodos MPLS, y forman parte un camino predeterminado para interconectar dos extremos remotos. La conmutación local basada en etiquetas reduce el tiempo de conmutación, ya que en lugar de encaminar el tráfico basándose en una dirección IP destino y la máscara de red más específica de la tabla de rutas, conmuta sobre una etiqueta cuyo interfaz de salida está asociado y forma parte de un camino MPLS ya establecido por toda la red. El hecho de tener longitudes fijas, posibilita la conmutación hardware, más rápida que si se enfocan en software.

Los siguientes elementos y conceptos relacionados entre sí conforman el protocolo:

3.1.4.1. FEC (Forwarding Equivalente Class)

Una FEC se define como el concepto que en una red MPLS hay que diferenciar y proporcionar un tratamiento diferente para la comunicación extremo a extremo del mismo, es decir, el determinante para asociar ese tráfico a una red privada virtual.

Realmente es un conjunto de paquetes que comparten unas mismas características para su transporte, así todos recibirán el mismo tratamiento en su camino hacia el destino. La asignación de un paquete a un determinado FEC se produce una vez el paquete entra en la red. Cada FEC puede representar unos requerimientos de servicio para un conjunto de paquetes o para una dirección fija.

Para el diseño de nuestra red VPLS jerárquica, entenderemos como una FEC al usuario o cliente completo

Especificación de diseño 6

3.1.4.2. LSR (Label Switch Router)

Los routers en una red MPLS se les denominan LSR en terminología propia del protocolo. Se hace también distinción entre los que se encuentran dentro de la nube y los que ejercen de acceso a los usuarios. Todos los LSRs utilizan protocolos de enrutamiento basados en IP, y en nuestro caso, como IGP utilizarán OSPF. Se podría construir utilizando otro protocolo de routing interno. En cambio, no todos los LSRs de una red MPLS tienen que utilizar el mismo protocolo de señalización de etiquetas, aunque lo recomendable es hacerlo de forma consistente.

Los elementos ubicados en la frontera de la red, es decir, los que reciben el tráfico de los clientes, asumirán ciertas funciones adicionales a los que se encuentran en el núcleo, como identificar la FEC y encapsular el tráfico hacia la red en MPLS. Añadirán las etiquetas correspondientes al inyectar tráfico a la red y extraerán las que reciban de la red al reenviar tráfico hacia usuarios. A estos elementos también se les denomina **LER (Label Edge Router).**

Los LSRs que instalados en el núcleo no realizan clasificación del tráfico en función de las características definidas en una FEC, sino exclusivamente en etiquetas previamente negociadas y establecidas en la tabla de reenvío con duplas interfaz de entrada y etiqueta frente a interfaz de salida y etiqueta.

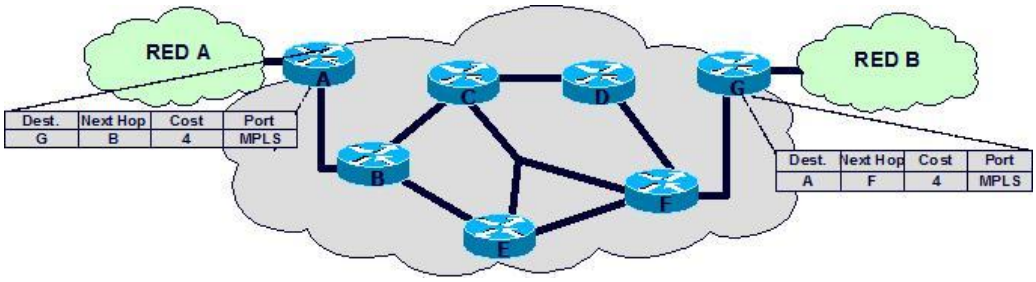

Imagen 9 - MPLS

Los caminos MPLS se señalizan desde un LER hacia un punto final definido o bien por el administrador de la red o bien por los protocolos de señalización elegidos.

3.1.4.3. LSP (Label Switched Path)

Los caminos MPLS a través de la red reciben el nombre de caminos etiquetados y conmutados. Los LSPs se crean a partir del intercambio de etiquetas entre nodos MPLS. Sobre los LSPs de MPLS fluirá el tráfico de los usuarios debidamente encapsulado, para que la conmutación en los nodos basada en las etiquetas del correspondiente FEC consiga que se establezca la comunicación final entre elementos remotos.

Los LSPs se pueden establecer de diferentes formas, y utilizando varios protocolos de señalización que se analizarán más adelante, pero básicamente se diferencian en la manera de crearse. Cada LSP puede constituirse de uno o varios caminos físicos correctamente señalizados, para contar con un camino principal y uno o varios secundarios. Estos últimos a su vez, pueden estar pre-señalizados de antemano o pueden señalizarse cuando sean necesitados.

Los paths de los LSP se definen de dos formas, o mediante una combinación de ambas. El administrador de la red puede definir manualmente los saltos que debe seguir el path hacia un determinado LER, o permitir que sea el IGP utilizado en la red quien indique en función de su tabla de rutas el camino óptimo. Ambas opciones pueden compartir path, y construir éste de forma **estricta** en algunos tramos, y de forma **loose** completando con la información de routing del protocolo interno. Los saltos estrictos, deben ser sobre interfaces directamente conectadas en cada salto MPLS.

En el caso de definir paths completamente estrictos, éste se verá afectado ante fallos simples en alguno de los saltos que utilice ya que no permite otros pasos. Se asegura que el tráfico sigue el camino deseado por el administrador, pero implica más trabajo manual y no deja que el IGP complete los saltos que puedan fallar con caminos alternativos.

Los LSPs son unidireccionales, por lo tanto es necesario un LSP para el tráfico de vuelta desde el otro extremo. Se recomienda que los paths definidos para ambos LSPs complementarios sean simétricos. Son necesarios siempre dos LSP entre dos nodos para establecer la comunicación y para completar la dupla de etiquetas en la tabla de reenvío. Los LSPs pueden crearse de dos formas diferentes, propiciando así la elección del protocolo de distribución de etiquetas adecuado. Los LSPs pueden ser solicitados desde un LER de acceso, hacia un extremo remoto

explícitamente. Es decir, desde un LER se solicita la señalización de un path hacia un nodo MPLS remoto. Cada nodo intermedio corre MPLS y es consciente que será un nodo de tránsito para ese LSP. El extremo final recibe la petición y genera el etiquetado correspondiente para este LSP, indicándole al nodo inmediatamente anterior la etiqueta con la que debe encapsular el tráfico que reenvíe sobre ese LSP. Así sucesivamente hasta que el nodo origen recibe la señalización del anterior y el LSP queda completamente señalizado. El mismo ejercicio debe realizarse invirtiendo los extremos para establecer el camino de vuelta. Esta forma de señalizar se denomina **bajo demanda,** y el protocolo recomendado es **RSVP.**

Imagen 10 - MPLS (2)

Otra forma de señalizarse los LSP es de forma automática o no solicitada, en la que al activar el protocolo de señalización correspondiente cada LSR distribuye etiquetas a sus vecinos para que éstos tengan ya la señalización necesaria para conmutar tráfico hacia ellos aunque no se hayan solicitado los LSPs explícitamente. Esta forma de señalización es soportada por el protocolo **LDP.** Este protocolo puede funcionar distribuyendo etiquetas de manera **explícita** o de forma **liberal.**

Para el diseño de nuestra red VPLS jerárquica, definiremos LSPs con caminos principales y secundarios, estos pre señalizados de antemano. Los path principales serán de tipo estricto, y los secundarios una combinación de tipo estricto y loose. Para la señalización de los LSPs utilizaremos el protocolo **RSVP**

Especificación de diseño 7

3.1.4.4. Túneles y etiquetas MPLS

Las etiquetas MPLS son de 32 bits, de los cuales, los 20 primeras son el identificativo de la misma, los 3 siguientes definen la prioridad del tráfico, el siguiente bit estará activo si se trata de la última etiqueta que se haya apilado, y los últimos 8 representan el contador de saltos, conocido como **TTL** (Time To Live)

El protocolo MPLS admite el apilamiento de etiquetas MPLS en las tramas que encapsula y reenvía. De esta forma, la red es capaz de conmutar tráfico basándose en la etiqueta exterior conservando las interiores. Normalmente se utilizan dos etiquetas, una exterior y otra interior que identifican y aíslan el tráfico de los clientes. La etiqueta exterior es la propia del LSP señalizado y utilizado para comunicar con los extremos remotos. Esa etiqueta es común para todos los usuarios ya que representa el salto negociado entre vecinos adyacentes. Los nodos no examinan las sucesivas etiquetas a no ser que detecten que son el extremo del túnel y sean conscientes de que ahí se acaba el path y deben examinar la siguiente etiqueta.

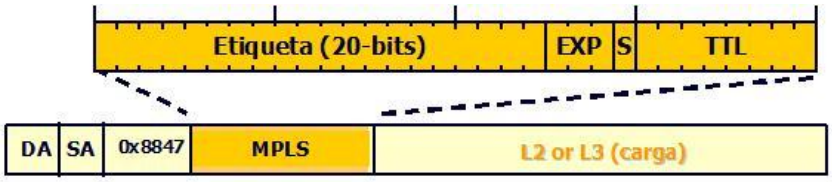

Imagen 11 - MPLS (3)

La etiqueta interior se asocia a la FEC definida en el LER de entrada en base a un criterio de entrada. Esta etiqueta también se negocia utilizando un protocolo de señalización siguiendo uno de los mecanismos definidos anteriormente, pero esta etiqueta no se intercambia en cada salto al ir encapsulada sobre otra etiqueta MPLS. Es decir, el tráfico recibido en un LER es clasificado conforma algún criterio definido que veremos más adelante, una vez identificada la FEC, se asocia la etiqueta correspondiente a esa FEC y al nodo destino, que no tiene porqué estar directamente conectado. Dicha trama, es a su vez encapsulada en otra trama MPLS para utilizar el LSP definido como túnel de red. La etiqueta asociada a la FEC, en nuestro caso al cliente, es encapsulada con la etiqueta del LSP establecido en la red para conectar al extremo remoto, pero conmutando de etiqueta salto a salto.

Todos los usuarios tendrán su FEC y criterio de clasificación definidos, y tendrán una tabla de reenvió de etiquetas definidas para conectar con otros nodos remotos, pero todos compartirán la tabla de etiquetas negociadas para el establecimiento del LSP. Se puede diferenciar por etiqueta de cliente, interior, y etiqueta de red, exterior.

En el siguiente gráfico se define una red MPLS mallada completamente con los túneles o LSPs establecidos. En segunda instancia se definen las adyacencias remotas entre LSRs para la comunicación extremo a extremo de cada FEC, o cliente.

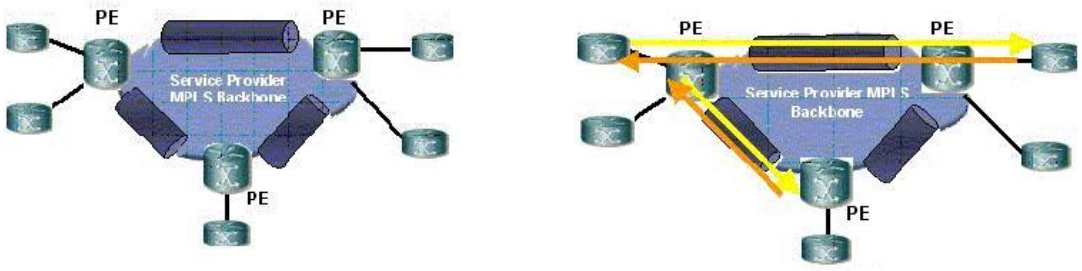

Imagen 12- MPLS (4)

Al tratarse de servicios VPLS de nivel 2, el criterio de clasificación a la entrada de la red, es decir, en los LER, se basará en la información de las tramas Ethernet que reciba del usuario, dirección física de destino, identificador de VLAN, etc. Si el servicio que se ofreciese fuese redes privadas virtuales de nivel 3, nos fijaríamos en las direcciones IP destino o información incluida en la cabecera IP.

Para el diseño de nuestra red VPLS jerárquica, mientras la señalización de los LSPs se ha definido mediante el protocolo RSVP, los peerings remotos para comunicar FECs y crear los túneles internos será utilizando el protocolo **LDP.** La modalidad del intercambio de etiqueta será bajo demanda configurada por el administrador de red. El criterio a aplicar en entrada para los LER serán **direcciones MAC destino.** En base a estas y a los interfaces por donde recibe el tráfico, determinará si se trata de un cliente u otro, y utilizará la tabla de etiquetas correspondiente.

Especificación de diseño 8

3.1.4.5. PHP (Penultimate Pop Hopping)

Esta funcionalidad del protocolo MPLS ahorra una etiqueta en el extremo final del túnel o LSP, evitando así una doble búsqueda en las tablas de reenvío del último LSR, es decir, un LER de salida. Como en este nodo, la etiqueta debe ser extraída obligatoriamente, delega en el penúltimo LSR del path esta acción, para que sólo llegue tráfico encapsulado con la etiqueta del cliente.

Los LSPs que se establecen bajo demanda, comienzan a señalizarse por el extremo final. Éste es conocedor de que ahí termina el túnel, por lo que negocia con el LSR inmediatamente anterior que no encapsule el tráfico ya se trata del último salto, y el LER de salida tendrá que examinar la siguiente etiqueta interior para determinar la salida de la del tráfico. De esta forma, ambos añaden en la entrada de reenvío una etiqueta que implícitamente representa una etiqueta nula o vacía, ya que no se encapsulará en el propio LSP.

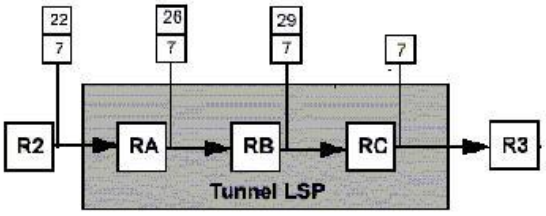

Imagen 13 - MPLS (5)

Contando con esta funcionalidad en la red, los LER de salida tendrán en su tabla de reenvío MPLS como etiquetas de entrada sólo los valores que representen implícitamente una etiqueta nula, porque se trata de los extremos finales del túnel, y los LSR anteriores habrán quitado la etiqueta exterior utilizada en el LSP

Para el diseño de nuestra red VPLS jerárquica, **no** será necesario utilizar la funcionalidad PHP

Especificación de diseño 9

3.1.4.6. RSVP (Resource reSerVation Protocol)

El **protocolo de reserva de recursos** (**RSVP**), descrito en el RFC 2205 es un protocolo de capa de transporte designado para reservar recursos a través de una red integrada de servicios, previo a MPLS. RSVP no es una aplicación de transporte, es más bien un protocolo de control de tráfico, como **ICMP**, o protocolos de encaminamiento. RSVP puede ser utilizado tanto por hosts como por routers para pedir o entregar niveles específicos de calidad de servicio para los flujos de datos de las aplicaciones. RSVP define como deben hacer las reservas las aplicaciones y como liberar los recursos reservados una vez que han terminado. Las operaciones RSVP generalmente dan como resultado una reserva de recursos en cada nodo a lo largo de un camino.

RSVP no es en sí mismo un protocolo de encaminamiento y fue diseñado para operar con los actuales y futuros protocolos de encaminamiento. Con este protocolo se reservan recursos para los flujos unidireccionales: un flujo de tráfico en una sola dirección desde el emisor a uno o más receptores. RSVP reserva recursos para un flujo que se identifica por la dirección de destino, el protocolo de identificación y, opcionalmente, el puerto de destino. En MPLS un flujo se define como un LSP.

La utilización de RSVP se vio motivada tras la aparición de MPLS. El modo de funcionamiento es **downstream-on-demand,** es decir, hay que configurar un LER para que genere una petición de reserva en un camino hacia un nodo MPLS remoto. Una solicitud de reserva se genera en el nodo de entrada a la red MPLS a través de un mensaje *RSVP Path Message.* Las etiquetas se irán distribuyendo dinámicamente a lo largo del camino, empezando desde el destino hasta acabar de nuevo en el nodo origen, a través de mensajes *RSVP Resv Message.*

El establecimiento de sesiones RSVP se hace atacando a direcciones de Interfaz en lugar de IPs de loopback ya que los LSP tienen muchos caminos, primarios y secundarios, y en un mismo nodo pueden coincidir varios LSPs.

3.1.4.6.1. RSVP-TE (Extensiones de Ingeniería de Tráfico)

Tanto en los mensajes de PATH como de RESV, se incluye cierta información determinante para el establecimiento de la señalización. Por ejemplo, la ruta seguida por el path se va almacenando en el objeto **Recorded Route Object,** de forma que un nodo puede evitar un bucle si en la lista de interfaces aparece uno propio, o por ejemplo, si el administrador de red lo conviene, decidir el camino completo desde el origen, forzando una ruta explícita por la red utilizando el **Explicit Route Object.**

La necesidad de forzar el camino del LSP, o la optimización en la utilización de los troncales de red, para reservar ancho de banda para cierto tráfico, o la inclusión de mecanismos de reenrutado para que los LSPs vuelvan a activar un camino más óptimo que el que esté utilizando sin afectar al servicio, fueron, entre otras, motivaciones para actualizar el protocolo con extensiones de Ingeniería de Tráfico.

Dado que los flujos de datos a través de LSPs se identifican completamente desde la etiqueta de origen aplicada a la entrada de la red, estos paths pueden tratarse como túneles. La aplicación de la Ingeniería de tráfico sobre estos túneles es clave para la explotación del protocolo MPLS. Las extensiones de RSVP se definen en el RFC 3209.

Además de optimizar el uso de recursos de la red, indicar una ruta explícita incluyendo o no reserva de ancho de banda, evitar congestión en la red o bucles, también se pueden definir varios LSPs entre dos nodos y aplicando políticas de tráfico locales distribuirlo entre ellos.

Todas estas mejoras en los LSPs se negocian durante la fase de establecimiento del path, incluyendo objetos en los mensajes de PATH y RESV que se intercambian los nodos MPLS desde el origen hasta el destino y viceversa.

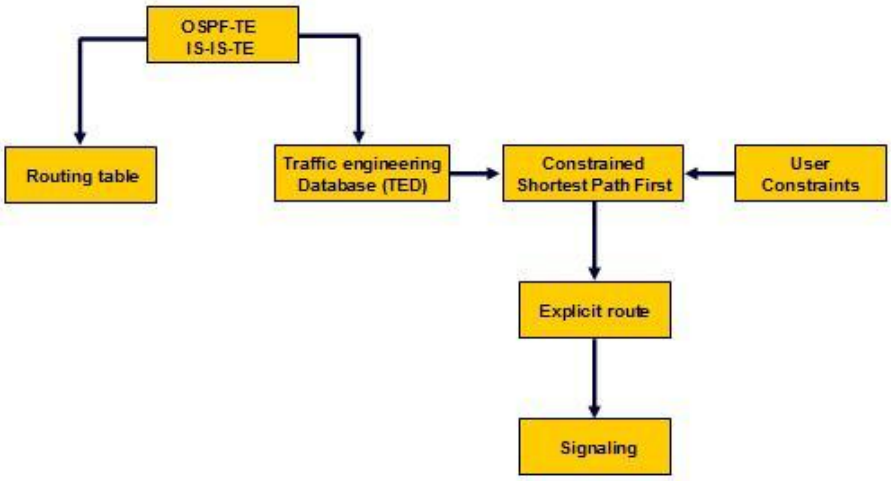

Imagen 14 - RSVP

En función de las optimizaciones que se quieran introducir en el proceso de señalización del LSP se tendrá que seguir el algoritmo expuesto arriba para la creación del mismo. Este proceso operacional lo realiza el nodo MPLS que actúe como LER de entrada a la red. Almacenará información de enrutamiento de los protocolos internos que esté utilizando como IGPs, en este caso OSPF, y junto las condiciones de usuario, reserva de ancho de bando, múltiples LSPs, o lo más habitual, una ruta estricta (incluyendo todos los saltos o la mayoría y dejando el resto tipo loose para que los determine el IGP utilizado) se generará un objeto ERO que represente al path.

Este RFC con las extensiones a RSVP, incluyó también la aplicación de timers en las sesiones entre vecinos a través de mensajes HELLO y sus correspondientes respuestas afirmativas (ACK). Ajustando los timers de los diferentes protocolos involucrados en el servicio VPLS se pueden detectar fallos en los enlaces rápidamente y forzar el tráfico por alternativas de backup.

Estos timers facilitan el control de las sesiones RSVP, ya que al ir sobre UDP no es fiable en cuanto a conexión extremo a extremo.

Las extensiones de Ingeniería de Tráfico hay que habilitarlas en los equipos para que éstos implementen lo recogido en el RFC 3209.

Para el diseño de nuestra red VPLS jerárquica, utilizaremos extensiones de tráfico sobre el protocolo **RSVP-TE,** para forzar los paths principales de los LSP por el camino que mejor creamos conveniente. No reservaremos ancho de banda para aplicaciones, los recursos de red serán compartidos por todos y serán las políticas de QoS definidas posteriormente las que regulen el tipo de tráfico en caso de congestión. Los timers de RSVP los fijaremos a **1 segundo.**

Especificación del diseño 1

3.1.4.6.2. Mecanismos de protección de LSPs

Los túneles RSVP definidos para el transporte de tráfico deben ser protegidos por otros caminos secundarios, de manera que ante fallos físicos en el camino principal, exista uno o varios caminos alternativos y el transporte MPLS no se pierda.

Existen varios métodos para proteger los LSPs, consiguiendo conmutaciones entre el camino principal y el secundario del orden de decenas de milisegundos. Esta disponibilidad, en milisegundos, es lo más rápido en cuanto a conmutaciones de rutas o caminos para tráfico en conmutadores de datos.

En cualquiera de las modalidades, los LSPs siempre conmutan al camino principal una vez que éste se haya restablecido correctamente. Y como implementación del protocolo MPLS, lo hacen de manera no disruptiva, lo que se conoce como ***"make before break".***

3.1.4.6.2.1. LSP - protección de camino

Los LSPs pueden definirse con uno o varios caminos secundarios. Éstos, como se explica en el apartado *LSP (Label Switched Path),* se pueden definir manual o dinámicamente, y además, se pueden establecerse de antemano y estar señalizados previamente a la conmutación de camino.

De esta forma, los LSPs cuentan con alta disponibilidad extremo a extremo, y ante indisponibilidad del enlace principal, se conmutará al secundario correspondiente.

3.1.4.6.2.2. MPLS Fast Reroute (MPLS FRR)

La funcionalidad avanzada de Fast Reroute para MPLS supone un complejo método de redundancia para los LSPs, pero altamente efectivo.

Los LER o equipos que originan los LSPs, incluyen esta opción en la petición de establecimiento del LSP. Si los nodos que reciben el mensaje no soportan dicha funcionalidad, ignoran este aspecto pero siguen señalizando el LSP principal.

Básicamente se trata de pre-señalizar un desvío para el LSP principal, de forma que ante fallos físicos en el camino, ya exista un desvío sólo para evitar ese problema. Existen dos modos de definición de los caminos de backup en MPLS FRR, protección de nodo y protección de enlace.

La protección de nodo implica que cada uno de los nodos por los que se establece el LSP, deben generar un desvío a sí mismos, apoyándose en los equipos que tenga alrededor, para así disponer de una alternativa señalizada previamente al posible fallo que le afecte.

La protección de enlace, se utiliza para asegurar un desvío al LSP si el enlace protegido falla. La siguiente imagen esquematiza ambas opciones de MPLS FRR.

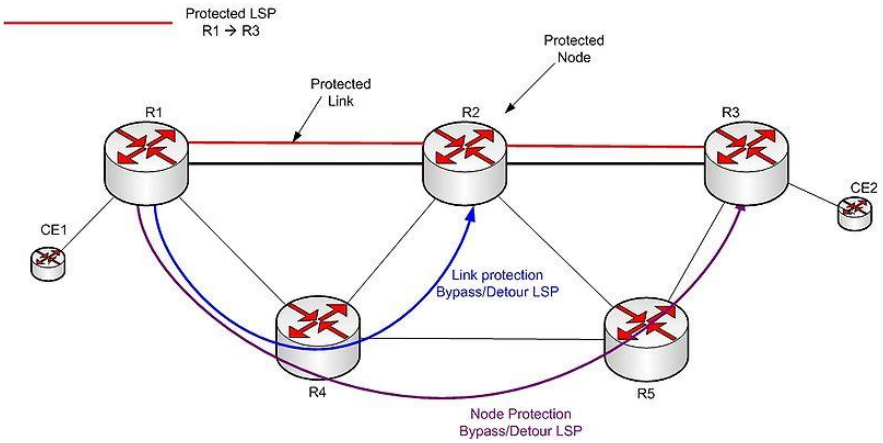

Imagen 15 - RSVP (2)

En el presente proyecto no se han incluido pruebas de mecanismos de FRR. Las versiones de software utilizadas y la configuración utilizada en los servicios VPLS han detectado problemas en el software con esta funcionalidad.

3.1.4.7. LDP (Label Distribution Protocol)

LDP es otro protocolo de señalización de etiquetas entre nodo MPLS, definido para el establecimiento de túneles entre ellos. Los routers que levanten una sesión LDP serán considerados **peers.** Este protocolo puede ser utilizado para construir y mantener bases de datos de etiquetas para conmutar tráfico en entornos MPLS.

LDP puede utilizarse para distribuir la etiqueta exterior, propia del LSP, o para hacerlo con la etiqueta interior, el identificador de circuito virtual que hace referencia al cliente final. Para la distribución de etiquetas internas, se levantarán sesiones dirigidas al nodo remoto, aunque no esté directamente conectado (Targeted LDP).

La descripción del protocolo, de su funcionamiento y características no influye en el objetivo de su operabilidad, ya que la etiqueta la negocia de la misma forma tanto para definir la etiqueta exterior como la interior.

En el caso de utilizarse para señalizar el LSP, tendríamos la doble encapsulación de MPLS basada en el mismo protocolo dos veces, es decir, LDP sobre LDP.

En el caso de utilizar RSVP-TE para la señalización de los LSP, tendremos LDP sobre RSVP como doble encapsulación en una red MPLS compartida por varios clientes cuyo tráfico debe estar aislado.

Como componente del servicio VPLS, suele utilizarse para fijar las etiquetas asociadas a las FEC de entrada, con la idea de formar VPNs de nivel 2 a través de túneles MPLS señalizados con RSVP, es decir, un túnel dentro de otro.

En este diseño utilizaremos LDP como protocolo de señalización de etiquetas extremo a extremo. La infraestructura MPLS será controlada con RSVP-TE para definir los LSPs, y la relación de vecindad origen – destino se mantendrá con sesiones LDP que directamente mapeen el tráfico de entrada de cada cliente, a su respectiva etiqueta LDP interna, y lo asocien a un LSP señalizado mediante RSVP que comunique con el mismo extremo remoto.

Las sesiones LDP son fiables puesto que utilizan el protocolo TCP para el establecimiento d las mismas, pero existen dos formas de iniciarlas. El funcionamiento básico y por defecto es **no solicitado** (downstream unsolicited), y de esta forma un router MPLS trata de descubrir vecinos LDP aunque no haya una FEC definida requiriendo un LSP para un flujo de datos. El LSR generará tramas multicast UDP para localizar más LSR en su red. Una vez localizados intentará establecer una sesión TCP para iniciar el intercambio de etiquetas aunque éstas no hayan sido solicitadas explícitamente.

Además, LDP contempla dos formas de almacenar dichas etiquetas negociadas con los LSR que corren LDP de su red. Puede almacenar todas las, de forma **liberal,** o sólo aquellas que estén activas, de manera **conservativa.** La primera es más rápida ante fallos y por lo tanto necesidad de convergencia al tener las etiquetas almacenadas en memoria, pero obviamente requiere más recursos

El modo de funcionamiento alternativo es similar a RSVP-TE, y a través de configuración, el equipo trata de levantar sesiones LDP sobre TCP con aquellos extremos indicados explícitamente,

es decir, **bajo demanda**. Para dichas sesiones se utilizan direcciones IP de loopback, y por lo tanto se utilizará la información correspondiente al IGP, en este caso OSPF, para alcanzar al vecino con el cual queremos establecer el *peering.*

La etiqueta se negocia extremo a extremo por medio de sesiones TCP, por lo que en estos casos no se está pretendiendo realizar conmutación de etiquetas en los saltos intermedios, es decir, sólo habrá un salto, no se establece un LSP salto a salto por cada LSR ya que ese propósito no desempañará RSVP-TE en este diseño. LDP va a negociar etiquetas extremo a extremo para definir el identificativo del circuito virtual.

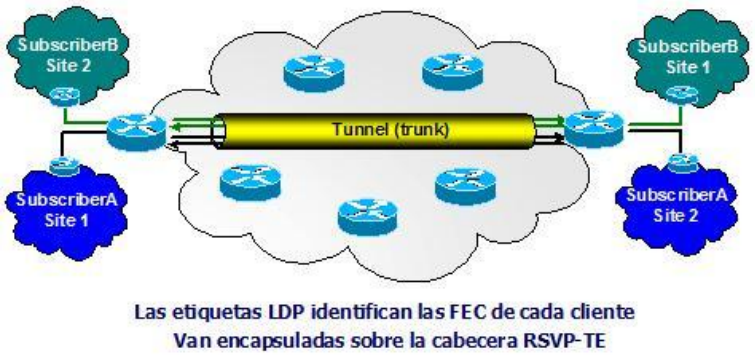

Imagen 16 - LDP

3.1.4.8. Enlace físico, Túnel MPLS, y Circuito Virtual (VC)

La red se sostiene sobre **enlaces físicos** los cuales deben soportar todo el tráfico de los usuarios. Es importante conocer de antemano la calidad y capacidad de los troncales que se van a disponer para decantarse por un diseño de red o por otro. Además de la calidad de servicio que se pretenda ofrecer, y su funcionamiento cuando los troncales entren en congestión, hay que prever el crecimiento que se seguirá a lo largo del tiempo, para que el diseño inicial cubra el mayor período posible sin necesidad de aplicar grandes cambios en la red.

Los enlaces físicos deben estar dimensionados apropiadamente, buscando un compromiso entre tráfico cursado de media, y tráfico máximo esperado en situaciones pico. No tiene sentido sobredimensionar la red con troncales poco aprovechados, cuando el patrón estadístico de la gran mayoría de redes de comunicaciones es bastante estable. Probabilísticamente no todos los usuarios transmiten a la vez, de ahí el negocio de las telecomunicaciones, pero tampoco hay que trabajar con troncales con ancho de banda insuficiente para ciertos intervalos de tiempo.

También es recomendable que dicha infraestructura de transmisión tenga la correcta redundancia física a nivel de diseño, ya que si los troncales comparten hardware posiblemente afecten a la plataforma de datos que soporten por encima. Es necesario que los enlaces físicos se establezcan utilizando hardware de transmisión diferente. Normalmente un nodo de conmutación MPLS al menos dispone de dos enlaces físicos o troncales de red para conectarse con otros nodos de la red. Sobre dichos troncales se definirán los LSPs de la nube MPLS, y seguramente un enlace soporte el path secundario del otro y viceversa, por lo que si los enlaces físicos se montan sobre el mismo hardware no se consigue la disponibilidad lógica de los LSPs.

En todo diseño de red se debe evitar los puntos únicos de fallo a nivel físico y a nivel lógico.

Una vez se tienen implementada la conectividad física, se sube de nivel en la pila de protocolos del servicio VPLS y se definen los enlaces lógicos que harán posible la conmutación de paquetes sobre la red.

El núcleo de la red está basado en MPLS, por lo tanto es necesario crear **LSPs** entre los distintos nodos que formen la red. Los LSP son **unidireccionales,** por lo que es necesario definir una pareja de LSPs en sentidos opuestos para conseguir que el tráfico sea bidireccional entre los dos extremos.

Además, los LSPs se definen entre dos nodos, por lo que dentro de una red MPLS es necesario completar un **full-mesh** de LSPs entre todos los nodos si se quieren integrar a todos en la red MPLS y si se quiere ofrecer conectividad con toda la red. Esto implica que los LSRs que sostienen la red sean capaces de tener separado el plano de control y señalización frente al plano de conmutación (tráfico real de clientes) y evitando que el exceso de carga en uno de ellos afecte al otro. Es decir, que deben instalarse equipos lo suficientemente potentes para soportar una elevada carga de señalización teniendo en cuenta la obligación de mantener un full-mesh de LSPs, y que dicha utilización de CPU no afecta a la pura conmutación que debe mantener para el tráfico real de datos, guiado por dicha señalización.

Por tanto, en una red MPLS simple, tenemos un número de LSPs unidireccionales igual a:

$$Nº\ LSPs = N \times (N - 1)$$

, siendo N el número de nodos MPLS existentes en la red a los que hay que dar conectividad total.

Con esta cifra, hay que plantearse inicialmente el tamaño de la red y la ambición de la misma plataforma, ya que un exceso de PEs MPLS puede suponer un problema de escalabilidad en cuanto a número carga de señalización en los equipos. Es de suponer que si se asume el diseño de una red MPLS plana, el problema del crecimiento está considerado.

Los LSPs proporcionan la conectividad entre PEs, de forma que los nodos de tránsito no tengan que realizar búsquedas basadas en direccionamiento IP en sus tablas de rutas normales, y conmuten el tráfico en base a etiquetas preestablecidas por los protocolos de señalización de los LSP, en este caso, RSVP-TE.

El tráfico que entra en la red se encapsula en MPLS para disminuir el retardo de conmutación. Pero no sólo se añade al tráfico la etiqueta RSVP-TE, como se describe anteriormente, las redes MPLS utilizan doble encapsulado para aislar el tráfico de cada cliente y proporcionar servicios VPNs. La etiqueta asociada al cliente viaja encapsulada en la etiqueta del LSP, y no varía en todo el túnel. Sólo se intercambia la etiqueta RSVP-TE de señalización del LSP, común para todos los clientes. La etiqueta asociada a la FEC de cada cliente permanece constante a lo largo del túnel MPLS, y sólo tiene sentido examinarla en el nodo LER de salida de la red, ya que el peering LDP establecido para tal negociación de etiquetas tiene lugar entre un LER de entrada y un LER de salida, es decir, dos PEs.

Dicha segunda etiqueta interna, sirve para identificar el **circuito virtual**, el cual transportará el flujo de datos del cliente definido para su VPN.

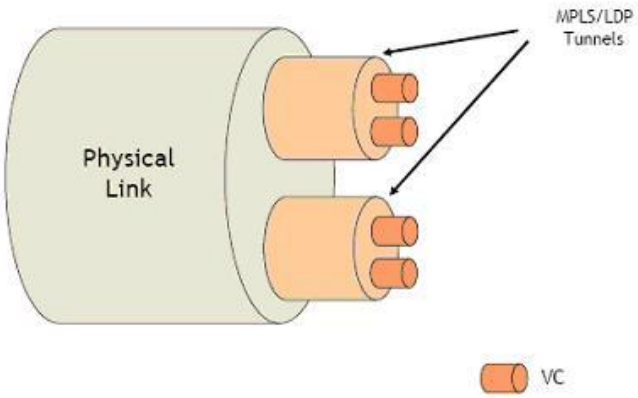

Imagen 17 - LDP (2)

Los circuitos virtuales se mantienen como se ha indicado sobre peerings LDP entre PEs remotos. Para el establecimiento de dicha adyacencia, se intercambia una serie de parámetros en la fase de negociación, como el **VC_ID,** identificador de circuito virtual, el cual debe ser el mismo en los dos extremos. De esta forma, se estable un peering unívoco por cada conexión de cliente entre dos PEs dados.

Resumiendo, se trata de una doble encapsulación MPLS sobre troncales físicos dedicados para las conexiones de nodos MPLS. En un primer lugar se implementa conectividad física y lógica cubriendo las capas 1 y 2 del modelo TCP/IP, en un segundo lugar se cubren las capas 3 y 4 con un protocolo de señalización de red MPLS, y finalmente se cubren las posteriores con un segundo protocolo de señalización sobre conexiones TCP.

3.1.4.9. Consideraciones sobre MTU (Maximun Transfer Unit)

La **unidad máxima de transferencia** (*MTU)* es un término que expresa el tamaño en **bytes** de la unidad de datos más grande que puede enviarse usando una determinada tecnología. Es muy importante conocer este valor para las diferentes tecnologías a utilizar ya que un desajuste en dicho valor puede causar la inoperatividad del servicio.

Es posible que en el origen, se conozca este dato y sea posible implementar fragmentación, de forma que el tráfico transmitido no supere el valor máximo fijado por la tecnología subyacente. Generalmente, se utilizan medios físicos basado en Ethernet, cuya MTU era inicialmente de 1500 bytes. Hoy día, se utilizan tramas de tamaño superior, denominadas **jumbo frames,** con valor máximo de hasta 9000 bytes.

4. Red VPLS Jerárquica

4.1. Descripción H-VPLS

Debido a las características de VPLS, las redes basadas en esta tecnología tienen escalabilidad limitada. El núcleo de la red se basa en un red IP/MPLS, por lo que necesitan un mallado completo de LSPs entre todos los PEs de la red. Esto resta posibilidades de crecimiento, tanto por la complejidad que adquiere una red de un tamaño considerable y un mallado completo de LSPs, como por la limitación de los equipos a la hora de manejar el tráfico de control de un gran número de LSPs de origen, destino, como de tránsito.

La solución a este problema radica en el principio fundamental de organización de redes cuya dimensión es mediana o grande, y viene a ser, dividir la red en regiones, áreas o subredes. Como con muchos protocolos, se ha definido para solventar el problema de mallado completo de VPLS la arquitectura de VPLS **jerárquico**. Dicha arquitectura, como su propio nombre indica, sugiere varios niveles de jerarquía para los diferentes equipos de la red. Definiendo una región central o **Core,** se conectarán las regiones inmediatamente inferiores en la jerarquía a dicha región sólo con uno o dos enlaces. Todas las regiones deberán pivotar sobre el Core para vascular tráfico de una a otra. Esta separación jerárquica reduce considerablemente la necesidad de mallar la red, ya que sólo recae la obligatoriedad del full-mesh de LSPs sobre cada región.

Las regiones actuarán como redes MPLS independientes, con LSPs que formen un mallado completo dentro de la misma región, y con LSPs que conecten con interfaces de equipos del Core, de forma que se obtenga salida hacia el Core y otras regiones para el tráfico de los clientes.

Planificando correctamente el crecimiento de la red por regiones, tanto en número de elementos por región, a los que hay que mallar completamente, como en número de regiones a definir, generalmente en función de distribuciones geográficas, se puede conseguir una optimización en el plano de control de los equipos, y por consiguiente del rendimiento de los mismos y de la propia red.

4.2. Mallado completo dentro de la región VPLS – enlaces mesh

Como se ha descrito anteriormente, se diseñará la red en base a regiones VPLS interconectadas entre sí por medio de una región Core.

Inicialmente no se deberían crear conexiones de clientes o accesos a la red en los equipos de Core. Técnicamente sería posible, pero administrativamente el resultado no es tan limpio como si se mantiene cierta jerarquía también con los accesos de los clientes. No se debe permitir que un equipo de Core, realice funciones de PE.

Las regiones VPLS, independientemente de su nivel jerárquico, estarán compuestas por LSR que señalizarán los LSPs que los interconectan. Dichos equipos implementan la arquitectura VPLS completamente, desde el nivel físico hasta los peerings LDP entre PEs sobre cada VPN de cliente. Además, implementan una funcionalidad para evitar el reenvío de tráfico entre nodos de la misma región, denominado **split-horizon.**

Esta regla, elimina en gran medida la posibilidad de bucles en una red de conmutación de tráfico. La idea básica de esta regla es que un equipo no reenvía tráfico por aquel interfaz por el que lo recibió, evitando así devolver el mismo tráfico al origen del mismo. La regla de split-horizon sobre VPLS requiere un análisis un poco más profundo.

Los LSPs se definen de forma estándar, señalizados según este diseño mediante el protocolo RSVP-TE. Los LSPs llevarán todo el tráfico entrante y saliente de la red, independientemente del cliente que sea, es decir, serán troncales dentro de la región. En cambio, las sesiones LDP establecidas entre los PEs que se definen extremo a extremo, entre direcciones de loopback, se definen como **mesh.**

Una sesión LDP mesh entre dos PEs, aplicará la regla de split-horizon sobre el servicio, es decir, sobre el túnel de cliente definido entre dos PEs. La aplicación de esta regla sobre el túnel de cliente o circuito virtual, implica que el tráfico que recibe un PE por dicho circuito no se reenviará de vuelta por el mismo. La regla de split-horizon, aplicada en ámbitos VPLS, no sólo incluye al propio túnel o SDP, sino a todos aquellos definidos como mesh dentro de la misma VPLS.

Los servicios VPLS son VPNs de nivel 2. Dichas VPNs se conectan entre sí por medio de SDPs o sesiones LDP entre PEs tunelizadas sobre LSPs RSVP-TE. Todos los SDPs de la misma región VPLS se definen como **mesh,** de forma que el tráfico que reciben los PEs, no se reenvía utilizando aquellos SDPs tipo mesh dentro de la VPLS.

De esta forma, el tráfico broadcast de un cliente, por ejemplo ARP, entra en la red por medio de un PE. Éste, al ser broadcast o al no haber cerrado el ciclo de aprendizaje de MACs, reenvía la trama por toda la VPLS de dicho cliente, es decir, lo reenvía por todos los interfaces de cliente que pudiera haber en ese PE y por todos los SDPs que conectan con el resto de PEs de la región. Cuando la trama llega al PE remoto, éste tampoco tiene un interfaz conocido para reenviarla puesto que la dirección MAC destino es broadcast, por lo tanto debería reenviarla por todos los interfaces de clientes y por todos los SDPs, pero éstos últimos son de tipo mesh también, por lo que la trama no vuelve a ser enviada a otros PEs.

Con esta aplicación de la regla split-horizon sobre el full-mesh de cada región, se evita que a un mismo nodo le llegue la misma trama por dos caminos. En el siguiente ejemplo se detalla el reenvío por inundación en la VPLS de un usuario de una trama broadcast, sobre la misma región VPLS

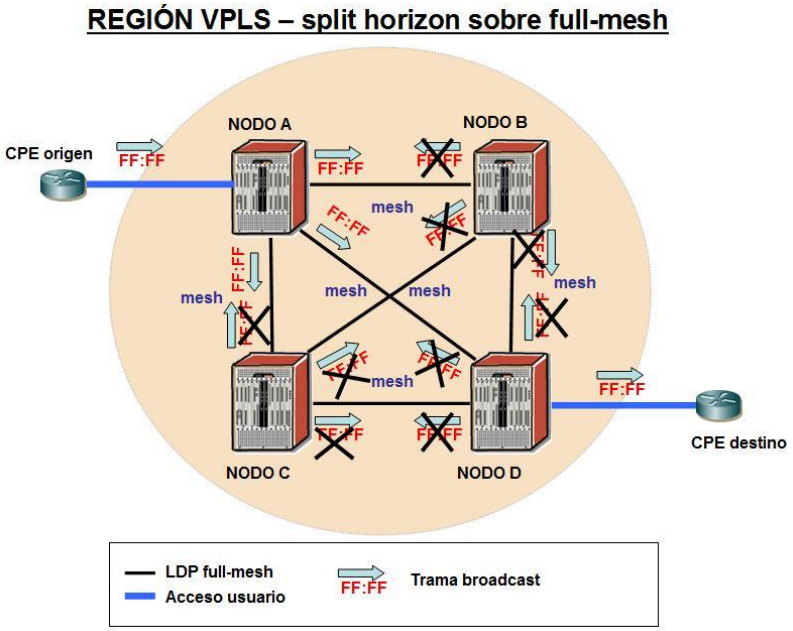

Imagen 18 - H-VPLS

En el diagrama, se puede observar el tratamiento de una trama broadcast de nivel 2 sobre una región VPLS con full-mesh de sesiones LDP entre los equipos que la componen. Un router origina la trama broadcast, por ejemplo ARP para descubrir la dirección IP destino de otro equipo conectado a la misma red. La trama llega al nodo A, y como la dirección MAC destino es broadcast, inunda la red enviando la misma trama por todos los interfaces de acceso y de red asociados a la VPLS del cliente, esto es, por los enlaces contra los nodos B, C y D definidos de tipo **mesh.** Los nodos B y C, reciben la trama y la tienen que reenviar puesto que se trata de una broadcast. No tienen interfaces de acceso de este cliente, por lo que sólo podrían reenviarlo por los interfaces de red, pero al haber recibido la trama por un enlace mesh, no pueden reenviarlo por otro enlace de tipo mesh, por lo que los nodos B y C no vuelven a reenviar la trama. El nodo D la recibe por un interfaz tipo mesh. Automáticamente se descarta la posibilidad de reenviarlo por la red ya que sólo tiene definidos enlaces mesh. Sólo puede enviar la trama por enlaces de acceso y así lo hace, haciendo que la trama llegue al router del cliente que se conecta a la red por el nodo D. Se ha evitado la posibilidad de bucles en la red, y se ha conseguido reenviar una trama broadcast sobre un medio definido como multipunto.

4.3. Conexión de una región VPLS con el Core – enlaces spoke

La conexión entre dos equipos conectados a la misma región de la red VPLS se explica en el apartado anterior. Se evita posibilidad de bucles lógicos aplicando la regla de split horizon sobre los peering LDPs para la VPLS definiéndolos como mesh. Pero es necesario proporcionar conectividad con la región Core, y por ende con las demás regiones VPLS a las que el cliente tenga algún equipo conectado.

Para proporcionar conectividad a una región con el Core, se define la asociación tipo **spoke.** Los SDPs o peering LDP se definen como spoke, y adoptan el comportamiento habitual de un switch de capa 2, y aplican la regla de split-horizon sólo sobre el enlace por el que han recibido el tráfico. Es decir, un equipo que recibe tramas por un enlace spoke, las puede reenviar por aquellos interfaces de acceso del cliente, y por todos los enlaces de red menos por el que lo recibió.

Se definirán como tipo spoke aquellos enlaces que soporten conectividad con el Core. De esta forma, todo el tráfico cuyo destino es otra región, será reenviado a través de la conexión spoke de cada región, puesto que la regla de split-horizon no se aplica sobre dicho enlace, ya que no es de tipo mesh. Asimismo, el tráfico a recibir por una región desde el Core, puede reenviarse por todos los enlaces tipo mesh de dicho nodo, y serán los PEs remotos quienes apliquen la regla de split-horizon sobre sus propios enlaces mesh.

En el siguiente esquema se reproduce la situación anterior, una trama broadcast de un cliente, cuyo objetivo es alcanzar la sede conectada en otra región remota. Dentro de la propia región VPLS a la que el equipo de cliente está conectado, se aplica la regla de split-horizon sobre los enlaces LDP tipo mesh, evitando el bucle lógico. Los nodos C y D, descartarán la trama puesto que no tienen salida para ella. En cambio, el nodo B, que recibe la trama por un enlace mesh, puede reenviarla por otro tipo spoke que tiene configurado contra la región Core.

De esta forma, la trama consigue llegar al Core, ya que la conectividad sobre enlaces spoke no está limitada. El Core se comporta como una región VPLS normal, con la excepción de que no tendrá interfaces de acceso configuradas. También implementa un full-mesh en cuanto a peerings LDP, por lo que la conectividad con todos los elementos de la región está asegurada y libre de bucles. La trama broadcast llegará a todos los elementos del Core, y sólo aquellos que tengan definidos enlaces spoke hacia otras regiones VPLS, podrían reenviar de nuevo la trama.

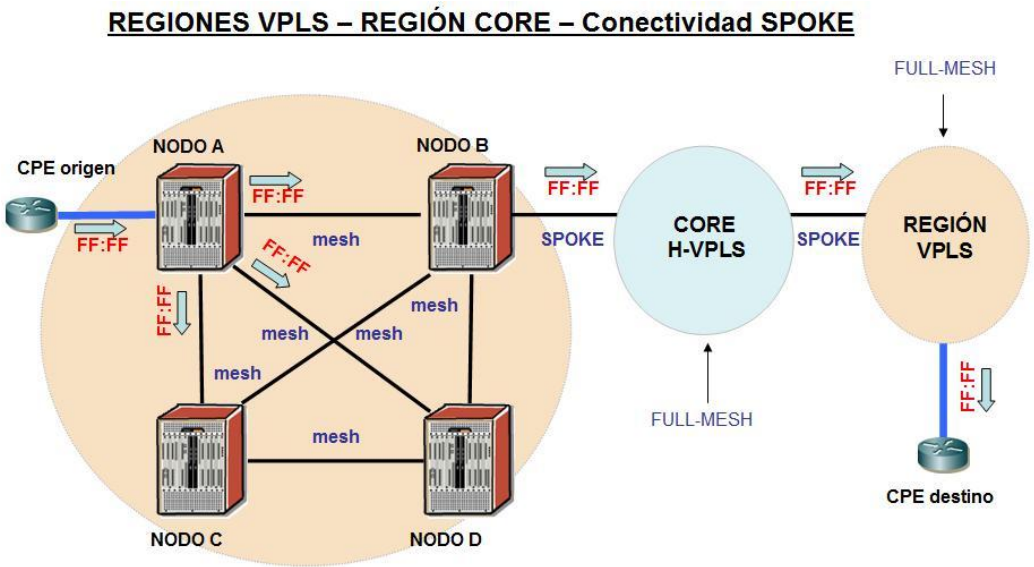

Imagen 19 - H-VPLS (2)

Al llegar la trama a la región remota, el tratamiento será el mismo y el PE que tenga la conexión spoke con el Core, inundará la región con la trama sobre todos sus enlaces tipo mesh contra el resto de PEs de la región. Finalmente, la trama será reenviada al exterior de la red por aquellos interfaces de acceso del cliente, consiguiendo transmitir la trama broadcast hacia un extremo final.

4.4. Conexión de una región VPLS con el Core redundada – RSTP

Resulta lógico que la conectividad a implementar entre una región y el Core, no sea soportada por un único enlace, y sean dos equipos los que lo proporcionen. Ante caída de uno de los enlaces o de los equipos, la otra conexión soportaría todo el tráfico hacia otras regiones.

Por definición del servicio VPLS jerárquico, estas dos conexiones han de ser definidas como **spoke.** Con lo cual, el flooding de tramas que ocurra en una región, tiene dos salidas tipo spoke hacia el Core. Como se ha explicado en los apartados anteriores, la región es full-mesh y evita toda opción de bucle dentro de la misma

La conexión al Core es por tanto doble, y de tipo spoke. El problema de posibilidad de bucle se traslada ahora a la conectividad lógica entre una región y el Core. Existen dos enlaces tipo spoke por los que el tráfico es reenviado al Core. En un equipo frontera del Core, al tráfico se le aplica la misma regla de split-horizon. Al recibirlo de una región por medio de un enlace spoke, el equipo puede reenviar el tráfico broadcast o desconocido por enlaces de red tipo mesh. Así, llegará a todos los equipos del Core puesto que tienen implementado un full-mesh entre ellos. Al llegar la trama a los demás equipos de Core, realizan la misma comprobación de la regla split-horizon. Las tramas se reciben por enlaces tipo mesh, con lo cual no pueden volverse a reenviar hacia equipos Core puesto que los unen enlaces de ese tipo, pero el segundo equipo frontera del Core soporta la segunda conexión tipo spoke, por lo que por ese enlace sí podría reenviar al trama broadcast. De esta manera, nos encontramos a un equipo del Core inyectando la misma trama sobre la región de origen. Al llegar a la región VPLS de origen de nuevo, se hará flooding por la misma devolviendo la trama al equipo origen generando así un bucle de nivel 2 que afectaría gravemente al servicio.

Esta situación no es deseable, por lo que es necesario implementar algún mecanismo de control lógico sobre la segunda conexión de la región VPLS y el Core, de forma que quede inhabilitada para el servicio de forma lógica y sólo esté activa en caso de fallo en la conexión principal.

En el siguiente esquema se representa gráficamente la posibilidad de bucle entre una región con doble conexión al Core

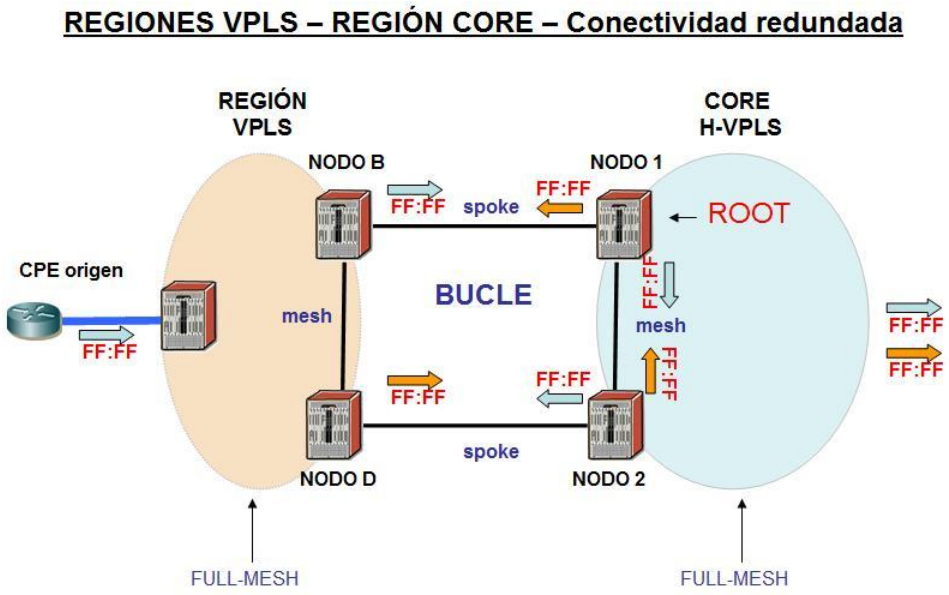

Imagen 20 - H-VPLS (3)

Dentro de la región se utilizan enlaces tipo mesh que evitan el reenvió de tramas entre PEs, pero la conexión doble tipo spoke hacia el Core provoca un comportamiento no deseado en el servicio. La aparición del bucle lógico descarta esta opción sin mecanismos de control adicionales sobre la conexión de redundancia. Es necesario deshacer lógicamente el bucle físico construido para dotar de redundancia a la conexión de las regiones con el Core VPLS.

4.4.1. Rapid Spanning Tree Protocol sobre conexiones redundadas al Core

En redes de nivel 2, el protocolo más utilizado para deshacer bucles físicos de forma lógica es Spanning Tree Protocol. STP es un protocolo de nivel de enlace de datos, nivel 2 en la capa OSI o TCP/IP, y está basado en un algoritmo de detección de bucles diseñado por Radia Perlman.

Su función es la de gestionar la presencia de bucles en topologías de red debido a la existencia de enlaces redundantes (necesarios en muchos casos para garantizar la disponibilidad de las conexiones, como por ejemplo, la conexión redundante de regiones VPLS contra una región Core). El protocolo permite a los dispositivos de interconexión activar o desactivar automáticamente los enlaces de conexión, de forma que se garantice que la topología está libre de bucles.

Los bucles infinitos ocurren cuando hay rutas alternativas hacia una misma máquina o segmento de red de destino. Estas rutas alternativas son necesarias para proporcionar redundancia, ofreciendo una mayor fiabilidad. Si existen varios enlaces, en el caso que uno falle, otro enlace puede seguir soportando el tráfico de la red. Los problemas aparecen cuando utilizamos dispositivos de interconexión de nivel de enlace, como switches o bridges. En el caso de una red VPLS, al implementar VPNs de nivel 2, los equipos de red se comportan del mismo modo que switches.

Cuando hay bucles en la topología de red, los dispositivos de interconexión de nivel de enlace reenvían indefinidamente las tramas broadcast o multicast, al no existir ningún campo TTL (Tiempo de Vida) en la Capa 2, tal y como ocurre en la Capa 3. Se consume entonces una gran cantidad de ancho de banda, y en muchos casos la red queda inutilizada. Un router, por el contrario, si podría evitar este tipo de reenvíos indefinidos ya que examina el tráfico en la cabecera IP, restando así un salto en el contador de vida del paquete. La solución consiste en permitir la existencia de enlaces físicos redundantes, pero creando una topología lógica libre de bucles superpuesta.

Este algoritmo cambia una red física con forma de malla, en la que existen bucles, por una red lógica en árbol en la que no existe ningún bucle, por medio de los paquetes propios del protocolo que se intercambia los equipos de red, BPDUs. (Bridge Protocol Data Units)

La idea reside en seleccionar, bien manualmente o bien dejando hacer al protocolo, a un equipo como nodo raíz. Este equipo será la referencia para todos los demás, y se deberán definir caminos

hacia él. La elección de nodo raíz se determina por el identificador único de cada equipo. El valor más bajo será el elegido como raíz.

Una vez definido el nodo raíz, cada elemento de la red calcula el coste de alcanzar al nodo raíz y escoge aquel con menor métrica. El puerto del enlace con menor métrica hacia el nodo raíz se denomina root port. En el caso de segmentos de red compartidos por más equipos, se deberá elegir al switch con un camino con menor métrica hacia el root. El puerto de dicho switch se denominará designated port. Existen ciertos criterios, basados en identificadores de switch únicos, o en direcciones MAC para deshacer posibles empates en caso de existir varios caminos hacia el nodo raíz desde diferentes puertos o diferentes switches de un mismo segmento. La métrica de los enlaces está tabulada en base al ancho de banda de los mismos.

En todo caso, y tras establecerse dichos caminos con mejor métrica hasta el root, todo puerto cuyo camino no conduce al nodo root con la menor métrica, es decir, los que nos designados o puertos raíz, permanecen en estado bloqueado. De esta manera, se deshace la topología en malla para convertirla en un árbol de expansión desde el nodo raíz hacia el resto de equipos de la red.

La siguiente figura muestra un sencillo ejemplo de cómo una red con bucle físico implementa STP para deshacerlo, seleccionando un nodo root, en este ejemplo el nodo C, y bloqueando en el resto de switches los puertos con peor métrica hacia el nodo raíz.

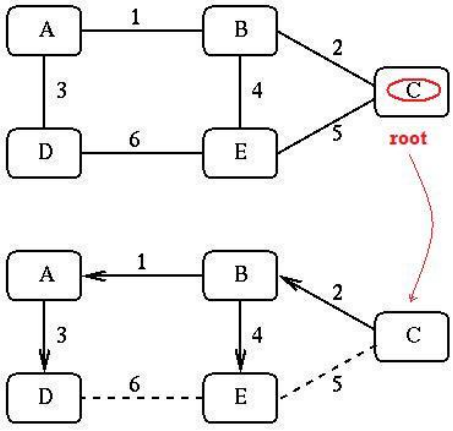

Imagen 21 - H-VPLS (4)

El protocolo tiene definidos varios estados para los puertos de los switches, escucha y aprendizaje en los que se procesan las BPDUs y se determina el árbol de expansión convergentemente en todos los equipos, y activo o bloqueado, que intuitivamente indican el estado final al que llegan después de dicha convergencia en el protocolo.

Si un cambio sucede en la red, un falla un enlace o un equipo, o se incorpora un nuevo elemento, se generan Topology Change Notifications (TCN BPDUs) inyectadas por equipos que no son root. Estas BPDUs se propagan rápidamente a todos los switches para que actualicen sus tablas de direcciones MACs y arranquen de nuevo el proceso de elección de root. Este proceso suele determinarse entre aproximadamente 30 y 50 segundos, lo cual supone un tiempo elevado, e incluso inestabilidad en la red si ésta sufre muchos cambios. El protocolo ha evoluciona en diferentes versiones del mismo, tales como Rapid STP, PerVLAN STP, o Multi-STP, e incluso combinaciones de nuevas extensiones. La variedad que utilizaremos para el diseño de la red VPLS jerárquica será RSTP, para proporcionar mayor rapidez de convergencia en caso de fallos de los enlaces que conectan regiones VPLS con el Core.

Las mejoras introducidas por RSTP frente al original están enfocadas a optimizar el tiempo de convergencia. En RSTP, se detecta la caída hacia el nodo raíz con un solo fallo en los paquetes hello que se intercambian los nodos, lo cual significa 2 segundos si no se alteran por el administrador de la red. También, los nodos no dejan de escuchar las BPDUs procedentes del camino hacia el root, y resumen la información referente a STP hacia los nodos consecutivos, facilitando a éstos la tarea de decidir qué puertos deben bloquear. RSTP mantiene además puertos de backup para el puerto raíz.

El compendio de estas novedades introducidas en STP provocan un tiempo de convergencia 30 veces menor que el original.

4.4.2. Rapid Spanning Tree Protocol aplicado en redes VPLS jerárquicas

La implementación de RSTP entre una región VPLS y el Core de la red afecta a los nodos de red que soportan la conexión redundada entre las mismas. Se han definido dos equipos de la región VPLS que harán de portavoces del tráfico del cliente cursando en dicha región. Estos dos equipos se conectan con otros dos del Core, para formar el enlace redundado. Los cuatro equipos correrán

una instancia del protocolo RSTP para deshacer el cuadrado físico que se construye entre ellos. El escenario extrapolado a STP es muy sencillo, ya que son sólo 4 elementos y un único bucle.

Fijaremos uno de los equipos Core como nodo raíz manualmente, como administradores de la red. La prioridad máxima en RSTP para seleccionar al nodo raíz es 0. La convergencia del protocolo designará puertos root en los demás nodos y bloqueará aquellos que no soporten el mejor camino hacia la raíz del árbol de expansión que se definirá.

La aflicción del protocolo sobre una red VPLS, no debe bloquear el puerto físico de la segunda conexión redundante. Lo debe deshabilitar de forma lógica, y no sobre el propio nivel de enlace como ocurre cuando el protocolo es aplicado en una red de nivel 2, debe hacerlo sobre el nivel de aplicación, esto es, sobre el protocolo LDP, ya que entiende el bucle desde el nivel MPLS hacia arriba.

RSTP en la conexión VPLS de una región, debe bloquear el establecimiento de peerings LDP entre PEs, de forma que la VPLS del cliente sólo tenga una salida lógica de la región, por el enlace que no ha sido bloqueado y cuya señalización con etiquetas LDP está activa. Si se bloquea el puerto físico o a nivel de enlace, se pierde señalización para el resto de protocolos involucrados en la arquitectura VPLS, es decir, se impiden las adyacencias OSPF, la señalización RSVP para el establecimiento de los path MPLS y sus caminos de backup a través de la Ingeniería del Tráfico.

Es importante destacar que RSTP sólo bloquea conexiones a nivel de aplicación en la pila de protocolos de VPLS.

Además, sólo bloquea conexiones definidas como tipo spoke. Los enlaces tipo mesh están libres de bucles, son intra-regionales y forman parte de un mallado completo con implementación de la regla split-horizon, por lo que no tiene sentido bloquear dichas conexiones. Por tanto, RSTP sobre enlaces redundados entre regiones VPLS y Core, sólo bloquean conexiones tipo spoke, aquellas que proporcionan la conectividad entre las dos regiones.

En el siguiente diagrama se escenifica al situación del diseño en la conexión redundante de una región VPLS hacia el Core, con instancias de RSTP corriendo en los nodos de red.

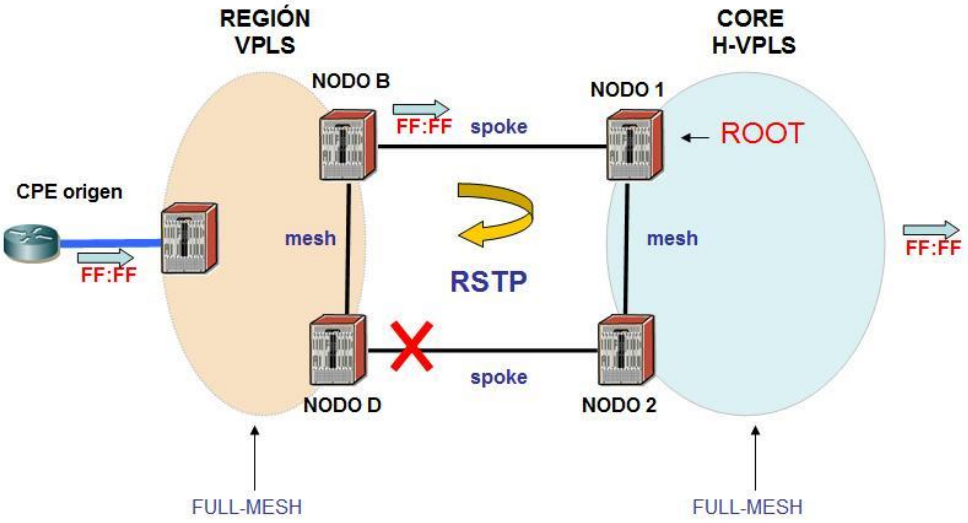

Imagen 22 - H-VPLS (5)

En el ejemplo, se ha definido manualmente al nodo 1 como nodo raíz de la región. Por definición del servicio, sólo se bloquearán conexiones tipo spoke. En el diseño planteado en este proyecto, se trataría de las conexiones entre región VPLS y Core. La convergencia del protocolo RSTP por tanto, bloqueará un interfaz lógico en la conexión más alejada del nodo raíz, en este ejemplo, la conexión nodo D – nodo 2. Este enlace sólo estará bloqueado a nivel LDP. El resto de protocolos no se deben ver afectados por RSTP.

El enlace bloqueado, impide el reenvío de tramas de un cliente por medio de ese enlace al Core, pero no impide la señalización LDP a través del mismo. Sólo se bloquea en el plano de datos, no en el plano de control.

En la figura, se cómo una trama broadcast llega al Core desde una región VPLS sólo por medio de uno de las conexiones. La segunda permanece bloqueada en cuanto a reenvío de tramas, por lo que el Core no devolverá el tráfico por dicha conexión spoke hacia la región originaria del mismo. Lo hará hacia otras regiones por medio de otras conexiones spoke. Si las conexiones desde el Core con regiones remotas fueran a su vez redundantes, una de los enlaces permanecería bloqueado lógicamente implementando la misma estructura de RSTP entre región y Core.

En caso de fallo en el enlace principal, automáticamente se desbloquearía la situación actual y el enlace secundario pasaría a ser el activo y el tráfico se bascularía sobre él. La convergencia en este paso la marca RSTP en menos de dos segundos.

4.4.2.1. Consideraciones sobre las conexiones entre región VPLS y Core

Lo anteriormente expuesto para dotar de redundancia a una conexión entre una región VPLS y el Core de la red es totalmente válido para el diseño de la red VPLS jerárquica de este estudio. Pero requiere de ciertas consideraciones especiales claves para el mantenimiento del servicio y el correcto funcionamiento de la solución definida.

Como requisito inicial se fijó la necesidad de que los enlaces físicos entre nodos de la red, no compartieran recursos hardware. De esta forma, evitamos posibles fallos del servicio a nivel físico. No siempre es posible este supuesto en la mayoría de los caso por razones presupuestarias, pero siempre será el objetivo deseable.

Por otra parte, la redundancia lógica debe cubrir todos los casos posibles, para mantenerse robusta y fiable y no afectar al servicio. En el diseño de la solución redundante con RSTP como protocolo de control ante bucles lógicos, se debe evitar con la máxima prioridad, la posible caída de los dos enlaces mesh que forman parte de la conectividad entre regiones. Estos son, el enlace mesh entre los dos equipos frontera de la región VPLS, y el enlace mesh entre los dos equipos del Core. Dichos enlaces existirán siempre puesto que forman parte del mallado total dentro de cada región.

RSTP sobre la arquitectura de VPLS, solo podrá bloquear enlaces tipo spoke. Si falla uno de los enlaces tipo mesh que cierran el bucle lógico entre regiones, el propio bucle se deshará y los dos enlaces spoke pasarán a estado activo, reenviarán tramas, y producirán el tan perjudicial bucle en la red.

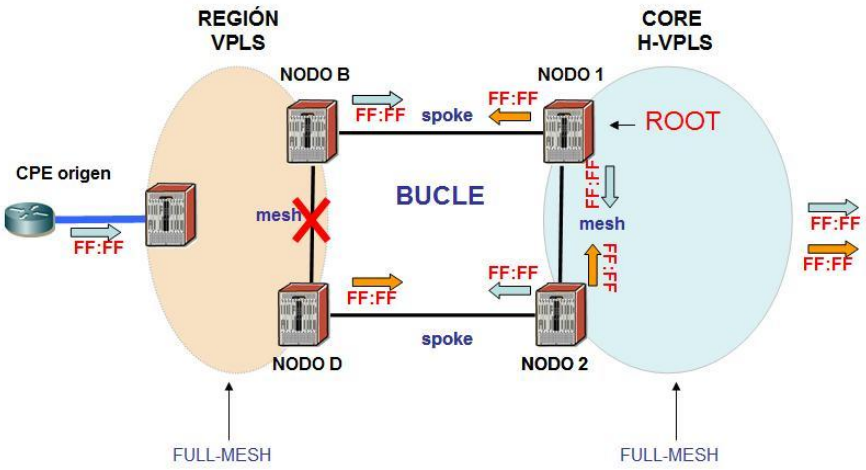

Imagen 23 - H-VPLS (6)

Por este motivo los enlaces mesh deben tener especial protección, ya que su puesta fuera de servicio provoca un bucle en una conexión redundada. Por ello, es preciso definir caminos de backup para el LSP que conecta los dos PEs. Dichos path deberían señalizarse de antemano para potenciar la convergencia a nivel MPLS, y que el fallo no se perciba en las capas superiores. Es decir, conmutaciones de path en los LSP de RSVP-TE manejan tiempos del orden de milisegundos, frente a los segundos que puede detectar RSTP. Si a nivel de aplicación los cortes no son percibidos, el protocolo RSTP no generará TCN BPDUs no actuará sobre las conexiones spoke bloqueadas y no se posibilitará la aparición de bucles lógicos que afecten a los servicios de los clientes

4.5. Doble conexión de una región VPLS con el Core redundada – reparto de carga

Las conexiones entre regiones VPLS y el Core de la red, se han definido de forma redundante, con los enlaces físicos entre las mismas, y libres de bucles lógicos con la aplicación de instancias RSTP entre los equipos que soportan dicha conectividad. Además, se afianzan los enlaces críticos en el mallado entre región y Core para evitar desbloqueos de conexiones spoke que sí generarían bucles lógicos no deseados.

El modelo es correcto, redundante y fiable, pero sólo utiliza activamente una de las conexiones físicas entre región VPLS y Core. Si se quiere ofrecer además reparto de carga entre los dos troncales, es necesaria la implementación de una solución adicional. Con el modelo definido en el capítulo anterior, una conexión permanece lógicamente bloqueada por RSTP evitando el reenvío de tramas en el plano de datos. La opción de conseguir reparto de carga entre la conexión región VPLS y Core para por duplicar la solución en dicha interconexión.

La idea básicamente consiste en duplicar las conexiones spoke entre el Core y las regiones. Dichas conexiones son las únicas susceptibles de ser bloqueadas por RSTP, motivo por el cual no obliga a duplicar las conexiones mesh que forman parte de la conexión redundante de cada región. Las conexiones spoke se definirán de forma lógica sobre los mismos enlaces físicos por partida doble.

Las conexiones spoke, son en esencia peerings LDP que señalizan la conmutación de etiquetas para un determinado cliente. Estas conexiones se incluyen en las instancias RSTP que se definen entre regiones y Core. Para proporcionar reparto de carga, se definirán dos instancias RSTP entre las regiones, seleccionando en cada una de ellas un nodo raíz diferente. Cada una de las conexiones spoke del mismo enlace, se asocia a una instancia RSTP diferente, de forma que un proceso RSTP se bloqueará un enlace spoke, y en otro el complementario, al tener como root a equipos diferentes.

Para el caso sencillo, se ha definido un nodo del Core como nodo raíz. En el caso doble, se definirá como root la otra pareja del Core que ofrece conectividad contra la región VPLS dada. Si cayese un enlace físico, inhabilitaría las dos conexiones spoke que lleva por encima de forma automática, y las instancias RSTP tendrían las otras dos conexiones spoke soportadas por el segundo enlace activas y cursando todo el tráfico de la región VPLS en concreto.

El siguiente esquema pretende sintetizar la solución redundante con reparto de carga para la conexión entre regiones y Core.

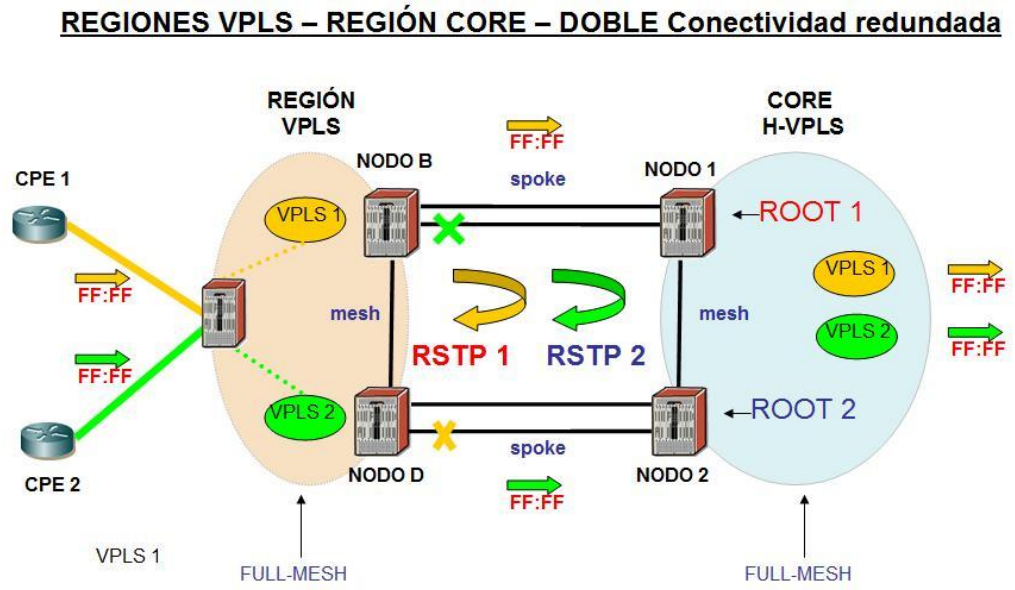

Imagen 24 - H-VPLS (7)

En la figura se puede observar dos equipos de diferentes clientes, generando tramas broadcast hacia el resto de elementos de sus respectivas VPNs La conexión contra el Core y por ende con el resto de regiones, se ofrece con dos enlaces spoke, pero la VPLS del cliente sólo se debe asociar a uno de ellos, es decir, la señalización LDP para ese tramo sólo debe hacerse sobre una única sesión entre los nodos de red. Como se expresa en el gráfico, hay dos instancias de RSTP corriendo sobre los enlaces, con dos nodos raíz diferentes que provocan el bloqueo de uno de los enlaces spoke involucrados en cada instancia.

Así, el enlace nodo B – nodo 1 está bloqueado por la instancia RSTP 2, cuyo root es el nodo 2. Por su parte, la instancia RSTP 1 bloquea el enlace nodo D – nodo 2 al tener como root al nodo 1.

El cliente 1 está asociado al enlace spoke gestionado por la instancia RSTP 1, por lo que su tráfico será cursado sobre el enlace superior. El segundo cliente está asociado al enlace spoke gestionado por la instancia RSTP 2, por lo que su tráfico será transportado por el enlace inferior. Al

duplicar el modelo de redundancia entra las regiones VPLS y el Core, se consigue el comportamiento esperado de reparto de carga.

Para que este reparto de carga sea equitativo, es necesario un estudio profundo sobre el tráfico de los clientes, su caudal, y el tipo de perfil para definir criterios de asignación a un enlace o a otro que se ajusten a medidas razonablemente equitativas.

4.6. Interconexión con otras redes

Debido a su condición de red de nivel 2, es altamente probable que la VPN del cliente requiera conectividad con otras redes o plataformas de servicios. Por ejemplo, conexión a redes MPLS de nivel 3, conexiones con usuarios sobre accesos ADSL concentrados en BRASes (Broadband Reamote Access Servers), conexiones con plataformas de voz sobre IP, etc.

Las redes VPLS se basan en Ethernet, por lo que es imprescindible que la arquitectura de las otras redes lo soporte. Existen métodos para encapsular ciertas tecnologías como ATM o FR en VPLS o MPLS, pero para la red definida en este proyecto basará el nivel de enlace con otras redes en Ethernet.

4.6.1. VPLS con transporte de etiqueta de VLAN

Dependiendo de las opciones que soporte cada suministrador de hardware, es posible implementar una opción u otra, en cuanto al transporte de la etiqueta de VLAN en la trama Ethernet de cada usuario final. Los pseudowires que conectan PEs remotos en una VPLS, pueden transportar transparentemente la etiqueta de VLAN original, dejando en el lado del cliente la responsabilidad de adecuar sus configuraciones para que la solución final sea consistente en cuanto a comunicación entre equipos pertenecientes a las mismas VLANS.

4.6.2. VPLS sin transporte de etiqueta de VLAN

La segunda opción, es eliminar la etiqueta de VLAN de la trama Ethernet, y transportar la propia trama sin VLAN tag, de forma que sea en el PE remoto donde se reconstruya la trama, y además de añadir los campos correspondientes a la parte física (preámbulo y FCS), se añade también en la cabecera Ethernet el identificativo de VLAN.

Si la configuración entre equipo de cliente y PE es consistente, y ambos tienen la misma definición de VLANs, éstas no tienen porqué ser las misma en ambos extremos. Se puede (y así se ha probado en el proyecto) utilizar una VLAN entre CE y PE, con significado local, transportar el tráfico encapsulado en MPLS, y recomponer la trama con otra VLAN en el extremo remoto donde el nuevo identificativo de VLAN es consistente y tiene significado local entre el PE y el CE remotos

4.7. Otras consideraciones

4.7.1. Escalabilidad de redes VPLS – PBB

La escalabilidad de una red que soporte servicios VPLS, es un factor muy a tener en cuenta. Las entradas en las tablas de MACs de cada uno de los servicios VPLS, consumen recursos de memoria, siendo ésta obviamente finita, y en algunos casos, pequeña.

Debido a la proliferación de accesos Ethernet, con mayor ancho de banda y menor coste, se han ido definiendo técnicas para la minimización de consumo de recursos por parte de los equipos de red. Los proveedores deben monitorizar el consumo de dichos recursos para evitar problemas en el futuro.

Una de las técnicas definidas para tratar de mitigar el problema, es PBB (Provider Backbone Bridges - IEEE 802.1ah-2008). Esta idea se basa en el stacking de etiquetas de VLAN o QinQ, pero yendo un poco más allá que, por ejemplo, apilar etiquetas para diferenciar domino de cliente y de proveedor. Con PBB se separa completamente el dominio del proveedor, proporcionando así más control sobre las direcciones MAC.

En el siguiente esquema se puede observar cómo sería la pila de protocolos si se utiliza PBB.

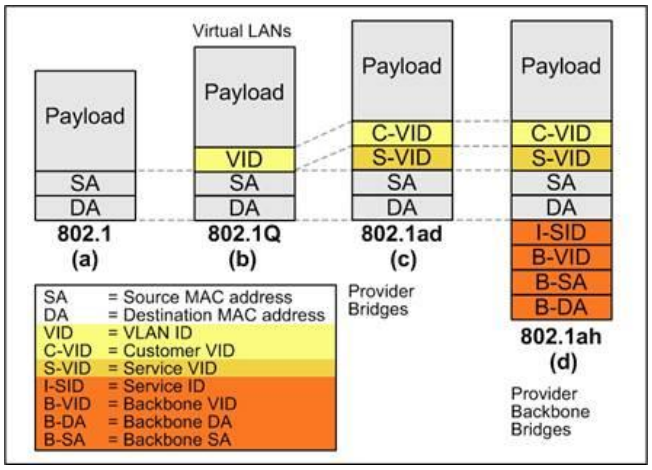

Imagen 25 - H-VPLS (8)

En la parte de red, el proveedor añadirá una nueva cabecera Ethernet con direcciones MAC propias de los equipos de red, de forma que dentro del Core sólo se manejen las MACs de los nodos de red, dejando las direcciones de clientes en el borde de la red.

En los accesos, todas las instancias VPLS de los clientes se mapean a una instancia VPLS de red, definida por el proveedor de servicios. Dicha VPLS encapsulará el tráfico añadiendo sus propias direcciones MAC de origen y destino, de forma que el resto de equipos de la red sólo deben conocer las MACs de los nodos de red, dejando a los equipos de borde la responsabilidad de mantener las entradas de los clientes. Así, las MACs de los usuarios finales son transparentes para los nodos de red.

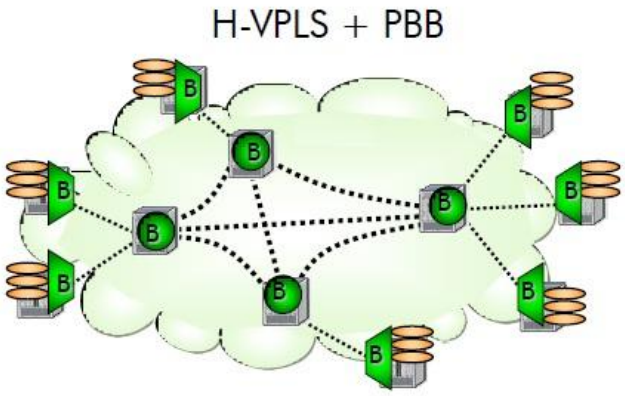

Imagen 26 - H-VPLS (9)

De esta forma, se reducen considerablemente los servicios, las direcciones MAC aprendidas en la planta, y los pseudowires para interconectar instancias VPLS entre PEs remotos.

En las pruebas realizadas durante este proyecto, no se ha incluido configuración de PBB. Sería una recomendación dentro del diseño, y no una especificación, condicionada a la necesidad real del operador de este tipo de redes y de los clientes que vaya a provisionar

4.7.2. Señalización LDP frente a señalización BGP

Otra consideración a mencionar dentro del ámbito de este proyecto es la posibilidad de utilizar BGP como protocolo de señalización para servicios VPLS. No se ha considerado inicialmente como una opción disponible, porque la gran mayoría de proveedores a nivel mundial utilizan el protocolo LDP como señalización en sus servicios VPLS. LDP está mucho más ampliamente extendido que BGP. Algunos de los motivos de la preferencia entre LDP y BGP se tratan de mencionar a continuación, pero no forma parte del diseño de la red VPLS objeto de este proyecto.

La señalización BGP para redes VPLS se recoge en la RFC 4761. La parte de transporte en este caso puede ser similar a las redes basadas en LDP. Se trata de transporte MPLS, por lo que se puede utilizar RSVP, LDP, o túneles GRE. Se requiere, eso sí, un mallado completo para el plano de transporte, ya que la jerarquía entendida en redes LDP no es compatible en este escenario.

Sobre las sesiones iBGP establecidas entre PEs de la red, se intercambia información de alcanzabilidad, con el siguiente formato:

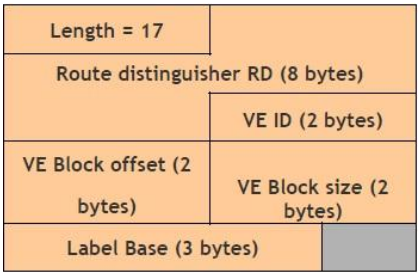

Imagen 27 - H-VPLS (10)

En estas actualizaciones se intercambian las etiquetas que identificaran los servicios VPLS en los PEs remotos. Se puede conseguir jerarquía en el plano de control, evitando por ejemplo un mallado de sesiones BGP entre todos los nodos, utilizando reflectores de rutas (BGP RRs).

La utilización de la RFC 4761 (BGP VPLS) en lugar de la RFC 4762 (LDP VPLS) puede implicar la utilización de determinados equipos de datos de un único fabricante, ya que la gran mayoría apuesta por señalización LDP, debido a las ventajas que ofrece para este tipo de redes, como por ejemplo indicación del estado del pseudowire, autodescubrimiento de instancias VPLS utilizando Radius, NMS, o incluso BGP autodiscovery.

5. QoS

La calidad de servicio es un apartado muy importante dentro de la planificación y el diseño de una red, especialmente, cuando se trata de operadores de red que ofrecerán recursos compartidos a diferentes usuarios.

Los mecanismos de calidad de servicio priorizan cierto tipo de tráfico frente a otros. Esto es esencial ya que hay aplicaciones que son más sensibles que otras a los posibles fallos de la red. Éstos pueden ir de retardos elevados, pérdida de paquetes, desordenamiento de los mismos en la entrega final, o errores de tramas. Dependiendo del tipo de aplicación que el usuario final utilice, deseará que se haga un tratamiento especial a ciertos tráficos.

Habitualmente, las redes se diseñan y planifican para que puedan cubrirse las necesidades de todos sus usuarios, pero éstos, estadísticamente no cursan tráfico al mismo tiempo, por lo que es casi necesario dimensionar la red en base al tráfico real cursado, y no a los máximos teóricos que cada cliente pueda alcanzar. Así, la suma total del ancho de banda contratado por todos los usuarios es muy superior al ancho de banda disponible en los troncales de la red.

Es indispensable, no obstante, monitorizar los troncales para prever crecimientos en determinadas áreas de la red, de forma que se pueda ir ampliando caudales en función de las necesidades reales de tráfico.

Los mecanismos de calidad de servicio, no entran en funcionamiento si no existe congestión en la red. El tráfico se clasificará y priorizará en cada salto, pero no se activarán mecanismos agresivos de descartes o encolamientos si no hay sobrecarga en los enlaces.

Para que un operador de red pueda clasificar, priorizar y encolar el tráfico de cada usuario de forma consistente, es necesario que los paquetes entrantes a la red estén marcados con sus correspondientes valores en los campos reservados a tal efecto de cada protocolo.

Cada protocolo tiene definidos una serie de campos cuyos valores determinan la prioridad de los paquetes y el tipo de tráfico que es.

5.1. QoS en protocolos.

En el caso de una red VPLS, las cabeceras de los protocolos serán principalmente Ethernet para el nivel 2, incluso pudiera ser MPLS, e IP para el nivel 3. Es posible también clasificar tráfico en base al número de puerto de la cabecera de nivel 4, pero a esto se recurriría en casos especiales y puntuales.

5.1.1. 802.1p

Los bits del campo de la cabecera Ethernet que se utilizan para calidad de servicio van incluidos en la etiqueta de VLAN, por lo que sería necesario utilizar 802.1q como método de encapsulación Ethernet. En dicha etiqueta, se reservan 3 bits para definir hasta 8 clases de servicio diferentes.

No está definida la manera de cómo tratar el tráfico que tiene asignada una determinada clase o prioridad, existe libertad a las implementaciones. IEEE, sin embargo, ha hecho amplias recomendaciones al respecto. Los 8 valores van del 0 al 7, siendo el 0 el de baja prioridad y suelen tener correspondencia con los valores de capas superiores.

El inconveniente es la necesidad de introducir en la cabecera Ethernet la etiqueta de VLAN, la cual consta de 4 bytes.

5.1.2. Tipo de Servicio

El segundo campo de la cabecera IP se definió para asumir la calidad de servicio asignada al tráfico IP en cada paquete. Este campo consta de 8 bits para definir el tipo de servicio. Originalmente utilizaba los tres bits más significativos, denominados Precedencia IP, pero estándares más recientes utilizan los 6 bits más significativos, para componer la tabla de servicios diferenciados.

5.1.2.1. Precedencia IP

Es decir, inicialmente, se trataban los 3 primeros bits del mismo modo que por ejemplo el campo 802.1p de la cabecera Ethernet, consiguiendo hasta 8 clases de servicio. Éstos valores si tenían un nombre o concepto asociado de forma que fuera común el uso de los mismos para aplicaciones parecidas.

Así, se tienen los siguientes valores de Precedencia IP del campo ToS de la cabecera IP:

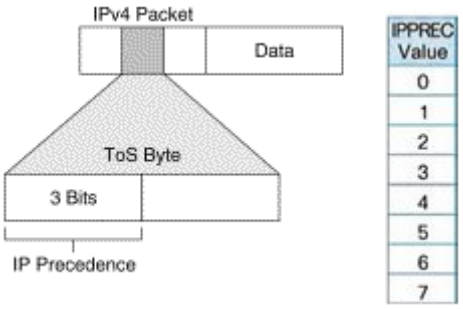

Imagen 28 - QoS

Los restantes bits del campo ToS tienen definiciones poco utilizadas en la práctica y pocas arquitecturas los tienen en cuenta. Esto corresponde a un estándar antiguo y en cada vez menos utilizado, ya que el modelo de DSCP permite mayor flexibilidad en la definición de clase de servicio.

5.1.2.2. DSCP

Con las especificaciones recogidas en la RFC 2474 (Definition of the DS Field in the IPv4 and IPv6 Headers) se añaden 3 bits más para la definición de calidad de servicio, de forma que se obtiene mayor granularidad a la hora de clasificar y priorizar tráfico.

Basándose en la arquitectura de Internet, éstos valores tendrán sentido en cada salto que tengan en su camino el paquete IP. Cada router examinará la cabecera IP y por lo tanto el campo ToS, del cual extraerá la información necesaria para actuar en cuanto a calidad de servicio se refiere.

El RFC 2474 actualiza nuevamente este campo utilizando Codepoints (DSCP) para definir el comportamiento en cada salto. Particularmente señala que el campo es denominado ahora DS (Differentiated Services) en reemplazo de ToS (Type of Service), cuyos seis primeros bits se utilizan para registrar el Codepoint (DSCP: differentiated services codepoint) respectivo y los dos últimos no son actualmente utilizados.

Especificaciones posteriores detallan y completan el uso de estos campos, el RFC 2597 (Assured Forwarding PHB Group) recomienda valores para utilizar para un grupo denominado Assured Forwarding, y el RFC 3246 (An Expedited Forwarding PHB) recomienda utilizar los Codepoints para EF. El cómo hacer uso de estas marcas se define en el RFC 2475 (An Architecture for Differentiated Services)

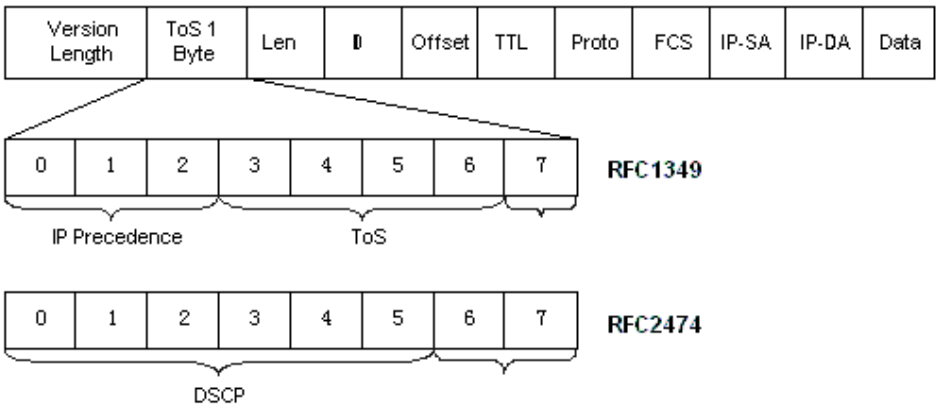

Imagen 29 - QoS (2)

5.1.3. EXP MPLS

De forma análoga a IP, la cabecera MPLS cuenta con un campo cuyos bits determinan la prioridad del tráfico encapsulado. Originalmente no contemplado, la calidad de servicio en la cabecera MPLS se señaliza en el campo Experimental, cuyo uso se reservó inicialmente, y ahora soporta la categorización de los paquetes. Se trata de 3 bits que proporcionan 8 diferentes clases de servicio, habitualmente mapeadas a las que se definen en Precedencia IP.

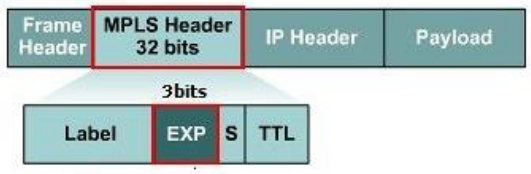

Imagen 30 - QoS (3)

5.2. QoS definida para la red H-VPLS

El diseño y la definición de la calidad de servicio para la red deben ir ligados a las necesidades reales de los usuarios y aplicaciones de los mismos. Al tratarse de operadores de red con enlaces troncales compartidos por los clientes, y con sobresuscripción en cuanto al posible tráfico entrante y la capacidad total de los troncales, es crítico definir un modelo ajustado a los diferentes tipos de tráfico, escalable, y con flexibilidad para no necesitar una monitorización en tiempo real del tráfico cursado.

En este ejercicio, se va a diferenciar entre la calidad de servicio dentro de la red, es decir, en el Core IP/MPLS del proveedor, y la calidad de servicio en el acceso, siendo ésta la priorización y administración del ancho de banda de los accesos de los usuarios.

Evidentemente, las capacidades de los enlaces troncales son de uno o varios órdenes superiores a las capacidades de los accesos. Aun así, la suma total del ancho de banda contratad por todos los clientes es muy superior a los propios troncales, por lo que ambos diseños deben estar muy relacionados.

5.2.1. QoS de red

Pese a disponer de 6 bits en el caso de DSCP para tipificar numerosas calidades de servicio, la experiencia denota que la calidad de servicio en el Core de la red, puede simplificarse a 3 o 4 valores, sin que las aplicaciones finales se vean afectadas. Los enlaces troncales de las operadoras de red no suelen estar cursando tráfico continuamente al 100% de su capacidad, y es frecuente doblar los enlaces cuando se alcanzan valores sostenidos del 70 o el 80 por ciento.

Los accesos tienen anchos de banda muy inferiores y su coste suele ser un factor clave en la rentabilidad del servicio para el cliente, por lo que habitualmente lo utilizan al máximo de su capacidad. Esto si requiere un ejercicio más fino en cuanto a clasificación y priorización, ya que ahí si se observa una congestión más frecuentemente.

Para el diseño de la red VPLS jerárquica se han definido tres tipos de caudales con diferente calidad de servicio en el Core de la red, más un cuarto necesario para el tráfico de control propio de los PEs y Ps, es decir, el tráfico de protocolos de control y señalización.

Los diferentes valores de QoS que marcará el usuario en su tráfico serán mapeados a las 3 clases de servicio de red. Éstas tendrán sus valores definidos acorde con las previsiones de tráfico y con la flexibilidad necesaria para cubrir ráfagas o picos sostenidos durante ciertos intervalos de tiempo.

5.2.1.1. Definición de caudales y arquitectura

Se definirán tres tipos de tráfico mayoritarios, y se ofrecerán como valor añadido al servicio básico de conectividad IP sobre VPNs Es decir, el capítulo de calidad de servicio aporta muchos matices a la hora de modelar el servicio a proporcionar a los clientes. Las necesidades de los mismos pueden afectar a su vez al proceso de modelado y formato comercial.

Los tres tipos de calidad de servicio de red soportarán tráfico definido como **best-effort,** tráfico más prioritario definido como **asegurado,** y tráfico definido como **prioritario** con un nivel más de prioridad que el resto.

El primer tipo de CoS cursará todo el tráfico cuyos valores de retardo, pérdidas y encolamiento no afectan al rendimiento de las aplicaciones que lo generan. Normalmente, este tipo de trafico suele ser tráfico de Internet, navegación, descargas y similares, utilizados conjuntamente por los usuarios del cliente. Incluso, ese caudal suele compartirse por todos los usuarios finales.

El segundo caudal supone un cierto nivel más de prioridad. Las aplicaciones de los clientes, servidores centrales de archivos, programas con movimientos en bases de datos remotas, correo electrónico, y muchas más opciones, requieren un tratamiento diferenciado respecto al tráfico de Internet, ya que los retardos y las tasas de error pueden afectar a su rendimiento.

El tercer caudal, más prioritario, suele reservarse para tráfico crítico, cuyos valores de retardo y tasas de error deben ser mínimos y para garantizar el rendimiento al usuario final. Fundamentalmente, se trata de tráfico de voz sobre IP, o videoconferencias. Estas aplicaciones de VoIP o multimedia son muy sensibles a retardos, reordenamiento de paquetes o tasas de error elevadas que hacen inviable la comunicación final. Por esto, siempre se dedica un porcentaje de la

red a este tipo de tráfico, y aquí si es necesario comprobar su rendimiento con mayor frecuencia. Esto, evidentemente implicará un coste añadido al servicio ofrecido.

Finalmente, al margen de las calidades definidas para el tráfico de los usuarios, es imprescindible contar con un volumen de tráfico especial y necesario generado por los propios equipos del operador. Los protocolos de routing y señalización apenas cursan cientos de kilobits por segundo, pero éste tráfico ha de protegerse frente a situaciones de congestión severa, ya que la pérdida de adyacencias, peerings y demás señalización puede causar la caída generalizada de la red. Sin plano de control es imposible enrutar el tráfico de los clientes. Aunque el volumen de los usuarios alcancen capacidades de Gigabit Ethernet, sin el tráfico de keepalives y hellos entre los routers del Core será imposible proporcionar el servicio.

Habitualmente los diferentes fabricantes de hardware para redes diseñan sus equipos de modo que el tráfico de control, que es generado por el propio equipo, marque por defecto los valores de DSCP, IP Precedente o Experimental MPLS con la máxima prioridad, de forma que el administrador de red no tenga la necesidad de definir políticas de calidad de servicio al respecto. Haya o no un diseño en la red donde se utilice el equipo de calidad de servicio, los paquetes del plano de control salen marcados con la máxima prioridad.

5.2.1.2. Clasificación y reescritura

El tráfico entrante a la red debe ser clasificado correctamente, no sólo para que los mecanismos de calidad de servicio sean coherentes, también para que sean eficaces. Es necesario acordar con el cliente los valores de QoS que marcará y utilizará para los servicios contratados. Según los caudales que quiera utilizar, y la diferenciación del tráfico que transita, se definirán unas políticas u otras en sus accesos.

5.2.1.3. Mapeo de calidad de servicio en IP/MPLS

La estructura de calidad de servicio definida para la red, se puede sintetizar en la siguiente tabla.

Al haber definido cuatro caudales de tráfico diferenciado en el Core de la red, se agruparán los valores de CoS para reducirlos a estos cuatro principales. El tráfico de Internet o no prioritario no

irá marcado y se tratará en modo defecto. Los valores de AF1 a AF4 se agrupan en una misma clase de servicio equivalente, el tráfico crítico, en su correspondiente clase prioritaria, el tráfico de control generado por los propios nodos de red para señalizar la pila de protocolos completa, en sus respectivas clases asignadas a tal efecto

Precedencia IP	DSCP binario	DSCP decimal	DSCP	DSCP forwarding class	MPLS EXP
0	000 000	0	BE	Best effort	0
1	001 010	10	AF11	Assured Forwarding	1
1	001 100	12	AF12	Assured Forwarding	1
1	001 110	14	AF13	Assured Forwarding	1
2	010 010	18	AF21	Assured Forwarding	2
2	010 100	20	AF22	Assured Forwarding	2
2	010 110	22	AF23	Assured Forwarding	2
3	011 010	26	AF31	Assured Forwarding	3
3	011 100	28	AF32	Assured Forwarding	3
3	011 110	30	AF33	Assured Forwarding	3
4	100 010	34	AF41	Assured Forwarding	4
4	100 100	36	AF42	Assured Forwarding	4
4	100 110	38	AF43	Assured Forwarding	4
5	101 110	46	EF	Expedited Forwarding	5
6	110 000	48	NC1	Network Control	6
7	111 000	56	NC2	Network Control	7

Imagen 31 - QoS (4)

5.2.1.4. QoS en acceso

En los accesos a la red VPLS, y dependiendo de las aplicaciones de cada una de las sedes o delegaciones de los clientes, así como el perfil asignado, el tráfico irá marcado con unos valores u otros. El tratamiento que recibirá en el acceso, con habitualmente mucho menor ancho de banda, dependerá de los valores de CoS que marquen los equipos de cliente. El tratamiento que recibirá desde los PEs hacia la red será el definido en la tabla del capítulo anterior, donde se mapean los valores de CoS de IP a valores de prioridad en la cabecera MPLS, según las reglas de calidad de servicio para este diseño.

Para el diseño de la red VPLS jerárquica, se definirán cuatro caudales diferenciados de tráfico en el Core, uno para tráfico no prioritario o de Internet, otro para tráfico de datos (que engloba todas las opciones de AF) un caudal prioritario para tráfico multimedia y de voz, y otro caudal para proteger el tráfico de control de los nodos de red

Especificación de diseño 10

6. Plan de pruebas y resultados

6.1. Escenario

Para comprobar el correcto funcionamiento de la red, y los servicios que se definirán y transportarán sobre la misma, se define el siguiente escenario:

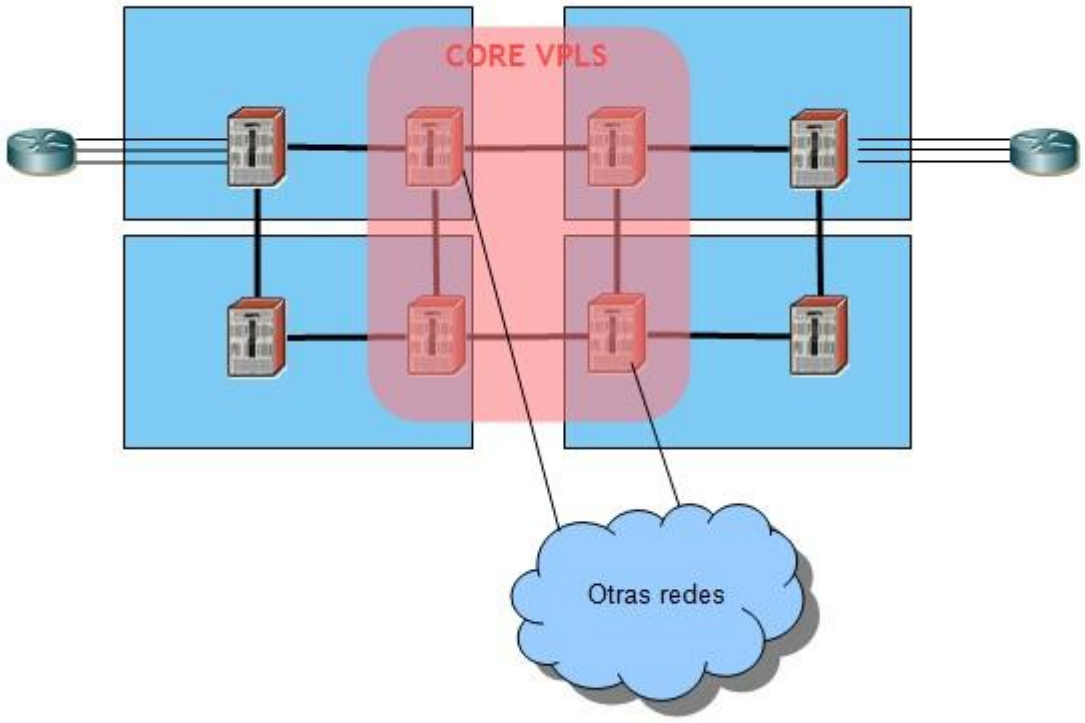

Imagen 32 - Escenario pruebas

6.2. Hardware utilizado

Los equipos utilizados para el escenario de pruebas cumpliendo el papel de PE y P son **Alcatel-Lucent Service Routers 7750**. La versión de software cargada en los equipos es la **TimOS 5.0.R4**.

Como equipos auxiliares se han utilizado **Cisco 2811**.

Las conexiones utilizadas han sido enlaces Ethernet, tanto Fast Ethernet como Gigabit Ethernet.

6.3. Plan de pruebas y resultados

6.3.1. Pruebas físicas

6.3.1.1. Detección de fallos sobre enlaces físicos.

Es de vital importancia en las redes basadas en Ethernet poder contar con mecanismos de detección de caída de enlaces físicos. Las redes de nivel metropolitano y nacional, cimentan su infraestructura sobre circuitos de transmisión óptica. Esto introduce una serie de equipos de transmisión (nivel 1) en el camino físico entre dos nodos remotos. El transporte de los circuitos Ethernet puede sufrir cortes, tanto unidireccional como bidireccionalmente. Es fundamental que el equipo de conmutación de datos, PE y P, detecten dichos fallos. Los interfaces de los puertos Ethernet se mantienen operativos en tanto en cuanto la señal física se mantenga. Esto puede implicar que un corte entre dos equipos de transmisión intermedios, no se provoque la caída física en el extremo final, es decir, el último equipo de transmisión y el equipo de datos. Por lo tanto, el equipo de datos mantendría el interfaz operativo, provocando así un efecto indeseado de sumidero de tráfico. El resto de la red seguirá enviando tráfico hacia ese destino, sin haber detectado el fallo en la capa física.

Existen varios mecanismos de detección de cortes sobre líneas de transmisión para la tecnología Ethernet. En este proyecto, se ha utilizado el protocolo 802.3ah, conocido como Ethernet First Mile. Ver capítulo 3.1.3.5 Detección de bucles físicos

Las pruebas se han realizado lanzando pings entre los equipos Cisco auxiliares. Los cortes se han generado a nivel de transmisión y a nivel Ethernet. El papel del interfaz Ethernet en cuanto al protocolo EFM se refiere, se ha intercambiado para determinar ambas opciones

El interfaz se puede definir como activo o pasivo. El primero envía PDUs hacia el vecino y espera contestación. El segundo responde a PDUs recibidas. El intervalo de transmisión para las PDUs configurado es de **200 ms**

Tipo de Corte	Tiempo detección EFM	Tiempo detección LSP	Tiempo conmutación tráfico	Recuperación del enlace
Unidireccional Fallo a nivel SDH EFM activo	< 1 seg	< 1 seg	< 1 seg	Sin pérdidas
Unidireccional Fallo a nivel SDH EFM pasivo	< 2 seg	< 1 seg	< 1 seg 1 paquete perdido (ping)	Sin pérdidas
Unidireccional Fallo a nivel Ethernet EFM activo	< 1 seg	< 1 seg	< 1 seg	Sin pérdidas
Unidireccional Fallo a nivel Ethernet EFM pasivo	< 1 seg	< 1 seg	< 1 seg	Sin pérdidas

Tabla resultados 1

6.3.1.2. Detección de micro-cortes sobre enlaces físicos.

Con la idea de simular pequeños y sucesivos cortes en los enlaces, se desconectan y conectan repetidas veces de forma inmediata las fibras conectadas tanto en los equipos de transmisión como en los de datos (PE, P)

Se comprueba que los cortes se detectan y que el tráfico conmuta correctamente a otro enlace. Después de la situación de inestabilidad, el enlace recupera correctamente y vuelve a estar operativo:

Tipo de Corte	Tiempo conmutación tráfico	Recuperación del enlace
Microcorte SDH	< 1 seg	Sin pérdidas
Microcorte Ethernet	< 1 seg	Sin pérdidas
Conmutación SDH a camino secundario	< 1 seg 4 paquetes perdido (plano de control)	Sin pérdidas

Tabla resultados 2

6.3.1.3. Detección de micro-cortes sobre puertos de acceso de cliente.

Siguiendo la misma operativa, se desconecta y conecta los enlaces de acceso a la red VPLS simulando cortes en los circuitos de los clientes.

El puerto detecta caída física, y recupera sin problemas. Se detecta que envía un mensaje a los nodos VPLS indicando la eliminación de la dirección MAC asociada al puerto que ha sufrido la caída:

Antes de la caída:

```
A:NODO1# show service id 300 fdb  detail

===============================================================================
Forwarding Database, Service 300
===============================================================================
ServId    MAC               Source-Identifier     Type/Age  Last Change
-------------------------------------------------------------------------------
300       00:00:0c:07:ac:01 sdp:3:300             L/0       10/22/2007 08:28:31
300       00:02:85:24:59:60 sdp:11:300            L/150     10/22/2007 09:13:08
300       00:06:d7:8c:c9:52 sdp:37:300            L/150     10/22/2007 09:02:52
300       00:1a:6c:1e:a9:38 sdp:3:300             L/15      10/22/2007 08:28:36
300       00:1b:53:b4:1f:00 sdp:11:300            L/0       10/22/2007 09:12:44
-------------------------------------------------------------------------------
No. of MAC Entries: 5
===============================================================================
```

Después de la caída, la dirección MAC del equipo de cliente se elimina y se comunica al resto de Pes *(send flush on failure)*:

```
A:NODO1# show service id 300 fdb  detail

===============================================================================
Forwarding Database, Service 300
===============================================================================
ServId    MAC               Source-Identifier     Type/Age  Last Change
-------------------------------------------------------------------------------
300       00:00:0c:07:ac:01 sdp:3:300             L/0       10/22/2007 08:28:31
300       00:06:d7:8c:c9:52 sdp:37:300            L/180     10/22/2007 09:02:52
300       00:1a:6c:1e:a9:38 sdp:3:300             L/0       10/22/2007 08:28:36
300       00:1b:53:b4:1f:00 sdp:11:300            L/0       10/22/2007 09:16:41
-------------------------------------------------------------------------------
No. of MAC Entries: 4
===============================================================================
```

6.3.1.4. Detección de bucle físico en el enlace

El protocolo EFM también es capaz de detectar bucle físico al recibir PDUs con la misma dirección MAC de origen. Puesto que este protocolo se corre entre equipos VPLS del proveedor, se puede asegurar la detección de los posibles bucles físicos o a nivel SDH.

Tipo de Fallo	Detección / Acción	Recuperación del enlace
Bucle a nivel SDH	Detectado / Puerto deshabilitado	Sin pérdidas
Bucle a nivel Ethernet	Detectado / Puerto deshabilitado	Sin pérdidas

Tabla resultados 3

6.3.1.5. Detección de bucle físico en el enlace de acceso del cliente

La detección de un posible bucle físico en el acceso del cliente es primordial ya que éste puede dejar inoperativa el resto de la VPN. El procedimiento es el siguiente:

- El equipo detecta movimiento de MACs entre los diferentes interfaces (Acceso y Core) del servicio durante 5 segundos, y al superar el umbral fijado, bloquea uno de los puertos involucrados en dicho movimiento de MACs.
- El puerto permanece bloqueado un tiempo fijado por configuración.
- Tras el vencimiento de ese temporizador, se desbloquea.
- Si el proceso persiste, volverá a bloquearse.
- Una vez bloqueado el interfaz tres veces, no volverá a habilitarse automáticamente (requerirá operación del proveedor)

Tipo de Fallo	Detección / Acción	Recuperación del enlace
Bucle a nivel Ethernet	Detectado / Puerto deshabilitado	Sin pérdidas

Tabla resultados 4

6.3.1. Pruebas lógicas

6.3.1.1. Comprobación del estado de un servicio VPLS

A continuación se listan una serie de comprobaciones para verificar que un servicio VPLS está correctamente configurado y operativo.

- Estado del puerto a nivel físico. Existe enlace y se detecta la línea:

```
A:NODO1# show port 5/1/3

===============================================================================
Ethernet Interface
===============================================================================
Description        : Puerto de acceso
Interface          : 5/1/3                       Oper Speed       : 100 mbps
Link-level         : Ethernet                    Config Speed     : 100 mbps
Admin State        : up                          Oper Duplex      : full
Oper State         : up                          Config Duplex    : full
Physical Link      : Yes                         MTU              : 1518
IfIndex            : 169967616                   Hold time up     : 0 seconds
```

- Estado de la tabla de MACs del servicio:

```
A:NODO1# show service id 300 fdb detail

===============================================================================
Forwarding Database, Service 300
===============================================================================
ServId    MAC               Source-Identifier     Type/Age  Last Change
-------------------------------------------------------------------------------
300       00:19:56:20:53:af sdp:9:300             L/0       09/28/2010 11:42:47
300       00:1a:6c:1e:a9:38 sdp:9:300             L/120     09/28/2010 12:02:13
300       00:1b:53:b4:1f:00 sap:5/1/3:300         L/160     10/01/2010 07:54:21
-------------------------------------------------------------------------------
No. of MAC Entries: 3
===============================================================================
```

- Negociación de etiquetas MPLS para la VPLS. Señalización tLDP:

```
A:NODO1# show service id 300 labels

===============================================================================
Martini Service Labels
===============================================================================
Svc Id     Sdp Binding        Type  I.Lbl                E.Lbl
-------------------------------------------------------------------------------
300        7:300              Spok  131064               131057
300        9:300              Mesh  131045               131048
-------------------------------------------------------------------------------
Number of Bound SDPs : 2
-------------------------------------------------------------------------------
```

- Establecimiento de túneles RSVP para el transporte:

```
*A:NODO1# show router mpls lsp

===============================================================================
MPLS LSPs (Originating)
===============================================================================
LSP Name                              To                 Fastfail     Adm   Opr
                                                         Config
-------------------------------------------------------------------------------
LSP_NODO1NODO2                     10.16.81.25            No           Up    Up
LSP_NODO1NODO3                     10.16.81.20            No           Up    Up
-------------------------------------------------------------------------------
LSPs : 2
```

- Asociación del transporte MPLS al servicio VPLS:

```
A:NODO4# show service id 300 all

===============================================================================
Service Detailed Information
===============================================================================
Service Id        : 300                Vpn Id            : 0
Service Type      : uVPLS
Description       : NVTs
Customer Id       : 9001
Last Status Change: 09/28/2010 12:16:39
Last Mgmt Change  : 09/28/2010 12:12:20
Admin State       : Up                 Oper State        : Up
MTU               : 1514               Def. Mesh VC Id   : 300
SAP Count         : 3                  SDP Bind Count    : 2
Send Flush on Fail: Enabled            Host Conn Verify  : Disabled


-------------------------------------------------------------------------------
Service Destination Points(SDPs)
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
 Sdp Id 13:300  -(10.16.81.22)
-------------------------------------------------------------------------------
Description     : NODO4-NODO5-Spoke
SDP Id            : 13:300                  Type              : Spoke
VC Type           : Ether                   VC Tag            : n/a
Admin Path MTU    : 9194                    Oper Path MTU     : 9194
Far End           : 10.16.81.22             Delivery          : MPLS

Admin State       : Up                      Oper State        : Up
Acct. Pol         : None                    Collect Stats     : Disabled
Managed by Service : 5003                   Prune State       : Pruned
Managed by Spoke  : 13:5003
Ingress Label     : 131047                  Egress Label      : 131021
Ing mac Fltr      : n/a                     Egr mac Fltr      : n/a
Ing ip Fltr       : n/a                     Egr ip Fltr       : n/a
Ing ipv6 Fltr     : n/a                     Egr ipv6 Fltr     : n/a
Admin ControlWord : Not Preferred           Oper ControlWord  : False
Last Status Change : 10/01/2010 06:40:42    Signaling         : TLDP
Last Mgmt Change  : 09/28/2010 12:12:20

Associated LSP LIST :
Lsp Name          : LSP_NODO4NODO5
Admin State       : Up                      Oper State        : Up
Time Since Last Tr*: 00h51m19s
```

- Eliminación de entradas en la tabla de MACs de un servicio: Al realizarse, la tabla correspondiente a una VPN limpia las direcciones MAC aprendidas y reinicia automáticamente el proceso de aprendizaje

Tipo de Comprobación	Resultado	Prueba con tráfico
Nivel físico	Correcto	Correcto
Nivel lógico (MAC)	Correcto	Correcto
Nivel de señalización LDP	Correcto	Correcto
Nivel de transporte MPLS	Correcto	Correcto
Nivel de servicio VPLS	Correcto	Correcto

Tabla resultados 5

6.3.1.2. Consistencia en el diseño del Core

Para comprobar la eficacia tanto del diseño como de los protocolos de señalización, se realizan una serie de pruebas para obtener las medidas de los tiempos de convergencia del plano de control.

- Fallo simple en enlace troncal: El transporte MPLS está definido con un camino primario y un secundario preseñalizado de antemano. Se produce un corte en uno de los enlace entre equipos del Core
- Fallo doble en enlaces troncales: Igual que el caso anterior, pero añadiendo un camino terciario tipo loose. Se produce un bucle de hasta 30 segundos, el tiempo que tarda en calcularse el tercer camino con la información de OSPF.

Tipo de Comprobación	Caminos activos LSP	Pérdidas
Fallo simple enlace troncal	Primario (estricto): caído Secundario (estricto): activo Terciario (loose): sin señalizar	< 1 seg
Fallo doble enlace troncal	Primario (estricto): caído Secundario (estricto): caído Terciario (loose): activo	< 30 seg

Tabla resultados 6

6.3.1.3. Comprobación de la conexión redundante al Core mediante RSTP

Como se ha especificado en el diseño de la red, la conexión de una región VPLS con el Core debe ser redundante. Para evitar bucles en dicha conexión doble, es necesario correr una instancia de RSTP entre los nodos de red para bloquear una de las dos conexiones, de forma que se cree un árbol STP manteniendo un camino bloqueado y deshaciendo así la posibilidad de bucle. Es este caso, se comprueba su funcionamiento, así como un cambio de topología.

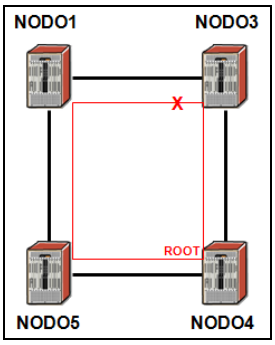

Imagen 33- Escenario pruebas (2)

- Se habilita una instancia de RSTP y se comprueba que se bloquea uno de los dos enlaces redundantes:

```
A:NODO4# show service id 5001 stp detail

===============================================================================
Spanning Tree Information
===============================================================================

-------------------------------------------------------------------------------
VPLS Spanning Tree Information
-------------------------------------------------------------------------------
VPLS oper state    : Up                        Core Connectivity : Up
Stp Admin State    : Up                        Stp Oper State    : Up
Mode               : Rstp                      Vcp Active Prot.  : Rstp

Bridge Id          : 00:00.00:16:4d:de:29:5b   Bridge Instance Id: 0
Bridge Priority    : 0                         Tx Hold Count     : 6
Topology Change    : Inactive                  Bridge Hello Time : 1
Last Top. Change   : 2d 21:59:37               Bridge Max Age    : 20

Root Bridge        : This Bridge
Primary Bridge     : This Bridge

Root Path Cost     : 0                         Root Forward Delay: 15
Rcvd Hello Time    : 1                         Root Max Age      : 20
Root Priority      : 0                         Root Port         : N/A


-------------------------------------------------------------------------------
Spanning Tree Sap/Spoke SDP Specifics
-------------------------------------------------------------------------------
SDP Identifier     : 11:5001                   Stp Admin State   : Up
Port Role          : Designated                Port State        : Forwarding
Port Number        : 2048                      Port Priority     : 128
Port Path Cost     : 10                        Auto Edge         : Enabled
Admin Edge         : Disabled                  Oper Edge         : False
Link Type          : Pt-pt                     BPDU Encap        : Dot1d
Root Guard         : Disabled                  Active Protocol   : Rstp
Last BPDU from     : 80:00.00:16:4d:d9:25:5b
Designated Bridge  : This Bridge               Designated Port Id: 34816
```

```
A:NODO1# show service id 5001 stp

===============================================================================
Stp info, Service 5001
===============================================================================
Bridge Id          : 80:00.00:16:4d:de:3d:5b   Top. Change Count : 7
Root Bridge        : 00:00.00:16:4d:de:29:5b   Stp Oper State    : Up
Primary Bridge     : 80:00.00:16:4d:d9:25:5b   Topology Change   : Inactive
Mode               : Rstp                      Last Top. Change  : 2d 22:05:33
Vcp Active Prot.   : Rstp
Root Port          : Vcp                       External RPC      : 10

===============================================================================
Stp port info
===============================================================================
Sap/Sdp Id         Oper-     Port-      Port-       Port-  Oper-  Link-  Active
                   State     Role       State       Num    Edge   Type   Prot.
-------------------------------------------------------------------------------
7:5001             Up        Alternate  Discard     2048   False  Pt-pt  Rstp
```

- Se provoca un cambio de topología definiendo un nodo diferente como raíz

Tipo de Comprobación	Resultado	Pérdidas
Convergencia RSTP	Correcto	No aplica
Cambio de topología RSTP	Correcto	< 1 seg

Tabla resultados 7

6.3.1.4. Tiempos de convergencia

Se realizan diferentes pruebas para reproducir posibles fallos en los distintos protocolos de señalización que articula los servicios VPLS.

- Se deshabilita el protocolo OSPF de interfaces de Core, se cambian métricas de los enlaces, se modifican timers en las adyacencias,
- Se deshabilitan caminos de los LSPs RSVP
- Se provocan fallos sobre la señalización LDP

Tipo de fallo	Resultado	Pérdidas recuperar
OSPF – Fallo IP loopback	Cae la adyacencia y sesión tLDP	No
OSPF – Fallo timers	Cae la adyacencia	No
OSPF – Cambio coste	No cae la adyacencia	No
RSVP – Camino principal	Conmutación a secundario	No
RSVP – Interfaz caído	Pérdidas de tráfico < 10 seg	No
RSVP – Equipo aislado	Sólo activo el tráfico local	No
LDP – Cambio timers	Servicio no afectado	No
LDP – Equipo tránsito deshabilitado	Conmutación a otro SDP	No
LDP – Equipo aislado	Sólo activo el tráfico local	No

Tabla resultados 8

6.3.1.5. Interacción con STP de cliente

Para asegurar el funcionamiento de la red ante tráfico STP de los posibles clientes, se comprueba la interacción de los nodos de red con tramas STP de las diferentes versiones del protocolo

Tipo de fallo	Resultado	Afectación
Spanning Tree	Si el acceso es 802.1q, las BPDUs se descartan	No
Per VLAN Spanning Tree	BPDUs conmutan transparentemente por la red	No
Rapid Spanning Tree	Si el acceso es 802.1q, las BPDUs se descartan	No

Tabla resultados 9

7. Especificaciones de diseño

La siguiente tabla recoge las especificaciones y recomendaciones fijadas a lo largo del proyecto:

Recomendación 1 OSCF como IGP	Por el tamaño de redes que se pueden llegar a administrar con **OSPF**, su apertura de código implementado por la mayoría de fabricantes y por las cualidades propias del protocolo, le elegimos como **IGP** para nuestra red VPLS jerárquica.
Recomendación 2 OSPF con un área	Para el diseño de nuestra red VPLS jerárquica definiremos un único un área OSPF, que por consiguiente será el área cero. Todos los nodos de la red serán **Backbone Routers,** no habiendo ASBR o ABR en ningún momento. Los enlaces utilizados para conectarse entre sí los equipos serán sobre Ethernet, simulando un punto a punto, por lo que en el establecimiento de la adyacencia se designarán un DR entre los dos vecinos.
Recomendación 3 Elementos del Core	Para el diseño de nuestra red VPLS jerárquica definiremos como **Core** de la red **cuatro nodos** que actúen como **P routers.** Estos cuatro elementos no tendrán conexiones de clientes o usuarios, estarán conectados entre sí formando un cuadrado y además, desempeñarán funciones jerárquicas.
Recomendación 4 Aprendizaje de MACs	Para el diseño de nuestra red VPLS jerárquica habilitaremos la funcionalidad de **aprendizaje de MACs**, contemplando la posibilidad de definir tamaños de tablas de MACs por VPLS en función del número de dispositivos de cada usuario
Recomendación 5 Detección de bucles	Para el diseño de nuestra red VPLS jerárquica habilitaremos la funcionalidad de **detección de bucles,** tratando de ajustar los umbrales de movimientos de MACs al máximo
Recomendación 6 Definición de FEC	Para el diseño de nuestra red VPLS jerárquica, entenderemos como una FEC al usuario o cliente completo
Recomendación 7 Tipo de LSPs	Para el diseño de nuestra red VPLS jerárquica, definiremos LSPs con caminos principales y secundarios, estos pre señalizados de antemano. Los path principales serán de tipo estricto, y los secundarios una combinación de tipo estricto y loose. Para la señalización de los LSPs utilizaremos el protocolo **RSVP**
Recomendación 8 Señalización VPLS	Para el diseño de nuestra red VPLS jerárquica, mientras la señalización de los LSPs se ha definido mediante el protocolo RSVP, los peerings remotos para comunicar FECs y crear los túneles internos será utilizando el protocolo **LDP.** La modalidad del intercambio de etiqueta será bajo demanda configurada por el administrador de red. El criterio a aplicar en entrada para los LER serán **direcciones MAC destino.** En base a estas y a los interfaces por donde recibe el tráfico, determinará si se trata de un cliente u otro, y utilizará la tabla de etiquetas correspondiente.

Recomendación 9 PHP no habilitado	Para el diseño de nuestra red VPLS jerárquica, **no** será necesario utilizar la funcionalidad PHP
Recomendación 10 Calidad de servicio	Para el diseño de la red VPLS jerárquica, se definirán cuatro caudales diferenciados de tráfico en el Core, uno para tráfico no prioritario o de Internet, otro para tráfico de datos (que engloba todas las opciones de AF) un caudal prioritario para tráfico multimedia y de voz, y otro caudal para proteger el tráfico de control de los nodos de red

8. Conclusiones

Las pruebas realizadas en laboratorio, así como las redes reales desplegadas analizadas bajo este proyecto, sitúan a las redes basadas en VPLS jerárquico como muy buenas opciones para solucionar el problema de las redes privadas virtuales.

El actual desarrollo de infraestructuras de red basadas en Ethernet, y la creciente demanda de ancho de banda son los principales argumentos a favor de este tipo de redes.

La relación coste / rendimiento debe ser analizada en detalle en base a los equipos necesarios, los servicios a prestar, y el rendimiento que se quiere realizar sobre la red. Cada caso desprenderá un balance particular. Los equipos de conmutación para Ethernet y multiservicio suelen ser relativamente más económicos que los routers.

La implementación de la pila de protocolos es eficiente, fiable y dinámica. También permite flexibilidad al intercambiar diferentes protocolos sin alterar el conjunto de la arquitectura. El plano de control y el plano de transporte pueden tratarse por separado. Algunos de los protocolos de señalización o de transporte pueden sustituirse flexiblemente.

Existen mecanismos que potencian la escalabilidad de estas redes, que no han sido objeto de este proyecto pero confirman la tendencia a desarrollar tecnología basada en estas arquitecturas.

Como aspecto más crítico a controlar y monitorizar en la red cabe destacar a la posibilidad de bucles en las líneas físicas, que afecten de forma general al servicio completo, ya que se trata de un domino de difusión emulado. A pesar de haber varias técnicas para detectarlos y aislarlos, no se puede olvidar el error humano en la operación diaria de la red.

Por otra parte, la calidad de servicio permite clasificar y priorizar el tráfico conforme a las necesidades de cada usuario, siendo de vital importancia en redes con mucha carga de tráfico.

9. Bibliografía

9.1. RFCs

Documento	Nombre	Fuente
Virtual Private LAN Services over MPLS	draft-ietf-ppvpn-vpls-ldp-0x.txt	http://tools.ietf.org/id/draft-ietf-ppvpn-vpls-ldp-00.txt
Virtual Private LAN Services Using LDP	draft-ietf-l2vpn-vpls-ldp-09	http://tools.ietf.org/html/draft-ietf-l2vpn-vpls-ldp-09
Signaling Standby State of Pseudowire Groups in H-VPLS	draft-pdutta-l2vpn-hvpls-standby-00	http://tools.ietf.org/html/draft-pdutta-l2vpn-hvpls-standby-00
LDP Specification	RFC 3036	http://www.ietf.org/rfc/rfc3036.txt
LDP Applicability	RFC 3037	http://www.ietf.org/rfc/rfc3037.txt
RSVP-TE: Extensions to RSVP for LSP Tunnels	RFC 3209	http://tools.ietf.org/html/rfc3209
Traffic Engineering (TE) Extensions to OSPF Version 2	RFC 3630	http://tools.ietf.org/html/rfc3630
Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)	RFC 4447	http://tools.ietf.org/html/rfc4447
Encapsulation Methods for Transport of Ethernet over MPLS Networks	RFC 4448	http://tools.ietf.org/html/rfc4448
Multiprotocol Extensions for BGP-4	RFC 4760	http://tools.ietf.org/html/rfc4760
Virtual Private LAN Service (VPLS) Using LDP Signaling	RFC 4762	http://tools.ietf.org/html/rfc4762

9.2. Manuales de configuración

Documento	Fuente
Alcatel-Lucent 7750 SR OS Interface Guide Alcatel-Lucent 7750 SR OS Routing Protocols Guide Alcatel-Lucent 7750 SR OS System Mgmt Guide Alcatel-Lucent 7750 SR OS System Basics Guide Alcatel-Lucent 7750 SR OS Services Guide Alcatel-Lucent 7750 SR OS Router Config Guide Alcatel-Lucent 7750 SR OS QoS Guide Alcatel-Lucent 7750 SR OS MPLS Guide	Información de Soporte, Documentación y Guías de configuración incluidas en el paquete de software de la versión utilizada

9.3. Información de proveedores de hardware

Proveedor	Sitio web
Cisco Systems	www.cisco.com
Alcatel - Lucent	www.alcatel-lucent.com
Juniper Networks	www.juniper.net
Huawei Technologies	www.huawei.com

10. Acrónimos

ATM	Asynchronous Transfer Mode.
BGP	Border Gateway Protocol.
BPDU	Bridge Protocol Data Unit.
CIR	Committed Information Rate.
COS	Class Of Service.
CPU	Central Processing Unit.
DSCP	Differentiated Services Code Point.
EIGRP	Enhanced Interior Gateway Routing Protocol.
FR	Frame Relay.
IP	Internet Protocol.
LACP	Link Aggregation Control Protocol.
LAN	Local Area Network.
LDP	Label Distribution Protocol.
LER	Label Edge Router.
LSR	Label Switching Router.
MAC	Media Access Control.
MAN	Metropolitan Area Network.
MBS	Maximum Burst Size.
MPLS	Multi Protocol Label Switching.
NAT	Network Address Translation.
OAM	Operation, Administration and Maintenance.
OSPF	Open Shortest Path First (Gateway Protocol).
PE	Provider Edge.
PVST	Per VLAN STP.
QinQ	Q in Q, VLANs in VLAN.
QOS	Quality Of Service.
RFC	Request For Comment.
RIP	Router information Protocol.
RSVP-TE	Resource Reservation Setup Protocol - Traffic Engineering.
SDH	Synchronous Digital Hierarchy.
SNMP	Simple Network Management Protocol.
SONET	Synchronous Optical Network.
STP	Spanning Tree Protocol.
SW	Software.
TCP	Transmission Control Protocol.
VC	Virtual Channel.
VLAN	Virtual LAN.
VoIP	Voice Over IP.
VPLS	Virtual Private LAN Services.
VPN	Virtual Private Network.