UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

DISEÑO DE REDES DE CAMPUS UTILIZANDO TECNOLOGÍA DE VIRTUALIZACIÓN

INFORME DE SUFICIENCIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ELÉCTRÓNICO

PRESENTADO POR:

VICTOR MANUEL TELLO ARAGON

PROMOCIÓN
1986- I

LIMA – PERÚ
2010

DISEÑO DE REDES DE CAMPUS UTILIZANDO TECNOLOGÍA DE VIRTUALIZACIÓN

SUMARIO

Las redes de comunicación se han convertido en el eje fundamental para que las Empresas construyan una plataforma tecnológica sobre la cual desarrollen sus negocios. Muchas de estas redes nacieron con el objetivo principal de brindar conectividad entre los usuarios basada en el protocolo IP y en la bien desarrollada tecnología Ethernet.

Sin embargo, en su mayoría el diseño de estas redes no contempla desde el inicio el requerimiento de alta disponibilidad, conocido ahora como resiliencia. Con la tendencia actual de integrar todos los servicios de red sobre el protocolo IP, no solo aplicaciones de Voz sobre IP o Telefonía IP, sino también ahora el transporte de vídeo seguridad, control de procesos industriales, almacenamiento, comunicaciones unificadas y otros, este requerimiento no puede seguir siendo ignorado.

En el presente informe, se presenta un planteamiento de mejora de las técnicas de diseño de redes enfocadas al ambiente de campus desde una perspectiva de sistemas, considerando como requerimiento vital la resiliencia a nivel de arquitectura. Para esto se propone un nuevo modelo de diseño de red de campus, utilizando el concepto de **virtualización**, el cuál será desarrollado en base a una tecnología específica disponible en el mercado, mostrando las mejoras en alta disponibilidad, recuperación frente a fallos y la simplificación que se logra con este enfoque. Finalmente analizaremos un caso práctico, los costos asociados de equipamiento y las ventajas que pueden obtener las Empresas al incorporar estas técnicas en el diseño de sus redes de campus.

INDICE

INTRODUCCION

La industria de las redes y las comunicaciones está cambiando debido a que la forma como nos comunicamos está cambiando. Ahora ya no simplemente hablamos, sino que enviamos texto (chat, SMS). Enviamos fotos desde nuestros celulares e inclusive recibimos vídeo (3G). El servicio de televisión por cable, ahora viene empaquetado con la telefonía y el acceso a Internet (Triple Play), y ya se habla ahora de redes inalámbricas de alta capacidad de cuarta generación con alcance metropolitano (Wimax, LTE). Estamos presenciando un nuevo mundo que requiere una nueva red de comunicaciones que se impulse sobre el poder unificador de la tecnología IP.

Esta realidad se está trasladando a las Empresas, quienes empiezan a comprender el tremendo potencial que significa vincular y potenciar sus procesos de negocio con el uso inteligente de esta tecnología. Por ejemplo, una tendencia de la que se habla mucho en estos días es la que se denomina como "Comunicaciones Unificadas" (Unified Communications, UC por sus siglas en inglés) la cual intenta integrar múltiples canales de comunicación (voz, e-mail, vídeo, mensajería instantánea, conferencia, presencia) en una sola interface con el fin de permitir la comunicación en tiempo real entre sus empleados, clientes, socios y proveedores, acelerando por ende la toma de decisiones y mejorando la productividad de la Empresa.

Sin embargo, toda esta visión requiere de la existencia previa de una infraestructura de red que pueda soportar estas nuevas aplicaciones y que brinde la adecuada confiabilidad, acorde a la criticidad que significa colocar todas nuestras aplicaciones y comunicaciones de cualquier tipo, sobre una sola plataforma.

En mi experiencia, las Empresas (independientemente de su tamaño) han intentado desde el inicio construir sus redes mayormente desde el punto de vista de la conectividad y la capacidad de transmisión (rendimiento), descuidando muchas veces los factores de confiabilidad y disponibilidad de las mismas.

Para el caso de la WAN, estos factores son atendidos hoy día mayormente por las Empresas de Telecomunicaciones, quienes ofrecen servicios de transmisión de datos (utilizando tecnología MPLS) con algún contrato de nivel de servicio (SLA) y/o algún esquema de redundancia, en el mejor de los casos.

Pero para el caso de las redes LAN construidas y manejadas por la propia Empresa, especialmente para la red de campus principal, esta prevención casi siempre

se reduce a la provisión de uno que otro elemento redundante en la red o algún contrato de mantenimiento correctivo, que no alcanza la talla necesaria para ofrecer la alta disponibilidad que se requiere hoy en esta infraestructura tan crítica.

El objetivo de este Informe es ofrecer un planteamiento de mejora a las técnicas de diseño de redes de campus que sirva a las Empresas (especialmente las de nuestro medio local) como una guía para la construcción efectiva y eficiente de esta parte de su infraestructura tecnológica, con énfasis en la confiabilidad y disponibilidad que se requieren hoy en día. Estos requerimientos se relacionan en la industria con el término conocido como RESILIENCIA, que podemos definir en forma general como la propiedad de un sistema para adaptarse y recuperarse frente a una falla crítica en alguno de sus componentes. Particularmente, he observado que la resiliencia se deja muchas veces de lado en la práctica de diseño, ya sea por la complejidad o el desconocimiento de las técnicas clásicas de diseño basadas en protocolos de capa 2 o capa 3, o por el alto costo asociado a una infraestructura redundante.

Para superar estos inconvenientes, el planteamiento que propongo en este Informe es la utilización de una tecnología de virtualización basada en el concepto de agrupamiento de conmutadores (**Switch Clustering**, por su expresión en inglés), el cual permite obtener una verdadera alta disponibilidad en el núcleo de la red de campus y a su vez, obtener un balanceo de carga desde el nivel de acceso sin involucrar el uso de las técnicas clásicas que complican innecesariamente el diseño.

CAPÍTULO I
PLANTEAMIENTO DE INGENIERÍA

1.1 Descripción del Problema

El problema sobre el que se centra el presente informe es lograr una mejora en las técnicas tradicionales de diseño de redes de campus, desde el punto de vista de la característica conocida como resiliencia.

Como se explicará en el capítulo "Marco Teórico Conceptual", las técnicas de diseño de redes de campus se apoyan en los protocolos de capa 2 (STP, RSTP) y de capa 3 (RIP, OSPF) para brindar alta disponibilidad y una capacidad automática de respuesta frente a fallos en algún componente de la red. Esta recuperación se produce en algunos casos en decenas de segundos o en mayor tiempo inclusive, lo cual es inadecuado para el funcionamiento de las aplicaciones actuales que trabajan en tiempo real. Es por eso necesario, ofrecer un enfoque alternativo que solucione esta debilidad.

1.2 Objetivo del Informe

El objetivo principal de este informe es ofrecer un planteamiento de mejora de las técnicas tradicionales de diseño de redes de campus en lo que respecta al requerimiento de resiliencia, mediante el uso de una tecnología de virtualización disponible en el mercado.

En mi experiencia profesional, el requerimiento de resiliencia es normalmente postergado o dejado de lado por las Organizaciones, debido al poco conocimiento de las técnicas de diseño asociadas a cumplir con este requerimiento o al mayor costo adicional que involucra añadir componentes redundantes a la red, en busca de lograr alta disponibilidad.

Mi objetivo es mostrar en este informe la utilización de una tecnología de virtualización conocida como **SMLT** (Split Multilink Trunking) la cual supera ambos obstáculos mediante una técnica de diseño más simple, disponible a costos más adecuados a nuestra realidad local.

1.3 Alcances del Informe

En primer lugar, se hará un uso de los conceptos de conmutación y enrutamiento en redes LAN basadas en protocolos TCP/IP cuando sea necesario, sin entrar en una

explicación teórica de los mismos, la cual se puede obtener revisando la bibliografía que se indica al final de este documento.

También debo decir que este informe no pretende cubrir en forma exhaustiva todas las técnicas y variantes de enfoque en lo que respecta al diseño de redes de campus, debido a lo extenso que significaría abordar tal tarea. Específicamente considero que los siguientes aspectos relativos al diseño de redes de campus, no serán cubiertos en este documento:

- Calidad de Servicio (QoS)
- Seguridad en Redes
- Manejo del tráfico multidireccionado (multicasting)
- Programación de equipos de red a detalle
- Técnicas de cableado estructurado u otros elementos propios del ambiente circundante.

Conviene aclarar que el planteamiento de diseño que propongo tiene un enfoque de sistemas, es decir está basado en bloques de construcción. Sin embargo, en un capítulo dedicado si se presenta un ejemplo de la programación de los equipos involucrados en el diseño a nivel de comandos, con la intención de mostrar objetivamente la simplicidad de la implementación.

Finalmente en la parte de costos, solo se consideran para el ejemplo de diseño mostrado los costos referentes a los equipos activos que conforman la red de campus.

1.4 Síntesis del Informe

- Capítulo 1: **Planteamiento del Problema de Ingeniería**. En este capítulo se describe el problema o situación que se desea mejorar en lo referente al diseño de redes de campus. Luego se presenta el objetivo principal del Informe y una descripción sintetizada de la estructura del informe.
- Capítulo 2: **Marco Teórico Conceptual**. En este capítulo, se realizará una descripción de las técnicas clásicas actuales de diseño de campus. Se revisará las ventajas de utilizar un enfoque de diseño basado en bloques, en cuanto a escalabilidad y disponibilidad.

  Se mostrará la forma de operación en caso de fallas en el bloque básico utilizando el protocolo STP y los puntos débiles de tal proceso.

  Luego se ampliará el tema con un repaso teórico del funcionamiento y operación del protocolo STP y de su nueva versión RSTP (Rapid Spanning Tree) y al final se realizará una crítica del modelo clásico presentado.

- Capítulo 3: **Virtualización en Redes de Campus**. Se planteará la necesidad de la utilización de tecnologías de virtualización que superen las desventajas del diseño clásico, como es el caso del agrupamiento de conmutadores.
  Específicamente, se examinará la utilización de la tecnología **SMLT** (**Split Multilink Trunking**) y su aplicación en el diseño de redes de campus con fuertes características de resiliencia. Se propone esta tecnología debido a su madurez en el mercado y su comprobada eficacia.
- Capítulo 4: **Ejemplo de Diseño de una Red de Campus**. Para ilustrar los conceptos y tecnologías explicados en capítulos anteriores, se mostrará un ejemplo de diseño de una red de campus para 1,000 usuarios utilizando la tecnología SMLT y dispositivos de red de mediana escala disponibles en el mercado.
  Se explicará las variadas consideraciones de diseño que deben tomarse en cuenta y al final, se presentará una plantilla de los comandos de programación utilizados para la puesta en funcionamiento de la red.
- Capítulo 5: **Costos y Presupuestos**. En esta parte se presenta una relación de los costos asociados a los componentes de red involucrados en el diseño presentado en el capítulo anterior, y una comparación relativa a otras opciones disponibles en el mercado.
- Capítulo 6: **Conclusiones**. Finalmente, se presentará una lista de las conclusiones más importantes obtenidas de la utilización del concepto de virtualización en el diseño de redes de campus y mi recomendación final.

CAPÍTULO II
MARCO TEÓRICO CONCEPTUAL

2.1 Objetivos

En este capítulo, presentare una revisión del enfoque clásico de diseño de redes de campus basado en el modelo jerárquico, el cual ya ha demostrado características de flexibilidad y facilidades de operación y mantenimiento que le siguen dando vigencia. Previamente hare una breve revisión de los conceptos y tecnologías de red asociadas a esta parte de la infraestructura. Al final, realizare un análisis crítico del modelo jerárquico para preparar la exposición que continuará en el capítulo siguiente

2.2 Panorama de las Redes de Campus

El término "campus" es utilizado en el mercado de las redes para referirse en forma genérica a un edificio o grupo de edificios conectados a una sola red empresarial que consiste de muchas redes LAN (Local Area Network).

La característica distintiva de un ambiente de campus es que la Empresa que lo posee, usualmente también es propietaria de los medios de transmisión (cables de cobre o fibra óptica) utilizados para la interconexión física. La tecnología que se ha impuesto por su flexibilidad es el Ethernet, desde su versión inicial a 10 Mbps, pasando por el Fast Ethernet hasta la adopción dominante en los centros de datos con Gigabit Ethernet y 10Gibabit Ethernet, y opciones de mayor capacidad actualmente ya casi por estas disponibles en el mercado.

2.2.1 Los retos iníciales

Originalmente, las redes de campus consistían en una simple red LAN a la cual los usuarios eran añadidos. Debido a limitaciones de distancia en la tecnología, los usuarios eran confinados a un edificio o a varios edificios próximos entre sí. El principal criterio de diseño era (y lamentablemente, lo sigue siendo en muchas empresas) solo la conectividad a la red.

A parte de la conectividad, el tema del rendimiento (performance) fue otro obstáculo histórico a superar en la redes de campus, lo cual se logro con la aparición de los dispositivos de red conocidos como conmutadores (switches, en inglés), que permitían dividir los dominios de colisión de tramas Ethernet para segmentar la red en forma lógica dentro de estos dispositivos y evitar que se inunde la red con tráfico innecesario.

Sin embargo, existía siempre de forma creciente el denominado tráfico de difusión (broadcast, en inglés) de capa 3 que necesariamente debía llegar a todos los usuarios, por ejemplo, para publicar un servicio de red. A medida que la cantidad de usuarios se incrementaba, también crecía este tráfico de difusión, lo cual reducía el ancho de banda disponible para los usuarios afectando el rendimiento esperado. Inclusive se llegaba al límite de presentarse tormentas de difusión, que podían hacer inutilizable la red. Para solucionar este tema, existían inicialmente dos enfoques:

- Utilizar los dispositivos de red conocidos como enrutadores (routers, en inglés) ubicados al borde de la red WAN, para segmentar lógicamente las redes LAN a nivel de capa 3, conectando varios conmutadores entre sí. Esto tenía el riesgo de afectar el rendimiento, dado que la capacidad de procesamiento de estos dispositivos era inferior a la de los conmutadores, porque las decisiones de enrutamiento estaban basadas en un algoritmo de software.
- Utilizar una funcionalidad de los conmutadores conocida como VLAN (red virtual de área local o "virtual LAN" en inglés), que permitía a estos dispositivos dividirse lógicamente de alguna manera (normalmente basado en el agrupamiento de puertos físicos) de tal forma que se constituían varios dominios de difusión, separados unos de otros, sin tener que utilizar dispositivos separados para cada dominio. Esta funcionalidad permitía inclusive extender el concepto de VLAN a varios conmutadores de la red. Sin embargo, de necesitarse la interconexión de tráfico entre estas VLAN, nuevamente era necesario la participación de los enrutadores.

2.2.2 **Conmutación de Múltiples Capas**

Con el escenario anterior, la industria se enfoco en desarrollar nuevos dispositivos, lo cual desencadeno en los conceptos de conmutación de capa 3 (enrutamiento basado en hardware) y conmutación de capa 4 (una variante del enrutamiento de capa 3 que considera la aplicación que se transmite).

Aparecieron entonces los denominados conmutadores multicapa, que combinaban las funcionalidades de conmutación de capa 2 de los dispositivos tradicionales con las nuevas funcionalidades de conmutación de capa 3 y 4 requeridas. De esta forma, el tráfico de campus era procesado a alta velocidad, mientras que al mismo tiempo se satisfacían estos nuevos requerimientos.

Esta combinación no solo resolvía los problemas de rendimiento, sino que evitaba que se formen cuellos de botella en la red. El concepto principal de la conmutación de múltiples capas era "enrutar la primera vez y conmutar todo lo demás". Cada fabricante enfoco el diseño de sus propios dispositivos en este objetivo, y la industria y el mercado se vieron altamente beneficiados.

2.2.3 El Tráfico en la Red

Por lo explicado hasta ahora, debe observarse que la conectividad y el rendimiento fueron los requerimientos originales en la redes de campus. Eso estuvo bien, pero independientemente de la tecnología subyacente, el principal reto que enfrenta un diseñador hoy día es que su red funcione efectiva y eficientemente con respecto a las aplicaciones. Para lograr esta meta, es necesario entender y manejar adecuadamente el flujo de tráfico a través de la red.

Los dispositivos conectados y las aplicaciones de software que corren en la red, todas generan tráfico de datos. Las aplicaciones más comunes son la transferencia de archivos y el correo electrónico, las cuales siguen un patrón conocido y tienen una ruta determinada. Las aplicaciones más complejas y demandantes que incluyen la videoconferencia o la transmisión unidireccional de vídeo, no son fáciles de predecir. Esto nos lleva a caracterizar el tráfico de una forma más genérica.

Históricamente, las redes LAN tuvieron como enfoque atender aplicaciones bajo un modelo de tráfico del tipo 80/20 (80% del tráfico es local y 20% del tráfico es remoto). La idea en ese entonces era localizar el tráfico dentro de grupos de trabajo y minimizar el tráfico hacia el centro de la red (backbone, en inglés) para evitar sobrecargarlo. Este objetivo de diseño se lograba reagrupando a los usuarios que compartían información y añadiendo servidores locales que replicarán información remota dentro de ese grupo.

Posteriormente, con el auge creciente del Internet en los años 90s esta tendencia en el tráfico se fue revirtiendo y el modelo de tráfico cambio al tipo 20/80 (20% del tráfico local y 80% del tráfico centralizado) con lo cual el modelo de diseño para las redes de campus se fue formando en base a las buenas prácticas impulsadas por los fabricantes líderes del mercado (Cisco Systems por ejemplo) que tomaban en cuenta esta tendencia hacía el tráfico centralizado, la aparición de dispositivos de red cada vez más potentes y la división por capas conocida del modelo de referencia OSI.

2.3 Modelo Clásico de Diseño

El modelo clásico para el diseño de redes de campus se inicio con los servicios y la inteligencia de red de capa 3 en el centro de la red y compartiendo ancho de banda a nivel de los usuarios. Con el desarrollo de la tecnología, los servicios de capa 3 se fueron repartiendo o distribuyendo a través de la red y la conmutación de capa 2 llego al nivel del usuario de forma masiva.

Es así que el modelo de diseño fue tomando un carácter jerárquico, es decir, fue dividiendo en capas las funciones y servicios que se realizaban en la red. Las capas que se consideraban eran las siguientes:

- Capa de Acceso

- Capa de Distribución
- Capa Núcleo

Esta simple definición se mantiene en el diseño de redes hasta la época actual con pocas variantes, mayormente impulsada por el fabricante Cisco quien tenía la preferencia de plantear este modelo dado su predominancia comercial en el mercado de los enrutadores utilizados en la capa de distribución (ver figura 2.1).

Sin embargo, el modelo jerárquico ha demostrado su utilidad como una buena práctica en el diseño de redes porque simplifica la gestión de la red y permite un crecimiento controlado. Estas características permiten a los diseñadores plantear un esquema de construcción basado en bloques, lo cual será analizado en la sección siguiente.

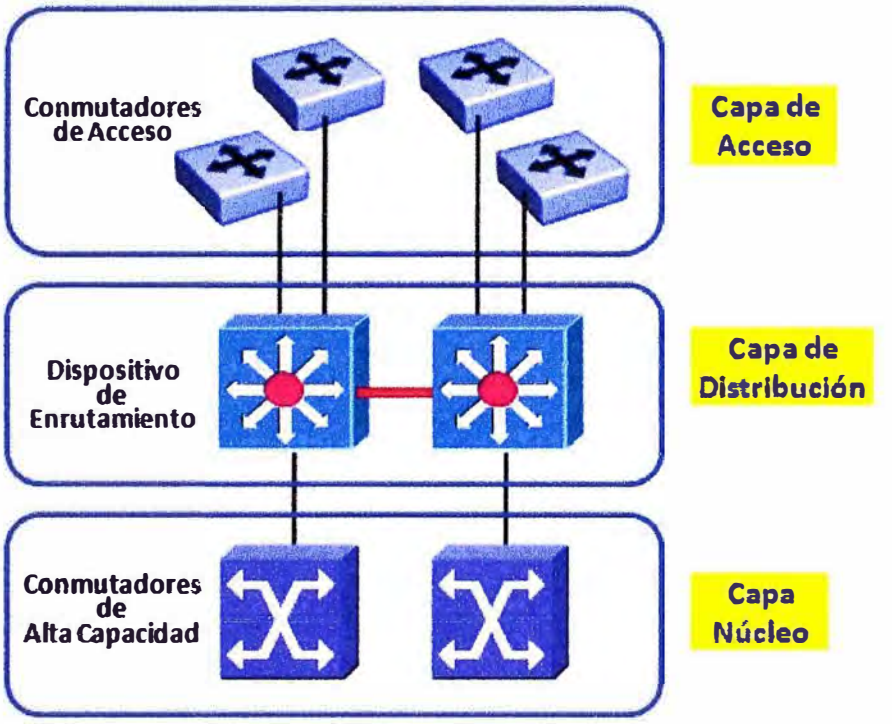


Fig. 2.1 Modelo Jerárquico

A continuación, se describen las 3 capas consideradas en el modelo jerárquico de diseño de redes de campus.

2.3.1 Capa de Acceso

La capa de acceso, como su nombre lo indica, es el punto o el nivel en el cual los usuarios ingresan o son permitidos de ingresar a la red. Esta capa puede proveer un nivel más sofisticado de control, cuando se le añaden características de filtrado o listas de acceso como las usadas en un enrutador. Sin embargo, la función clave sigue siendo proveer la conectividad de entrada a la red para los usuarios finales, y otras funciones adicionales como por ejemplo:

- La segmentación del ancho de banda en capa 2.
- La membresía a una VLAN.
- El filtrado de tráfico por dirección MAC.

El principal criterio para la elección de los dispositivos usados en esta capa es que brinden estas funciones a bajo costo y con una alta densidad de puertos de red.

2.3.2 Capa de Distribución

La capa de distribución marca la división entre la capa de acceso y la capa núcleo, y es donde se realiza un gran procesamiento y manipulación de los paquetes que viajan por la red. En el ambiente de campus, la capa de distribución puede representar una variedad de funciones, tales como:

- Concentración y enrutamiento entre VLANs.
- Acceso a servicios de red para grupos de trabajo.
- Definición de dominios de difusión.
- Seguridad

La función de esta capa es brindar la conectividad basada en políticas, es decir, en reglas inteligentes de manejo y procesamiento los paquetes de red.

2.3.3 Capa Núcleo

A la capa núcleo se le denomino como la espina dorsal (backbone, por su término en inglés) de la red, asignándole como principal propósito el conmutar el tráfico de la forma más rápida posible. A esta capa se le retiro intencionalmente, toda función relativa a la manipulación y procesamiento de paquetes, y en general cualquier tarea que pudiese afectar la función de conmutar tráfico en capa 2, lo más rápido posible.

La capa núcleo quedo como responsable de conectar “bloques” o regiones de la red entre sí, sin agregar mayor sofisticación que la alta rapidez y rendimiento en lograr dicha tarea. Veremos esto con mayor detalle en la próxima sección.

2.4 Bloques de Construcción

El modelo jerárquico explicado anteriormente no reflejaba por sí solo una metodología. Para lograr esto, era necesario brindar un enfoque en el diseño de redes que imitará la construcción de una obra civil en Ingeniería y que se basará en bloques repetibles interconectados entre sí como ladrillos, que simplificarán este proceso de construcción. El fabricante Cisco fue el primero en proponer este enfoque, y sin convertirse en una marca registrada, fue seguido por sus competidores y por la industria en general que lo adopto como una buena práctica.

El enfoque basado en bloques de construcción, aunque simple en su concepción, ha probado su utilidad y eficacia, al permitir el diseño de redes de campus de escala

diversa, especialmente aquellas de tamaño muy grande que pueden llegar a los miles y decenas de miles de usuarios.

Actualmente, el fabricante Cisco ha impulsado en su bibliografía un nuevo modelo de diseño de redes empresariales (incluyendo campus y la MAN/WAN inclusive) denominado ECNM (Enterprise Composite Network Model). El ECNM es en mi opinión, una generalización de los bloques de construcción que explicare más adelante, porque mantiene los conceptos básicos que me interesa resaltar en este informe y por tanto, no afecta las conclusiones que planteo al final del informe. Para el interesado en revisar el ECNM de Cisco, le pido referirse a la bibliografía de este fabricante que presento al final del documento.

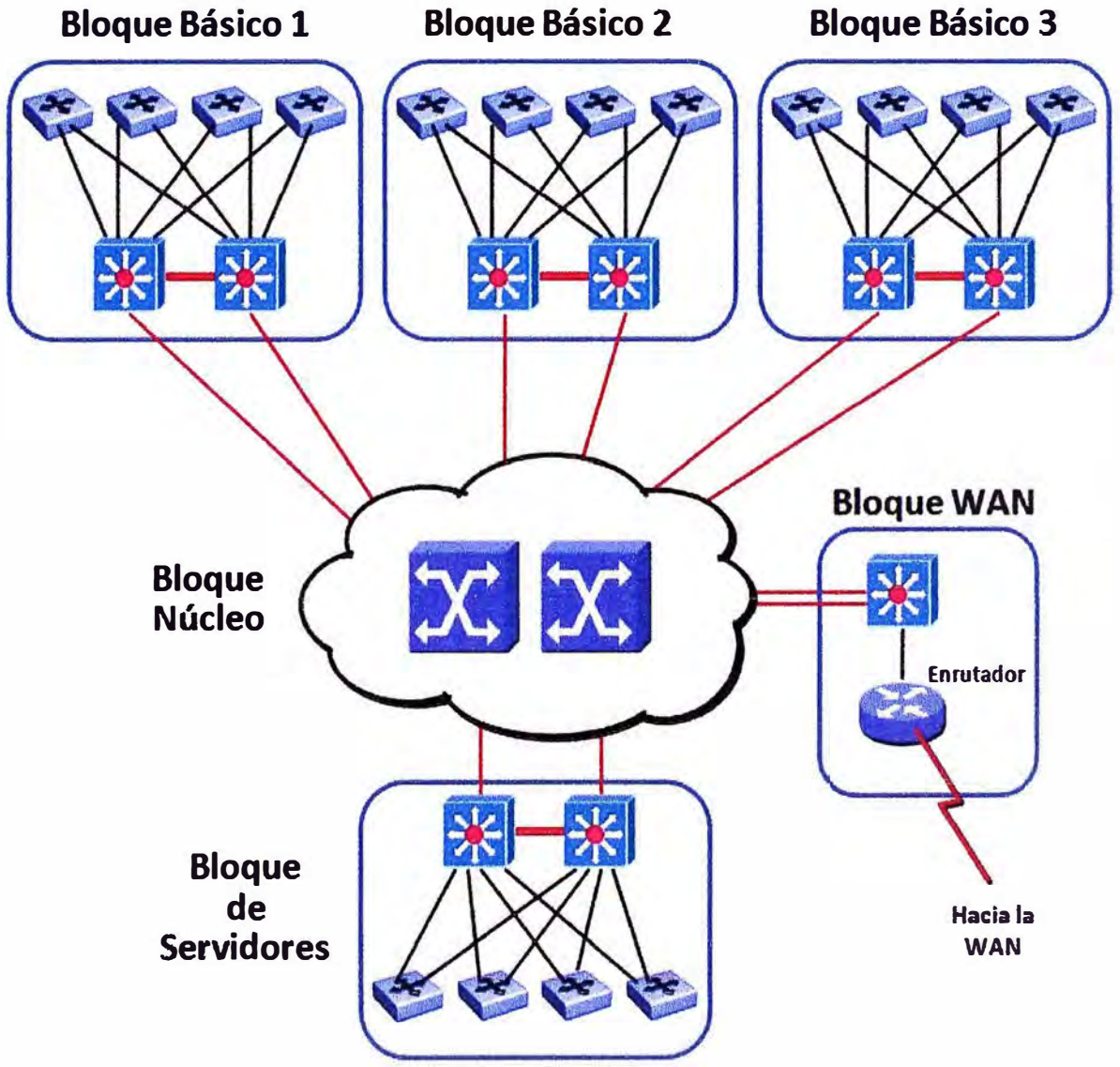


Fig. 2.2 Bloques de Construcción en el Diseño de Redes de Campus

Los bloques de construcción más comúnmente usados en el diseño de redes son los siguientes:

- Bloque Básico
- Bloque Núcleo
- Bloque de Servidores
- Bloque WAN

En la figura 2.2 anterior se presenta un esquema de conectividad de estos bloques, que refleja su posicionamiento en un diseño típico. Los bloques de servidores y WAN son casos especiales del bloque básico, y por lo tanto no serán analizados con detalle. Mi análisis se centrará en los dos primeros bloques según explico a continuación.

2.4.1 Bloque Básico

2.4.1.1 Descripción

El bloque básico contiene una combinación balanceada y escalable de servicios de capa 2 y capa 3, es decir, capacidad de conmutación y enrutamiento. El bloque básico de la figura 2.3 siguiente, previene que todo el tráfico de difusión así como otros problemas de nivel de red, atraviesen el bloque núcleo y alcancen otros bloques básicos.

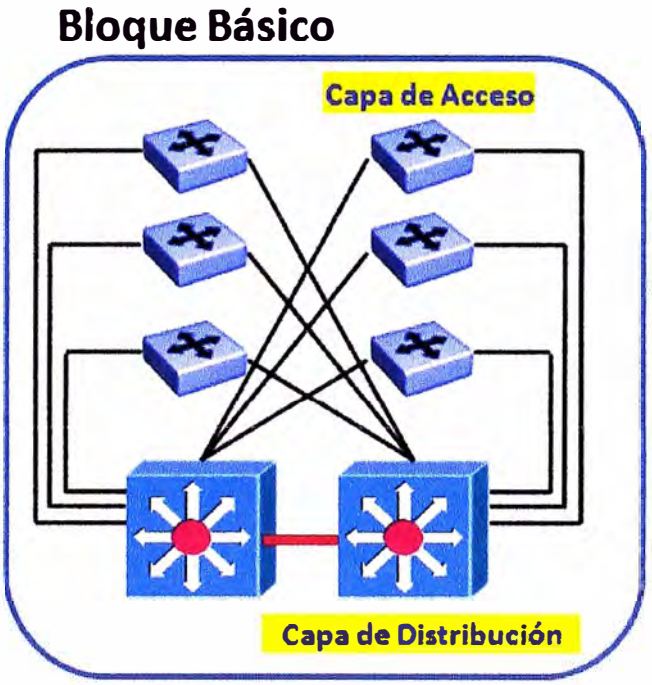


Fig. 2.3 Bloque Básico

Los conmutadores de capa 2 (ubicados en los cuartos de cableado) conectan los usuarios a la red al nivel de la capa de acceso y proveen ancho de banda dedicado en cada puerto. Estos dispositivos se interconectan entre sí al nivel de la capa de distribución. El dispositivo de la capa de distribución provee esta conectividad y actúa como punto de concentración de todos los conmutadores. Este dispositivo normalmente es un conmutador multicapa.

La capa de distribución también provee funcionalidad de capa 3, lo cual significa que hace enrutamiento de paquetes y otros servicios de red. La capa de distribución protege al bloque básico contra fallas producidas en otras partes de la red. Por ejemplo, si un bloque básico experimenta una tormenta de difusión, el conmutador multicapa previene que la tormenta se propague al núcleo y a otras partes de la red. Cada bloque es protegido de las fallas que ocurren en otros bloques. De cualquier modo, el bloque donde se produce la falla aún experimenta problemas hasta que el dispositivo que genera ese tráfico es encontrado y reparado. Actualmente, los dispositivos conmutadores cuentan con umbrales de filtrado de este tráfico de difusión que ayudan a reducir la probabilidad de que existan estos problemas.

2.4.1.2 Características

Los conmutadores de la capa de acceso pueden soportar una o más subredes de capa 3. Una subred debe residir dentro de un dominio de difusión de capa 2. Esto significa que la regla general en el bloque básico es que todas las estaciones que pertenecen a una misma VLAN deben tener asignadas direcciones de red (direcciones IP) dentro de la misma subred.

La propiedad de una VLAN de aislar el tráfico de difusión es la característica que permite que las VLANs sean identificadas una a una con las subredes. Por ejemplo, el protocolo ARP (Address Resolution Protocol, en inglés) se propaga solo dentro de la VLAN donde se origino el pedido.

Todas las subredes deben terminar en un dispositivo de capa 3, tal como el motor de enrutamiento existente dentro de un conmutador multicapa. Para que una trama de capa 2 pueda llegar a conectarse a dispositivos en otras VLANs, debe necesariamente atravesar ese motor de enrutamiento. En el modelo propuesto en el bloque básico, las VLANs no se deben extender más allá del conmutador multicapa de la capa de distribución.

Adicionalmente, el bloque básico plantea que existan conexiones redundantes desde los conmutadores de acceso hacia los conmutadores de distribución, para mantener la resiliencia en la red. El protocolo STP (o su versión mejorada, el protocolo RSTP) es el encargado de permitir que estos enlaces redundantes existan y a la vez, que se prevengan los indeseables lazos de red que comentaremos más adelante, cuando se analice con detalle este protocolo. Al menos, una instancia independiente de este protocolo deberá operar por cada bloque básico considerado en la red.

Finalmente, es una buena práctica que exista una sola VLAN dedicada para la administración de los dispositivos en todos los bloques de la red (normalmente la VLAN por defecto, VLAN 1). Dado que la regla del bloque básico es que no se propaguen las

VLANs a través del núcleo, para lograr que exista una sola VLAN de administración será necesario crear una conectividad directa entre la VLAN 1 de cada bloque básico, lo cual se considera una excepción aceptable en la topología de los bloques de construcción.

2.4.1.3 Dimensionamiento

El tamaño del bloque básico, en términos de la cantidad de dispositivos que puede contener, es flexible. Sin embargo, el número de conmutadores que colapsan en la capa de distribución depende de factores limitantes como:

- Diferentes tipos y patrones de tráfico, lo que se refleja en diferentes tipos y capacidades de los enlaces de subida (uplinks, en inglés).
- La capacidad de conmutación multicapa de los dispositivos ubicados en la capa de distribución. Esto depende mucho del dispositivo elegido y se mide normalmente en MPPS (millones de paquetes por segundo).
- Numero de puertos de concentración en la capa de distribución.
- Número de usuarios o grupos de usuarios por cada conmutador de la capa de acceso.

El bloque básico fue creado pensando en soportar una cantidad aproximada de 2,000 usuarios por bloque, lo cual es una cantidad muy grande para el tamaño de Empresas en nuestro país. Actualmente, los dispositivos de red existentes tienen suficiente capacidad de procesamiento para soportar estos niveles. Luego, la decisión del tamaño de bloque básico, en la práctica, está más basada en costos que en características técnicas, dado que cuanto más grande es un bloque básico, más sofisticados y caros son los dispositivos requeridos para su construcción. La decisión de balance es dividir una red en varios bloques básicos (de entre 200 a 500 usuarios cada uno, como máximo), en vez de pretender juntar todo en un solo gran bloque.

2.4.2 Bloque Núcleo

2.4.2.1 Descripción

Un bloque núcleo es requerido cuando existen dos o más bloques básicos. El bloque núcleo es responsable de transferir tráfico entre los bloques básicos a nivel de capa 2, sin hacer el uso de operaciones de procesamiento intensivo como es el caso del enrutamiento. Todo el tráfico que va y viene desde los bloques básicos (ú otra variante de estos, como el bloque de servidores o el bloque WAN), debe pasar a través del núcleo.

Debido a este modelo, el núcleo soporta mucho más tráfico que cualquier otro bloque, y por ende, los dispositivos núcleo deberán tener una muy alta capacidad de conmutación y alto rendimiento. En el pasado, se utilizaron varias tecnologías de red en el núcleo, pero para este informe, solo analizaremos un núcleo Ethernet (Gigabit

Ethernet, 10 Gigabit Ethernet o superior). En la figura 2.4 siguiente, se muestra un bloque núcleo que interconecta a dos bloques básicos.

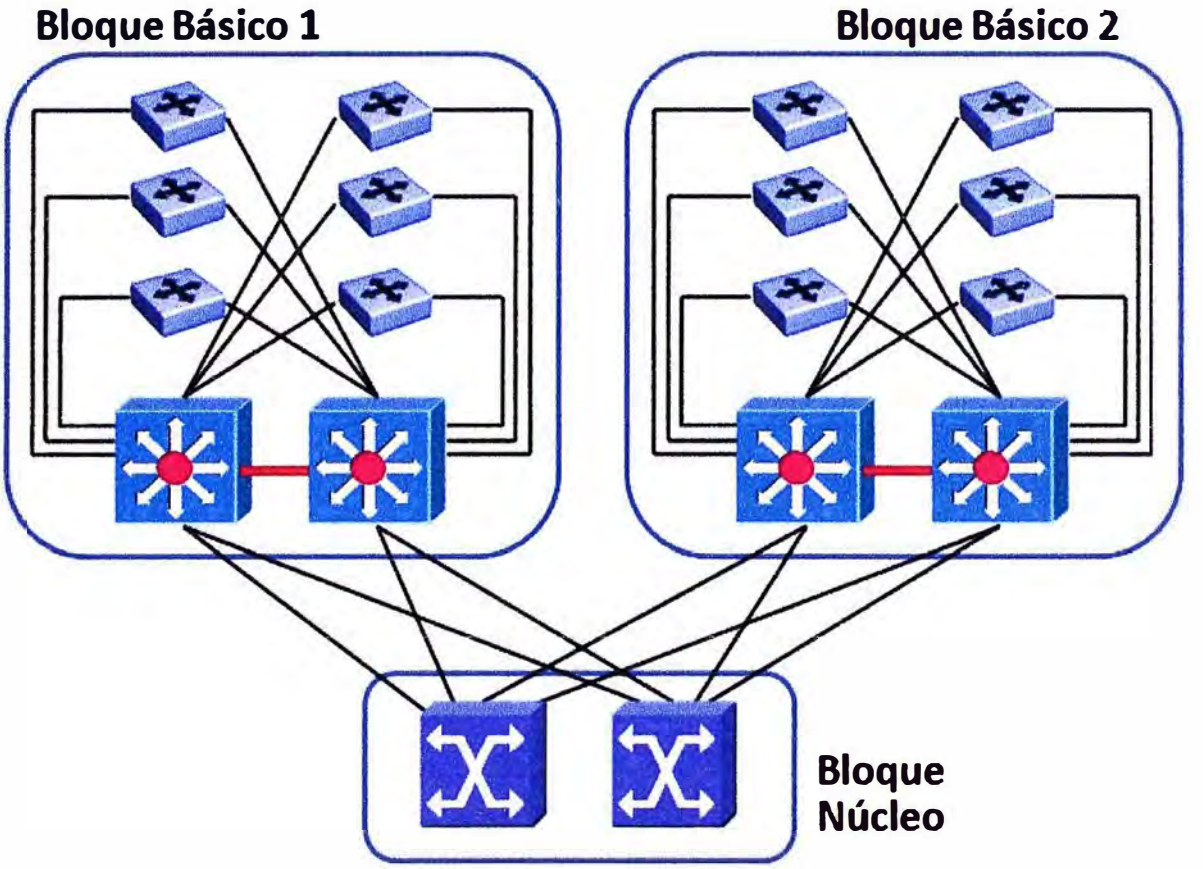


Fig. 2.4 Interconexión con el Bloque Núcleo

Debido a que un conmutador de distribución (ubicado en la parte superior de un bloque básico) proporciona funcionalidad de capa 3 (es decir, se comporta como un enrutador), entonces una subred de núcleo (que equivale a una VLAN de núcleo) deberá existir para interconectar todos los conmutadores de distribución con el conmutador de núcleo. Cada conmutador de distribución deberá enrutar el tráfico entre la subred (VLAN) de núcleo y las subredes (VLANs) locales que son propias de cada bloque básico.

El bloque núcleo podría consistir de una sola subred, sin embargo, para brindar resiliencia y para efectos de tener un balanceo de carga en capa 3, al menos dos subredes deben ser configuradas. Debido a que las VLANs terminan en el dispositivo de la capa de distribución, los enlaces que van al núcleo no son enlaces de troncal de VLANs, y por tanto el tráfico es enrutado hacia el núcleo. Luego, los enlaces de núcleo no transportan múltiples subredes por cada uno.

Normalmente el núcleo está constituido de dos conmutadores cada uno soportando una diferente subred. El medio entre los conmutadores de distribución y los

conmutadores de núcleo debe ser capaz de soportar la cantidad de tráfico manejado por todos los conmutadores de distribución a la vez.

Si los dos conmutadores de núcleo estuvieran enlazados entre sí (ambos en una sola subred), esta conexión debería ser capaz para soportar la cantidad de tráfico que se transmite entre los bloques básicos conectados a ambos lados.

El diseño del núcleo debiera considerar la utilización promedio de los enlaces y estar preparados para permitir el crecimiento futuro del tráfico. Existen dos tipos de diseño para el bloque núcleo: núcleo colapsado y núcleo dual.

2.4.2.2 Núcleo Colapsado

El núcleo colapsado existe cuando las funciones de capa de distribución y capa de núcleo del modelo jerárquico son realizadas en un mismo dispositivo. Un diseño colapsado es lo prevalente en las redes de campus pequeñas. Aunque las funciones de cada capa están en el mismo dispositivo, están se mantienen independientes en su operación. La figura 2.5 siguiente muestra un ejemplo de un núcleo colapsado.

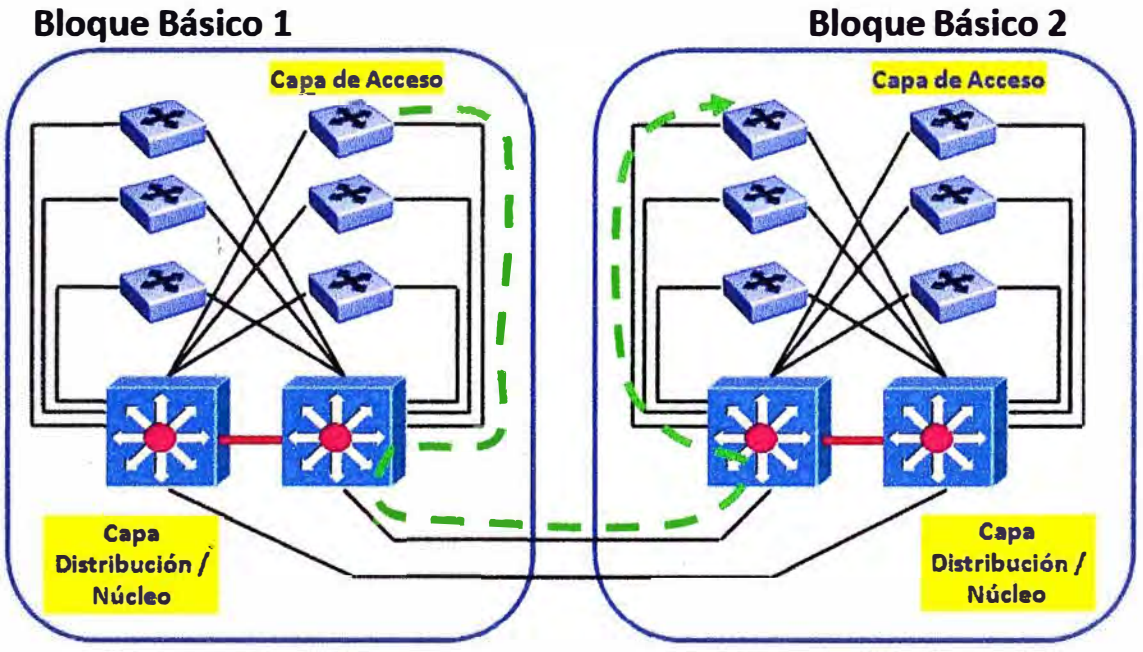


Fig. 2.5 Interconexión de Núcleo Colapsado

En este diseño, cada conmutador de la capa de acceso tiene un enlace redundante a cada conmutador de la capa de distribución. En cada conmutador de acceso puede haber configurada más de una VLAN (cada VLAN corresponde a una subred). De cualquier forma, todas las VLANs terminan en los conmutadores de distribución.

Los enlaces redundantes de subida solo ofrecen resiliencia de capa 2 entre los conmutadores de acceso y distribución. El protocolo STP es el encargado de bloquear uno de los enlaces redundantes para efectos de prevenir los lazos.

La redundancia en capa 3 es provista por los 2 conmutadores de distribución que en conjunto corren un protocolo como HSRP desarrollado por Cisco o el VRRP que corresponde al RFC 3768. El objetivo de estos protocolos es proteger a los usuarios de una falla en el nodo de salida por defecto (default gateway, en inglés) utilizado por el protocolo IP para la transferencia de datos hacia redes externas. En este caso, se protege todas las subredes ubicadas debajo de los 2 conmutadores de distribución. HSRP (o VRRP) crean el concepto de "enrutador virtual" o "enrutador fantasma", al cual se asigna la dirección IP del nodo de salida por defecto, de tal forma que uno de los conmutadores asume este rol y en caso de que este falle, el segundo lo reemplaza de forma automática. El protocolo VRRP será revisado con más detalle en el capítulo siguiente.

2.4.2.3 Núcleo Dual

Una configuración de núcleo dual es necesaria cuando existen 2 o más bloques básicos en el campus y se requiere que existan conexiones redundantes hacia el núcleo.

La figura 2.6 siguiente muestra una configuración dual en la cual el núcleo contiene solamente 2 conmutadores en el centro. Estos conmutadores no están conectados entre ellos para evitar que se produzcan lazos de red.

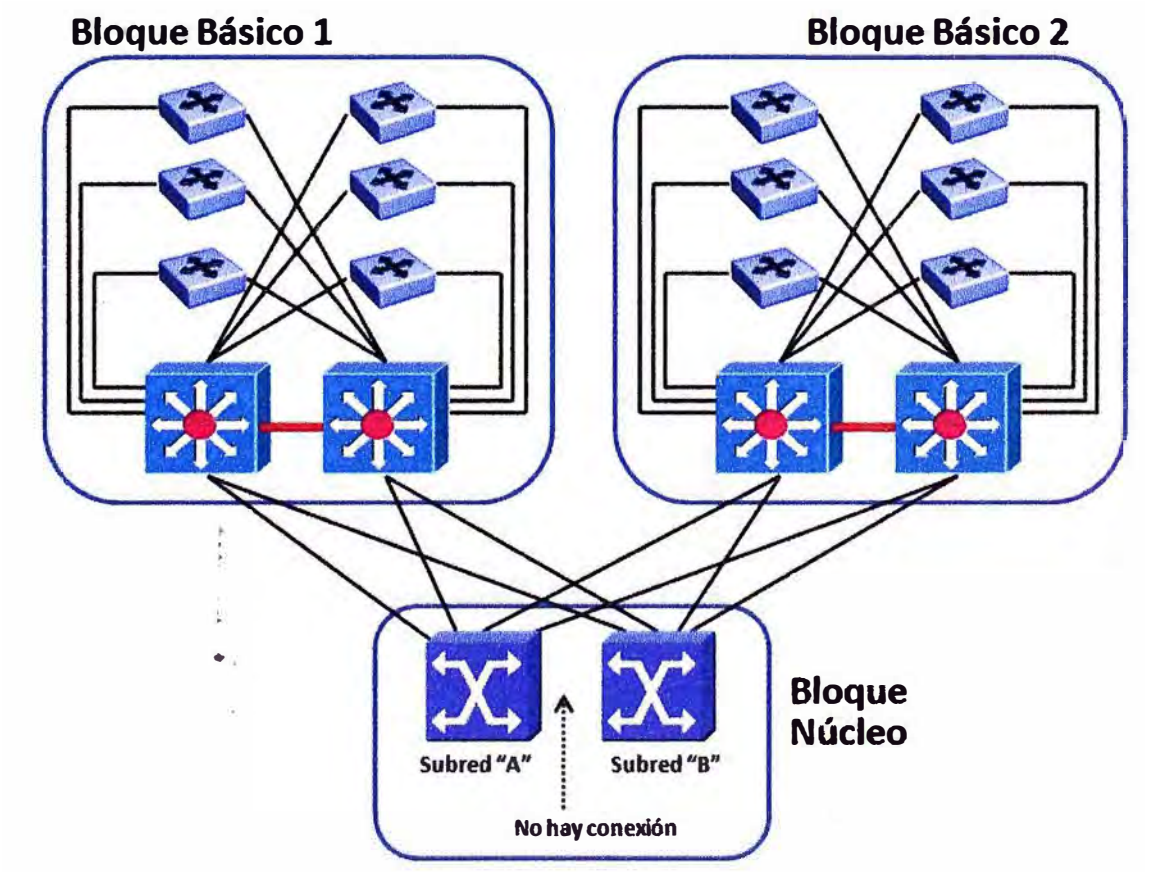


Fig. 2.6 Interconexión con Núcleo Dual

Una topología de núcleo dual provee a la capa de distribución, dos rutas de igual costo en capa 3 y dos veces el ancho de banda respecto a usar un solo enlace. En este caso, cada conmutador de núcleo transporta un número simétrico de subredes a la función de capa 3 de cada conmutador de distribución.

Cada bloque básico esta redundantemente enlazado a ambos conmutadores de núcleo, lo que permite tener rutas de capa 3 distintas pero con el mismo peso. Si un dispositivo de núcleo falla, la convergencia no es un problema, debido a que las tablas de ruteo en los conmutadores de distribución ya tienen una ruta de capa 3 establecida hacia el otro dispositivo de núcleo.

De esta forma, el protocolo de enrutamiento de capa 3 (que no corre en el núcleo) selecciona el enlace que se utiliza para atravesar el núcleo. Obsérvese que, el protocolo HSRP trabaja mirando hacia dentro de cada bloque básico y el protocolo STP no es necesario en el núcleo, porque no hay enlaces redundantes entre los conmutadores ubicados allí.

2.4.2.4 Dimensionamiento

Debido a que los dispositivos de capa 3 (ubicados en la capa de distribución) aíslan el núcleo, el protocolo de enrutamiento entre ellos es utilizado para mantener el estado actual de la red. A medida que el protocolo de enrutamiento envía sus actualizaciones y cambios a los otros enrutadores a través de la red, la topología de la red también puede cambiar.

Cuanto más enrutadores se conecten al núcleo, tanto más tiempo tomará que estas actualizaciones y cambios se propaguen a través de la red y cambie la topología. También puede suceder que uno o más enrutadores se conecten a la WAN o al Internet, lo cual añade más fuentes de actualizaciones y cambios topológicos. El tipo de protocolo de enrutamiento que se use en los dispositivos de capa 3, determinar el número de dispositivos de distribución que pueden ser conectados al núcleo.

TABLA N° 2.1 Comparación entre Protocolos de Enrutamiento

Protocolo de Enrutamiento	Max. # de Pares de Enrutamiento	# de Subredes en Núcleo	Max. # de Bloques soportados
OSPF	50	2	25
EIGRP	50	2	25
RIP	30	2	15

La tabla 2.1 siguiente muestra estos datos para los protocolos de enrutamiento más utilizados y el número máximo de enrutadores emparejados (peer routers, en inglés) para los cuales estos protocolos pueden mantener información de estado.

Todos los bloques básicos (incluyendo variantes como un bloque de servidores o un bloque WAN) deben ser incluidos en la cuenta del número máximo de bloques básicos indicado en la tabla anterior. Los datos que aparecen son los máximos teóricos para el número total de parejas soportadas por cada protocolo de enrutamiento. En la práctica, el número máximo de parejas debería ser un número cercano a 15.

2.5 Escalabilidad del Bloque Núcleo

En la sección anterior, se determino que el bloque núcleo basado en conmutadores de capa 2 tenía limitaciones. Vamos a revisar esto con más detalle y analizar a continuación que sucedería con la escalabilidad si se utilizan conmutadores multicapa en su reemplazo.

2.5.1 Núcleo de capa 2

Utilizar conmutadores de capa 2 en el núcleo es una solución muy costo-efectiva y de alto rendimiento, pero tiene limitaciones que debemos analizar.

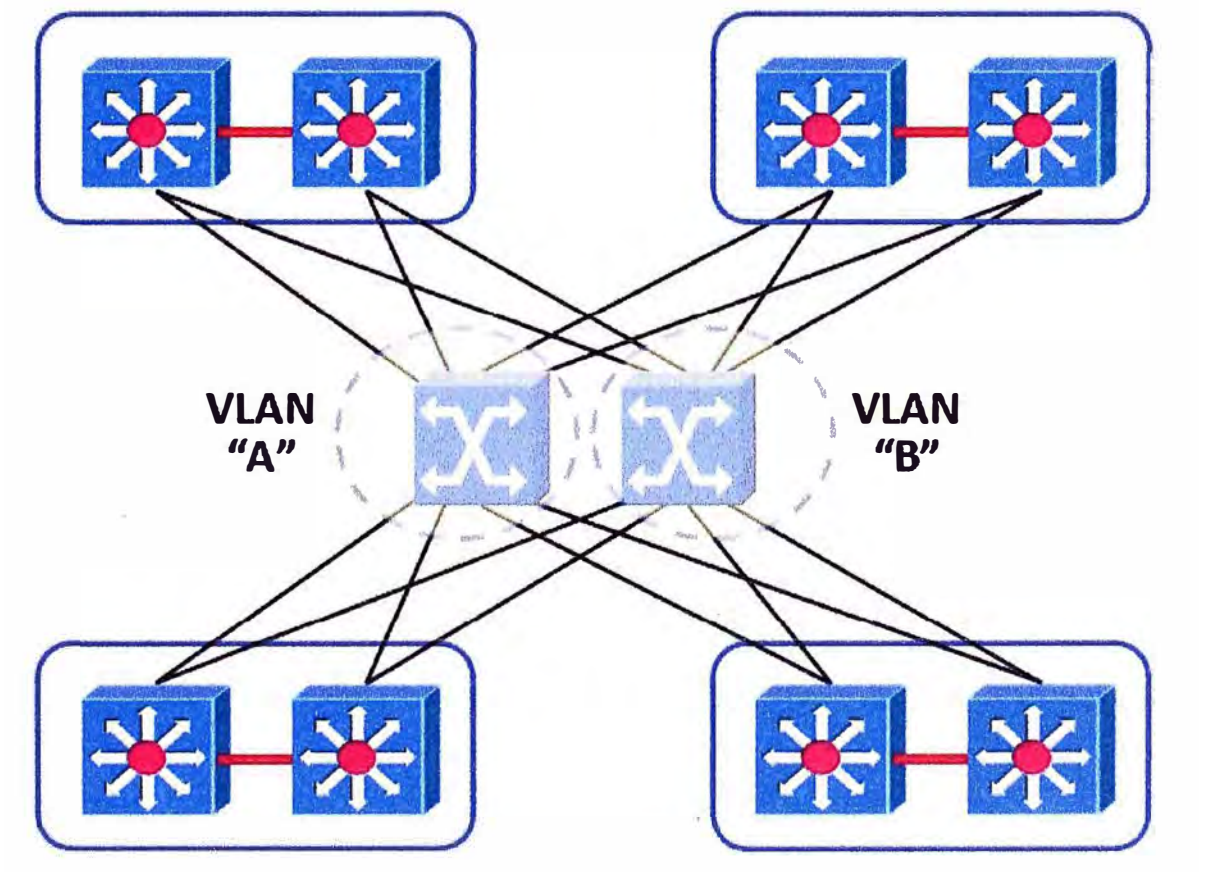


Fig. 2.7 Núcleo de Capa 2

El protocolo STP representa un límite práctico al escalamiento de un núcleo de capa 2. A medida que se incrementa el número de dispositivos de núcleo, también es necesario incrementar el número de enlaces desde los conmutadores de distribución a fin de mantener la redundancia. Debido a que el protocolo de enrutamiento determina el

número de rutas de capa 3 de igual costo, el número de conmutadores de núcleo independientes es también limitado.

Si se conectan los conmutadores de núcleo entre sí se crearían lazos, y aunque se utilice el protocolo STP en el núcleo, esto afectaría seriamente el rendimiento entre los bloques básicos. Luego, el mejor escenario corresponde a un núcleo de 2 conmutadores de capa 2 sin lazos en su topología.

En la figura 2.7 anterior se tiene 8 conmutadores de distribución que se conectan al núcleo. Estos tienen conexiones a cada uno de los conmutadores del núcleo, los cuales no tienen conexiones entre sí. Al no existir lazos en el núcleo, no hay necesidad de utilizar el protocolo STP. Con dos conmutadores de núcleo, solo hay dos rutas posibles de igual costo en capa 3 hacia todas las VLANs de destino.

2.5.2 Núcleo de capa 3

La mayoría de los diseños exitosos siguen el modelo clásico de diseño explicado hasta ahora, con funciones de capa 2 en la capa de acceso, funciones de capa 3 en la capa distribución y funciones de capa 2 nuevamente en la capa núcleo.

Sin embargo, hay situaciones donde es necesario usar funciones de capa 3 en el núcleo, por las siguientes razones:

- Acelerar la convergencia.
- Añadir el balanceo de carga automático.
- Eliminar las limitaciones de emparejamiento.

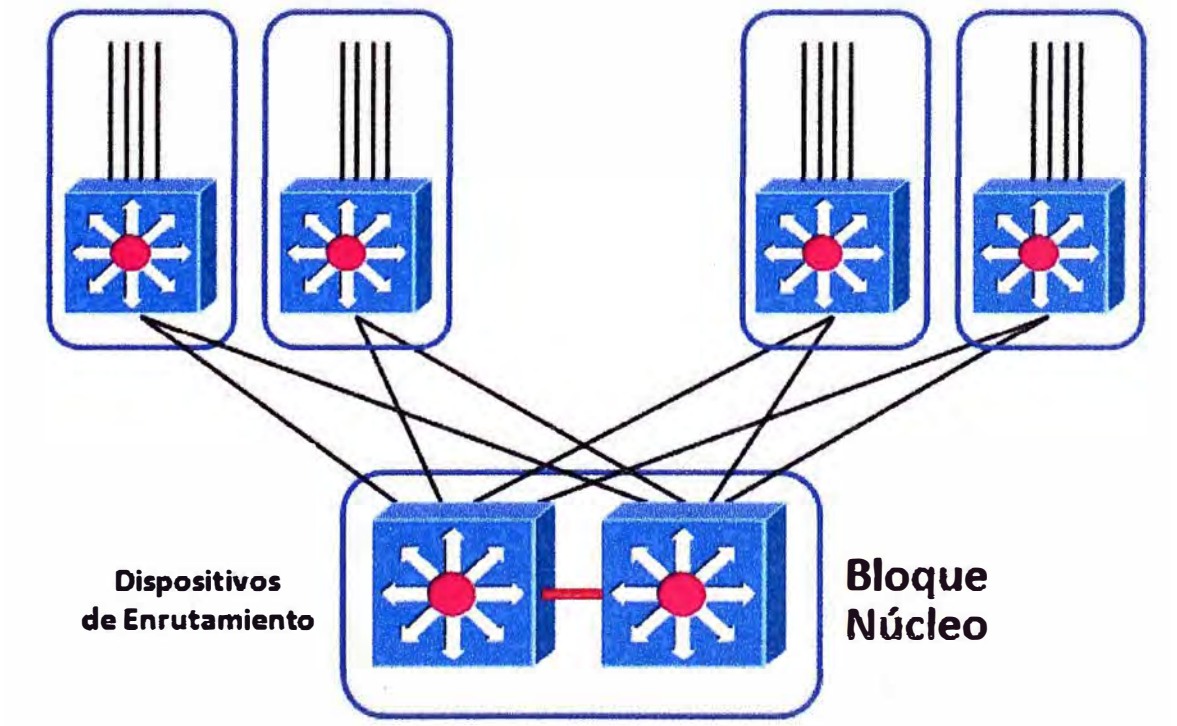


Fig. 2.8 Núcleo de Capa 3

La figura 2.8 anterior muestra un núcleo de capa 3. Cada conexión desde los dispositivos de la capa de distribución es una subred separada, y los dispositivos del núcleo son responsables ahora de hacer enrutamiento entre estas subredes.

2.5.2.1 Acelerar la Convergencia

A medida que se incrementa el número de bloques básicos, cada dispositivo de distribución debe ser conectado al núcleo. Debido a que existe un límite en el número de bloques básicos a ser conectados a un núcleo dual de capa 2, incrementar el número de conexiones significaría incrementar el número de dispositivos de capa 2 en el núcleo, y para mantener la redundancia, estos deberían estar interconectados entre sí. Al interconectar estos dispositivos para mantener la redundancia, sería necesario crear lazos, lo cual a su vez implica utilizar el protocolo STP como se explico antes.

El protocolo STP puede llegar a tener un tiempo de convergencia de más de 50 segundos. Si hay una falla en el núcleo de la red, la convergencia STP podría deshabilitar el núcleo por más de minuto, lo cual es inaceptable.

Con la implementación de dispositivos de capa 3 en el núcleo, el protocolo STP se vuelve innecesario. Con esta alternativa, los protocolos de enrutamiento son usados para mantener la topología de red. Sin embargo, debe recordarse que aún los protocolos de enrutamiento pueden tomar entre 5 a 10 segundos para lograr su convergencia.

2.5.2.2 Añadir el balanceo de carga automático

Con el balanceo de carga lo que se busca es lograr una mejor distribución del tráfico entre los múltiples enlaces que proveen redundancia. Con múltiples dispositivos de capa 2 en el núcleo interconectados entre sí, la única forma de balancear carga por enlaces es habilitar varias instancias de STP por cada VLAN (esto es conocido como PVSTP en los equipos Cisco), lo cual significa escoger diferentes topologías STP por cada VLAN y hacer un balance manual de las VLANs por cada enlace. Claramente, esto es muy complicado y hace más dificultoso el diseño.

Con dispositivos de capa 3 en el núcleo, los protocolos de enrutamiento pueden hacer el balanceo de carga de forma más natural, aprovechando rutas de igual costo entre la subred de origen y la subred de destino.

2.5.2.3 Eliminar los problemas de emparejamiento

Otro problema que se suscita en un diseño de núcleo de capa 2 en una red grande, es el referente al emparejamiento entre enrutadores. El emparejamiento asegura que el protocolo de enrutamiento que corre en los enrutadores, si mantiene la información de estado y disponibilidad de los enrutadores vecinos. En este escenario, cada dispositivo de distribución se vuelve una pareja con cada otro dispositivo de distribución en la red.

Luego, la escalabilidad se vuelve una traba en esta configuración debido a que cada dispositivo debe mantener el estado de todos los otros dispositivos.

Con la implementación de dispositivos de capa 3 en el núcleo, una jerarquía de enrutamiento es creada, y ahora el dispositivo de distribución no tiene que hacer emparejamiento con todos los otros dispositivos de esta misma capa. Este tipo de núcleo es preferido en las redes de campus en las cuales la red soporta varias decenas de bloques básicos.

Sin embargo, este diseño tiene un punto débil por el lado del costo involucrado. Debido a que el objetivo principal del núcleo es transportar los paquetes tan rápido y eficientemente como sea posible, agregar a esto la capacidad de enrutamiento, se vuelve un tema de elección de los dispositivos de mayor capacidad posible, lo cual se refleja al final en un incremento muy significativo de los costos del diseño.

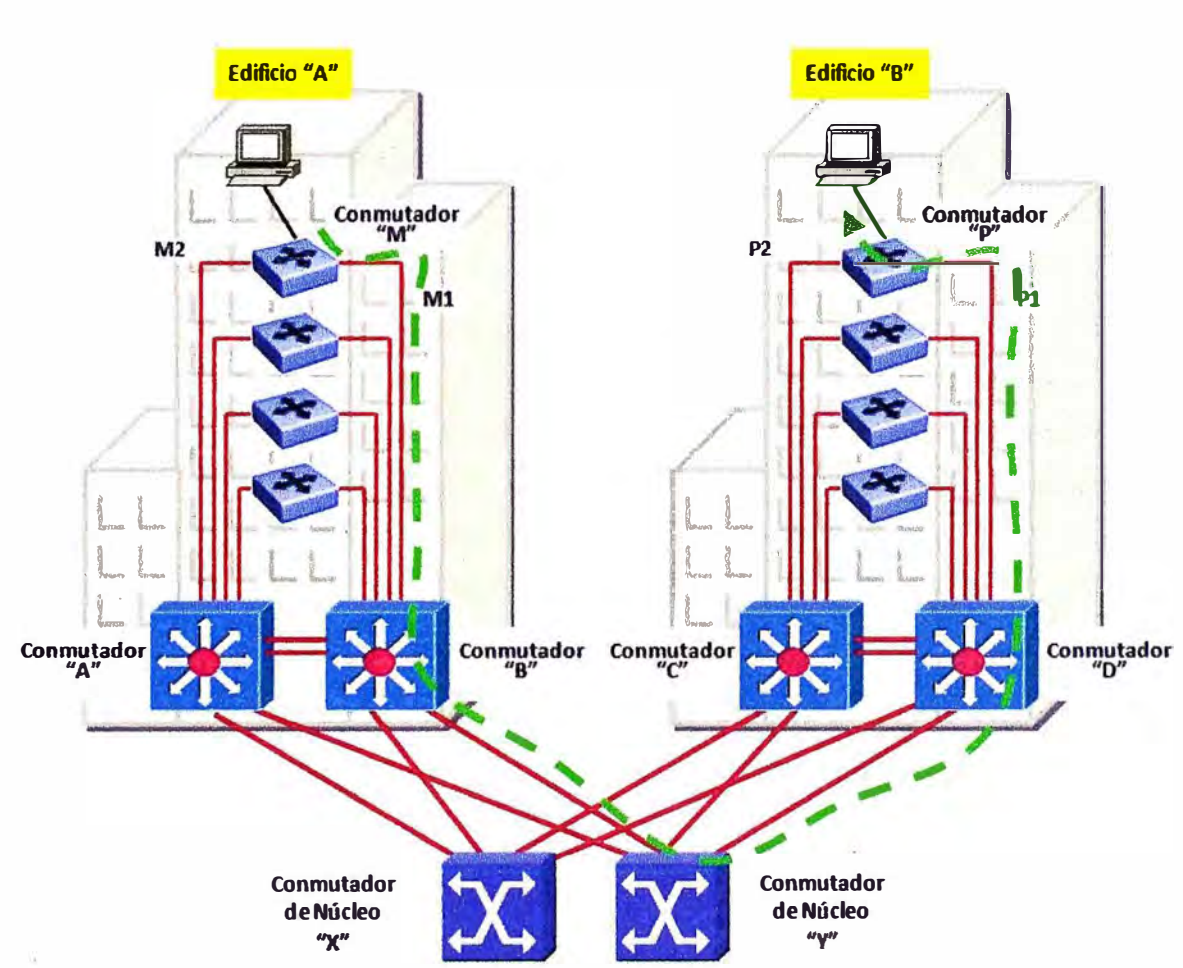


Fig. 2.9 Ejemplo de Diseño de Red de Campus

2.6 Análisis de la Disponibilidad

En la figura 2.9 anterior se muestra un ejemplo de diseño de campus que servirá para analizar la disponibilidad de la red en caso de una falla, y como es que todos los

servicios y protocolos de red revisados en este capítulo interactúan entre sí para lograr que la red tenga resiliencia.

El ejemplo presenta dos edificios A y B, cada uno con 10 pisos y 1,000 usuarios finales. Todos los usuarios de un piso son conectados horizontalmente a un conmutador de acceso en el cuarto de cableado. Cada conmutador de acceso es enlazado verticalmente a un conmutador de la capa de distribución.

Si un enlace entre un conmutador del cuarto de cableado y un conmutador de la capa de distribución se corta, entonces hasta 100 usuarios de un piso podrían perder su conexión al núcleo de la red. Para prevenir esto, cada conmutador de acceso tiene 2 enlaces, dirigidos uno a cada uno de los 2 conmutadores de distribución del edificio. El protocolo STP bloquea el enlace redundante (podría ser cualquiera de ellos) a fin de prevenir que exista un lazo.

El balanceo de carga a través del núcleo es logrado mediante la inteligencia del protocolo de enrutamiento de capa 3 que corre en los conmutadores de distribución. En la figura, hay 4 rutas de igual costo en capa 3 entre los 2 edificios. Las cuatro rutas son: AXC, AYD, BXC y BYD. Estos 4 caminos son considerados iguales por el protocolo de enrutamiento. Nótese que todos los caminos desde ambos edificios hacía el núcleo son de un salto simple.

2.6.1 Escenario 1: Falla en un Enlace

En este escenario, un usuario conectado al conmutador "M" de acceso, quiere transferir datos al usuario conectado al conmutador "P" del otro edificio. Supongamos que por configuración, el rol de enrutador activo le corresponde al conmutador multicapa B, y que el protocolo HSRP está activado entre los conmutadores A y B.

Como se muestra en la figura 2.9 anterior, el camino lógico para la transmisión de datos entre "M" y "P" es como sigue:

- El conmutador "M" transmite los datos sobre el enlace M1 hacia el conmutador multicapa "B".
- El conmutador multicapa "B" enruta estos datos fuera de su red hacia la subred del conmutador de núcleo "Y".
- El conmutador de núcleo "Y" transfiere los datos desde el enlace BY hacia el enlace YD, con destino al conmutador multicapa "D".
- El conmutador multicapa "D" transfiere finalmente los datos sobre el enlace P1 hacia el conmutador de acceso "P", y este a su vez a la estación final.

Si el enlace M1 falla, como se muestra en la figura 2.10 siguiente, entonces el enlace M2 se vuelve el enlace primario (no se revisa por el momento el tiempo que tarda lograr esto), y el camino de datos desde "M" hacia "P" es ahora como sigue:

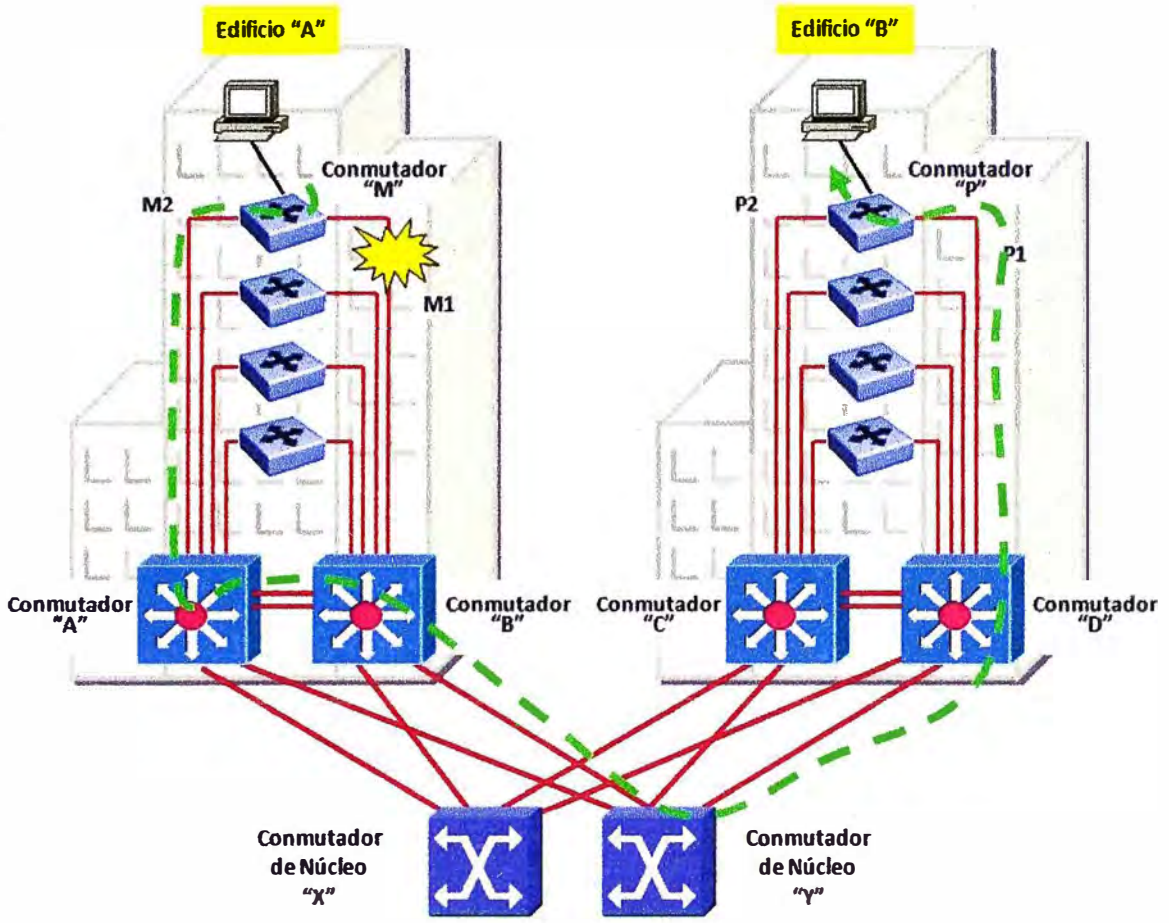


Fig. 2.10 Falla en Enlace M1

- El conmutador "M" transmite los datos sobre el enlace M2 hacia el conmutador multicapa "A".
- El conmutador multicapa "A" transmite estos datos hacia el enrutador HSRP activo, que sigue siendo el conmutador multicapa "B".
- El conmutador multicapa "B" enruta estos datos fuera de su red hacia la subred del conmutador de núcleo "Y".
- El conmutador de núcleo "Y" transfiere los datos desde el enlace BY hacia el enlace YD, con destino al conmutador multicapa "D".
- El conmutador multicapa "D" transfiere finalmente los datos sobre el enlace P1 hacia el conmutador de acceso "P", y este a su vez a la estación final.

2.6.2 Escenario 2: Falla en un Conmutador de Distribución

En este escenario 2, asumimos que existe una falla en uno de los conmutadores de distribución lo cual significa que hasta 1,000 usuarios finales pueden perder su conexión al núcleo. Para ofrecer redundancia en este caso, también se aprovecha la existencia de los enlaces redundantes entre la capa de acceso y la capa de distribución, como lo plantea el modelo de diseño clásico.

En la figura 2.11 siguiente, los motores de enrutamiento de ambos conmutadores de distribución están interconectados usando el protocolo HSRP. Esta configuración permite un nivel de redundancia para el servicio de nodo de salida por defecto (default Gateway, en inglés) referido al protocolo IP utilizado en la capa de acceso.

En el escenario de que el conmutador multicapa "B" falle, el camino de datos entre "M" y "P" es el siguiente:

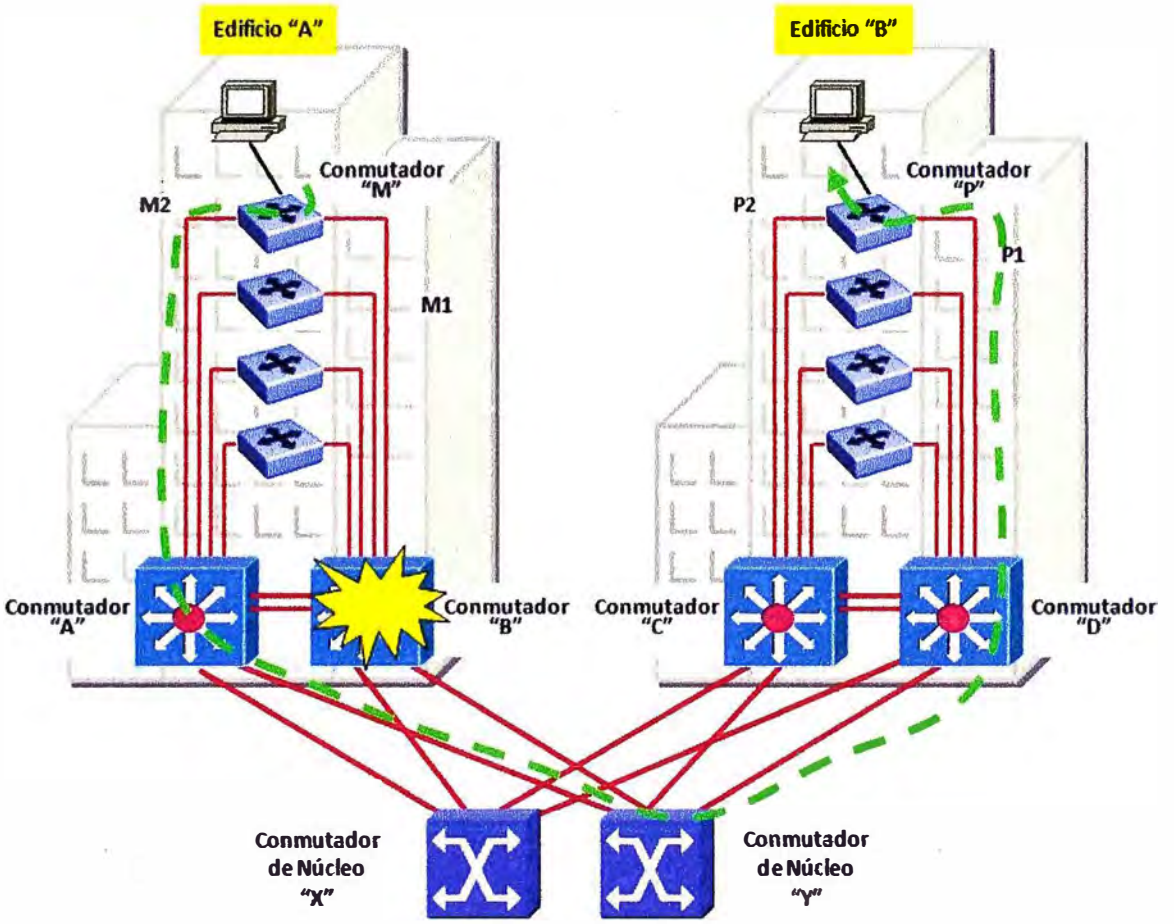


Fig. 2.11 Falla en Conmutador "B"

- El conmutador "M" cambia la transmisión de datos hacia el enlace M2 con dirección al conmutador multicapa "A".
- Debido a que el conmutador multicapa "B" esta fuera de servicio, el conmutador multicapa "A" asume el rol de enrutador HSRP activo (esto toma un tiempo que no se analizará en este momento).
- El conmutador multicapa "A" enruta los datos fuera de su red hacia la subred del conmutador de núcleo "Y".
- El conmutador de núcleo "Y" transfiere los datos desde el enlace AY hacia el enlace YD, con destino al conmutador multicapa "D".

- El conmutador multicapa "D" transfiere finalmente los datos sobre el enlace P1 hacia el conmutador de acceso "P", y este a su vez a la estación final.

2.6.3 Escenario 3: Falla adicional en un Conmutador de Núcleo

En el último escenario, asumimos que existe una falla adicional a la anterior, cuando se produce una caída en uno de los conmutadores de núcleo. Para prevenir este escenario, el núcleo consiste de dos conmutadores de capa 2, lo que significa que cada enlace desde un conmutador de distribución al núcleo es del mismo costo en capa 3. Esta topología ofrece protección contra esta falla y adicionalmente, balanceo de carga en capa 3 manejado por el protocolo de enrutamiento como se explico anteriormente.

En la figura 2.12 siguiente, veamos lo que sucede con el transporte de datos cuando se produce una falla total en el conmutador de núcleo "Y":

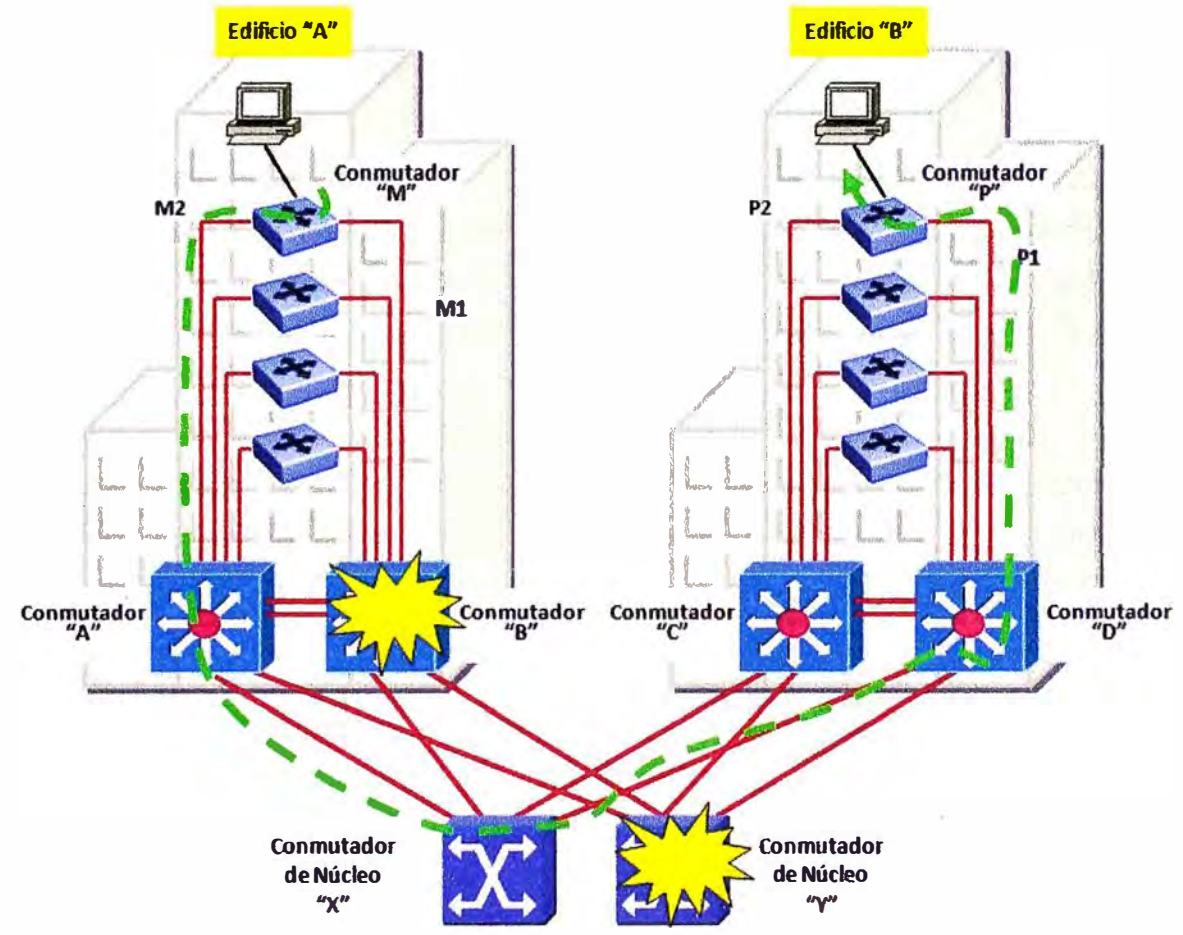


Fig. 2.12 Falla en Conmutador "B" y Conmutador "Y"

- El conmutador "M" sigue transmitiendo los datos por el enlace M2 con dirección al conmutador multicapa "A". Debido a que el conmutador multicapa "B" sigue fuera de servicio, el conmutador multicapa "A" sigue asumiendo el rol de enrutador HSRP activo.

- Debido a que el conmutador de núcleo "Y" se encuentra ahora fuera de servicio, el conmutador multicapa "A" enruta los datos fuera de su red hacia el otro conmutador de núcleo "X".
- El conmutador de núcleo "X" transfiere los datos desde el enlace AX hacia el enlace XD, con destino al conmutador multicapa "D".
- El conmutador multicapa "D" transfiere finalmente los datos sobre el enlace P1 hacia el conmutador de acceso "P", y este a su vez a la estación final.

Del análisis de los escenarios de falla anteriores, se deduce que para efectos de lograr una adecuada resiliencia a nivel de la red de campus, el protocolo STP o su versión mejorada el protocolo RSTP, juegan un papel clave en el diseño.

Vamos a analizar a continuación con detalle y en forma comparativa la operación de estos protocolos, revelando sus ventajas y desventajas para extraer al final algunas conclusiones respecto a su utilidad en el diseño de redes de campus.

2.7 Protocolos STP y RSTP

2.7.1 Antecedentes

El protocolo Spanning Tree (Spanning Tree Protocol, o STP por sus siglas en inglés) fue originalmente desarrollado al final de los años 80 por la empresa DEC, y fue posteriormente estandarizado por la fuerza de tarea del Internet del IEEE (IEEE-802.1 Internet Work Task Force) con el nombre IEEE-802.1D.El propósito del STP fue y sigue siendo prevenir los lazos en las redes conmutadas (bridged networks, en inglés) mediante el uso de conexiones redundantes.

A diferencia de los enrutadores, los conmutadores no tienen un mecanismo para descartar los paquetes redundantes que dan vueltas por la red. Esto causa problemas tales como la duplicación de tramas unidireccionadas (unicast en inglés) y la multiplicación de tramas multidireccionadas (multicast).

Cuando dos o más conmutadores están conectados en un lazo por la red LAN, ellos pueden multiplicar las tramas multicast, enviándolas vuelta tras vuelta hasta que la red termine por bloquearse. La incapacidad para tolerar lazos activos es una restricción fundamental que debe ser tomada en cuenta cuando se diseña redes LAN conmutadas.

Sin embargo, las conexiones redundantes son esenciales cuando se diseña pensando en alta disponibilidad, y justo es en este caso donde se utiliza el protocolo STP. Al bloquear las conexiones redundantes, STP permite que exista un solo camino primario activo entre todos los nodos. Si una falla en un dispositivo o en un enlace causa que este camino primario se vuelva inutilizable, STP habilitará un segundo camino alternativo. Los conmutadores, mediante el uso del protocolo STP, aseguran que no existan estos lazos en la red.

2.7.2 La necesidad de RSTP

El STP es un protocolo que ha existido por largo tiempo y aún podría seguir siendo utilizado, pero para los estándares de las aplicaciones actuales es muy lento. Su lentitud es el resultado de su tiempo de convergencia, el cual puede tomar desde 30 a 50 segundos.

La mayoría de las aplicaciones actuales requieren redes LAN con alta disponibilidad que puedan recuperarse en menos de un segundo. Dado que STP es tan lento, no resulta práctico para las aplicaciones de hoy y por tanto un protocolo más rápido fue requerido.

El grupo de trabajo IEEE 802.1w desarrollo un nuevo protocolo con rápida reconfiguración denominado STP rápido (Rapid Spanning Tree o RSTP por sus siglas en inglés). Más que un protocolo nuevo, RSTP es realmente una versión mejorada y más rápida de STP que tiene la propiedad de ser interoperable con el anterior.

Los administradores familiarizados con la operación de STP pueden rápidamente aprender el nuevo algoritmo dado que en ambos, la terminología y los parámetros básicos se han mantenido sin cambios.

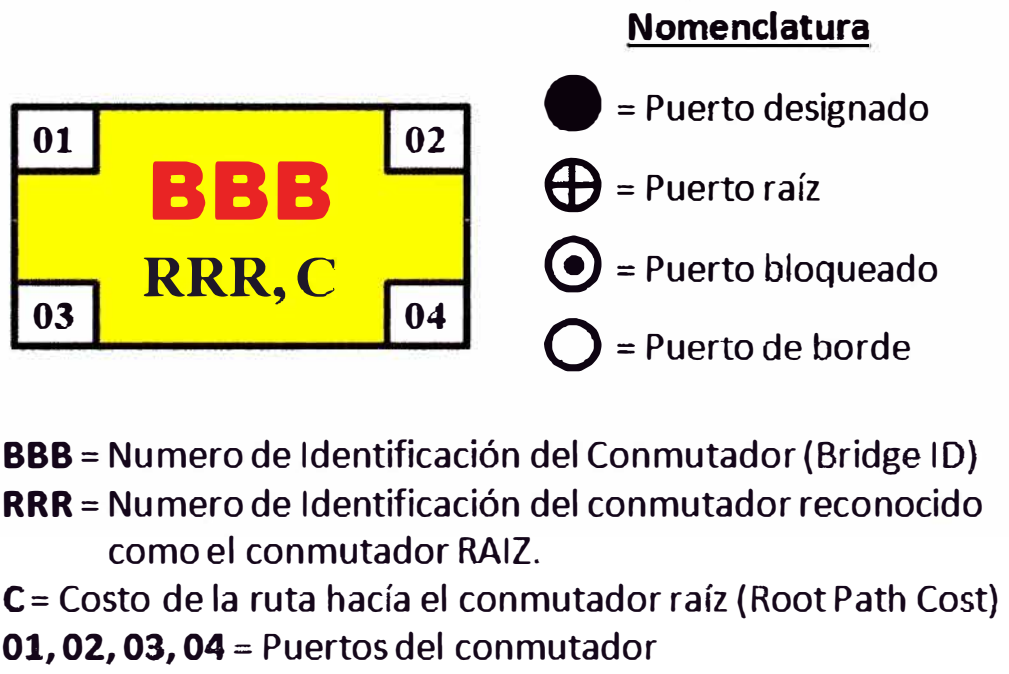


Fig. 2.13 Explicación de la Nomenclatura

2.7.3 Conceptos básicos en RSTP

Antes de revisar como STP y RSTP pueden operar, es necesario entender algunos conceptos básicos. En la figura 2.14 siguiente, se muestra una red LAN simple con 4

conmutadores que servirá para el entendimiento de estos conceptos. En esta figura se asume que el protocolo (STP o RSTP) todavía no ha comenzado su operación y que solo se refleja las conexiones físicas entre los conmutadores.

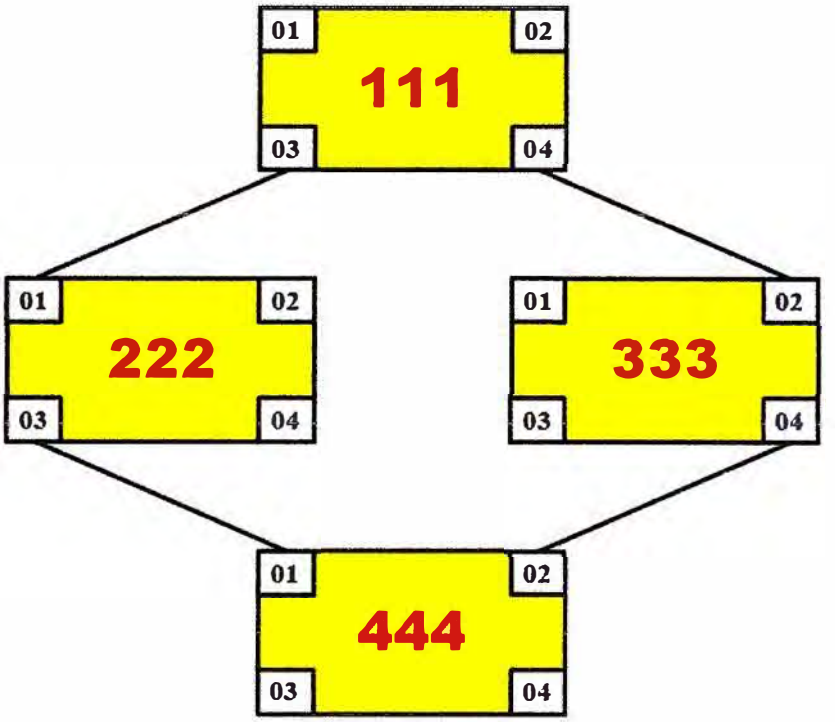


Fig. 2.14 Interconexión en Red LAN de Ejemplo

A esta figura se le aplica la nomenclatura explicada en la figura 2.13 anterior, para que los términos y objetos puedan ser entendidos durante la explicación que presentare en las páginas siguientes.

2.7.3.1 Topología de árbol invertido

La topología STP puede ser vista como un árbol invertido, que incluye un componente raíz (denominado el conmutador raíz o "root bridge" en inglés), las ramas (los segmentos LAN y sus conmutadores asociados) y las hojas (las estaciones terminales). En un árbol, no hay partes desconectadas que puedan considerarse como parte del árbol; esto es, el árbol abarca y "conecta" la totalidad de sus hojas. Adicionalmente, no hay lazos en un árbol. Si Usted traza un camino desde una hoja hacia otra hoja cualquiera, se encontrará que hay uno y solo un camino posible. Esto es también cierto en una topología de red LAN generada por STP. STP organiza y conecta los conmutadores en una topología libre de lazos, mientras que a su vez no se deja ningún segmento de red aislado.

En la figura 2.15 siguiente, se dibuja una red libre de lazos con conexiones activas que se extienden desde la raíz hacia las ramas. En este ejemplo, el conmutador 111 ya ha sido seleccionado como el conmutador raíz.

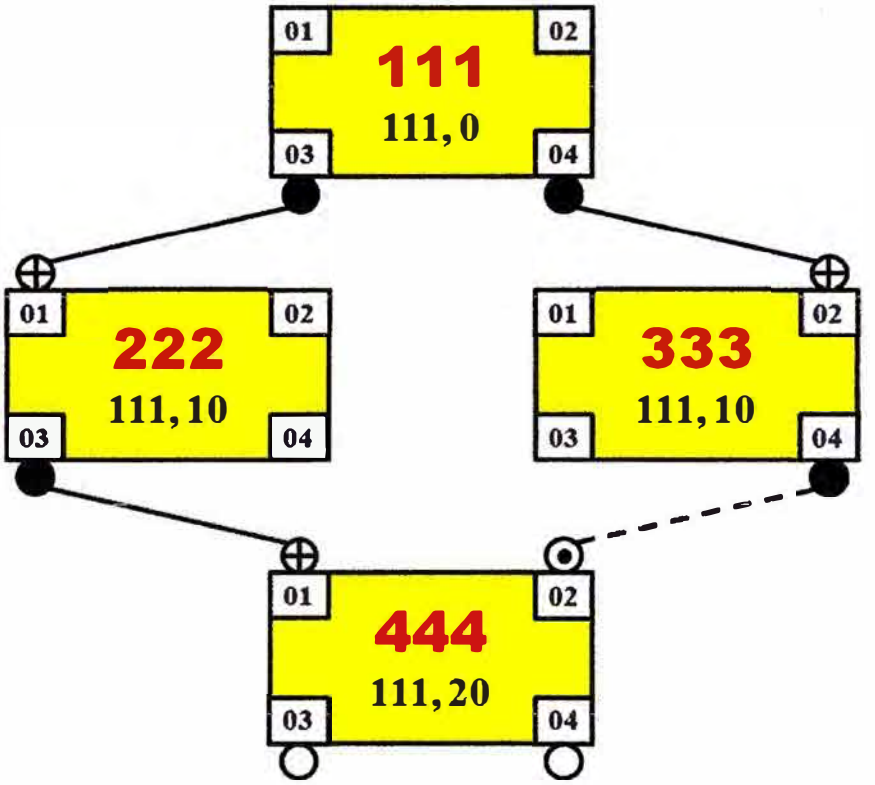


Fig. 2.15 Estado de Puertos con STP

De la misma forma como un árbol tiene una raíz, desde la cual todas las ramas se esparcen, STP tiene también un conmutador raíz. En una red LAN dada, solamente puede existir un conmutador raíz. Aunque cualquier conmutador puede ser el conmutador raíz, solamente el conmutador con el mejor número de identificación (Bridge ID en inglés) es el que se convierte en conmutador raíz.

El número de identificación está conformado por 2 números: un valor de prioridad (seleccionado por el administrador que configura la red) y el otro valor que corresponde a su dirección MAC.

El conmutador con el valor numérico de prioridad más bajo se convierte en el conmutador raíz. Cuando todos los conmutadores tienen el mismo valor de prioridad, el que tiene el valor de dirección MAC más bajo tiene la preferencia y es el que se convierte en conmutador raíz.

2.7.3.2 Conmutadores designados

Una forma simple de prevenir lazos en la red es asegurar que solamente un conmutador sea responsable por reenviar tráfico que viene desde el conmutador raíz hacia cualquier enlace o segmento dado (rama). En tanto que solamente exista un camino activo desde la raíz hacia cualquier estación terminal (hoja), entonces no existirá la posibilidad de que haya lazos en la topología LAN.

Al dispositivo responsable por reenviar tráfico hacia ese enlace se le conoce como conmutador designado (designated bridge en inglés) para ese enlace o segmento.

2.7.3.3 Estados de Puerto

Hay 3 estados operacionales internos asignados por RSTP a cada puerto de un conmutador:

- Descarte: Los puertos en estado de descarte no participan en la topología activa y no aprenden direcciones MAC.
- Aprendizaje: Los puertos en estado de aprendizaje si aprenden direcciones MAC pero no reenvían tráfico del usuario.
- Reenvío: Los puertos en estado de reenvío si participan completamente tanto en el reenvío de datos como en el aprendizaje MAC.

Los estados RSTP difieren en algo de los estados STP tradicionales. La tabla 2.2 siguiente presenta estas diferencias.

TABLA N° 2.2 Comparación de Estados en STP vs. RSTP

Estado de Puerto en STP	Estado de Puerto en RSTP	Se incluye el puerto en la topología activa?	Esta el puerto aprendiendo direcciones MAC?
Deshabilitado	Descarte	No	No
Bloqueado	Descarte	No	No
Escucha	Descarte	No	No
Aprendizaje	Aprendizaje	No	Si
Reenvío	Reenvío	Si	Si

2.7.3.4 Roles de Puerto

Un rol de puerto es una función que STP o RSTP asigna a cada puerto dentro de la topología. STP asigna uno de los 3 roles siguientes: puerto raíz, puerto designado o puerto bloqueado. RSTP en cambio divide el rol de puerto bloqueado en los roles de puerto de respaldo y puerto alternativo.

- Puerto Raíz: Es el puerto más cercano al conmutador raíz en términos de costo de ruta. Cuando un conmutador tiene múltiples rutas que lo conectan a la raíz, la mejor ruta se determina en base a lo siguiente:
  - El "vector de prioridad de mensaje" que se transporta en el interior de los paquetes de mensaje del protocolo conocidos como BPDU (Bridge Protocol Data Unit), y
  - El identificador del puerto receptor del mensaje BPDU que viene del conmutador vecino.

- Puerto Designado: Es aquel que envía el mejor BPDU en el segmento al cual está conectado. Todos los conmutadores conectados a un segmento compartido dado, escuchan los BPDUs de los otros y acuerdan que el conmutador que envía el mejor BPDU es el conmutador designado para tal segmento. La figura2.15 anterior muestra los puertos designados para toda la red.
- Puerto de Respaldo: Es aquel conectado a la misma red LAN que el puerto designado para esa red. STP bloquea a los puertos de respaldo dado que el puerto designado es el que tiene la mejor ruta desde esa red LAN hacia el conmutador raíz. En ese sentido, el puerto de respaldo es la contingencia del puerto designado. La figura 2.16 siguiente muestra la ubicación del puerto de respaldo.

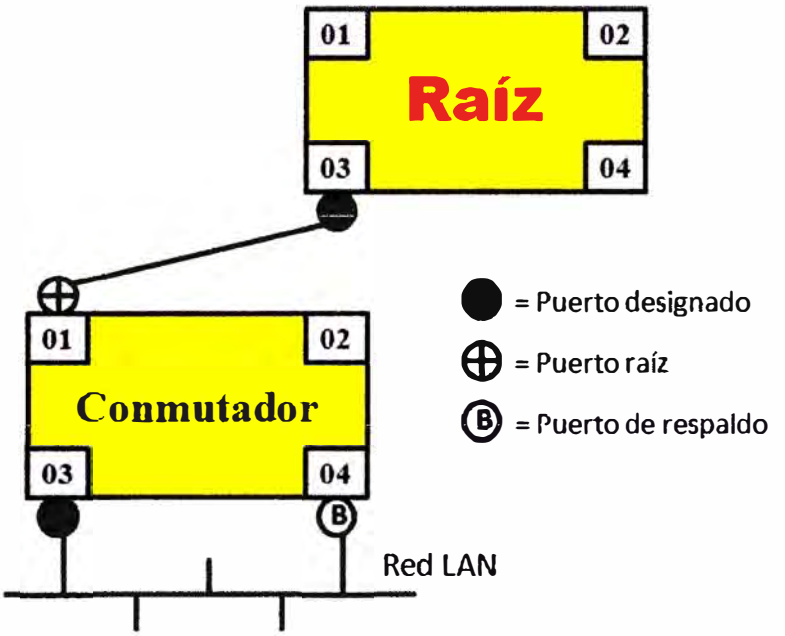


Fig. 2.16 Ubicación de Puerto de Respaldo

- Puerto Alternativo: Es aquel que proporciona una conexión redundante hacia el conmutador raíz y se puede volver un nuevo puerto raíz en el evento de que, el actual puerto raíz pierda su conexión al conmutador raíz. En ese sentido, el puerto alternativo es la contingencia del puerto raíz.

  La figura 2.17 siguiente muestra la ubicación del puerto alternativo. En muchos casos, el puerto alternativo se puede convertir en puerto raíz y cambiar al estado de reenvío sin ningún retardo.

2.7.3.5 Formato del BPDU

Los conmutadores aprenden e intercambian información acerca de cada uno, mediante el envío de pequeños paquetes de mensaje denominados BPDUs (Bridge Protocol Data Unit – Unidad de Datos del Protocolo de Conmutadores). El fin de este

intercambio es obtener información que les permita determinar la topología del árbol de expansión (spanning tree) correspondiente a la red LAN donde pertenecen.

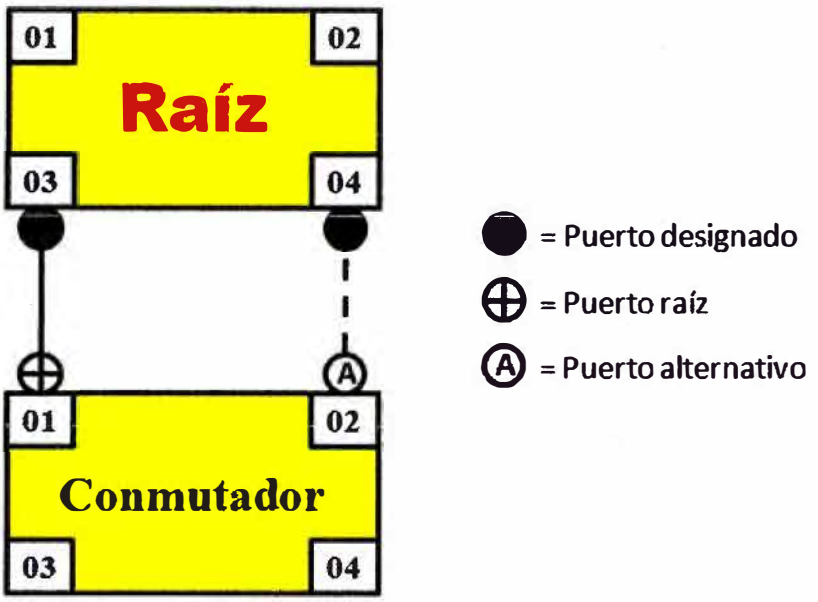


Fig. 2.17 Ubicación de Puerto Alternativo

Existen algunas diferencias entre los BPDU de STP y los BPDU de RSTP que se analizan a continuación.

- BPDUs de STP: STP utiliza dos tipos de BPDUs: BPDUs de configuración (tipo 1) y BPDUs de cambio de topología (tipo 2). Los BPDUs de configuración se generan periódicamente en el conmutador raíz, utilizando un período denominado tiempo de saludo (hello time). Este primer tipo de BPDU lleva toda la información requerida para calcular la topología STP. Los otros conmutadores escuchan los BPDUs de configuración en sus respectivos puertos raíz y los reenvían en sus puertos designados a los segmentos correspondientes.

Los BPDUs de cambio de topología (tipo 2) son enviados en dirección hacia el conmutador raíz (ósea, hacia arriba en el árbol invertido) por el conmutador que detecto un cambio en la topología. Cuando el conmutador raíz recibe un BPDU tipo 2, entonces debe proceder a informar a los otros conmutadores de que un cambio ha ocurrido en la topología actual.

Para realizar esto, el conmutador raíz coloca la bandera TC (Topology Change) en "1" lógico dentro de cada BPDU tipo 1 que le tocar enviar, por un período de tiempo especificado como la suma de los temporizadores "Forward Delay" y "Max Age". Cuando otro conmutador recibe un BPDU tipo 1 con la bandera TC activa, este procede a cambiar su tiempo de envejecimiento (aging time) de largo a corto a fin de limpiar las entradas de la Tabla MAC más rápidamente.

- BPDU de RSTP: RSTP utiliza solamente un tipo de BPDU. Este es similar al BPDU de configuración (tipo 1) usado en STP, con la excepción del valor que aparece en el campo "Tipo" (2 para RSTP y cero para STP) y el campo de "Banderas" (flags) que transporta información adicional propia del protocolo RSTP. En este campo, los BPDUs de STP usan solamente dos bits (banderas):
  - TC (Topology Change – Cambio de topología), y
  - TCA (TC Acknowledgement – Reconocimiento de cambio de topología)
  - RSTP por su parte utiliza este campo de la forma siguiente:
  - Cuatro bits (banderas) para codificar el rol y el estado del puerto que origina el BPDU, y
  - Dos bits (banderas) para manejar el mecanismo propuesta/acuerdo que se explica más adelante.

La figura 2.18 siguiente muestra gráficamente la diferencia entre los campos de "Banderas" usados en STP y RSTP.

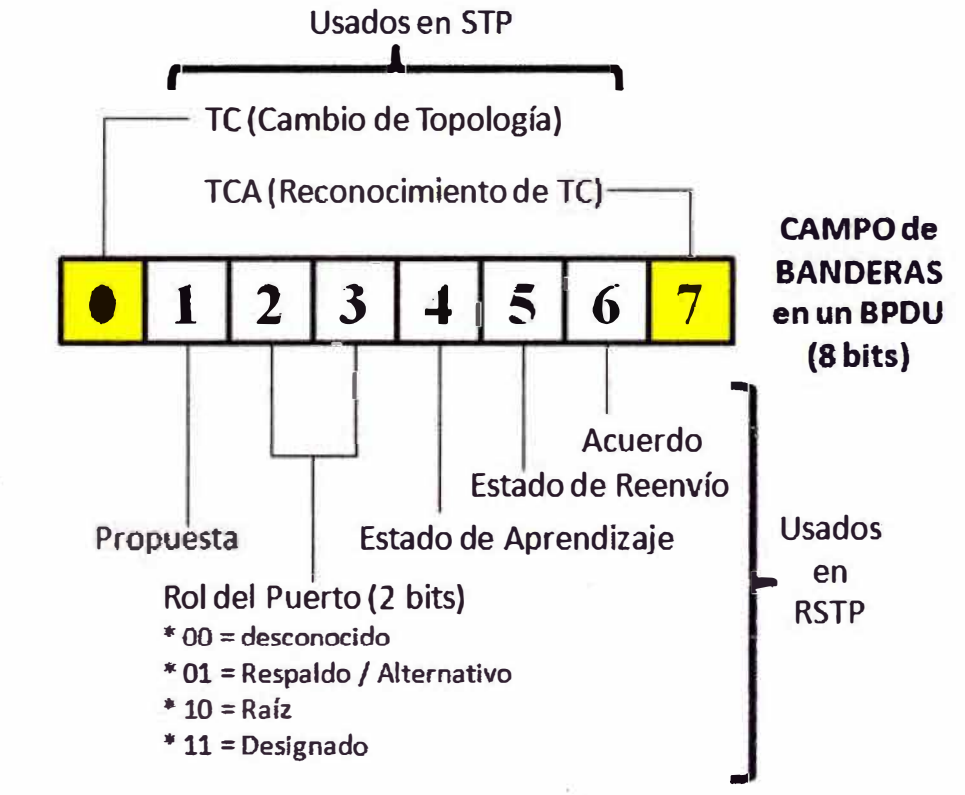


Fig. 2.18 Campo de Banderas en STP y RSTP

2.7.4 Diferencias entre RSTP y STP

En muchos aspectos STP y RSTP trabajan de la misma forma. Ambos reducen una red de conmutadores a una topología de árbol invertido expandido a fin de eliminar lazos. Ambos algoritmos reactivan las conexiones redundantes en el evento de que un enlace, o un componente, fallen.

La principal diferencia es el tiempo de convergencia. Mientras que a STP le puede tomar entre 30 a 50 segundos volver a converger, RSTP lo puede hacer en mucho menos tiempo. En una red diseñada cuidadosamente (no de cualquier forma) con RSTP, el tiempo de re-convergencia podría llegar a ser menos de un segundo.

Entonces aunque el cálculo computacional del árbol invertido expandido es idéntico entre STP y RSTP, si hay diferencias en la conducta de los dos algoritmos que analizamos a continuación:

2.7.4.1 Rápido envejecimiento de la Tabla MAC

Como parte de su operación normal, los conmutadores LAN escuchan el tráfico de red para aprender que estaciones terminales (direcciones MAC) existen en que puertos, y entonces almacenan entradas MAC-puerto en una base de datos interna. Después, cuando una trama arriba con destino hacia una de estas direcciones MAC, esta trama es conmutada al puerto apropiado de acuerdo a la información existente en su base de datos. La base de datos formada por estas entradas MAC-puerto es denominada comúnmente como Tabla MAC.

La Tabla MAC necesita ser re-aprendida cada vez que hay un cambio en la topología de red o de lo contrario, las tramas pueden ser enviadas a puertos equivocados (es decir, puertos que están impedidos de transmitir la trama a su destino final). Un cambio en la topología, tal como una falla en un enlace, puede causar que algunas estaciones de la red se conecten (involuntariamente) a conmutadores diferentes. En esta situación, aun cuando ninguna estación ha sido movida físicamente, para las conmutadores si parece como si estas estaciones hubiesen sido llevadas de una parte de la red para ser conectadas en otra parte. Para que el tráfico alcance a estas estaciones y todo opere de forma transparente, los conmutadores necesitan eliminar la información pasada que tienen y re-aprender las nuevas ubicaciones de las estaciones.

Los conmutadores que trabajan con STP no limpian sus Tablas MAC cuando ellos detectan un cambio topológico. En su lugar, ellos envían BPDUs de cambio de topología (tipo 2) en dirección al conmutador raíz, lo cual sirve para informar luego a todos los conmutadores de que un cambio ha ocurrido. Ahora, puede tomar varios segundos antes de que el BPDU tipo 2 alcance el conmutador raíz, y varios segundos más antes que los BPDUs tipo1 (con la bandera TC activada) generados por el conmutador raíz, alcancen los otros conmutadores en la red. Y aun así, los conmutadores STP que reciben esta alerta no limpian la información pasada inmediatamente. En su lugar, estos cambian su temporizador de envejecimiento de un valor largo a un valor corto (por defecto, este valor corto corresponde al temporizador “Forward Delay” de 15 segundos). Recién cuando ha

pasado este tiempo de envejecimiento reducido, si las entradas de la Tabla MAC no son refrescadas, ellas son removidas de la base de datos.

RSTP utiliza un mecanismo más eficiente para purgar la información pasada. Primero que todo, cada conmutador que detecta un cambio topológico envía BPDUs de RSTP con la bandera TC activada. Segundo, el conmutador que detecto el cambio, purga todas sus entradas inmediatamente. Finalmente, cada conmutador RSTP que recibe un BPDU con la bandera TC activada, purga también sus propias entradas pasadas y reenvía este BPDU de RSTP hacia otros conmutadores, creando un efecto de propagación en cadena.

La mejora explicada anteriormente es tremenda. En vez de esperar a que STP realice todo el proceso explicado antes, los conmutadores con RSTP ahora purgan las entradas pasadas de la Tabla MAC inmediatamente y le piden a otros conmutadores que hagan lo mismo.

2.7.4.2 BPDUs utilizados como pulsos de vida

Los conmutadores STP no generan sus propios BPDUs. Ellos esperan recibirlos en sus puertos raíz y entonces reenviarlos (después de procesarlos en su algoritmo) a sus puertos designados. Si un conmutador STP no recibe un BPDU por un tiempo igual al temporizador "Max Age" (cuyo valor por defecto es 20 segundos), entonces declara que el conmutador raíz ha fallado (asume el peor caso) y se nombra asimismo como el conmutador raíz, para luego comenzar por su cuenta un nuevo proceso de selección del conmutador raíz.

RSTP también difiere en este aspecto. Primero, cada conmutador envía sus propios BPDUs (los genera por si mismo) ya sea que reciba o no alguno en su puerto raíz. Su comportamiento es tal que, en vez de esperar por el temporizador "Max Age", un conmutador RSTP espera recibir un BPDU de su vecino hasta un tiempo máximo igual a 3 tiempos de saludo (hello time). Si el BPDU no es recibido dentro de ese tiempo, el conmutador asume que ha perdido la conexión con su vecino y ya no se asocia la perdida de esta conexión con la caída del conmutador raíz, dado que en RSTP, todos los conmutadores envían sus propios BPDUs. Obviamente, si un conmutador detecta una pérdida del enlace físico en un puerto que sirve de enlace, inmediatamente asume que la conexión con su vecino está perdida también, sin esperar ningún tiempo.

2.7.4.3 Puertos de borde

El algoritmo STP tiene un gran cuidado en asegurarse de que no existan lazos en la red. Esto explica porque los puertos del conmutador no envían tráfico hasta que todos los conmutadores se ponen de acuerdo en una determinada topología de red LAN. Sin

embargo, hay situaciones donde los puertos se conectan directamente a las estaciones finales, y por tanto no hay manera de que puedan crear lazos. RSTP reconoce esto y por eso permite configurar manualmente puertos a los cuales se sabe de antemano, que no se conectan otros conmutadores, y a los cuales se denomina puertos de borde.

Los puertos de borde no necesitan pasar a través de los estados regulares de transición del STP, sino que van directamente al estado de reenvío. Si un conmutador detecta un BPDU en un puerto de borde, automáticamente este retorna a ser un puerto normal de STP (non-edge port).

2.7.5 **Operación del protocolo RSTP**

La principal diferencia en la operación de STP y RSTP, es que RSTP ya no confía en temporizadores largos para re-converger después de que se produce un cambio en la topología. Para lograr esto, el algoritmo RSTP hace lo siguiente:

- Monitorea los estados operacionales de la capa MAC y retira los puertos que no están funcionando.
- Procesa los BPDUs inferiores para detectar los cambios en la topología.
- Mantiene el rastro de los puertos que proveen caminos alternativos hacia el conmutador raíz. Si el puerto raíz falla, RSTP puede rápidamente retirar ese puerto y hacer que un puerto alternativo sea su nuevo puerto raíz. Este nuevo puerto raíz puede ser colocado inmediatamente en el estado de reenvío sin ningún retardo.
- Cuando los conmutadores están conectados vía enlaces directos punto-a-punto, ellos usan un mecanismo de sincronización en vez de los temporizadores clásicos de STP, para cambiar un puerto designado al estado de reenvío.

Para ilustrar esto, veamos nuevamente la figura 2.15 mostrada anteriormente. Allí se presenta una red LAN compuesta por 4 conmutadores con sus números de identificación (111, 222, 333 y 444), los cuales están conectados vía una topología en anillo.

En esa configuración sucede lo siguiente:

- El conmutador 111 es elegido como el conmutador raíz.
- Los puertos que conectan los conmutadores 222 y 333 hacia el conmutador raíz se vuelven puertos raíz.
- El conmutador 444 tiene 2 caminos hacia el conmutador raíz:
  - Uno vía el conmutador 222 y
  - El otro vía el conmutador 333.
  - Se elige el puerto 01 como su puerto raíz (debido a que tiene el mejor vector de prioridad de puerto) y al mismo tiempo se bloquea el puerto 02.

2.7.6 Análisis de una Falla

Para saber que sucede cuando falla la conexión entre el conmutador 111 y el 222, analizaremos primero el caso con STP (ver figura 2.19):

2.7.6.1 Análisis para el caso STP

Después de una falla en el enlace, los conmutadores 222 y 444 continúan esperando por un tiempo igual al temporizador "Max Age" (20 segundos por defecto) antes de decidir ambos, que su camino hacia el conmutador raíz no está funcionando. Durante ese tiempo, el conmutador 444 descarta los BPDUs que vienen del conmutador raíz 111 y que son recibidos en el puerto 02 (que se encuentra en estado de bloqueo) porque los asume como inferiores.

Finalmente, después de que el temporizador "Max Age" expira, el conmutador 444 elimina la información de protocolo en el puerto 01, porque se da cuenta que tiene un camino hacia el conmutador raíz a través del puerto 02. Elige internamente este puerto 02 para que sea su nuevo puerto raíz y lo publica (reenviando un BPDU de configuración) hacia el conmutador 222 a través del puerto 01 que se convierte ahora en un puerto designado.

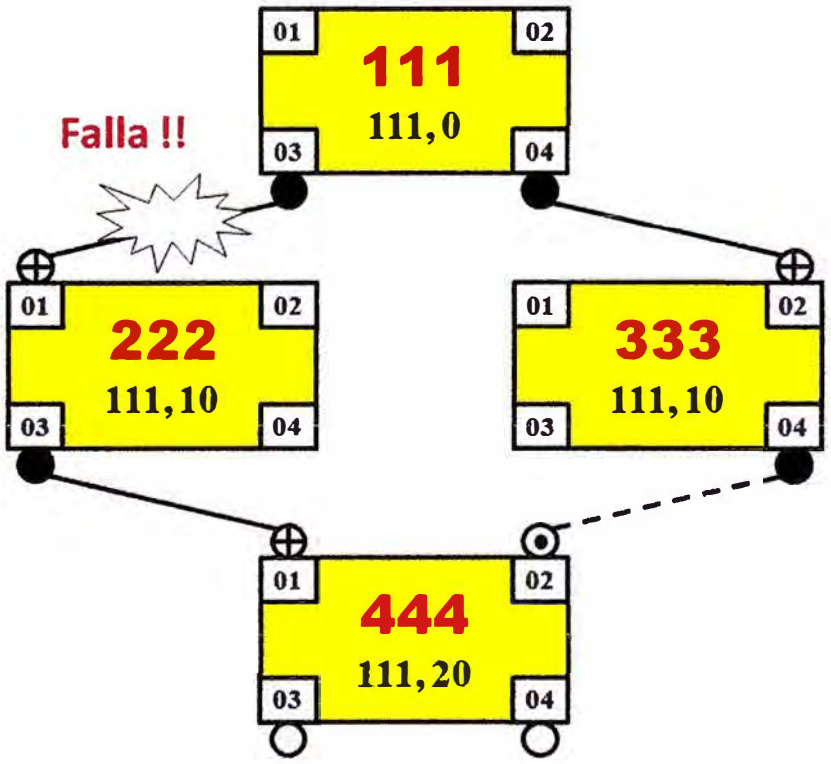


Fig. 2.19 Respuesta de STP a falla en Enlace

Pero a fin de asegurar que todos los conmutadores en la red LAN estén de acuerdo con esta nueva topología, el conmutador 444 no reenvía tramas de datos en los puertos 01 y 02 por un tiempo adicional, sino que hace pasar ambos puertos a través de los

estados de escucha y aprendizaje (15 segundos en cada uno) antes de colocarlos en el estado de reenvío.

Por su lado, el conmutador 222 nota (por el mensaje BPDU que recibió antes del conmutador 444) que su nueva ruta hacia el conmutador raíz es ahora a través del conmutador 444 y hace que su puerto 03 sea el nuevo puerto raíz, el cual también necesariamente, requiere pasar a través de los estados de escucha y aprendizaje.

El tiempo total de caída en la red como consecuencia de esta falla es percibido desde las estaciones conectadas a los conmutadores 222 y 444, según lo siguiente:

- Tiempo Total = 20 segundos (temporizador "Max Age") + 15 segundos (estado de escucha) + 15 segundos (estado de aprendizaje) = 50 segundos.

La nueva topología resultante después de la caída es la que se muestra en la figura 2.20 siguiente.

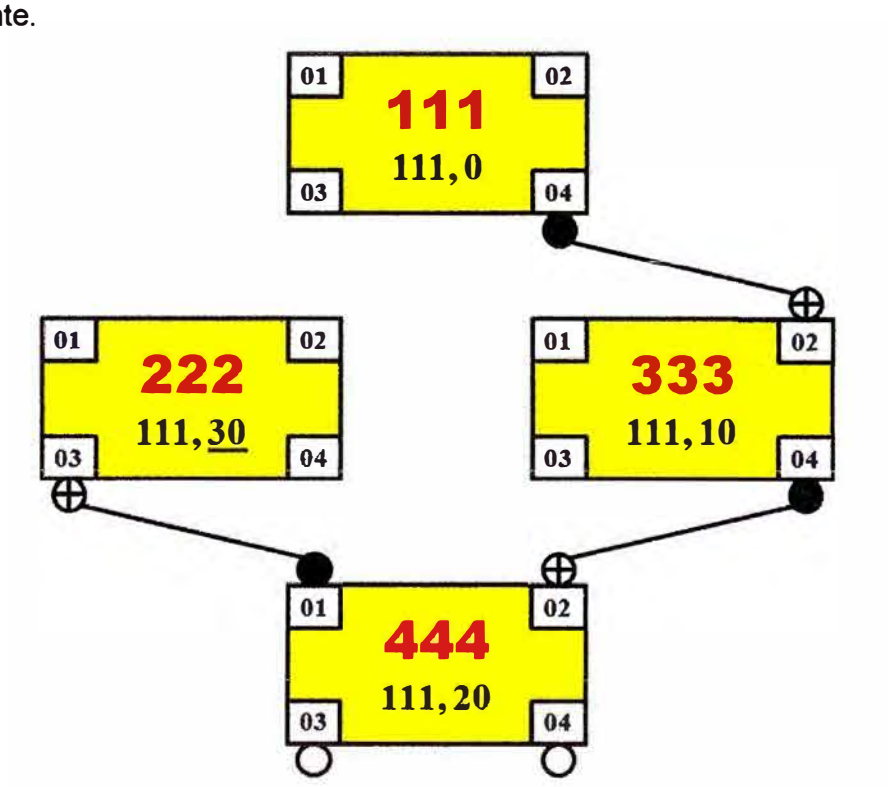


Fig. 2.20 Nueva Topología después de Convergencia

Para saber que sucede cuando el enlace entre el conmutador 111 y el conmutador 222 se restaura, veamos lo siguiente:

- Tan pronto como el conmutador 222 recibe BPDUs de configuración desde el conmutador raíz 111 en el puerto 01, procede a retirar internamente su puerto 03 y a convertir nuevamente al puerto 01 en su puerto raíz. Este puerto, de cualquier modo, tiene que pasar nuevamente a través de los estados de escucha y aprendizaje antes de que pueda enviar tramas de datos.

- El conmutador 222 hace al puerto 03 su puerto designado y lo anuncia, reenviando un BPDU de configuración hacia el conmutador 444.
- Algo similar sucede en el conmutador 444. Tan pronto como recibe el BPDU de configuración desde el conmutador 222 en el puerto 01, procede a retirar internamente el puerto 02 y hacer nuevamente al puerto 01 su puerto raíz. Pero este puerto también necesita pasar primero a través de los estados de escucha y aprendizaje. El puerto 02 permanecerá en el modo de bloqueo, ya que el protocolo STP determina que este ofrece una conexión redundante hacia el conmutador raíz (es decir, recibe BPDUs de configuración inferiores a los que recibe en el puerto raíz).

Como consecuencia de la recuperación del enlace, desde la perspectiva de las estaciones conectadas a los conmutadores 222 y 444 se percibe nuevamente una caída en la red, cuyo tiempo total corresponde a lo siguiente:

- Tiempo Total = 15 segundos (estado de escucha) + 15 segundos (estado de aprendizaje) = 30 segundos.

2.7.6.2 Análisis para el caso RSTP

Cuando el conmutador 222 pierde su conexión al conmutador raíz, inmediatamente decide por si mismo que él es el nuevo conmutador raíz (porque no tiene ningún otro puerto alternativo en ese momento) y comienza a publicar esto en su propio BPDU al conmutador 444. El conmutador 444 reconoce a los BPDUs recibidos del conmutador 222 como inferiores, y concluye inteligentemente que su conexión al conmutador raíz de la red a través del conmutador 222, no se encuentra funcionando.

Entonces, la reacción del conmutador 444 es activar inmediatamente su camino secundario a través del puerto 02 (él si tiene un puerto alternativo y por tanto, lucha por mantener al conmutador 111 como la raíz de su topología), hace internamente a este puerto 02 su nuevo puerto raíz e inmediatamente lo coloca en su estado de reenvío, sin esperar nada. Al mismo tiempo, hace internamente al puerto 01 su puerto designado, pero no puede activarlo inmediatamente (debido a que tiene una definición topológica pendiente con el conmutador 222) y por tanto, publica la nueva ruta activa que tiene hacia el conmutador raíz 111, enviando un BPDU de RSTP al conmutador 222, es decir intenta revirar o revertir la situación topológica planteada por este previamente. El conmutador 222 recibe, entiende y acepta esta información (analiza el BPDU recibido y se da cuenta que él no puede ser la raíz) y por tanto hace al puerto 03 su nuevo puerto raíz.

En una nueva etapa y resuelto el tema de la definición topológica entre ambos conmutadores, el conmutador 444 procede a realizar un intercambio de mensajes de protocolo conocido como “operación de sincronismo” con el conmutador 222 para ahorrar tiempo y hacer pasar al puerto 01 directamente hacia el estado de reenvío. Esta

operación requiere un intercambio previo de BPDUs de RSTP, donde ya no se utiliza temporizadores de por medio, y por ende esto sucede de forma muy rápida. Por lo explicado, el tiempo que pasa hasta que todos los conmutadores RSTP se ponen de acuerdo en la nueva topología podría tomar menos de un segundo.

Para saber que sucede cuando el enlace entre el conmutador 111 y el conmutador 222 se restaura, veamos lo siguiente:

- Cuando el conmutador raíz (111) detecta un enlace en el puerto 03, entonces comienza el proceso de sincronismo con el conmutador 222 para llevar este puerto nuevamente hacia el estado de reenvío. Esto requiere que el conmutador 111 envíe primero un BPDU al conmutador 222 con la bandera de "Propuesta" activada.
- El conmutador 222 reconoce este BPDU como superior (es decir, se ha recibido en el camino más corto hacia la raíz) y reacciona generando previamente una señal SYNC. Este mensaje SYNC es una señal que se envía a todos sus puertos designados (no los de borde) para que pasen al modo de bloqueo.
- Luego, el puerto 01 envía un BPDU de retorno al conmutador raíz con la bandera "Acuerdo" activada, indicando que el conmutador 222 está de acuerdo con la nueva topología planteada. Esto es una indicación al conmutador 111 de que el puerto 01 puede pasar al estado de reenvío inmediatamente, sin esperar a ningún temporizador.
- Como consecuencia de la reacción anterior ante el mensaje de sincronismo, la ruptura del lazo está ahora entre los conmutadores 222 y 444.
- A continuación, el conmutador 222 repite el proceso de sincronismo (como si fuera una ola que se propaga) para pasar su puerto designado 03 hacia el estado de reenvío. Envía por tanto un BPDU con la bandera de "Propuesta" activada hacia el conmutador 444, el cual retira al puerto 02 de su rol de puerto raíz y hace internamente que el puerto 01 sea su nuevo puerto raíz.
    - Luego, también reacciona a este mensaje enviando previamente un SYNC en simultáneo a todos sus puertos designados (no los de borde) para llevarlos al modo de bloqueo y retoma luego un BPDU de respuesta al conmutador 222 con la bandera de "Acuerdo" activada.
    - Esto informa al conmutador 222 de que puede proceder a colocar al puerto 03 en el estado de reenvío.

Como consecuencia de la recuperación del enlace, debe notarse que la ruptura final del lazo ha sido movida hacia el enlace que conecta los conmutadores 444 y 333. También debe notarse que el proceso completo con RSTP no requiere temporizadores (ni en la ruptura ni en la recuperación del enlace) y puede tomar menos de un segundo hasta que todos los conmutadores se pongan de acuerdo en la nueva topología.

2.7.6.3 Tiempo de Convergencia

Pese a todas las mejoras mencionadas, hay que notar que el tiempo de convergencia inicial en RSTP después de que todos los conmutadores son conectados y encendidos por primera vez, es similar al que se consigue en STP.

De cualquier modo, una vez que la red se vuelve estable y todos los conmutadores se ponen de acuerdo en la topología LAN existente, cualquier cambio subsiguiente (la falla en un enlace por ejemplo) es propagado rápidamente sin la necesidad de los temporizadores tradicionales de STP.

Dependiendo de la complejidad de la red, el tiempo que se requiere para establecer una nueva topología en RSTP puede variar desde decenas de milisegundos hasta varios segundos.

2.8 Desventajas del Modelo Clásico

El marco teórico presentado en este capítulo, ha servido para mostrar los fundamentos tecnológicos que dan soporte a la actual metodología de diseño de redes de campus basada en el modelo jerárquico y el enfoque de bloques de construcción que fue explicado con detalle.

Como ya se ha podido apreciar, este modelo tiene desventajas que desalientan o hacen complicado el diseño de redes de campus, por las razones que resumo a continuación:

- El protocolo RSTP que da el fundamento a la característica de resiliencia dentro del bloque básico, aun cuando ha mejorado significativamente en su tiempo de convergencia respecto al protocolo STP, todavía mantiene tiempos de convergencia de varios segundos en el caso de que se produzcan fallas en alguno de los conmutadores raíz configurados en el protocolo.
- El protocolo RSTP y los conceptos relativos a este, son complejos de implementar, más todavía cuando existen varias VLANs en un bloque básico lo cual es cada vez más común en esta época.
    - Esto es debido a que es necesario activar una instancia del protocolo RSTP por cada VLAN existente, haciendo que el diseñador tenga que ser muy cuidadoso y minucioso al momento de configurar los equipos.
    - Este tedioso proceso desalienta el incorporar la resiliencia como una característica básica en el diseño, o peor aún, puede generar errores que desestabilicen la red al momento que es necesario introducir cambios en la topología.

- Si la red de campus es grande, la necesidad de utilizar conmutadores de núcleo con capacidades simultaneas de capa 3 y alto rendimiento, hace que el diseño sea costoso y nuevamente se desalienta al diseñador en su intento de añadir resiliencia en el diseño.
- La utilización de un protocolo de enrutamiento en la capa núcleo (normalmente OSPF para asegurar el máximo rendimiento), unido a la complejidad inherente ya mencionada del protocolo RSTP, hace que en su conjunto, el mantenimiento del diseño de la red de campus sea una tarea delicada y no susceptible a cambios constantes, lo cual no corresponde a la necesidad actual de las Empresas que utilizan su infraestructura cada vez más de modo cambiante.

Estas desventajas que he mostrado en el modelo de diseño clásico de redes de campus, hace que proponga una alternativa de mejora que explicaré a continuación en el siguiente capítulo.

CAPÍTULO III
VIRTUALIZACIÓN EN REDES DE CAMPUS

En el capítulo anterior se presento el modelo jerárquico clásico y el enfoque de bloques de construcción, ambos aplicados al diseño tradicional de redes de campus. Aunque estos mecanismos han sido y siguen siendo muy útiles en el diseño, se observo al hacer una revisión del bloque básico que la utilización del protocolo RSTP para obtener la deseada característica de resiliencia en el modelo, no es completa porque existen siempre eventos que causan una disrupción importante en el servicio a los usuarios y adicionalmente, agrega una complejidad creciente cuando se requiere mantener varias VLANs en la red, como es la práctica común en esta época de convergencia de aplicaciones y servicios.

Por lo anterior, en este capítulo voy a presentar una mejora al modelo clásico de diseño, tomando lo mejor de este como es el enfoque basado en bloques, pero simplificando las reglas de construcción mediante el uso de una tecnología de virtualización llamada SMLT. Esta tecnología me va a permitir prescindir totalmente de la utilización del protocolo RSTP (con lo cual retiro la complejidad en el proceso de diseño) y a la vez fortalecer la característica de resiliencia, que es el objetivo de diseño en el cual se centra el presente informe.

3.1 Cambiando el Modelo Clásico

Las mejoras que propongo para el modelo de diseño que se reviso en el capítulo 2, se basan en un conjunto de reglas que paso a explicar a continuación.

3.1.1 Regla 1: Modelo de 2 capas

El modelo jerárquico de diseño estaba basado en 3 capas, principalmente porque en sus inicios los dispositivos de la capa de distribución no tenían la capacidad para realizar las funciones de capa 3 y capa 4 en ese nivel y a su vez, realizar conmutación de muy alta capacidad en capa 2. Esto ya no es así en la actualidad, dado que ya existen dispositivos que realizan conmutación multicapa hasta un nivel que supera los 100 MPPS y más.

Adicionalmente se asumía que solo ciertos dispositivos de la capa núcleo tenían la cantidad y tipo de puertos de alta capacidad (10G Ethernet por ejemplo) como para ser

ubicados en esa posición. Actualmente los dispositivos de la capa de distribución también cuentan con estas características.

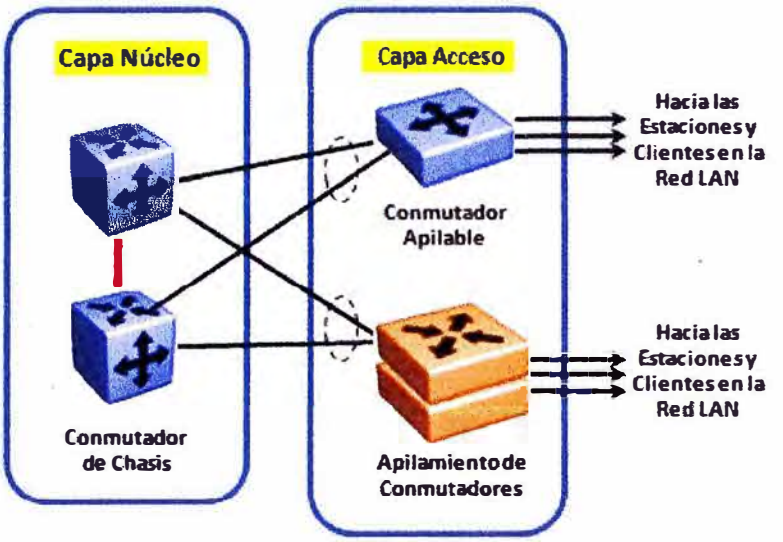


Fig. 3.1 Modelo de 2 Capas

Por lo explicado, el primer cambio propuesto para el modelo jerárquico es el siguiente:

- Las capas del modelo deben reducirse en lo posible de tres a dos, quedando solo una capa núcleo (que resume las capas núcleo y distribución del modelo clásico en una sola) y una capa de acceso. Esto puede apreciarse en la figura 3.1.
    - Existen situaciones especiales en las que debido a las distancias involucradas dentro del campus o las rutas propias del cableado, puede ser necesario extender el modelo a 3 capas. Esto se aprecia en la figura 3.2.

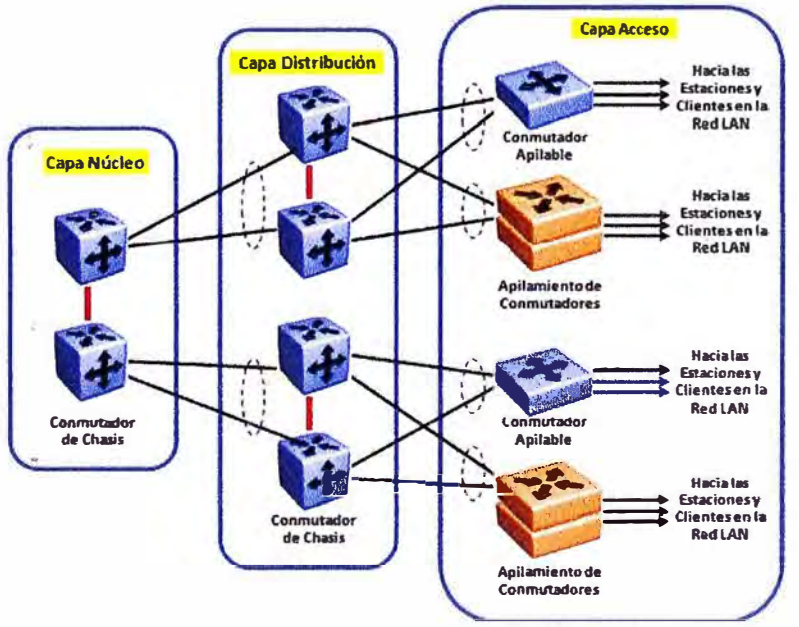


Fig. 3.2 Modelo de 3 Capas

- Sin embargo, el modelo de 2 capas debe ser utilizado con preferencia y siempre que sea posible, porque reduce la inversión en equipamiento, simplifica la arquitectura de la red y facilita la incorporación de la resiliencia en el diseño.

- Se debe utilizar solo el bloque básico como el único bloque de construcción para el diseño de campus. El bloque núcleo se hace innecesario al incorporarse sus funciones dentro del bloque básico. Esto hace que la conectividad entre bloques se haga de forma directa, como será explicado más adelante.

3.1.2 **Regla 2: Enrutamiento solo en el Núcleo**

Dentro del bloque básico reducido (2 capas) y desde la perspectiva de las funciones de red, existen dos opciones que pueden ser consideradas en el diseño:

- Realizar funciones de capa 2 en la capa de acceso y funciones de capa 3 en la capa núcleo/distribución.
- Realizar funciones de capa 3 en la capa de acceso y funciones de capa 3 en la capa núcleo/distribución.

Aunque existen dispositivos que permiten configurar ambas opciones, la primera opción es la preferida porque simplifica el diseño, facilita la administración y permite incorporar la resiliencia de una forma más sencilla con la utilización de tecnología de virtualización que vamos a explicar más adelante en este capítulo. Esto significa que se mantienen las VLANs de capa 2 en la capa de acceso y se hace enrutamiento entre VLANs en la capa núcleo.

La arquitectura simplificada de dos capas indicada en la regla 1, soporta que existan funciones de capa 2 o capa 3 en la capa de acceso. Como una regla práctica, el planteamiento propuesto es el siguiente:

- Si se están concentrando menos de 3,000 dispositivos en el acceso (la gran mayoría de los casos), es recomendable utilizar conmutación de capa 2 entre la capa de acceso y el núcleo.
- Si existen más de 3,000 dispositivos, sería necesario utilizar enrutamiento de capa 3 entre el acceso y el núcleo, porque esto permite distribuir las tablas ARP que puede soportar cada equipo en el núcleo y simplifica la creación de subredes. Sin embargo, para esta cantidad de dispositivos la recomendación será siempre dividir la infraestructura en bloques básicos más pequeños.

3.1.3 **Regla 3: Agrupamiento de Enlaces**

La tercera regla es la más importante que propongo como cambio en el modelo clásico de diseño de redes de campus. La propuesta es la utilización del concepto de agrupamiento de enlaces para conseguir la característica de resiliencia (que es objetivo

principal de este informe) dentro del bloque básico reducido, en reemplazo del uso del protocolo RSTP (IEEE 802.1w) que se usa en combinación con la característica PVRSTP (para el caso de los conmutadores Cisco) o el protocolo MST (IEEE 802.1s) para el caso de dispositivos de otras marcas.

Explicare a continuación en qué consiste la técnica de agrupamiento de enlaces y como se vincula con la tecnología de virtualización SMLT, las cuales en conjunto modifican completamente la forma de diseño de redes de campus.

3.2 Agrupamiento de Enlaces

El agrupamiento de enlaces consiste en una técnica para vincular lógicamente varios enlaces Ethernet entre dos dispositivos conmutadores, de tal forma que se comporten como si fueran uno solo 8 (ver figura 3.3). Esta configuración es necesaria para evitar que exista un lazo entre estos dos dispositivos.

Cuando esta técnica se aplica entre la capa de acceso y la capa núcleo del bloque básico, la ventaja que se obtiene es que se incrementa la resiliencia y el ancho de banda de la conexión entre ambas capas.

La técnica de agrupamiento de enlaces fue desarrollada originalmente por el fabricante Cisco quien la denomino "Etherchannel" y posteriormente le siguieron otros enfoques propietarios muy similares de fabricantes como Nortel (ahora Avaya), hasta que finalmente la IEEE especifico un estándar para esta técnica.

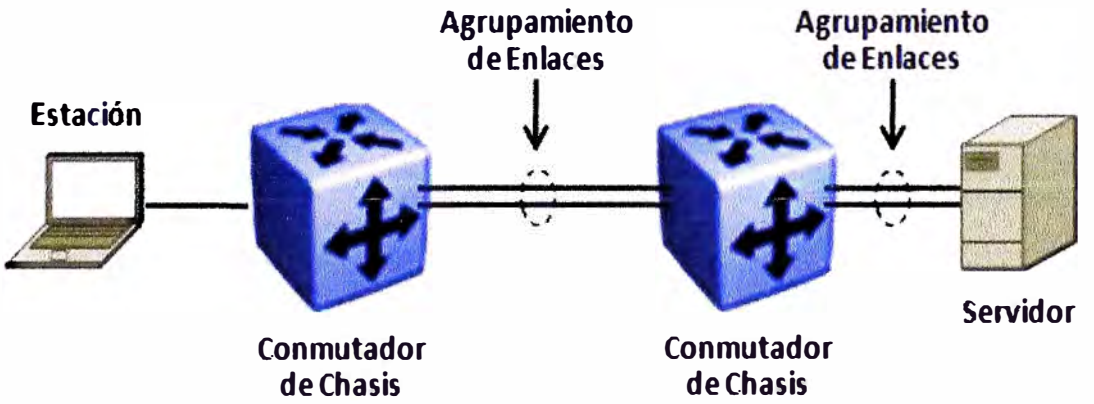


Fig. 3.3 Agrupamiento de Enlaces

3.2.1 MLT

MLT (MultiLink Trunking, por su denominación en inglés) es la técnica propietaria del fabricante Nortel (ahora Avaya) para realizar agrupamiento de enlaces Ethernet entre un par de conmutadores o entre un conmutador y un servidor que tenga varias interfaces de red. Al aplicarse MLT en un grupo de enlaces Ethernet, se dice que se genera un grupo MLT.

Cuando se produce una falla física en cualquiera de los enlaces del grupo MLT, MLT hace que el tráfico correspondiente a ese enlace cambie inmediatamente en menos de 1 segundo hacia otro de los enlaces restantes. Después de que el enlace es recuperado, este se integra nuevamente al grupo MLT, también en menos de 1 segundo. Los puertos de un grupo MLT son estáticos en su configuración, esto significa que no hay un protocolo especial de capa 2 entre los conmutadores.

Por ejemplo, MLT puede hacer que 4 enlaces separados de 100 Mbps parezcan como una sola troncal de 400 Mbps. Si uno de los enlaces falla, el ancho de banda agregado se reduce a 300 Mbps, pero la troncal misma se mantiene activa y continúa reenviando tráfico mientras se mantenga al menos un enlace físicamente activo.

3.2.1.1 DMLT

DMLT (Distributed MLT, por su denominación en inglés) es una variante de MLT que añade la capacidad de terminar el extremo opuesto de los enlaces físicos de un grupo MLT en:

- Diferentes conmutadores del tipo apilable, que se encuentran conectados entre sí para formar una pila unificada de conmutadores (ver figura 3.4).
- Diferentes módulos de un conmutador tipo chasis.

Con esta capacidad, DMLT incrementa la resiliencia de los enlaces de subida que van desde el cuarto de cableado hacia el núcleo de la red. Por tanto, una falla de un conmutador apilable o de un módulo de un conmutador tipo chasis, no interrumpiría la comunicación entre ambas capas de la red.

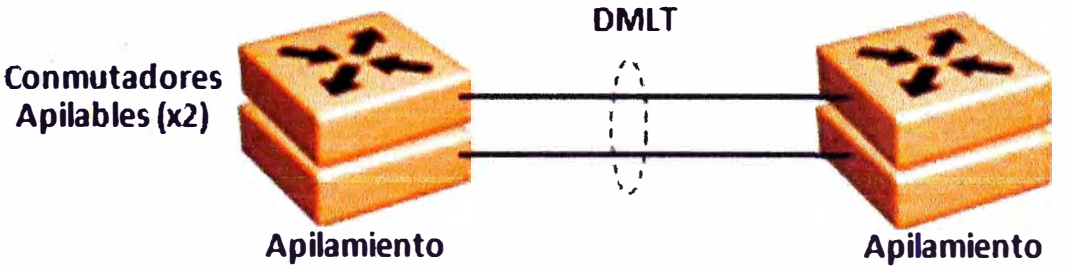


Fig. 3.4 Agrupamiento con DMLT

3.2.1.2 Operación de MLT

Los protocolos de capas superiores ven a un grupo MLT (también llamado troncal) como si fuera una sola interface lógica. Por ejemplo, si el protocolo RIP aprende una nueva ruta en uno de los puertos del grupo, la interface del siguiente salto es la troncal misma. El puerto específico escogido para realizar el reenvío de un paquete IP es transparente para RIP. Esta regla de aprendizaje también aplica a las direcciones MAC

de capa 2 aprendidas en un puerto de la troncal. La dirección MAC es identificada en la base de datos del conmutador como si fuera aprendida sobre la troncal MLT, no sobre un puerto específico (ver figura 3.5).

El algoritmo de MLT escoge un puerto físico para el reenvío mediante un cálculo interno realizado sobre las direcciones IP del paquete a transmitir. Debido a que muchas aplicaciones requieren que los paquetes de una sesión dada arriben en secuencia, MLT utiliza consistentemente el mismo camino para cualquier par de direcciones IP fuente-destino. MLT no mantiene un rastreo de estas sesiones (no opera en capa 4), simplemente aplica el mismo algoritmo de selección de camino a cada paquete IP, por lo cual la misma combinación de direcciones fuente-destino produce siempre la misma selección de enlace.

El algoritmo de reenvío de MLT fue creado para compartir la carga de tráfico entre los enlaces que conforman una troncal, y al mismo tiempo asegurar que los paquetes no se desordenen ni lleguen fuera de secuencia.

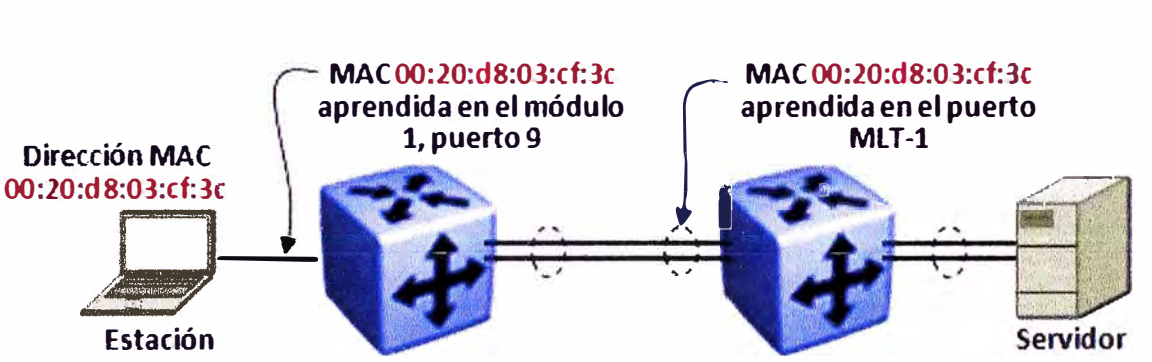


Fig. 3.5 Aprendizaje por MLT

3.2.2 IEEE 802.3ad

El estándar IEEE para el agrupamiento de enlaces es denominado 802.3ad. Este estándar utiliza el protocolo LACP (Link Aggregation Control Protocol, en inglés) para juntar múltiples enlaces físicos en un solo grupo lógico denominado LAG (Link Aggregation Group, en inglés). El cumplimiento con este estándar asegura la interoperabilidad de los agrupamientos de enlaces entre diferentes fabricantes de equipos como conmutadores o tarjetas de red.

3.2.2.1 Comparación con MLT

Aunque los agrupamientos basados en IEEE 802.3ad y MLT proveen servicios similares, existen diferencias que menciono a continuación:

- MLT es estáticamente definido mientras que con 802.3ad el agrupamiento es dinámico porque se realiza con el uso del protocolo LACP.

- LACP detecta dinámicamente si algún enlace (configurado como latente) puede ser agregado a un grupo (porque hay otro enlace que fallo previamente) y lo hace cuando esto es requerido.
- LACP (como en DMLT) soporta la formación de grupos a través de diferentes conmutadores apilados entre sí o a través de diferentes módulos en un conmutador tipo chasis, a fin de ofrecer mayor resiliencia frente a una falla de estos componentes.
- LACP realiza una verificación lógica de la conexión entre dos conmutadores directa y físicamente conectados. Esto significa que, en el raro evento de que un conmutador se vuelva inoperativo lógicamente (todo el equipo o solo en algún puerto) pero que sin embargo mantenga el estado de sus enlaces como operativos a nivel físico, los datos que se transmiten pueden perderse sin que exista forma de advertirlo. La característica de verificación lógica permite remover automáticamente el puerto "congelado" del grupo LAG, asegurando que no se transmite ningún dato al vacio. MLT no cuenta con esta característica por sí solo.

3.2.2.2 Limitaciones de LACP

Aunque 802.3ad cuenta con la ventaja de ser un estándar, tiene algunas limitaciones que es conveniente señalar:

- El protocolo LACP fue diseñado para operar entre dos conmutadores directa y físicamente conectados y no trabaja en un modo de extremo a extremo, si es que por condiciones de la red existen dispositivos intermedios entre los dos conmutadores. Esto sucede por ejemplo, cuando existe de por medio una red de anillo óptico o un servicio tipo Metro Ethernet provisto por un operador de servicios.
- Los tiempos de recuperación del protocolo LACP son más altos que los usados en MLT. El valor por defecto es un tiempo de muestreo de 30 segundos y se considera que con 3 muestreos fallados el puerto asociado debe ser deshabilitado. Esto llevaría a un tiempo de recuperación de hasta 90 segundos, que es demasiado para conseguir una resiliencia adecuada. Algunos fabricantes pueden reducir el tiempo de muestreo a 1 segundo, de esta manera reduciendo el tiempo de recuperación a 3 segundos que es algo más aceptable.

3.2.3 VLACP

Para salvar la limitación de MLT de trabajar en modo estático, el fabricante Nortel desarrollo el protocolo VLACP (Virtual LACP), el cual ofrece un mecanismo real de detección de fallas de extremo a extremo entre conmutadores directamente conectados o que tengan conectividad a través de una red intermedia (ver figura 3.6).

Esta característica añade un gran nivel de resiliencia y flexibilidad para el diseño de campus y por tanto será utilizado en el ejemplo que se presenta en el capítulo siguiente, en conjunto con MLT, DMLT y SMLT.

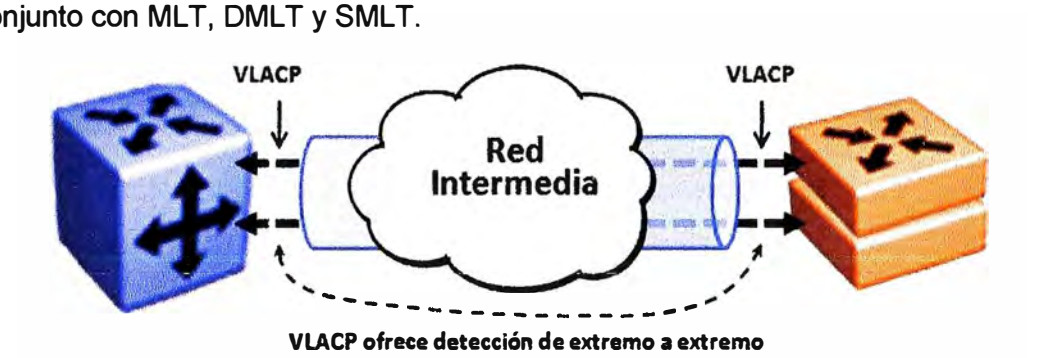


Fig. 3.6 Configuración usando VLACP

No debe confundirse VLACP con LACP del estándar IEEE 802.3ad. El primero no requiere la operación del segundo. LACP ofrece capacidades de agrupamiento de enlaces y detección de fallas punto a punto, mientras que VLACP solo provee la detección de fallas de extremo a extremo. Cabe mencionar que el fabricante Cisco ofrece un mecanismo similar denominado UDLD (UniDirectional Link Detection), que está orientado al mismo objetivo de detección de fallas en interfaces ópticas.

Entre las características más importantes de este protocolo VLACP, se puede mencionar las siguientes:

- Diseñado para operar de extremo a extremo, sin interesar si los conmutadores están directamente conectados o si existe una conexión intermedia entre ellos.
- Solo consiste en un mecanismo de detección del tipo pulso de vida, sin capacidad de realizar agrupamiento de enlaces.
- Se configura individualmente por cada puerto del conmutador.
- Presenta una muy ligera carga de procesamiento.
- La operación de VLACP consiste en enviar tramas de datos periódicamente por cada enlace donde está configurado. Si estas tramas no son recibidas en un enlace, este enlace es deshabilitado en un periodo de tiempo que es configurable.
- Puede correr independientemente como un protocolo de puerto a puerto, o encima de la operación de MLT, DMLT o SMLT.

Como veremos más adelante, VLACP es una característica crítica cuando se quiere desplegar redes con resiliencia que usan tecnología de virtualización como SMLT. VLACP puede detectar fallas de extremo a extremo como se ha explicado y puede deshabilitar enlaces de un conmutador, que pueden tener conectividad de enlace (capa 1) pero que por alguna razón no pueden procesar tráfico debido a una falla lógica interna. Esto será discutido con más detalle en la sección siguiente.

3.2.4 Distribución del Tráfico

Todas las variaciones de agrupamiento de enlaces (MLT, DMLT, SMLT, 802.3ad) utilizan un algoritmo de cálculo (hashing) que distribuye el tráfico a través de los enlaces físicos de una troncal. Normalmente el tráfico es distribuido por sesión, de tal forma que los paquetes nunca arriban a su destino en desorden.

- Debe observarse que el algoritmo de cálculo no necesariamente tiene que ser idéntico en ambos lados de la troncal, y por ende MLT/DMLT/SMLT y 802.3ad son interoperables con la mayoría de fabricantes, tanto en el caso de conmutadores como de tarjetas de red.
- Como regla general, los algoritmos están basados en las direcciones IP fuente-destino o las direcciones MAC fuente-destino.

3.2.5 Criterios de Diseño

A continuación presento una serie de criterios respecto al agrupamiento de enlaces que serán utilizados en el modelo de diseño propuesto:

- Los enlaces de subida que vienen desde el cuarto de cableado (capa de acceso) deben usar DMLT en lo posible para terminar los enlaces en diferentes conmutadores del apilamiento o en diferentes módulos si se utiliza un conmutador de chasis.
- Todos los enlaces en un grupo MLT deben ser:
  - De la misma velocidad y bidireccionales (dúplex).
  - Preferiblemente del mismo tipo de medio (Ejemplo: 1000BaseT, 1000BaseSX, 1000BaseLX).
- Debe deshabilitarse el protocolo STP/RSTP en ambos lados de los puertos pertenecientes a un grupo MLT/DMLT. Esto será ampliado en la sección siguiente referente a SMLT.
- Todos los enlaces de subida y troncales (puerto por puerto) deben usar VLACP. VLACP es usado preferentemente en reemplazo de LACP porque ofrece menos carga de procesamiento y menos tiempo de recuperación (menos de 1 segundo, frente a 3-4 segundos de LACP).
- Utilizar LACP cuando se están conectando conmutadores de diferentes fabricantes (Nortel y Cisco, por ejemplo).
- Utilizar LACP si se quiere aprovechar la propiedad de enlaces sobrantes de respaldo.
  - LACP soporta hasta un máximo de 8 enlaces activos. Los adicionales enlaces que se configuren son puestos en estado de espera.
  - El estado activo o en espera se determina por un parámetro de prioridad configurable por peso.

- No es necesario habilitar VLACP en una troncal LACP. Esto añadiría carga al procesador del equipo y no traería ningún beneficio.

3.3 SMLT

3.3.1 Introducción

En esta sección se mostrará como el agrupamiento virtualizado de conmutadores (switch clustering en inglés) utilizando la tecnología SMLT ofrece la característica de resiliencia necesaria en el diseño de una red de campus convergente (ver figura 3.7).

La utilización de enlaces redundantes que reenvían tráfico simultáneamente sin utilizar el protocolo RSTP permite lograr esto. La recuperación automática en menos de un segundo y la simplicidad en el diseño de red que se logra sin RSTP reducen el costo de administración de la red, y aseguran que aplicaciones como el video por demanda o la telefonía IP funcionen correctamente en el caso que se produzca una falla en la red.

El agrupamiento virtualizado de conmutadores también brinda la posibilidad de realizar actualizaciones en el núcleo de la red sin necesidad de detener el tráfico. Esto es posible cuando existen conexiones redundantes hacia un núcleo de 2 conmutadores, pudiendo desconectarse totalmente uno de los equipos sin afectar el tráfico circulante a efectos de realizar una actualización de mantenimiento. Luego este equipo retirado puede ser devuelto a su posición original sin que se afecte la operación. Repitiendo la operación con el segundo equipo, es posible lograr una actualización completa del núcleo sin la necesidad de prescindir de los servicios de red y en la hora más conveniente para el administrador.

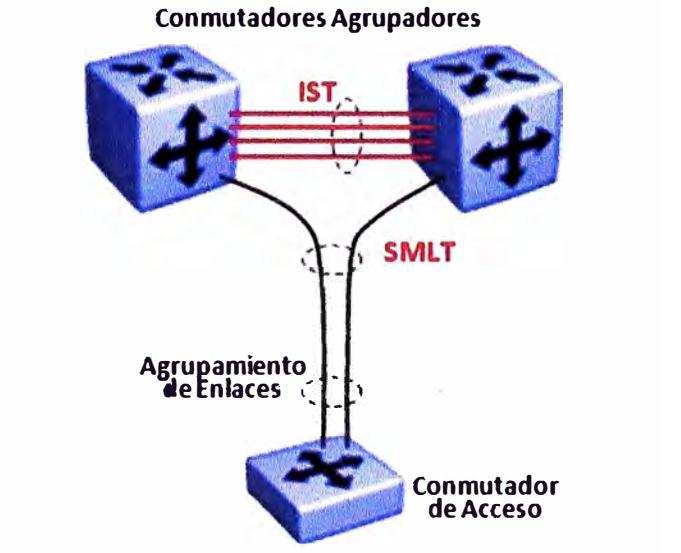


Fig. 3.7 Configuración con SMLT

Una ventaja adicional del agrupamiento virtualizado es que permite trabajar transparentemente con cualquier dispositivo que soporte algunas de las formas de

agrupamiento de enlaces que se reviso en páginas anteriores. Esto incluye dispositivos de múltiples marcas, servidores o equipos especiales.

3.3.2 Conceptos

SMLT es una forma mejorada del agrupamiento de enlaces, desarrollada originalmente por el fabricante Nortel (ahora Avaya) la cual permite combinar (o virtualizar) dos conmutadores en un solo dispositivo lógico que realiza funciones de red de capa 2.

SMLT conecta una troncal MLT de un conmutador en un extremo, a un par de conmutadores en el otro extremo, de tal forma que los enlaces de esta troncal se reparten entre estos dos equipos. A estos 2 equipos de llegada se les denomina conmutadores SMLT o conmutadores agrupadores, según se ve en la figura 3.1.

El uso de SMLT hace que RSTP sea innecesario de configurar en una red de campus y por lo tanto, vuelve obsoleta toda la complejidad de su operación que fue estudiada en el capítulo anterior.

SMLT provee caminos redundantes y sincronización de tablas MAC a través de un par de conmutadores agrupadores, lo cual resulta en un rápido mecanismo de recuperación y un real balanceo de carga para el tráfico que viene de la capa de acceso.

3.3.3 Modo de Operación

En SMLT, los dispositivos agrupadores son conectados entre sí por una troncal MLT que transporta un protocolo especial propietario denominado IST (Interswitch Trunk en inglés). Esta troncal IST transporta una VLAN dedicada sobre la cual se maneja el tráfico de control entre ambos equipos. Para el caso de los dispositivos de acceso, solo se utiliza MLT u otra técnica de agrupamiento de enlaces.

SMLT combina los 2 dispositivos agrupadores en una forma tal que aparentan ser un solo dispositivo lógico virtualizado. Desde el lado del acceso, el enlace conformado por múltiples puertos, es tratado como si terminará en el mismo equipo en el lado del núcleo. Por tanto, el conmutador de acceso reenvía el tráfico sobre el grupo MLT o grupo LAG utilizando sus propias reglas de cálculo y distribución de tráfico.

Los dos conmutadores agrupadores del núcleo utilizan la troncal IST para compartir información aprendida en capa 2, de tal forma que se comporten como si fueran un solo equipo.

Cualquier dirección MAC nueva aprendida por la troncal SMLT, inicia una actualización a través de la troncal IST hacia el otro conmutador. El resultado es que las bases de datos de ambos conmutadores están sincronizadas para todas las VLANs que se transportan por la troncal SMLT.

En el ejemplo de la figura 3.8 siguiente, la dirección MAC de la estación "A" es aprendida por el conmutador de acceso en el puerto 1 de la VLAN 2, la cual también se extiende hasta el núcleo de la red para efectos de enrutamiento. Cuando un paquete desde la estación "A" es reenviado hacia el núcleo, solo un enlace del grupo LAG es escogido basado en las reglas del 802.3ad.

El conmutador de núcleo aprende la dirección MAC sobre el enlace SMLT dentro de la VLAN 2 y luego, utiliza la VLAN de control de la troncal IST para informar al otro conmutador SMLT que esta dirección MAC ha sido aprendida por el enlace SMLT. El resultado es que las bases de datos de ambos conmutadores incorporan la dirección MAC y reenvían tráfico de retorno hacia la estación "A" utilizando su propio enlace.

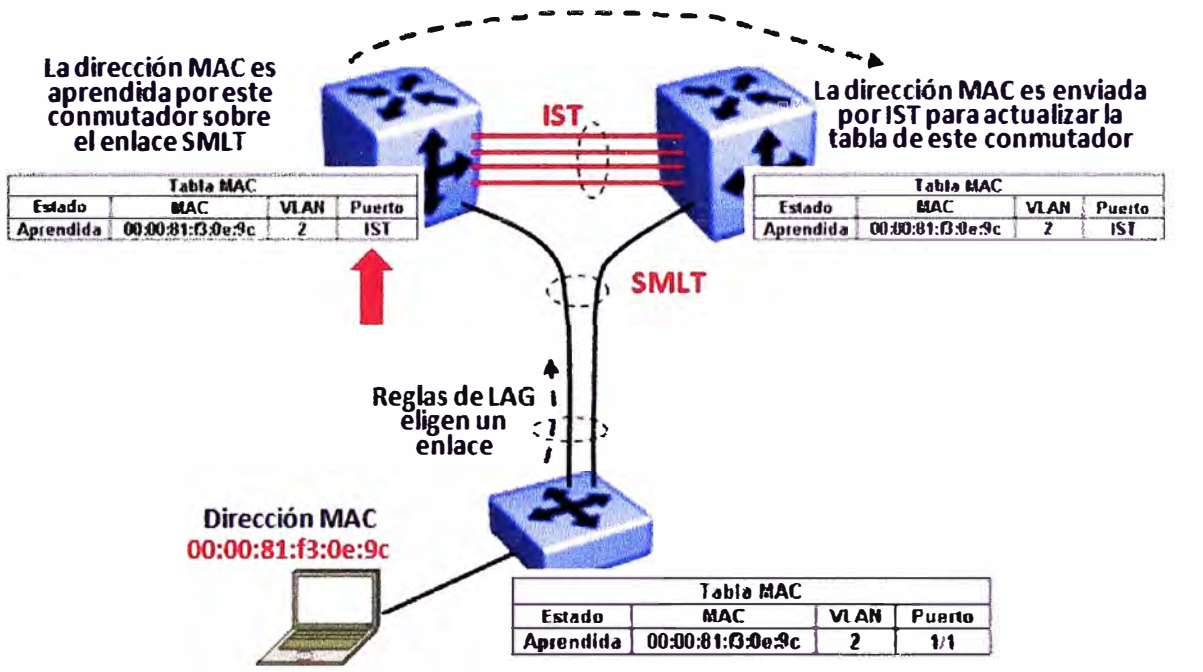


Fig. 3.8 Aprendizaje con SMLT

3.3.4 Ventajas de SMLT

- Ofrece alta escalabilidad
  - Actualmente existe dispositivos conmutadores de núcleo de mediana escala, que soportan hasta 31 troncales SMLT de subida y 8 puertos por cada troncal SMLT. Equipos de mayor capacidad llegan hasta las 127 troncales SMLT.
- Mejora la confiabilidad de la red en capa 2, porque ofrece múltiples caminos activos desde el acceso hasta al núcleo.
- Elimina la posibilidad de que exista un solo punto de falla en el núcleo.
- Ofrece interoperabilidad.
  - Un par de conmutadores SMLT pueden conectarse a dispositivos de otras marcas, como Cisco, utilizando la técnica "Etherchannel" por ejemplo.

- SMLT también es interoperable con 802.3ad.

- Elimina la necesidad de utilizar STP o RSTP.
  - En la figura siguiente se muestra una red con y sin SMLT.
  - Sin SMLT, es necesario usar RSTP para bloquear uno de los enlaces redundantes.
  - Con SMLT, se permite que la red utilice el ancho de banda completo de cada uno de los enlaces físicos disponibles.
- Ofrece una rápida recuperación ante fallas (menor a 1 segundo) y repartición dinámica de la carga de tráfico sobre los enlaces.

3.3.5 **Escenarios de Falla**

Si se utiliza SMLT y uno de los enlaces de subida falla (por un cable o una falla en una interface), el conmutador de acceso continúa transmitiendo datos hacia el núcleo.

Considérese la situación mostrada en la figura 3.9 siguiente, donde se presenta una falla en un enlace. La estación "A" se está comunicando con el servidor "D" y la estación "B" con el servidor "C" sobre los enlaces de la troncal SMLT indicados.

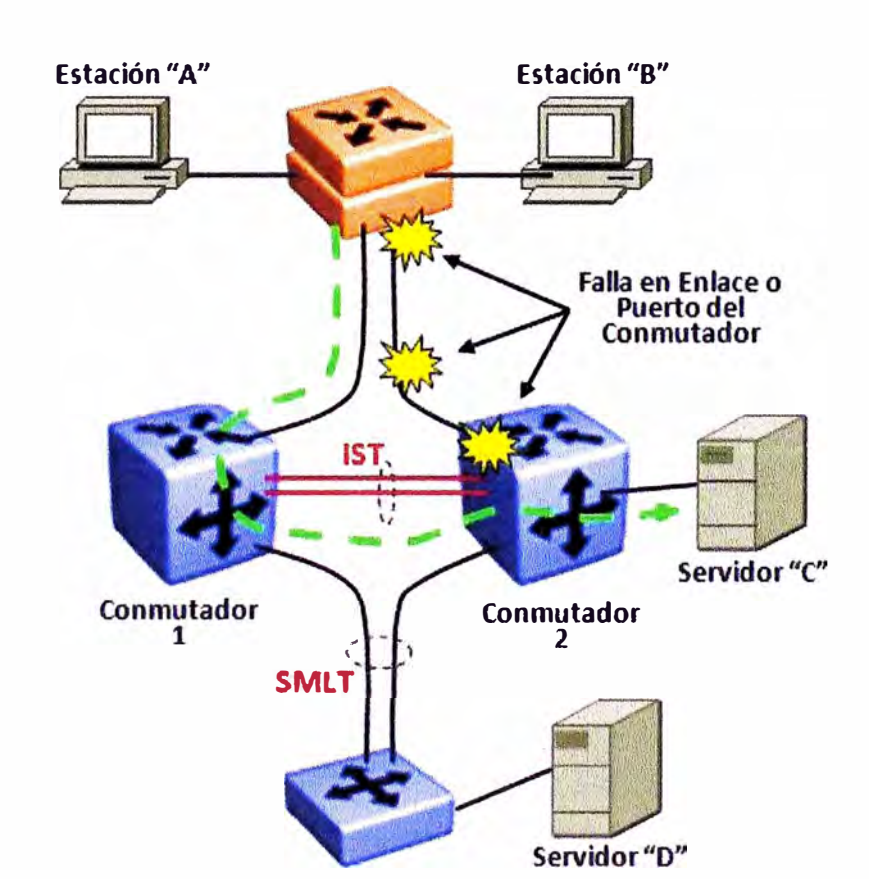


Fig. 3.9 Falla en Enlace con SMLT

Si el enlace de la derecha se cae debido a una falla, la estación "B" todavía se puede comunicar con el servidor "C", porque el tráfico que viene de subida cambia al

enlace de la izquierda hacia el conmutador 1, y de allí cruza la troncal IST para llegar al conmutador 2 y finalmente al servidor "C". Este cambio sucede muy rápido, en menos de 1 segundo.

Considérese ahora la situación mostrada en la figura 3.10 siguiente, donde se presenta una falla total del conmutador 1. La estación "A" ahora se comunica con el servidor "D" a través del enlace de la derecha. Nótese que si hubiera fallado el conmutador 2, la estación "B" no hubiera podido recuperar la comunicación al servidor "C", dado que este no tiene un enlace redundante hacia el núcleo.

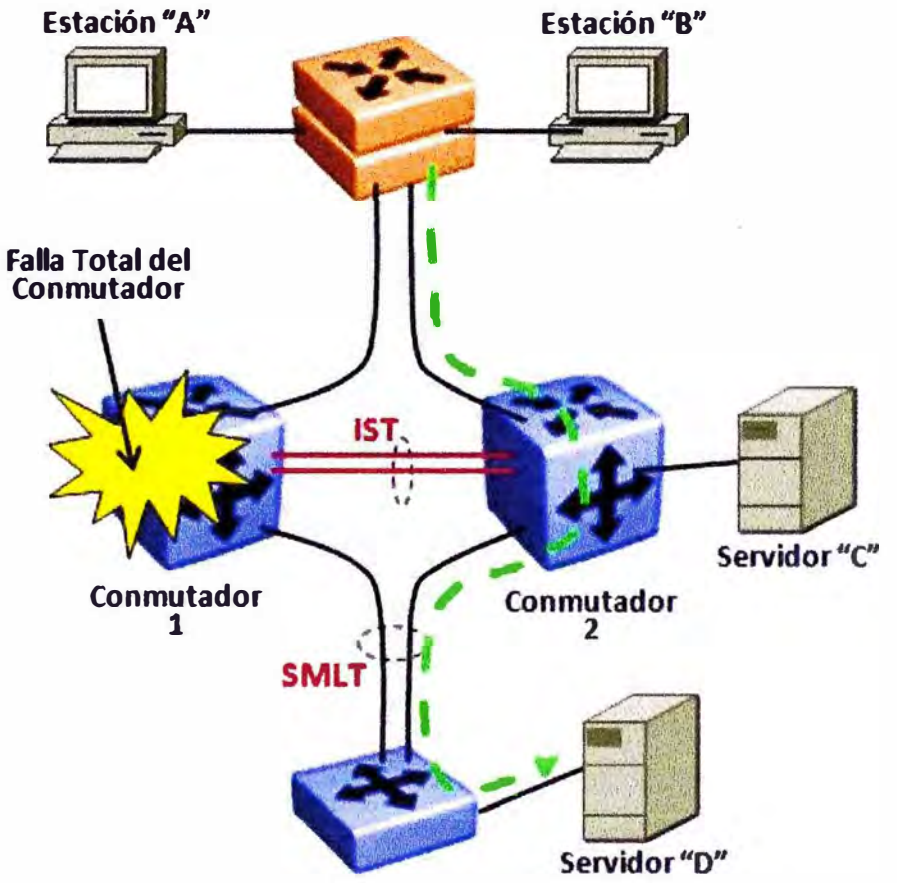


Fig. 3.10 Falla en Conmutador con SMLT

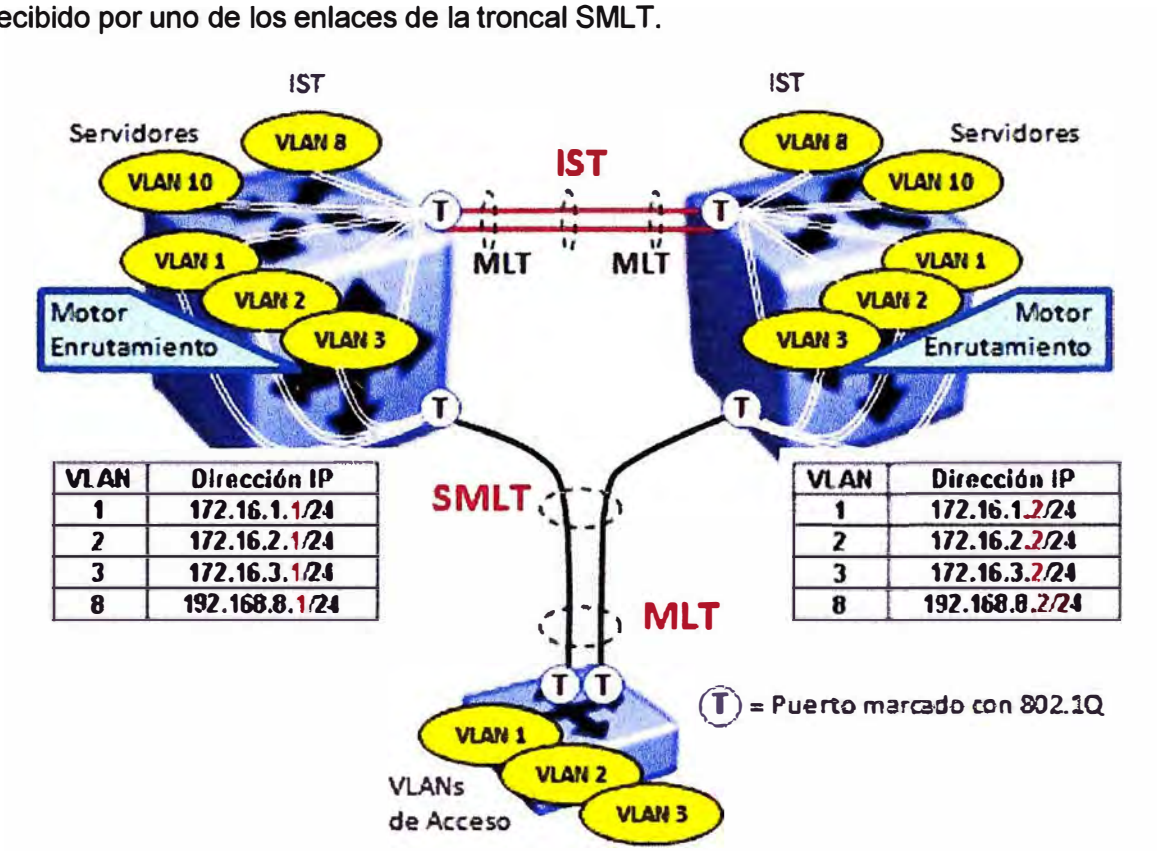
3.3.6 Configuración SMLT detallada

Para un análisis más detallado de la topología SMLT en un bloque básico de construcción, vamos a referimos a la figura 3.11 siguiente:

3.3.6.1 Membresía por VLAN

Sobre una troncal SMLT si es posible configurar múltiples VLANs. Para esto se requiere que ambos conmutadores tengan la misma configuración de VLANs, de tal forma que se produzca un reenvío confiable y predecible. Las direcciones MAC aprendidas sobre una VLAN en un conmutador SMLT son reproducidas hacia la base de datos de la VLAN en el otro conmutador.

En el ejemplo mostrado, las VLANs 1, 2 y 3 son transportadas a través de la troncal SMLT. Todos los conmutadores tienen estas VLANs previamente configuradas. El tráfico desde el conmutador de acceso es enviado por uno de los enlaces del grupo MLT y es recibido por uno de los enlaces de la troncal SMLT.



VLAN	Dirección IP
1	172.16.1.1/24
2	172.16.2.1/24
3	172.16.3.1/24
8	192.168.8.1/24

VLAN	Dirección IP
1	172.16.1.2/24
2	172.16.2.2/24
3	172.16.3.2/24
8	192.168.8.2/24

Fig. 3.11 Configuración SMLT detallada

3.3.6.2 Soporte de 802.1Q

Los puertos de una troncal SMLT pueden ser etiquetados (tagged) con 802.1Q o no. Para lograr consistencia, todos los puertos deben ser configurados de igual forma. La marca con 802.1Q permite el transporte de múltiples VLANs y el beneficio añadido de transportar información de prioridad con 802.1p.

En el ejemplo, todos los enlaces entre conmutadores son configurados para llevar la etiqueta 802.1Q, según se indica con una letra "T".

3.3.6.3 Configuración IST

Todas las VLANs que son miembros de la troncal SMLT, deben ser propagadas a través del enlace entre los conmutadores agrupadores. Este enlace también debe incluir la comunicación IST que se establece sobre interfaces IP sobre una misma VLAN de control existente en ambos conmutadores.

En el ejemplo, VLAN 1, VLAN 2 y VLAN 3 son miembros de la troncal entre los conmutadores, porque son miembros de la troncal SMLT. La VLAN 8 es configurada

como la VLAN IST. Se requiere configurar direcciones IP en esta VLAN de cada lado para efectos de establecer la comunicación.

El enlace entre conmutadores también podría incluir una VLAN que no fuera parte de SMLT. En este caso, se observa que la VLAN 10 corresponde a este caso.

3.3.6.4 Flujo de Tráfico

Las reglas usadas por cada conmutador SMLT para el flujo de tráfico son las siguientes:

- Después de que un paquete es recibido, se realiza una revisión de la base de datos MAC propia del equipo. Si una entrada existe y si fue localmente aprendida (por SMLT) o a través de la troncal IST, el equipo lo reenvía a su destino en su propio puerto local. Esto significa que el paquete no debe ser enviado a través de la troncal IST a menos que no exista otra conexión hacia el destino.
- Los paquetes desconocidos o paquetes de difusión son inundados a todos los puertos miembros de la VLAN.
- Para propósitos de compartición de carga en un ambiente SMLT de varios enlaces, el conmutador SMLT utiliza el mismo algoritmo de distribución de tráfico que MLT.

3.3.6.5 Limitaciones de IST

Se ha dicho antes que para configurar SMLT, es necesario crear una troncal IST entre los dos conmutadores agrupadores. Estos conmutadores utilizan el protocolo IST para compartir información de capa 2 (bases de datos de reenvío) entre ambos. Esto les permite enviar información idéntica hacia el conmutador de borde, de tal forma que este piensa que está conectado a un solo equipo.

Existen ciertas limitaciones en la creación de la comunicación IST:

- Un conmutador de núcleo puede participar en solo una troncal IST a la vez. Esto significa que no se pueden conectar en cadena tres o más conmutadores utilizando IST.
- El protocolo IST es propietario y requiere que se establezca sobre una troncal MLT necesariamente. No es posible usar 802.3ad.
- Se recomienda que la troncal IST (que es una troncal MLT especial realmente) utilice enlaces Gigabit Ethernet como mínimo.

3.4 Nuevo Modelo de Diseño

La utilización de la tecnología de virtualización SMLT explicada en la sección anterior cambia radicalmente el modelo de diseño clásico en redes de campus. A continuación presentare los tipos de topologías de núcleo SMLT que se pueden crear y los nuevos criterios de diseño de campus que se derivan como consecuencia.

3.4.1 Topologías SMLT

Existen 3 tipos de topologías (ver figura 3.12) que surgen de la utilización del concepto de agrupamiento de conmutadores utilizando la tecnología SMLT. El uso de cada topología dependerá del diseño particular de la red donde se requiera manejar la resiliencia.

- Topología Triangulo: Este es el caso más simple, donde solo hay un par de conmutadores de núcleo agrupados virtualmente y la conexión desde la capa de acceso se hace utilizando SMLT.
- Topología Cuadrado: Existen dos pares de conmutadores de núcleo agrupados virtualmente e interconectados entre ellos utilizando SMLT. Este tipo de núcleo puede ser escalado añadiendo más pares de conmutadores.
- Topología Malla Completa: En este caso se expande la topología cuadrado mediante la adición de conexiones adicionales entre los 2 pares de conmutadores del núcleo, de tal forma que cada conmutador tenga al menos una conexión a cualquier otro conmutador del cuadrado. Este tipo de núcleo también puede ser escalado añadiendo más pares de conmutadores.

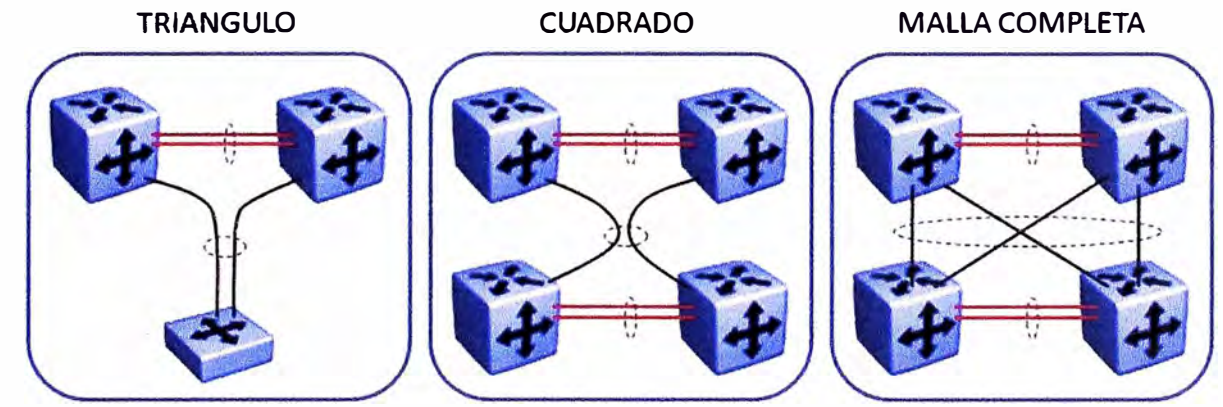


Fig. 3.12 Topologías SMLT

3.4.2 Modelo de Referencia

Con el objetivo de identificar varios aspectos del diseño de redes de campus con SMLT, presentamos en la figura 3.13 siguiente un modelo de referencia que sirve como ejemplo para reunir en una sola vista los conceptos revisados hasta esta parte.

En la figura se muestra una capa núcleo basada en seis conmutadores, donde aparecen las tres topologías revisadas en la sección anterior. Obsérvese que esta cantidad de equipos no es un requisito para crear un agrupamiento de conmutadores.

La conexión de la capa de acceso hacia la capa núcleo se hace utilizando SMLT de 2 o más enlaces según se requiera. En el núcleo se utiliza SMLT entre pares (resiliencia

de capa 2) o RSMLT (Routed SMLT en inglés) que es una variante de SMLT para ofrecer resiliencia de capa 3. RSMLT sería estudiado más adelante.

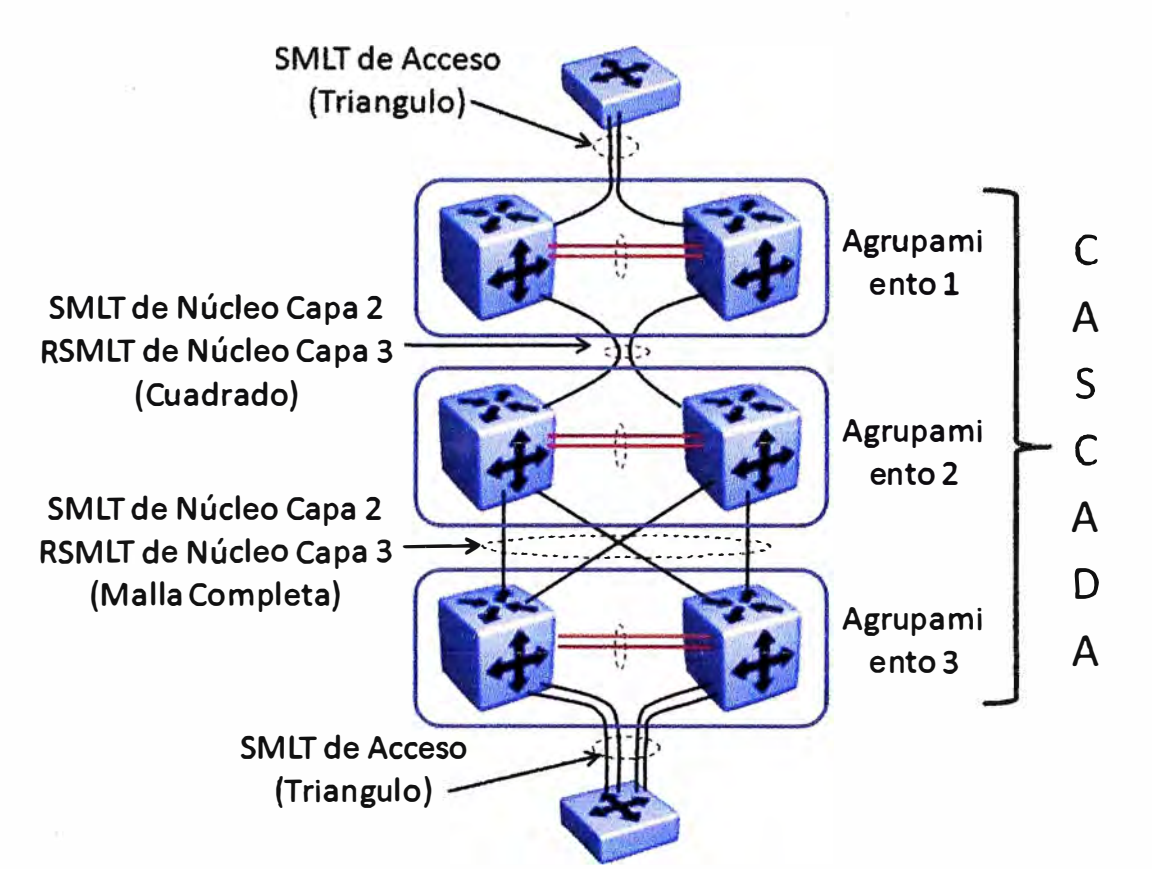


Fig. 3.13 Modelo de Diseño con SMLT

En la interconexión de estas capas, puede tomarse la precaución de activar algún mecanismo de capa 1 para la prevención de lazos, dado que STP/RSTP están desactivados. Esto puede suceder como consecuencia de una conexión errada o la desactivación de algún grupo MLT o LAG. Los fabricantes utilizan algunos mecanismos de este tipo, que dependen del tipo de equipo utilizado, y que por ser específicos de cada plataforma no serán revisados a detalle en este informe.

3.4.2.1 Diseño de 2 Capas

Con el diseño básico de 2 capas, los conmutadores de acceso se conectan directamente en el núcleo. El núcleo es un agrupamiento resiliente de conmutadores con un mínimo de 2 conmutadores y la suficiente cantidad de puertos disponibles para soportar la conexión dual a todos los conmutadores de borde.

Una troncal IST enlaza los 2 conmutadores virtualmente para simular un solo conmutador virtual. La troncal IST es un componente crítico del diseño y por eso debe ser altamente resiliente. La arquitectura de este agrupamiento y el flujo de tráfico a través del

núcleo es tal que no existe realmente un alto volumen de tráfico a través de la troncal IST, de tal forma que la resiliencia es más importante que el ancho de banda total.

La arquitectura es muy flexible y puede adecuarse a la mayoría de escenarios de diseño. La recomendación es habilitar funcionalidades de capa 2 en la capa de acceso y capa 3 en la capa núcleo. Esto no limita a que en algunos casos se pueda extender la capa 3 hasta el nivel de acceso en la red.

3.4.2.2 Diseño de 3 Capas

Con el diseño de 3 capas, la capa de distribución es insertada entre el acceso y el núcleo. La capa de distribución es necesaria si es que:

- Existe una distribución previa de las conexiones de fibra de planta externa que favorece u obliga a diseñar con esta topología.
- Existen múltiples edificios sobre un mismo campus que necesitan ser interconectados previamente, en cuyo caso todos estos forman la capa de distribución que se conecta hacia un núcleo centralizado.

Las consideraciones para la conexión entre la capa de acceso y la capa de distribución son las mismas indicadas en el diseño de 2 capas. Para la conexión entre la capa de distribución y el núcleo hay dos opciones:

- Utilizar SMLT (resiliencia en capa 2)

  En este diseño, la conectividad entre la distribución y el núcleo imita la del acceso al núcleo del diseño de 2 capas. La principal diferencia yace en la capacidad de interconectar con malla completa la capa de distribución con el núcleo.

  Una topología de malla completa provee el más alto nivel de resiliencia posible, mientras que se mantiene una recuperación frente a fallas en menos de 1 segundo. Una topología cuadrado o malla completa es mandatorio para mantener la resiliencia y el ancho de banda a plena capacidad entre la distribución y el núcleo.
- Utilizar RSMLT (resiliencia en capa 3)

  Si es requerido habilitar enrutamiento entre las capas de distribución y núcleo, la solución es habilitar RSMLT (SMLT ruteado) para seguir manteniendo una recuperación frente a fallas menor a 1 segundo mientras que a su vez, se mantiene un protocolo de enrutamiento IGP (como RIP ú OSPF) en el núcleo.

  RSMLT es una extensión de la tecnología de virtualización SMLT para proveer un concepto de enrutador activo-activo para redes SMLT que tienen el enrutamiento habilitado entre las VLANs del núcleo.

  En el caso de que falle un conmutador que hace enrutamiento, RSMLT previene que el reenvío de paquetes hacia la capa 2 siga funcionando mientras el protocolo de enrutamiento converge en el nivel de la capa 3. Esto permite que el reenvío de tráfico

no se detenga en el evento de que se produzca cualquier falla, evitando el corte del servicio para el usuario. RSMLT será explicado con más detalle en una sección más adelante.

3.5 VRRP

VRRP (Virtual Router Redundacy Protocol, en inglés) es un método estándar definido en RFC 2338, que se utiliza para mantener el enrutamiento de salida desde una red LAN cuando el nodo de pasarela por defecto (default gateway) tiene una falla.

Lo que hace el protocolo VRRP es detectar inmediatamente la falla y reasignar la función de enrutamiento IP hacia otro enrutador de respaldo en la misma LAN. VRRP opera transparentemente para el usuario final y no requiere de una configuración especial en cada estación.

3.5.1 Modos de Operación

VRRP está diseñado para eliminar el punto de falla único que se produce cuando se asigna estáticamente una dirección IP para el nodo de pasarela por defecto, en una estación de la red. VRRP utiliza el concepto de "dirección IP virtual", la cual es compartida lógicamente por dos o más enrutadores conectados en capa 2 a la misma subred. Si se elige la "dirección IP virtual" como la dirección de la pasarela por defecto en las estaciones terminales, VRRP ofrece un mecanismo de recuperación dinámico en el evento de que se produzca una falla en el equipo que controla esa dirección, según se puede apreciar en la figura 3.14 siguiente.

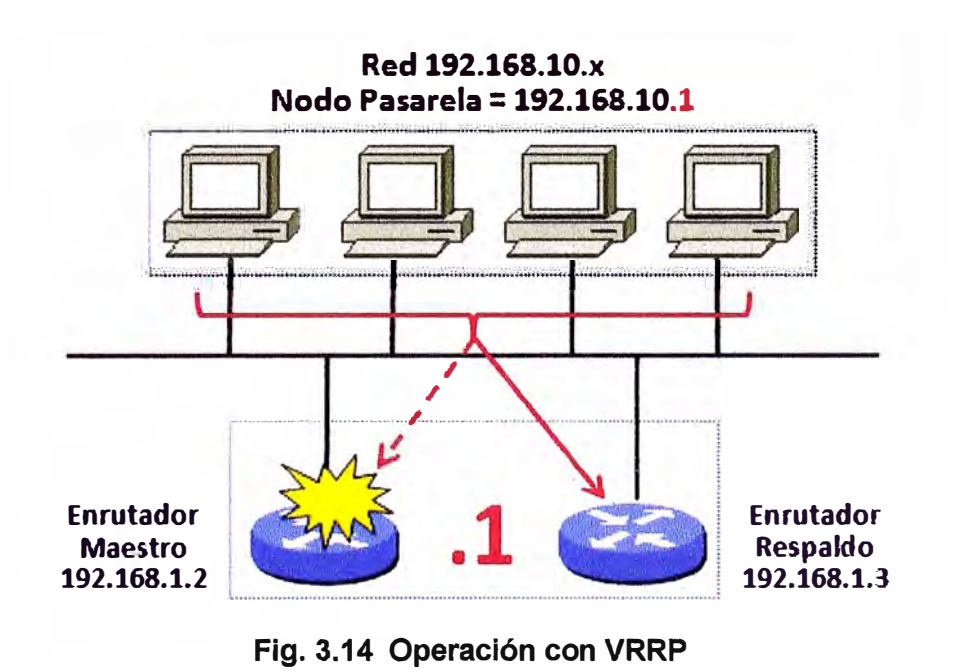


Fig. 3.14 Operación con VRRP

El enrutador VRRP que controla la dirección IP virtual es llamado enrutador maestro. Es el encargado de reenviar los paquetes que llegan a esta dirección virtual. Si

este enrutador maestro falla, existe un proceso de elección y recuperación que permite redirigir la responsabilidad del enrutamiento a otro enrutador VRRP en la misma subred.

En la figura 3.15 siguiente, las primeras 2 estaciones se configuran para tener una dirección de pasarela igual a la del enrutador VRRP 1, y las otras 2 estaciones hacen lo mismo para el enrutador VRRP 2. Esto no solo tiene el efecto de compartir la carga manualmente, sino que ofrece una redundancia completa. En el evento de que una de las interfaces de los enrutadores falle, el otro puede asumir la responsabilidad por ambas direcciones. Un enrutador actúa como respaldo del otro.

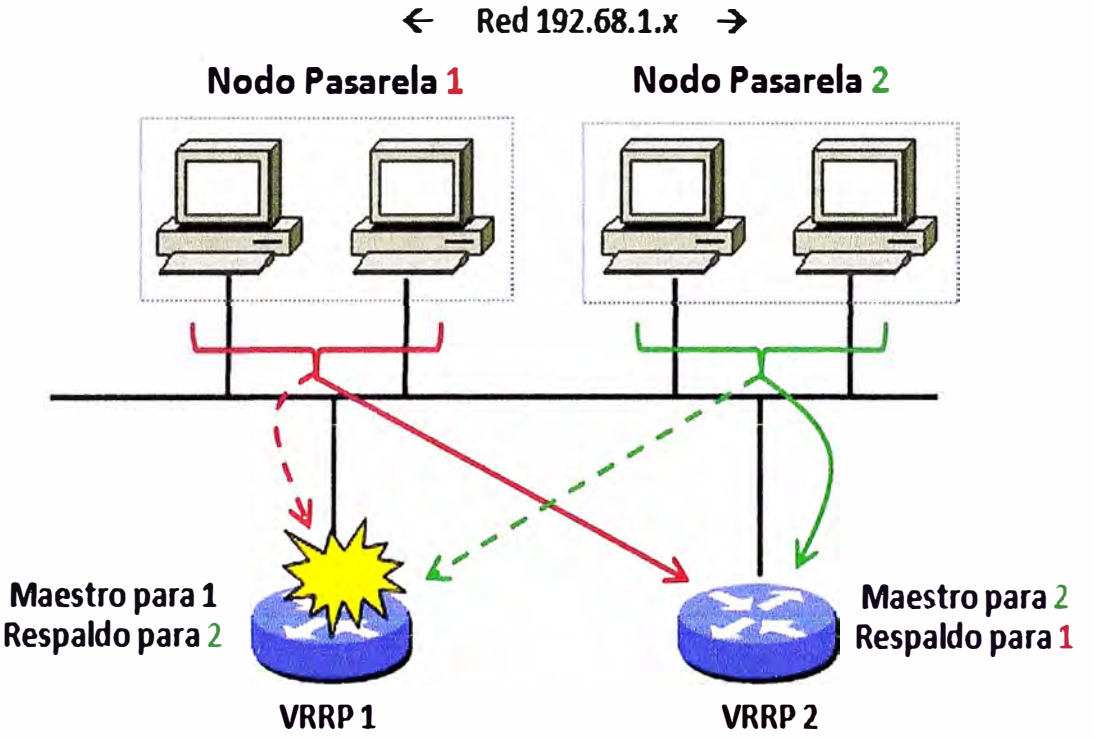


Fig. 3.15 Compartición de carga con VRRP

3.5.2 Nomenclatura

- Enrutador VRRP: Un dispositivo enrutador que corre el protocolo VRRP en conjunto con otros enrutadores.
- Enrutador Virtual: Un objeto ficticio que actúa como pasarela por defecto para las estaciones de una subred IP. Cara enrutador virtual se caracteriza por un número VRID (Identificador) y una "dirección IP virtual" que es ficticia. Este enrutador se comporta como un enrutador fantasma. Un enrutador VRRP puede ayudar a crear uno o más enrutadores virtuales.
- Dueño de la Dirección IP: Es posible que uno de los enrutadores VRRP tenga una dirección IP real que coincida con la dirección IP virtual. En este caso especial, el enrutador es dueño de la dirección IP virtual que ya no es ficticia.

- Dirección IP Primaria: Si el enrutador VRRP tiene varias direcciones IP, esta es la dirección IP real escogida que se utiliza como origen para todos los paquetes de publicación VRRP.
- Enrutador Virtual Maestro: Es el enrutador VRRP físico que asume la responsabilidad del reenvío de paquetes que se dirigen al enrutador virtual lógico. Este por ejemplo, responde pedidos ARP para esta dirección IP virtual.
- Enrutador Virtual de Respaldo: Es el enrutador VRRP físico que asume la responsabilidad del Enrutador Virtual Maestro en caso este falle. Puede ser más de uno, en caso se tenga una secuencia de respaldos en cadena.

3.5.3 **Parámetros**

- VRID: Es el identificador del enrutador virtual. Es un valor configurable entre 1 a 255.
- Prioridad: Es el valor utilizado por cada enrutador VRRP para la elección del enrutador maestro. El valor 255 es reservado para cuando el enrutador es dueño de la dirección IP virtual (siempre debe ser elegido). El valor cero es reservado para indicar que no se participa en la elección. El rango restante (1 – 254) está disponible para ser escogido por los enrutadores VRRP que respaldan al enrutador virtual. Cuanto más alto el valor, mayor preferencia en el respaldo. El valor por defecto es 100.
- Intervalo de publicación (Adv_Interval): Es el tiempo que existe entre cada publicación VRRP que envía el enrutador maestro. El valor por defecto es 1 segundo.
- Tiempo de sesgo (Skew_Time): Es un tiempo adicional menor a 1 segundo, que es utilizado en el cálculo de la no disponibilidad del enrutador maestro. Su valor corresponde a la siguiente fórmula: (256 – Prioridad) / 256.
- Intervalo de Maestro Caído (MD_Interval): Es el intervalo de tiempo que tiene que pasar para que el enrutador de respaldo declare al enrutador maestro como no operativo. Su valor corresponde a la siguiente fórmula: MD_Interval = (3*Adv_Interval) + Skew_Time)
- Modo de Preeminencia: Controla si es que un enrutador de respaldo que se conecta en la red y que tiene un valor de prioridad mayor al enrutador maestro existente, toma preeminencia sobre este. Los valores son “Verdad” y “Falso”. El valor por defecto es “Verdad”, es decir que cualquier interface VRRP con un valor mayor de prioridad se vuelve la interface maestra.
- Temporizador de Saludo Rápido (Fast Hello Timer): Es una funcionalidad propietaria de Nortel que permite ofrecer un rango de poleo VRRP mas corto, debajo del segundo, en el rango de 200 a 1,000 milisegundos.

- Temporizador de Espera (Holddown Timer): Es el tiempo que se retarda en activarse a preeminencia del enrutador maestro sobre el enrutador de respaldo, después de que el enrutador maestro se recupera de una caída y vuelve a ponerse disponible. El valor por defecto es cero (no se espera nada). A este temporizador, el fabricante Cisco le llama retardo de preeminencia (preempt delay).

3.5.4 Interface Crítica

En la figura 3.16 siguiente, se aprecia que las estaciones que utilizan el enrutador pasarela VRRP 2 para salir fuera de la subred, tienen que redirigir primero su tráfico hacia el enrutador pasarela VRRP 1, en caso de que la interface 3 este fuera de servicio. Para ahorrar a la red de este salto innecesario, con VRRP es posible configurar la interface 3 como una interface crítica.

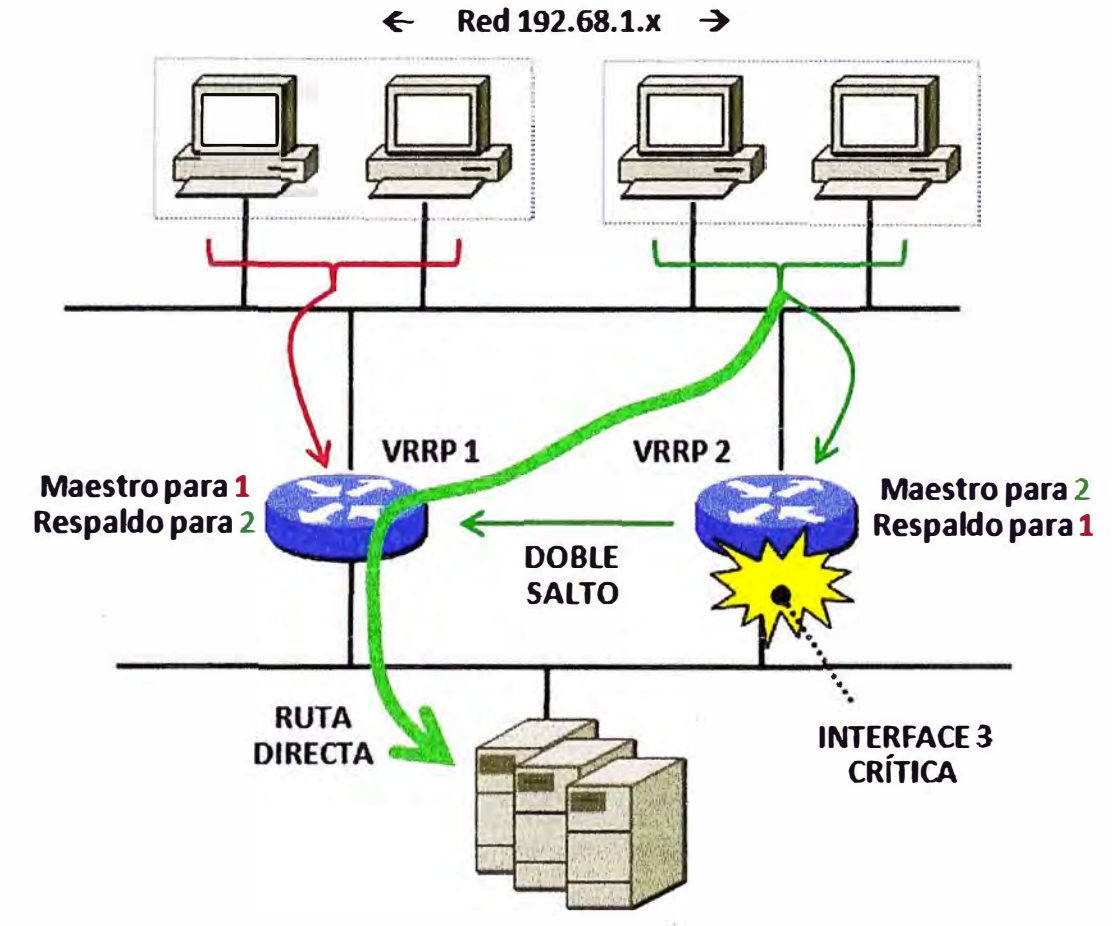


Fig. 3.16 Caso con Interface Crítica

En ese caso, si la interface 3 se vuelve inoperativa, el conmutador multicapa VRRP 2 (que es representado en la figura con funciones de enrutador) cambia al estado de respaldo y habilita al conmutador multicapa VRRP 1 para que se vuelva al estado maestro. Cuando la interface 3 se vuelve activa otra vez, el conmutador multicapa VRRP 2 reasume el estado maestro para su dirección VRRP.

3.5.5 Funcionalidad VRRP-BM

El fabricante Nortel (ahora Avaya) desarrollo una extensión propietaria al protocolo VRRP, denominada VRRP-BM (VRRP Backup Master), la cual es necesaria para coordinar adecuadamente el trabajo de este protocolo dentro del esquema de virtualización con SMLT que se reviso anteriormente.

Con el protocolo VRRP estándar, el conmutador multicapa maestro maneja todo el enrutamiento que va dirigido a la dirección IP VRRP (ya sea virtual o real). En la figura 3.17 siguiente, se ha configurado un enrutador virtual VRRP para soporte de resiliencia en capa 3 utilizando SMLT en forma simultánea. Nótese que el conmutador 1 superior es el enrutador maestro para la dirección IP virtual 10.10.10.1.

Fig. 3.17 Operación con VRRP-BM

Si el conmutador de acceso envía un paquete que necesita ser ruteado y a la vez utiliza SMLT para enviar el paquete al conmutador multicapa 2 inferior que tiene configurado el protocolo VRRP, este debe reenviar primero el paquete a través de la troncal IST, de tal forma que el conmutador maestro pueda reenviarlo adonde corresponda. Es decir, uno de los conmutadores no hace enrutamiento.

Con la funcionalidad VRRP-BM, el conmutador 2 inferior si puede reenviar el paquete recibido desde el conmutador de acceso, obviando la necesidad de enviarlo primero a través de la troncal IST. Es decir, ambos conmutadores realizan el enrutamiento en un modo activo-activo.

De esta forma, si la red utiliza SMLT y enrutamiento VRRP, la funcionalidad VRRP-BM permite que ambos conmutadores agrupadores puedan hacer enrutamiento. Esta funcionalidad puede ser extendida a más conmutadores que participan en un mismo grupo VRRP.

Obviamente, el tráfico IP es enviado a su destino si existe una ruta y una entrada en la tabla de enrutamiento del conmutador, de lo contrario el tráfico es descartado. También es posible utilizar en simultáneo, la funcionalidad de interface crítica explicada antes, para que si esa interface del conmutador 2 falla, el tráfico se derive directamente al conmutador maestro sin un salto a través de la troncal IST.

3.5.6 **Consideraciones de Diseño**

Para utilizar efectivamente VRRP dentro de una red de campus que utiliza la tecnología SMLT, se recomienda lo siguiente:

- No se debe escoger la dirección IP virtual (para una subred determinada) como una de las direcciones IP reales utilizadas por los conmutadores multicapa. Se debe utilizar una tercera dirección, como por ejemplo:
  - Dirección Conmutador 1 = x.x.x.2 (real)
  - Dirección Conmutador 2 = x.x.x.3 (real)
  - Dirección IP Virtual = x.x.x.1 (ficticia)
- Debe configurarse el temporizador Master_down (que es igual al intervalo Master_down) a un valor lo suficientemente alto para que el protocolo IGP que se corre en el núcleo tenga tiempo para reconverger y actualizar la tabla de ruteo.
  - Se recomienda un valor igual a 1.5 veces el tiempo de convergencia del protocolo IGP.
  - Para el caso de OSPF y asumiendo sus temporizadores por defecto, este valor del temporizador Master_down sería igual a 90 segundos.
- Para reducir el tiempo de convergencia de VRRP, el fabricante Nortel (ahora Avaya) habilita un temporizador rápido (Fast Hello Timer) que logra una recuperación en menos de 1 segundo. Sin esto, la recuperación sucede como mínimo en 3 segundos aproximadamente. Esta funcionalidad debe usarse con cuidado porque añade mayor tráfico de control y recarga el trabajo de la CPU del equipo.
- Implementar la funcionalidad VRRP-BM para una configuración del tipo activo-activo entre los conmutadores agrupadores.
- Configurar el conmutador VRRP maestro con un valor de prioridad de 200, para cada bloque básico en la red.

3.6 RSMLT

SMLT es una tecnología de virtualización que asegura que el tráfico de una VLAN dentro de un bloque básico este siempre disponible, mediante la provisión de conmutadores y enlaces redundantes. RSMLT (Routed Split Multilink Trunking, en inglés) es una extensión de SMLT que asegura que el tráfico de una VLAN pueda ser enrutado

siempre hacia fuera del bloque básico, mediante la redundancia en la función de enrutador que se realiza en los conmutadores de núcleo.

Esta extensión que se logra con RSMLT se podrá apreciar visiblemente en las próximas páginas y tendrá las siguientes características:

- Un par de enrutadores (conmutadores multicapa) de núcleo son configurados con RSMLT, como una funcionalidad adicional y encima de una configuración SMLT de núcleo (no del bloque básico).
- Cada enrutador en el par, responde activamente a la dirección MAC del otro enrutador para todo el tráfico que viene y termina sobre la troncal SMLT de núcleo que termina en estos dos enrutadores.
- RSMLT ofrece un mecanismo de recuperación de extremo a extremo para que el tráfico de origen de una VLAN de acceso (que es ruteado a través de una VLAN de núcleo que utilice SMLT en topologías triangulo, cuadrado o malla completa), pueda llegar hacia otra VLAN destino ubicada en otro bloque básico.

Debe observarse que RSMLT permite configurar un par de conmutadores de núcleo para que activamente, compartan el tráfico destinado a la dirección MAC del otro, esto es para el tráfico que tiene que ser ruteado.

Por teoría se sabe que si la trama recibida no tuviera como destino la MAC de uno de los conmutadores, entonces sencillamente dicha trama debe ser conmutada en capa 2 y no hay necesidad de hacer enrutamiento. En este sentido, RSMLT se parece a VRRP, pero como mostraremos más adelante, es más eficiente y opera mucho más rápidamente porque se basa en SMLT.

3.6.1 Conceptos Clave

RSMLT puede ser un poco confuso de entender al principio, por lo que es necesario revisar y entender primero algunos conceptos que sirvan para enfocar su utilidad en el diseño resilliente de redes de campus de gran envergadura:

- RSMLT debe habilitarse necesariamente sobre un par de conmutadores de núcleo (no uno, ni más de dos), sobre los cuales termina alguna troncal SMLT que transporta una o varias VLANs de núcleo.
- RSMLT permite que ambos conmutadores lean activamente los paquetes que vienen direccionados en las tramas con las direcciones MAC de cada uno y que vienen por una troncal SMLT. En contraste, los enrutadores que operan con VRRP solo atienden las tramas que vienen con la dirección MAC virtual (una sola).
- RSMLT asegura que el tráfico de origen desde una VLAN de acceso pueda salir de su bloque básico, cruzar una VLAN de núcleo y llegar a su destino en otra VLAN de

acceso ubicada en otro bloque básico, todo esto a pesar de que se produzca una falla en uno de los conmutadores de núcleo del bloque básico de destino.

- El proceso anterior si es posible de lograr utilizando algún protocolo de enrutamiento de capa 3 en el núcleo, pero no con la rapidez que se obtiene con SMLT para la convergencia en capa 2. Es decir, SMLT mejora dramáticamente el tiempo de convergencia a menos de 1 segundo dentro del bloque básico, pero a nivel del núcleo seguimos teniendo tiempos de convergencia de varias decenas de segundos, lo cual no es coherente.
- El fabricante Nortel (ahora Avaya) desarrollo RSMLT para proveer un mecanismo de recuperación rápido en menos de 1 segundo, para el tráfico que es ruteado a través del núcleo de la red, independientemente del tipo de protocolo de enrutamiento que se esté usando en esa capa. Eso significa que:
  - Es posible utilizar cualquier protocolo IGP en la VLAN de núcleo.
  - No es necesario realizar ningún ajuste especial en el protocolo IGP para acelerar la resiliencia, porque RSMLT se encuentra a cargo de ofrecer la resiliencia requerida.
  - No es necesario habilitar VRRP en el núcleo.
  - Es posible tener varias VLANs de núcleo con RSMLT habilitado en cada una de ellas.

3.6.2 **Requisitos**

- RSMLT solamente puede ser configurado en VLANs que atraviesen una troncal IST.
- RSMLT debe ser habilitado en cada VLAN donde se desee habilitarlo.
- RSMLT solo puede ser configurado en VLANs que tienen asignada una dirección IP, es decir VLANs que pueden ser ruteadas.
- RSMLT puede habilitarse solo sobre una configuración SMLT existente previamente.

3.6.3 **Forma de Operación**

- RSMLT opera independientemente por cada VLAN donde es configurado.
- RSMLT requiere que exista primero una VLAN subyacente (de núcleo normalmente) que utilice SMLT y un protocolo de enrutamiento, para que los conmutadores multicapa del núcleo compartan sus rutas entre ellos.
- Un par de conmutadores agrupadores en SMLT trabajan como un par de enrutadores redundantes en RSMLT. Esto puede ser configurado por cada subred IP.
- Ambos conmutadores RSMLT intercambian y mantienen las direcciones MAC de sus enrutadores internos y activamente se reenvían tráfico de control entre cada uno, en forma similar a VRRP.

- La idea es que si uno de los enrutadores falla, el otro se mantiene reenviando el tráfico del enrutador caído por un tiempo determinado.
- Existen dos temporizadores en RSMLT:
  - Temporizador de sostenimiento (Hold-up Timer, en inglés) que es el tiempo durante el cual un conmutador multicapa se mantiene reemplazando las tareas de reenvío de su conmutador pareja.
  - Temporizador de espera (Hold-down Timer, en inglés) que es el tiempo durante el cual un conmutador multicapa espera a que su conmutador pareja se recupere y pueda reasumir sus propias tareas de reenvío.

3.6.4 Escenarios de Falla

Revisaremos a continuación el escenario de una red de campus con 2 bloques básicos que utilizan un núcleo del tipo malla completa, para ver gráficamente como operan los conceptos de RSMLT explicados anteriormente.

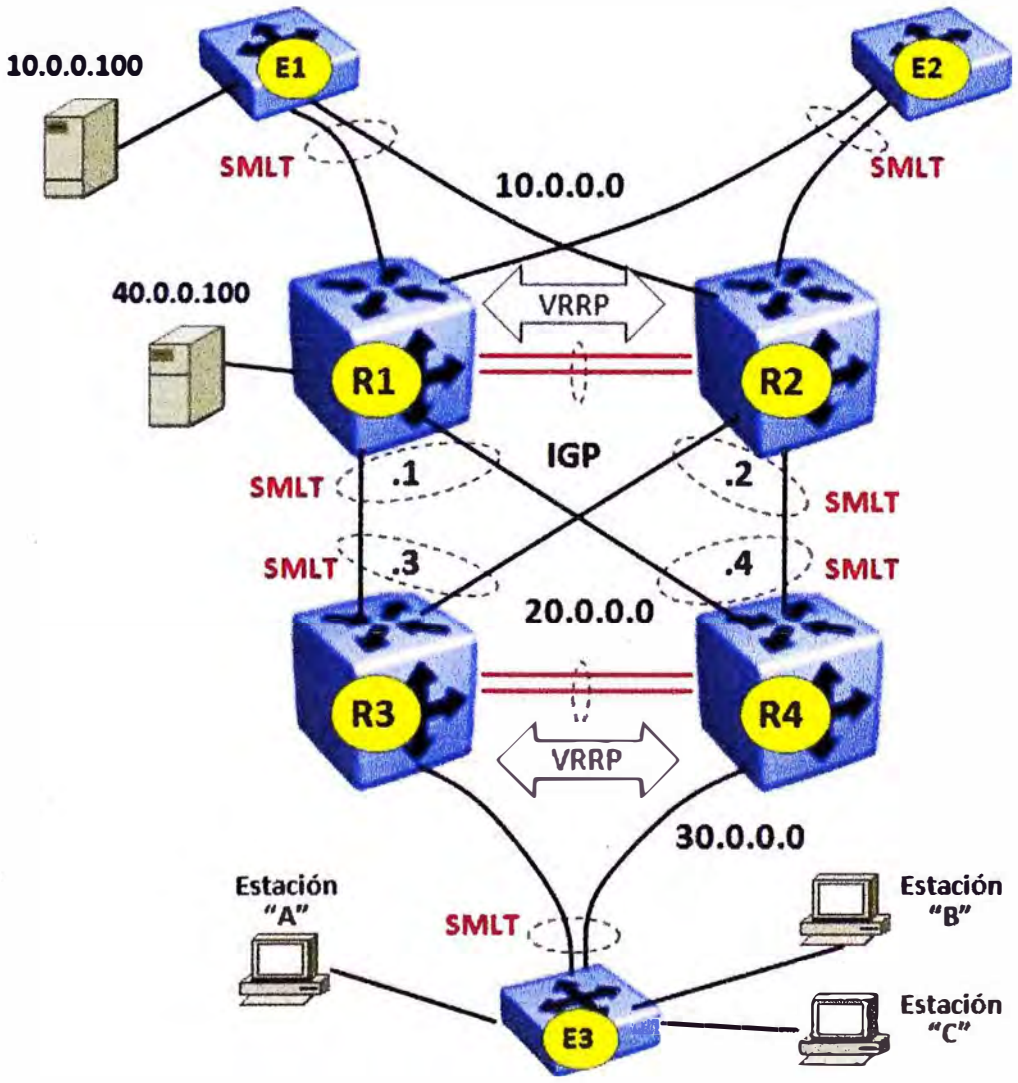


Fig. 3.18 Configuración Inicial con RSMLT

En la figura 3.18 anterior, se muestra que:

- Los conmutadores R1 y R2 comparten un SMLT hacia el conmutador E1 que se conecta al servidor 10.0.0.100. La conexión de este servidor está en la VLAN 10.0.0.0. Esto cumple con el primer requisito de que debe existir una VLAN SMLT previamente.
- La VLAN tiene una dirección IP (es decir es enrutable) con lo cual se cumple el segundo requisito.
- Por lo anterior, si es posible habilitar RSMLT en los conmutadores R1 y R2, específicamente para la VLAN 10.0.0.0. Si se hace esto, ambos conmutadores activamente harán enrutamiento y conmutación sobre sus propias direcciones MAC y sobre las del otro conmutador pareja, para todo el tráfico direccionado hacia la red 10.0.0.0.
- No es posible habilitar RSMLT en la VLAN 40.0.0.0, porque el servidor 40.0.0.100 solo tiene una conexión al conmutador R1. El conmutador R2 no puede enrutar el tráfico que recibe de los conmutadores R3 y R4, que tenga como destino el servidor 40.0.0.100. En su lugar, el conmutador R2 solo reenvía tal tráfico en capa 2 al conmutador R1 a través de la troncal IST, y de allí recién se llega al servidor.
- Si no se habilita RSMLT en la VLAN 20.0.0.0, el conmutador R2 no responderá al tráfico que va a la dirección 20.0.0.1 del conmutador R1, y viceversa. Si se desea habilitar RSMLT en un conjunto de VLANs, es necesario habilitarlo en cada una de ellas en forma individual.

En la misma figura 3.18 anterior, voy a mostrar cómo opera RSMLT a partir de una situación inicial:

- R1 y R2 soportan una troncal IST. Ambos están corriendo SMLT hacia el conmutador E1 y E2 en la red 10.0.0.0.
- R3 y R4 soportan una troncal IST. Ambos están corriendo SMLT hacia el conmutador E3 en la red 30.0.0.0.
- Los 4 conmutadores de núcleo están corriendo entre ellos una configuración SMLT de malla completa.
- R1 y R2 están corriendo RSMLT para la VLAN 10.0.0.0 de acceso. También están corriendo RSMLT para la VLAN 20.0.0.0 de núcleo.
- Se asume que en condiciones iníciales, R1 está a cargo de la red 10.0.0.0, es decir todo el tráfico que se dirige a esa red utiliza la dirección MAC de R1.
- R1 y R2 están corriendo VRRP para soportar la pasarela por defecto en 10.0.0.1 para las estaciones ubicadas en la red 10.0.0.0. Ellos no corren VRRP para soportar ninguna pasarela en la red 20.0.0.0 (no tendría sentido).

- Similarmente, R3 y R4 están corriendo VRRP para soportar la pasarela por defecto en 30.0.0.1 para las estaciones ubicadas en la red 30.0.0.0. Ellos no corren VRRP para soportar ninguna pasarela en la red 20.0.0.0 (no tendría sentido).

3.6.4.1 Estado Inicial

En la figura 3.19 siguiente, vamos a revisar como se distribuye el tráfico en esta configuración inicial:

- Las estaciones A, B y C están transmitiendo datos hacia el servidor en 10.0.0.100. Debido a los valores de sus direcciones IP, el tráfico desde la estación A va por la troncal SMLT hacia R3, mientras que el tráfico desde las estaciones B y C va por la troncal SMLT hacia R4.

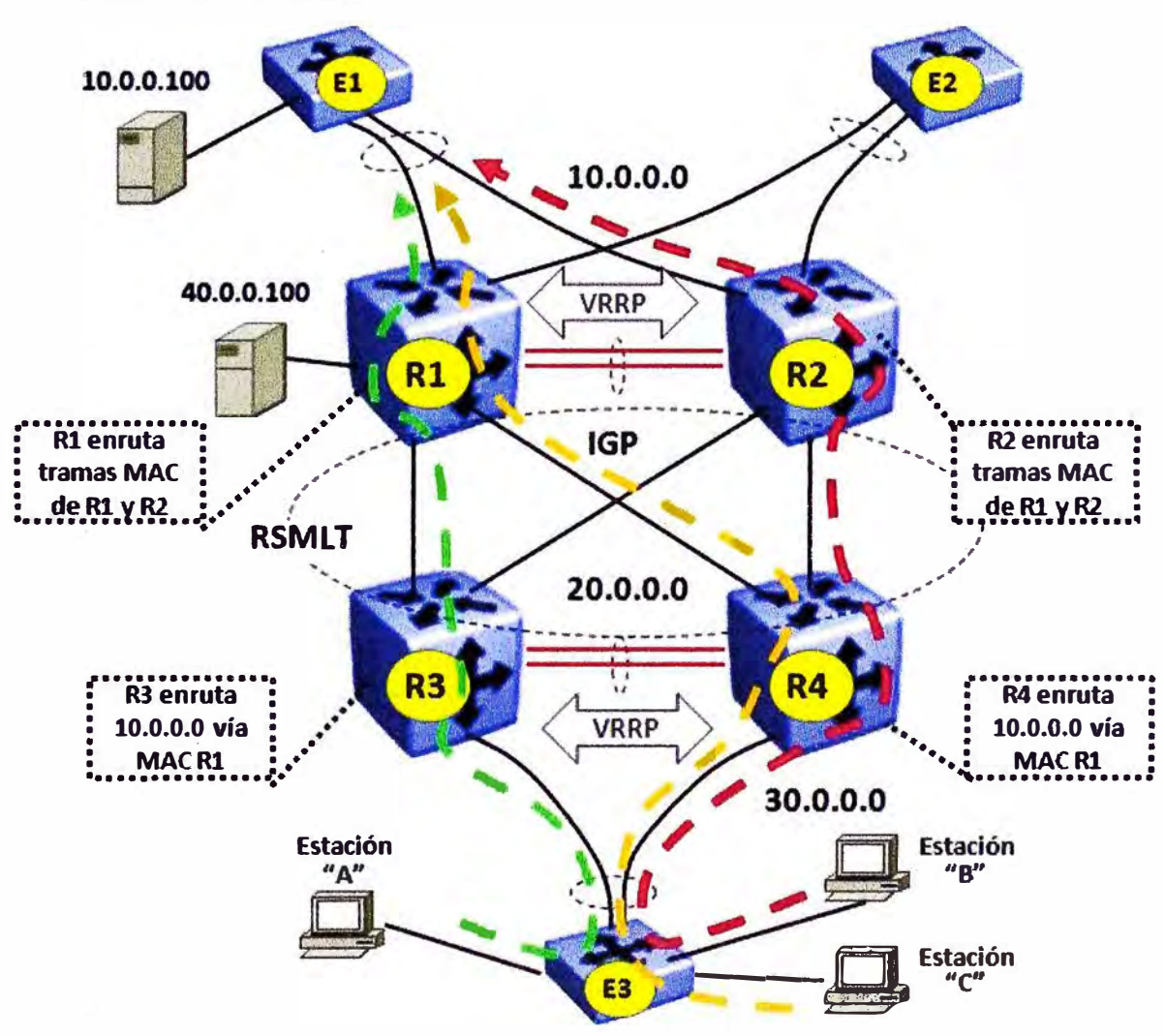


Fig. 3.19 Estado Inicial del Tráfico con RSMLT

- R3 y R4 tienen a R1 en sus tablas de enrutamiento como el siguiente salto para la red 10.0.0.0, así que ambos reenvían el tráfico hacia R1. Para R3, esto es por un enlace directo. Para R4 sin embargo, se vuelve un poco más complicado.

- R4 tiene una troncal SMLT hacia R1 y R2, así que desde su punto de vista, a él le parece que tiene una sola troncal con R1 con 2 enlaces físicos y por tanto, puede enviar el tráfico por cualquiera de estos.
    - R4 entonces recibe el tráfico desde la estación B y C, aplica el algoritmo de distribución de tráfico de SMLT y reenvía tráfico de la estación C hacía R1 por su enlace directo, pero el tráfico de la estación B por su enlace hacia R2.
    - En esta situación y si solo se tuviera SMLT, R2 reenviaría a continuación el tráfico recibido de la estación B por la troncal IST hacia el R1, y este finalmente reenvía todo el tráfico recibido hacia la red 10.0.0.0.
- Sin embargo, con RSMLT ambos conmutadores R1 y R2 están haciendo enrutamiento mutuo por cada uno de la red 10.0.0.0 sobre las troncales SMLT.
    - Esto significa que R4 envía tráfico a R2, con una dirección MAC destino de R1, y este activamente enruta ese tráfico (como si fuera suyo) sin enviarlo a través de la troncal IST.
- Obsérvese que R3 y R4 están enrutando el tráfico hacia el conmutador R1 físico, no hacia un enrutador virtual soportado por R1 y R2.
    - No hay tal concepto de enrutador virtual con RSMLT.

3.6.4.2 Falla en un Conmutador del Núcleo

Paso 1:

Voy a analizar que sucede si el conmutador R1 falla. En la figura 3.20 se muestran las etapas iníciales del proceso y el cambio de rutas:

- R3 y R4 se dan cuenta que sus enlaces físicos con R1 están inoperativos, y desactivan esos enlaces de su algoritmo de distribución.
- R3 y R4 cambian todo su tráfico hacia R2.
- Con RSMLT, R2 continúa respondiendo al tráfico de la dirección MAC de R1 y hace el enrutamiento que le toca en su reemplazo.
    - Para R3 y R4 es como si R1 estuviese todavía operando.

Paso 2:

- La red 20.0.0.0 en el núcleo comienza a reconverger. Con RIP esto puede tomar hasta 30 segundos. Con OSPF es mucho menos.
- El temporizador de sostenimiento en RSMLT se activa y R2 continua respondiendo al tráfico que llega para la dirección MAC de R1 hasta que la red 20.0.0.0 reconverga. Este valor de temporizador en R2 debe ser configurado de acuerdo al tipo de protocolo IGP existente en el núcleo.

- Hasta que la reconvergencia termine, tanto R3 como R4 piensan que aun se están comunicando con la red 10.0.0.0 a través de R1.
  - Desde su punto de vista, no ha existido ninguna interrupción en el acceso a esa red.

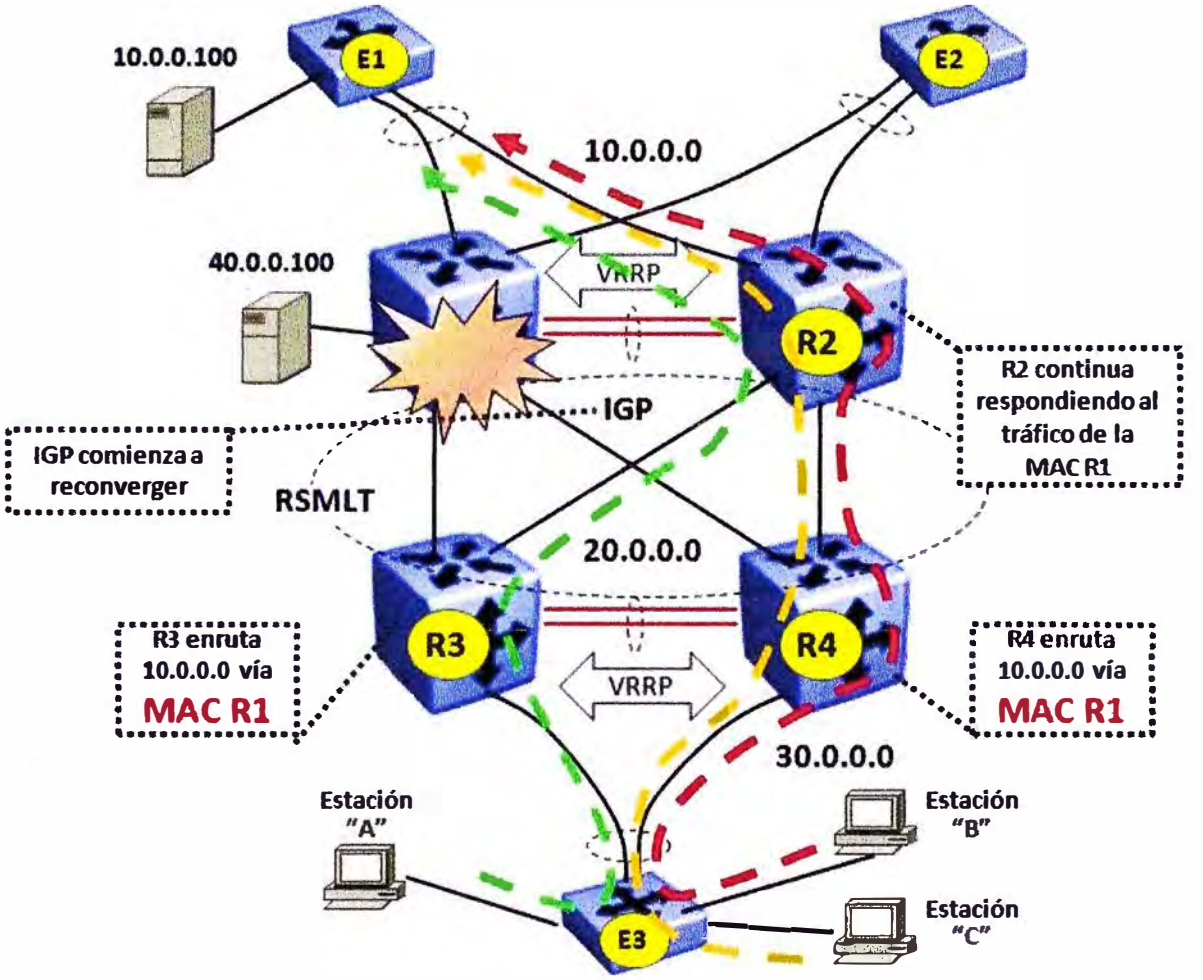


Fig. 3.20 Cambio del Tráfico con RSMLT

Paso 3:

- Después de que la red del núcleo reconverge, R2 queda como el dueño de la red 10.0.0.0. Ver figura 3.21.
- R3 y R4 tienen ahora a R2 como su siguiente salto en sus tablas de ruteo para la red 10.0.0.0, y por ende, envían el tráfico hacia esa red con la dirección MAC de R2.
- R2 ya no necesita responder al tráfico direccionado a la MAC de R1, y a continuación el temporizador de sostenimiento expira.

3.6.4.3 Recuperación del Conmutador de Núcleo

Paso 1:

- Cuando R1 se recupera, RSMLT permite que al inicio el equipo se comporte como un conmutador solamente. Ver figura 3.22.

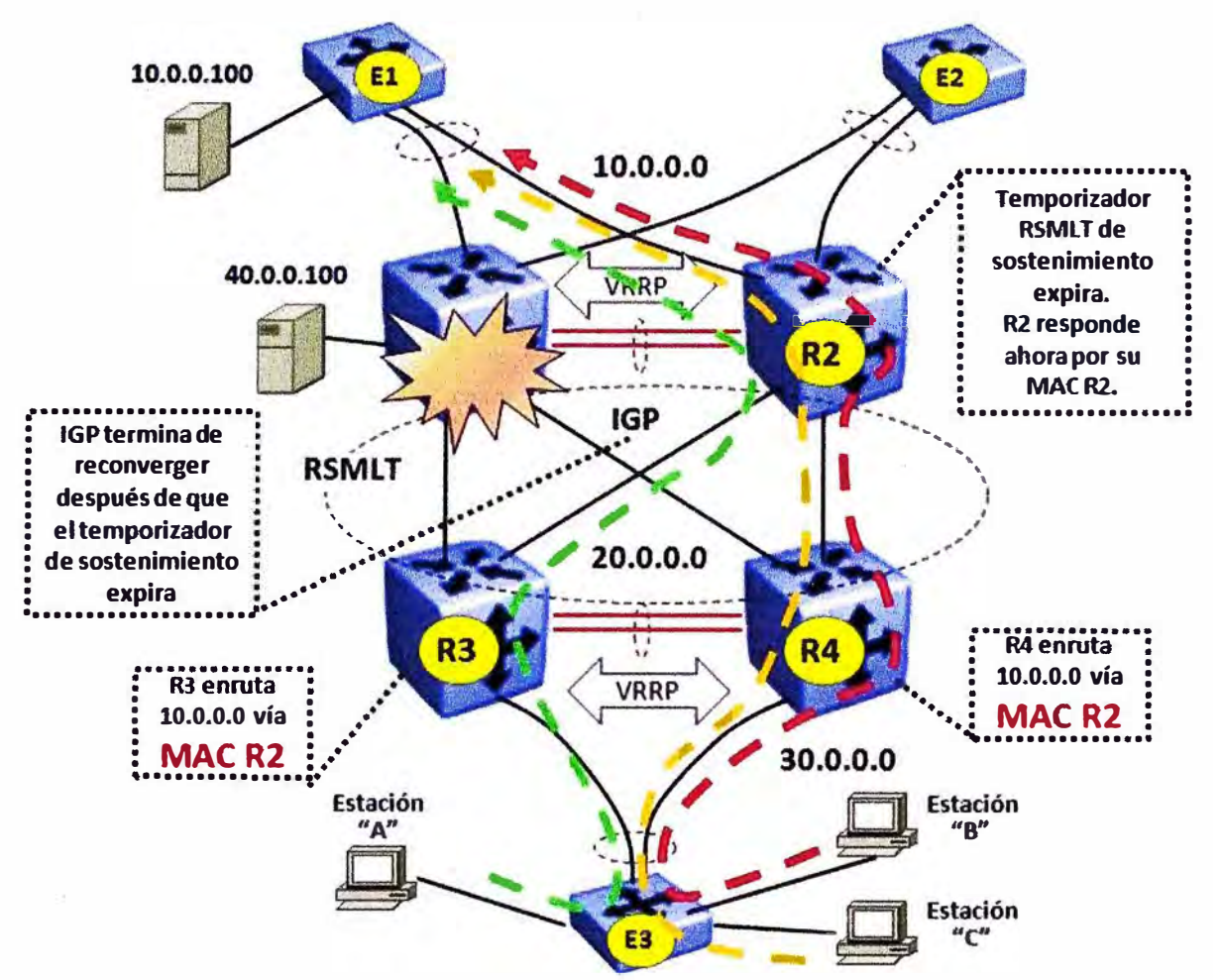


Fig. 3.21 Situación Final con RSMLT

- R3 y R4 reconocen que los enlaces físicos a R1 están disponibles de nuevo y los incluyen dentro de su algoritmo de distribución en SMLT.
- R3 comienza a reenviar tráfico de la estación A hacia R1.
- R4 comienza a reenviar tráfico de la estación B hacia R1.
- Hasta que el temporizador de espera de RSMLT en R1 no expire, R1 no puede enrutar su tráfico todavía. R1 solo conmuta en capa 2 el tráfico que recibe por la troncal IST hacia R2, el cual finalmente lo deriva a su destino.
- Aunque R1 está listo para enrutar tráfico de R2 nuevamente, obsérvese que en los primeros momentos, no existe tráfico que utilice la dirección MAC de R1 porque la red del núcleo aún no ha convergido.

Paso 2:

- Cuando la red del núcleo reconverge, R3 y R4 pueden o no pueden comenzar a enviar tráfico a R1 dependiendo de las prioridades que existan en el protocolo IGP.
- Si R3 y R4 comienzan a enviar tráfico a R1, este todavía no puede enrutar ese tráfico hasta que el temporizador de espera de RSMLT haya expirado, pero si puede

conmutar ese tráfico sobre la troncal IST con destino a R2, quien a su vez enruta todo el tráfico direccionado a R1.

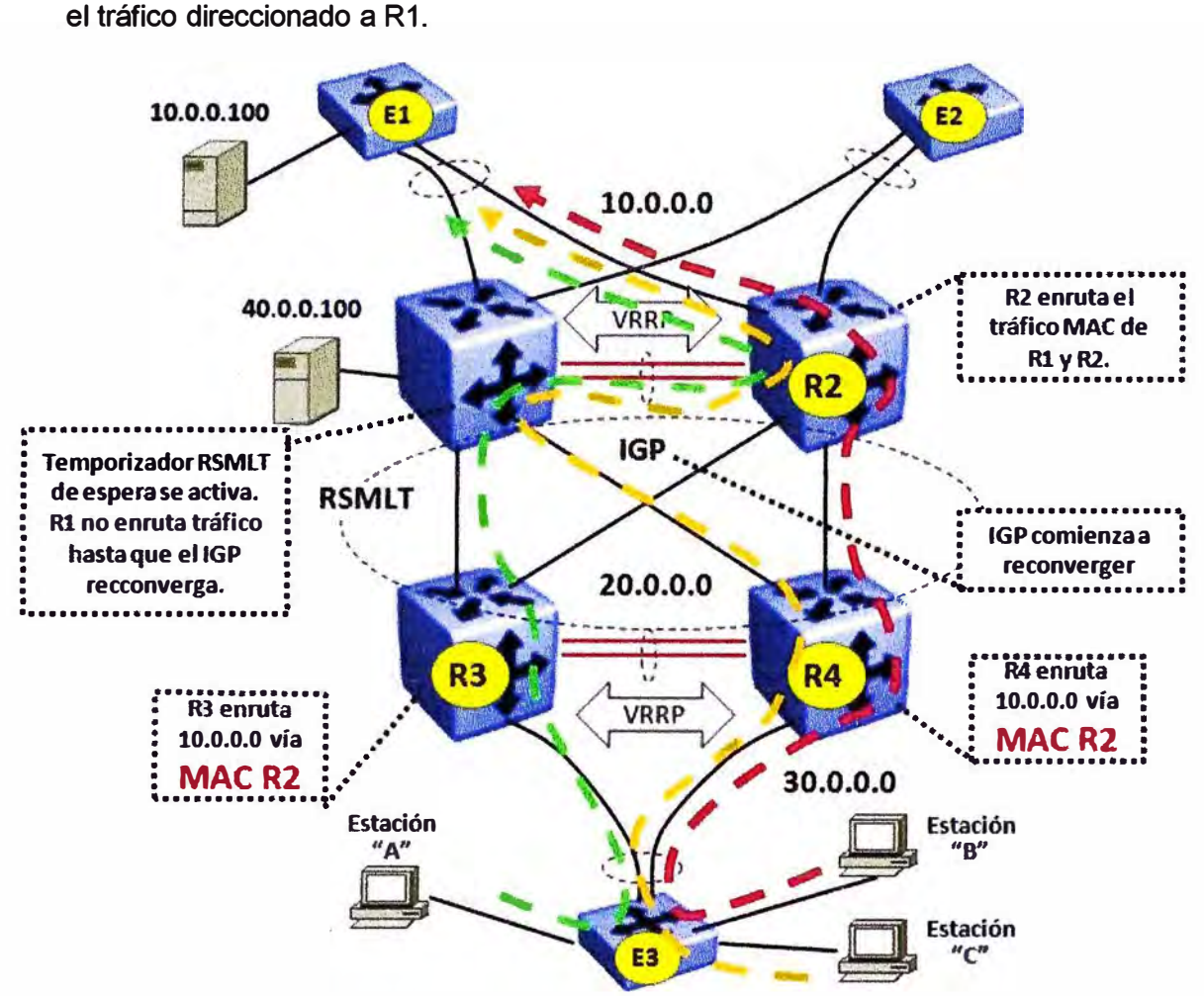


Fig. 3.22 Recuperación del Conmutador

Paso 3:

- Cuando la red del núcleo reconverge y el temporizador de espera de RSMLT expira en R1 y R1 ya puede comenzar a activamente enrutar el tráfico dirigido a su propia dirección MAC o a la del R2. Ver figura 3.23.
- En esta configuración, si se utiliza RIP en el núcleo, R3 y R4 continúan listando a R2 como la dirección del siguiente salto para la red 10.0.0.0.
- Si se corre OSPF en el núcleo, las tablas de ruteo de R3 y R4 si reflejan las prioridades que se hayan establecido previamente.

3.6.4.4 Comentarios Finales

- RSMLT es una característica soportada por el fabricante Nortel (ahora Avaya) solo en los conmutadores multicapa tipo chasis de gama alta. Esto enfoca su utilización en redes de muy alta cantidad de usuarios donde el costo y su característica de resiliencia pueden ser justificados.

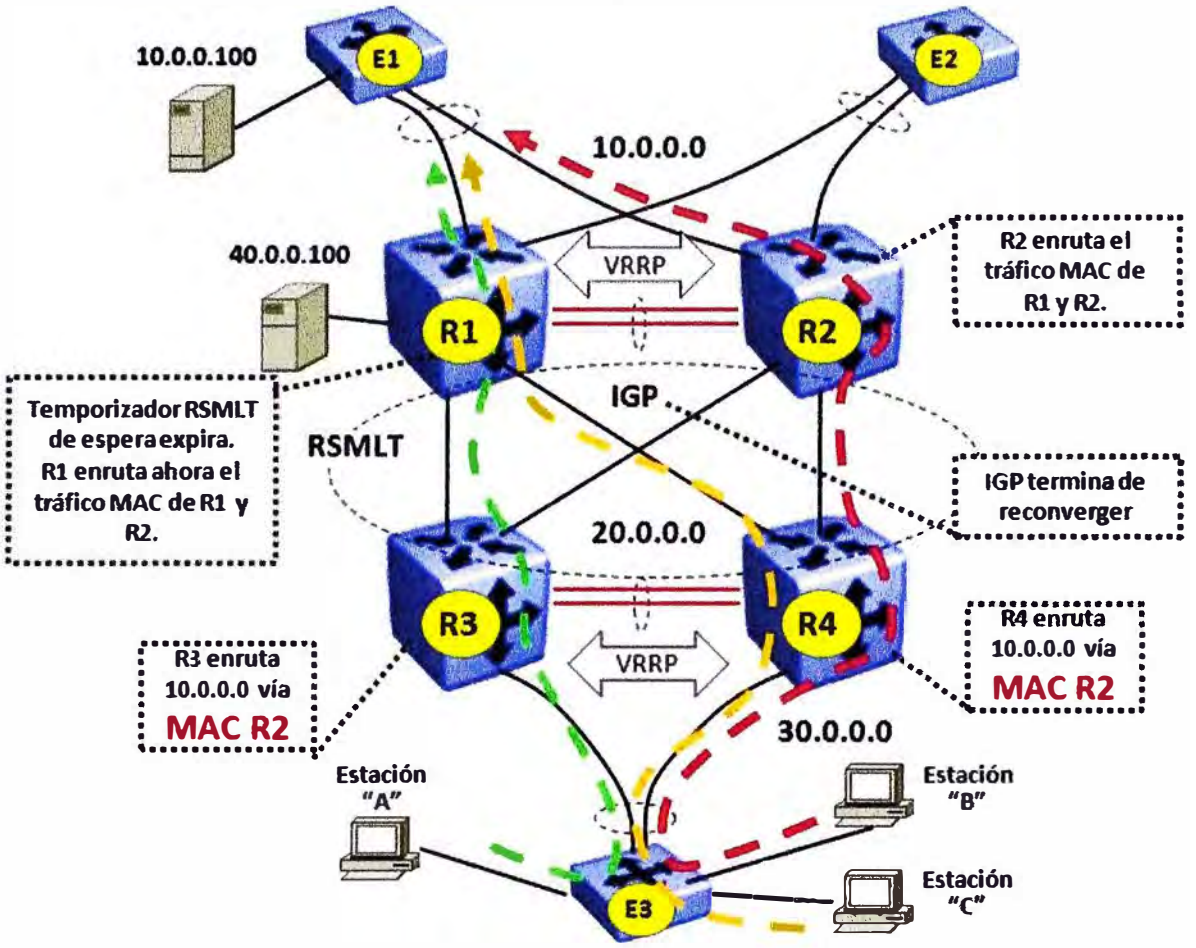


Fig. 3.23 Situación después de la Recuperación

- La rápida convergencia de RSMLT se logra mediante la provisión de una continuidad en el plano de transmisión de datos, manteniendo las direcciones MAC de ambos conmutadores virtualmente vivas, después de que uno de los dos equipos ha quedado inoperativo.
- Es posible utilizar RSMLT en lugar de VRRP para la parte de la capa de acceso y no solo en el núcleo. Las ventajas son la rápida convergencia y la menor carga de procesamiento que utiliza RSMLT. El punto en contra de esta elección viene por el lado de los costos asociados a los conmutadores de chasis que actualmente soportan RSMLT.
- Para la conectividad del núcleo que utiliza RSMLT:
  - Debe buscarse que las redes destino estén directamente accesibles desde los conmutadores que participan en RSMLT. Esto asegura un tiempo de recuperación menor a 1 segundo.
  - El tiempo de sostenimiento de RSMLT debe configurarse al menos a 1.5 veces el valor del tiempo de convergencia del protocolo IGP en el núcleo.

- Por ejemplo, si se usa los temporizadores por defecto de OSPF, el tiempo de sostenimiento sería de 60 segundos. Para RIP, el valor sería 180 segundos.

- Para la conectividad del acceso que utiliza RSMLT:
  - Puede usarse RSMLT o VRRP pero no ambos al mismo tiempo.
  - RSMLT se configura por cada VLAN, como en el núcleo.
  - El tiempo de sostenimiento en este caso debiera ser 9999 (interpretado como infinito) de tal forma que el conmutador que queda vivo pueda reenviar tráfico indefinidamente en reemplazo de su pareja.

CAPÍTULO IV
EJEMPLO DE DISEÑO DE UNA RED DE CAMPUS

La tecnología de virtualización SMLT ha estado disponible en el mercado desde el año 2001 aproximadamente y cuenta con miles de Clientes a nivel mundial que la han aprovechado en el diseño de sus infraestructuras de red. En el Perú, conozco personalmente a muy importantes empresas como Operadores de Telecomunicaciones, Bancos y Grupos Comerciales, que ya han implementado esta tecnología en el centro mismo de sus redes empresariales.

Para describir el procedimiento a seguir para el diseño de una red de campus que utiliza la tecnología SMLT, voy a tomar el caso de una Empresa local del sector comercial que como decenas de otras empresas nacionales, están preocupadas en modernizar su infraestructura para soportar su negocio con sus aplicaciones convergentes, y que requieren que dicha infraestructura sea muy confiable, simple de escalar y altamente resiliente, sin que esto signifique una muy alta inversión que haga inviable el proyecto, lo cual lamentablemente sucede muy a menudo en nuestro medio.

Estoy manteniendo el nombre de esta Empresa en reserva y modificando algunos datos particulares, lo cual es irrelevante para este informe porque mi intención es mostrar un procedimiento lo suficientemente general como para que pueda ser aplicado por algún Jefe o Gerente de Tecnología que lea este documento y se encuentre en una situación similar.

4.1 Situación Inicial

La Empresa de mi ejemplo (a la cual llamare XYZ) cuenta con una amplia red de tiendas a nivel nacional interconectadas por un servicio WAN basado en tecnología MPLS. Los sistemas de información se encuentran centralizados en su sede principal en Lima, desde donde se administra toda la infraestructura y servicios existentes.

La sede principal está constituida por un campus que agrupa a varios edificios que se encuentran interconectados entre sí por enlaces de fibra óptica propios que se han venido incorporando en la medida que el crecimiento del negocio lo ha requerido.

En cada uno de estos edificios existen diversos cuartos de cableado que concentran la conectividad de los usuarios finales así como de otros dispositivos, como: impresoras, teléfonos IP, cámaras IP y dispositivos de control de acceso.

La actual topología de la red de campus es del tipo backbone colapsado, donde cada cuarto de cableado cuenta con un solo enlace de fibra óptica que se concentra en el centro de datos. El centro de datos es un ambiente dedicado, donde se encuentran almacenados todos los servidores corporativos de la Empresa XYZ.

En ese ambiente existen varios gabinetes que soportan y protegen físicamente a los servidores y conmutadores centrales de la red. Los conmutadores centrales son del tipo chasis, basados en tecnología obsoleta y no cuentan con un servicio de mantenimiento correctivo en caso de fallas.

Debido al crecimiento del negocio y la falta de personal de soporte técnico especializado, la red de campus carece del ordenamiento físico y lógico que permita su administración adecuada.

Esto se ha hecho evidente para la alta Gerencia debido a varias fallas sucesivas en la red por razones técnicas (cortes o fallas en los enlaces de fibra por ampliaciones de obra civil) y errores humanos (reconexión inadecuada de equipos debido a traslados o movimientos internos), lo cual ha impactado en la productividad e imagen de la Empresa XYZ frente a sus clientes finales, proveedores y socios del negocio.

Por lo sucedido, la Empresa XYZ decidió iniciar un proceso de migración de su red porque entendió la importancia de contar con una infraestructura de red de campus altamente confiable y resiliente, que brinde un soporte sólido a sus procesos de negocio basados cada día más en aplicaciones convergentes y centralizadas, las cuales deben estar disponibles en todo momento.

A continuación, voy a presentar el procedimiento para el diseño de la nueva red de campus basado en datos específicos y en la utilización de la tecnología SMLT.

4.2 Requerimientos Específicos

En la tabla 4.1 siguiente se presenta una distribución de la cantidad de usuarios en la red por cada área de la Empresa XYZ, así como el tipo de puertos requeridos en cada caso.

Se puede observar que hay tres tipos de puertos requeridos:

- Puertos Gigabit Ethernet (10/100/1000 Mbps) de cobre.
- Puertos Fast Ethernet (10/100 Mbps) de cobre
- Puertos Fast Ethernet con la facilidad PoE (Power over Ethernet, en inglés) para alimentar directamente dispositivos terminales como:
  - Teléfonos IP
  - Cámaras IP
  - Puntos de acceso inalámbricos.

TABLA N° 4.1 Requerimiento de Puertos por Área

	Puertos de Red requeridos			
Area de la EMPRESA	**10/100/ 1000**	**10/100**	**10/100 PoE**	**Total**
Centro de Datos	45			45
Administración	48	96	96	240
Operaciones	48	24	48	120
Distribución	42		40	82
Logística		48	48	96
Ventas	44		48	92
Laboratorio	22	22	48	92
Ingeniería	46		24	70
Finanzas	47		24	71
Recursos Humanos		70	22	92
	342	260	398	**1000**

En el Centro de Datos se agrupan un total de 45 servidores con interfaces de red 10/100/1000 Mbps duales, lo cual permite realizar un agrupamiento de enlaces utilizando IEEE 802.3ad. La distancia entre los cuartos de cableado y el centro de datos no supera los 250 metros en ningún caso.

Se requiere que el diseño sea simple de entender y escalable en la medida que la Empresa XYZ vaya incorporando nuevos cambios por la dinámica de su negocio. El núcleo de la red debe ser totalmente redundante y debe permitir balancear

dinámicamente la carga de tráfico proveniente de los cuartos de cableado a través de enlaces redundantes que se concentren en el núcleo.

El costo en este diseño era el segundo criterio más importante. La Empresa XYZ no deseaba invertir en tecnología altamente costosa y sobredimensionada para sus necesidades, sino que deseaba crear una arquitectura de red que le permitiera ir creciendo y amortizando la inversión a medida que lo fuera necesitando.

4.3 Descripción de la Solución

Para cumplir con los requerimientos anteriores, la solución elegida fue utilizar una arquitectura de red de campus que utilizase la tecnología SMLT por las razones siguientes:

- Ofrece una alta disponibilidad y resiliencia con tiempos de recuperación en caso de fallas en el núcleo menores a 1 segundo.
- Descarta la utilización de tecnología compleja como RSTP y aprovecha el concepto simple de agrupamiento virtual de conmutadores que se logra con SMLT.
- Explota en forma simultánea los enlaces de subida que se conectan desde los cuartos de cableado brindando alta disponibilidad y balanceo de carga dinámico sin realizar configuraciones complejas.

La topología de red que se planteo para este caso fue de 2 capas en base a un solo bloque básico de conmutación en capa 2, dado que se contaba con dispositivos conmutadores de mediana escala que tenían la capacidad de manejar los 1,000 puertos de red requeridos en el proyecto. Los dispositivos conmutadores de núcleo elegidos para realizar la función de conmutadores agrupadores SMLT fueron del tipo multicapa, ya que era necesario que realicen enrutamiento de capa 3 y sirvan como frontera al resto de la red WAN (esta parte de la red no se muestra en el diseño).

4.4 Elección de los Componentes

Los dispositivos elegidos para el diseño fueron los conmutadores de las familias Avaya ERS 4500 (antes Nortel ERS 4500) y Avaya ERS 5600 (antes Nortel ERS 5600). A continuación presentare una breve descripción de las características de cada familia y sus principales funcionalidades. Las hojas técnicas detalladas de estos productos pueden ser revisadas en el Anexo "A" al final de este informe.

4.4.1 Conmutadores ERS 4500

La familia Avaya ERS 4500 es un conjunto de conmutadores apilables de capa 2 que ofrecen alto rendimiento y capacidad para manejar aplicaciones convergentes en la capa de acceso. A continuación sus principales características:

- Dos puertos traseros dedicados para apilamiento en cadena.

- Soporte de puertos Gigabit Ethernet en fibra mediante interfaces ópticas tipo SFP (Small Form-Factor Pluggable Transceiver, en inglés) para 2 y 4 puertos.
- Soporte de puertos 10 Gigabit Ethernet en fibra mediante interfaces ópticas tipo XFP (un tipo similar al SFP) hasta 2 puertos (solo en 2 modelos).
- Capacidad de apilamiento de hasta 8 conmutadores o 400 puertos con conmutadores del mismo modelo.
- Procesamiento de paquetes desde 6.6 hasta 138 MPPS (millones de paquetes por segundo).
- Capacidad de conmutación por cada equipo: Desde 48 hasta 184 Gbps.
- Capacidad de conmutación en el apilamiento: 320 Gbps.

En la actualidad existen 10 modelos de conmutadores disponibles en esta familia que se diferencian por la cantidad y el tipo de puertos soportados, según lo que se muestra en la tabla 4.2 siguiente:

TABLA N° 4.2 Modelos de Equipos ERS 4500

Familia ERS 4500	Características
ERS 4526FX	24 puertos 100BaseFX + 2 puertos 10/100/1000/SFP combo
ERS 4526T	24 puertos 10/100 + 2 puertos 10/100/1000/SFP combo
ERS 4526T-PWR	24 puertos 10/100 tipo PoE + 2 puertos 10/100/1000/SFP combo
ERS 4550T	48 puertos 10/100 + 2 puertos 10/100/1000/SFP combo
ERS 4550T-PWR	48 puertos 10/100 tipo PoE + 2 puertos 10/100/1000/SFP combo
ERS 4524GT	24 puertos 10/100/1000 con 4 puertos SFP compartidos
ERS 4524GT-PWR	24 puertos 10/100/1000 tipo PoE con 4 puertos SFP compartidos
ERS 4548GT	48 puertos 10/100/1000 con 4 puertos SFP compartidos
ERS 4548GT-PWR	48 puertos 10/100/1000 tipo PoE con 4 puertos SFP compartidos
ERS 4526GTX	24 puertos 10/100/1000 con 4 puertos SFP compartidos + 2 puertos XFP adicionales
ERS 4526GTX-PWR	24 puertos 10/100/1000 tipo PoE con 4 puertos SFP compartidos + 2 puertos XFP adicionales

Los puertos 10/100/1000/SFP combo son puertos adicionales que tienen dos salidas posibles, una 10/100/1000 Mbps para cable de cobre y otra Gigabit Ethernet para fibra óptica (utilizando una interface SFP). La primera salida es la salida por defecto, y la segunda se activa cuando se conecta la interface SFP correspondiente.

Los puertos SFP compartidos son puertos alternativos que comparten su salida con los últimos puertos principales 10/100/1000 que vienen en el equipo. Estos se activan solo cuando se conecta la interface SFP en la posición correspondiente.

4.4.2 Conmutadores ERS 5600

La familia Avaya ERS 5600 es un conjunto de conmutadores multicapa apilables que ofrecen muy alto rendimiento y capacidad para manejar aplicaciones convergentes en la capa núcleo. A continuación sus principales características:

- Dos puertos traseros dedicados para apilamiento en cadena.
- Soporte de puertos Gigabit Ethernet en fibra mediante interfaces ópticas tipo SFP (Small Form-Factor Pluggable Transceiver, en inglés) desde 6 hasta 24 puertos.
- Soporte de puertos 10 Gigabit Ethernet en fibra mediante interfaces ópticas tipo XFP (un tipo similar al SFP) hasta 8 puertos (solo en 1 modelo).
- Capacidad de apilamiento de hasta 8 conmutadores o 400 puertos con conmutadores del mismo modelo.
- Procesamiento de paquetes desde 100 hasta 173 MPPS (millones de paquetes por segundo).
- Capacidad de conmutación por cada equipo: 384 Gbps.
- Capacidad de conmutación en el apilamiento: 640 Gbps.

En la actualidad existen 5 modelos de conmutadores disponibles en esta familia que se diferencian por la cantidad y el tipo de puertos soportados, según lo que se muestra en la tabla 4.3 siguiente:

TABLA N° 4.3 Modelos de Equipos ERS 5600

Familia ERS 5600	Características
ERS 5632FD	24 puertos 10/100/1000/SFP compartidos + 8 puertos XFP adicionales
ERS 5650TD	48 puertos 10/100/1000 + 2 puertos XFP adicionales
ERS 5650TD-PWR	49 puertos 10/100/1000 tipo PoE + 2 puertos XFP adicionales
ERS 5698TFD	96 puertos 10/100/1000 con 6 puertos SFP compartidos + 2 puertos XFP adicionales
ERS 5698TFD-PWR	96 puertos 10/100/1000 tipo PoE con 6 puertos SFP compartidos + 2 puertos XFP adicionales

4.5 Diseño de la Capa de Acceso

- Apilamiento de conmutadores
  - Los conmutadores ERS tienen la capacidad de soportar apilamiento nativo en cadena lo que permite que se comporten como una sola unidad lógica.
  - Aprovechando esta característica, se configurará una troncal DMLT (Distributed MLT) con 2 o 4 enlaces desde cada cuarto de cableado, para prevenir la situación de que falle uno de los conmutadores apilables que tiene un enlace hacia el núcleo.

- Definición de VLANs
  - Se definirán 3 VLANs dedicadas para datos, voz y vídeo respectivamente, distintas a la VLAN 1 que viene por defecto. Esto es para evitar que un nuevo dispositivo que viene con VLAN 1 pueda generar inadvertidamente un lazo en la red. Todos los puertos se dejan en auto-negociación.
  - Los puertos de enlace al núcleo se configuran como troncal de VLANs para marcar las tramas según 802.1Q. Esta troncal debe transportar todas las VLANs excepto la VLAN 1.
  - La VLAN de administración (aquella por donde se ingresa a la interface de administración del equipo) será la misma VLAN de datos. Deberá asignarse una dirección IP de administración al conmutador.

- Agrupamiento de Enlaces

  Cada cuarto de cableado conteniendo conmutadores apilables debe tener una troncal MLT hacia el núcleo de la red. Las reglas para crear los grupos MLT desde los conmutadores ERS 4500 serán:
  - Crear un grupo MLT con el identificador 1 y nombrarlo "Core".
  - El identificador en cada apilamiento de conmutadores puede ser del mismo valor. Este valor solo tiene significación local y no se propaga por la red.

- Tramas no marcadas
  - Para prevenir que se genere un lazo cuando se conecta un conmutador ERS 4500 que viene con parámetros por defecto o que está mal configurado, hacia los enlaces de subida al núcleo, se utilizará la funcionalidad "Discard Untaggged Frames" (descartar tramas no marcadas) en los puertos que pertenecen al grupo MLT.
  - En otros modelos de mayor escala de la familia ERS existen mecanismos más elaborados que no se revisan en este informe.

- Uso del protocolo STP
  - Para prevenir que se genere un lazo a través de los puertos que van conectados directamente a los puertos de usuarios finales, se utilizará la funcionalidad "FastStart" del protocolo STP.
  - Estos serán los únicos puertos donde será habilitado STP.

4.6 Diseño de la Capa Núcleo

- Licencia de Software Avanzado

  Los conmutadores ERS 5600 requieren una licencia de software avanzado para habilitar las siguientes funcionalidades:

- OSPF (Open Shortest Path First)
- SMLT
- VRRP

Una licencia avanzada es requerida para cada conmutador de núcleo o apilamiento de conmutadores de núcleo. Para el diseño propuesto, se requieren dos licencias de software avanzado a fin de utilizar SMLT, VRRP y opcionalmente RIP o OSPF en la integración con la red WAN.

- Agrupamiento Virtual de Conmutadores

  La base del diseño se fundamenta en la tecnología de virtualización SMLT. Esta tecnología ofrece un diseño activo-activo en el núcleo que es completamente resiliente y que tiene un tiempo de recuperación menor a 1 segundo. Para configurar SMLT en los conmutadores ERS 5600, se seguirá el siguiente procedimiento:

  - Crear una VLAN separada para la troncal IST y no habilitar ningún protocolo de capa 3 en esta VLAN.
  - Utilizar una subred privada reducida (mascara de 30 bits) para la VLAN IST. Esta VLAN no deberá ser publicada fuera del núcleo y se utilizará solo para propósito de la comunicación IP entre ambos conmutadores del núcleo.
  - Verificar que todas las VLANs que participan en SMLT estén definidas en ambos conmutadores y que estén marcadas con 802.1Q en ambos extremos de la troncal IST.
  - Todos los enlaces de subida con SMLT deben estar marcados con 802.1Q para facilitar que se añada nuevas VLANs en el acceso sin impactar a los demás usuarios.
  - El protocolo STP debe ser manualmente deshabilitado en todos los puertos que participan en SMLT en los ERS5600, incluyendo la troncal IST. Debe hacerse lo mismo en el extremo del conmutador de acceso.
  - Los puertos asignados a un grupo MLT son indexados comenzando desde el número cero. La posición más baja de un puerto (conmutador 1, puerto 1) para un enlace MLT tiene un índice de cero. El segundo enlace MLT en la segunda posición más baja tiene un índice de uno, y así sucesivamente. El índice es usado por el algoritmo MLT para asignar un flujo sobre uno de los enlaces de un grupo MLT.
  - Lo recomendable es que la posición más baja de un enlace en un conmutador se empate con la posición más baja del otro conmutador. Esto ayuda a que el algoritmo MLT siempre resuelva un flujo de ida y vuelta sobre el mismo enlace entre ambos conmutadores.

- Tramas no marcadas

  Se debe habilitar la funcionalidad de "Discard Untagged Frames" en los puertos SMLT e IST, para prevenir que se genere un lazo cuando se conecta un conmutador ERS 5600 que viene con parámetros por defecto o que está mal configurado.

- VLACP

  La funcionalidad VLACP que se reviso en el capítulo anterior, ofrece un mecanismo de detección de fallas de extremo a extremo, que ayuda a prevenir problemas potenciales causados por malas configuraciones o cortes en el medio de fibra entre los conmutadores ERS5600 del núcleo.

  Debe usarse VLACP en todos los enlaces IST utilizando una dirección MAC reservada y multidireccionada (01-80-C2-00-00-0F) para prevenir que algún enlace se conecte inadvertidamente a un puerto incorrecto, que uno de los conmutadores se reinicie con valores por defecto o que se haga una configuración incorrecta. VLACP se configurará con un valor de temporizador alto.

- Uso del Protocolo STP

  STP debe ser deshabilitado en todos los puertos SMLT/IST. Se recomienda que todos los otros puertos queden habilitados con STP para prevenir la creación de un lazo por la conexión de un cable o dispositivo externo.

- VLANs

  Todas las VLANs configuradas en los conmutadores de borde deben terminar en el núcleo. Esto significa que hay que configurar las VLANs de voz, datos y vídeo en ambos conmutadores ERS 5600 del núcleo.

  Otros requisitos de diseño para el manejo de VLANs son los siguientes:

  - Cualquier otro servicio que sea ligado al núcleo, necesitará una VLAN adicional. En nuestro caso, todos los servidores serán colocados en una VLAN que solamente existirá en los conmutadores ERS 5600.
  - Por defecto, la VLAN 1 existe y todos los puertos pertenecen a ella. Esta VLAN no puede ser eliminada. La recomendación es que se use la VLAN 1 solo como repositorio de los puertos sin uso, y no para tráfico de datos.
  - Los valores de los temporizadores de la tabla MAC se dejan en sus valores por defecto (300 segundos).

- Enrutamiento

  Los conmutadores ERS 5600 automáticamente rutean entre VLANs, si:

  - La funcionalidad "ip forwarding" se encuentra habilitada. Esto es necesario en una configuración SMLT.
  - Las VLANs están configuradas con una dirección IP y una máscara de red.

La necesidad por agregar un protocolo de enrutamiento de capa 3, dependerá del ambiente que rodea a la red de campus. El enrutamiento entre el núcleo de conmutadores y un enrutador de WAN puede ser configurado con rutas estáticas, RIP u OSPF, como se aprecia en la figura 4.1 siguiente.

- La decisión por una de ellas dependerá de la estrategia de enrutamiento en la red WAN.
- Esta parte no es revisada en el presente informe.

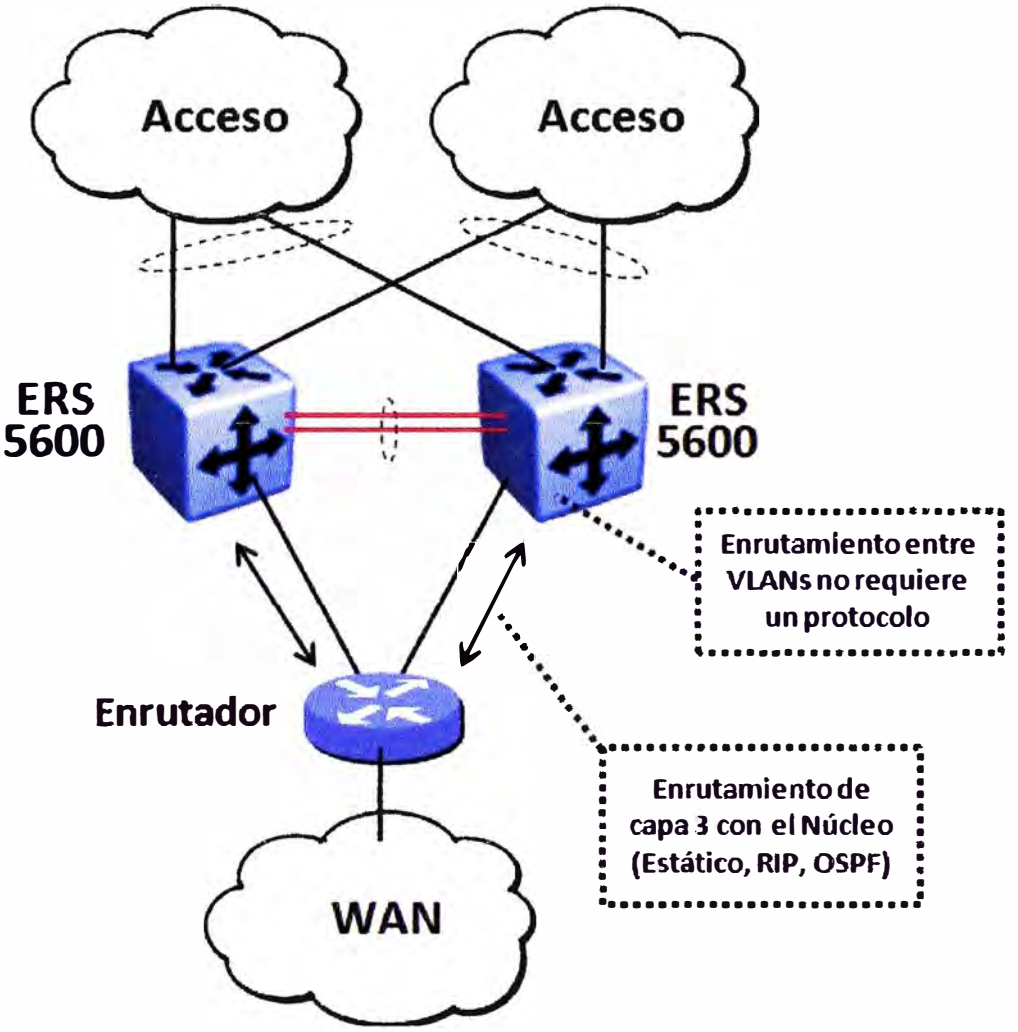


Fig. 4.1 Enrutamiento con la WAN

- VRRP

  Como se explico en el capítulo anterior, VRRP ofrece redundancia para el nodo de pasarela IP por defecto que requieren los usuarios finales. Este protocolo debe ser configurado por cada VLAN que alberga estaciones finales en la red.

  Junto con VRRP, la funcionalidad BM (Backup Master) debe ser habilitada en el agrupamiento de conmutadores para ofrecer enrutamiento activo-activo y reenvío directo del tráfico IP.

Las consideraciones de diseño para el uso de este protocolo son las siguientes:

- Habilitar VRRP-BM en cada VLAN.
- Configurar la prioridad de VRRP a 200 para el conmutador VRRP que va a ser designado como maestro. Dejar el valor por defecto (100) para el conmutador VRRP de respaldo.
- Alternar entre los conmutadores del núcleo el rol de maestro por cada VLAN que se configure.
- No seleccionar una dirección IP virtual que coincida con una dirección IP real de uno de los conmutadores. En su lugar, utilizar una tercera dirección, como por ejemplo:
  - Dirección IP Física de VLAN en conmutador 1 = x.x.x.2
  - Dirección IP Física de VLAN en conmutador 2 = x.x.x.3
  - Dirección IP Virtual de VLAN = x.x.x.1
- El temporizador "hold-down" de VRRP debe ser establecido en un valor lo suficientemente largo, para dar tiempo a que el protocolo IGP (en caso se utilice uno) pueda reconverger y actualizar la tabla de enrutamiento de los conmutadores.
  - Lo recomendado es colocar este temporizador a un valor mínimo de 1.5 veces el tiempo de convergencia del protocolo IGP.
  - Para OSPF, este valor debiera ser igual a 90 segundos si se utilizan temporizadores OSPF con valores por defecto.

- Relevo de DHCP

  El protocolo DHCP es utilizado para ofrecer direcciones IP dinámicas a las estaciones que se conectan a una red. Cuando se utilizan múltiples VLANs, la funcionalidad de relevo DHCP (DHCP relay en inglés) es necesaria para reenviar la información DHCP desde una VLAN de usuarios hacia el servidor DHCP que se ubicará en una VLAN distinta de servidores.

  Las consideraciones de diseño son:

  - Habilitar la funcionalidad de relevo DHCP para cada VLAN, utilizando la dirección IP física de la VLAN como el agente DHCP de relevo.
  - Cada agente DHCP de relevo puede soportar hasta 10 direcciones IP correspondientes a servidores DHCP.

4.7 Distribución de los Equipos

A continuación presentamos la tabla 4.2 con la distribución de los equipos seleccionados para el diseño.

En la tabla anterior se han tomado las siguientes consideraciones:

- Dado que las distancias entre los cuartos de cableado y el centro de datos no superan los 250 metros, se utilizara los siguientes elementos para la conexión en la capa física:
  - Tipo de Fibra Óptica
    - Multimodo 50um
    - Ancho de Banda Modal = 2000 MHz-Km
  - Enlaces de 1 Gigabit Ethernet
    - Tipo de SFP = 1000BaseSX
    - Distancia Máxima = 550 metros
  - Enlaces de 10 Gigabit Ethernet
    - Tipo de XFP = 10GBase-SR
    - Distancia Máxima = 300 metros
- Para distribuir la carga de tráfico proveniente desde y hacia los conmutadores de acceso, se utiliza DMLT con las siguientes reglas prácticas:
  - Hasta 80 puertos de usuario en el apilamiento, se utiliza 2 enlaces Gigabit Ethernet 1000Base SX.
  - Hasta 100 puertos de usuario en el apilamiento, se utiliza 4 enlaces Gigabit Ethernet 1000Base SX.
  - Hasta 400 puertos de usuario en el apilamiento, se utiliza 2 enlaces 10 Gigabit Ethernet.
- Para cada conmutador ERS 5632FD se debe habilitar la mitad del total de puertos para enlaces SMLT que aparecen en las columnas correspondientes de la tabla anterior, lo cual resulta en:
  - 12 puertos 1000BaseSX (habilitados con interfaces SFP).
  - 2 puertos 10GBase-SR (habilitados con interfaces XFP).
- Para la conexión dual 10/100/1000 de todos los servidores hacia el núcleo se utilizan dos conmutadores ERS 5650TD. Estos conmutadores son apilados en forma nativa con los conmutadores ERS 5632FD para formar una sola unidad lógica.
- Finalmente, la troncal IST se forma entre los conmutadores apilados en el núcleo de la forma siguiente:
  - 1 puerto 10GBase-SR (habilitado con interface XFP) en cada conmutador ERS 5632FD para la conexión entre ambos.
  - 1 puerto 10GBase-SR (habilitado con interface XFP) en cada conmutador ERS 5650TD para la conexión entre ambos.

- La troncal IST se forma creando un grupo MLT con los dos puertos 10GBase-SR indicados anteriormente.

4.8 Diagrama de la Red de Campus

Con la descripción anterior, se presenta a continuación el diagrama de la red de campus (figura 4.2) que muestra la topología general y la aplicación de las consideraciones de diseño que fueron explicadas anteriormente.

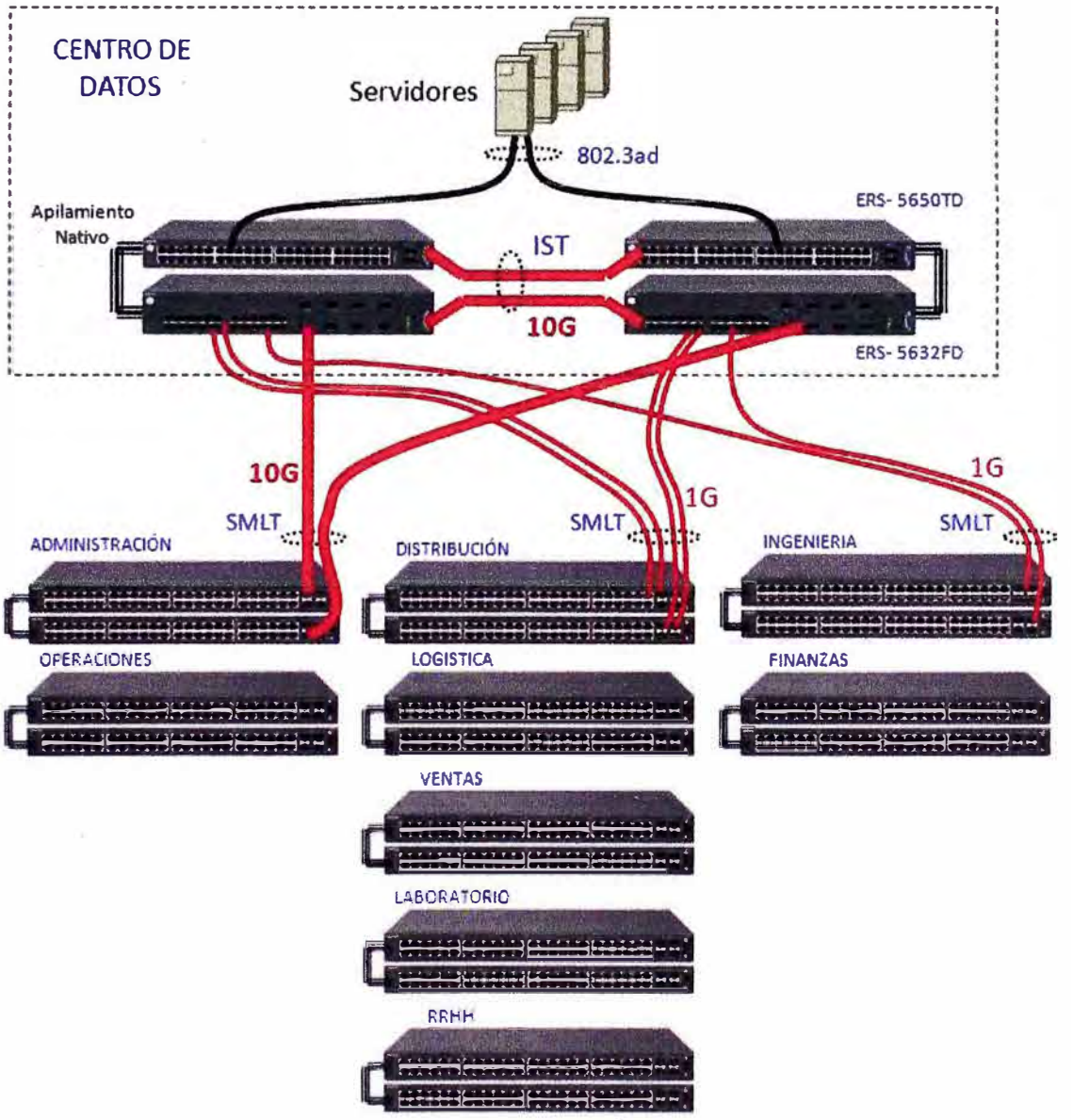


Fig. 4.2 Diagrama de la Red de Campus

4.9 Configuración de los Equipos

En esta sección voy a presentar las plantillas de configuración de los conmutadores de núcleo y de acceso que han sido utilizadas en el diseño. Voy a tomar como referencia la figura 4.3 detallada a nivel de puertos que sigue a continuación, para obtener los

parámetros utilizados en los comandos de configuración que corresponden a cada equipo.

Aunque esta plantilla solo muestra la configuración de algunos equipos de la capa de acceso, la misma es plenamente replicable para todos los equipos en los cuartos de cableado, donde solo se cambian los parámetros de uso que son dependientes de la ubicación, el direccionamiento IP y el equipo mismo.

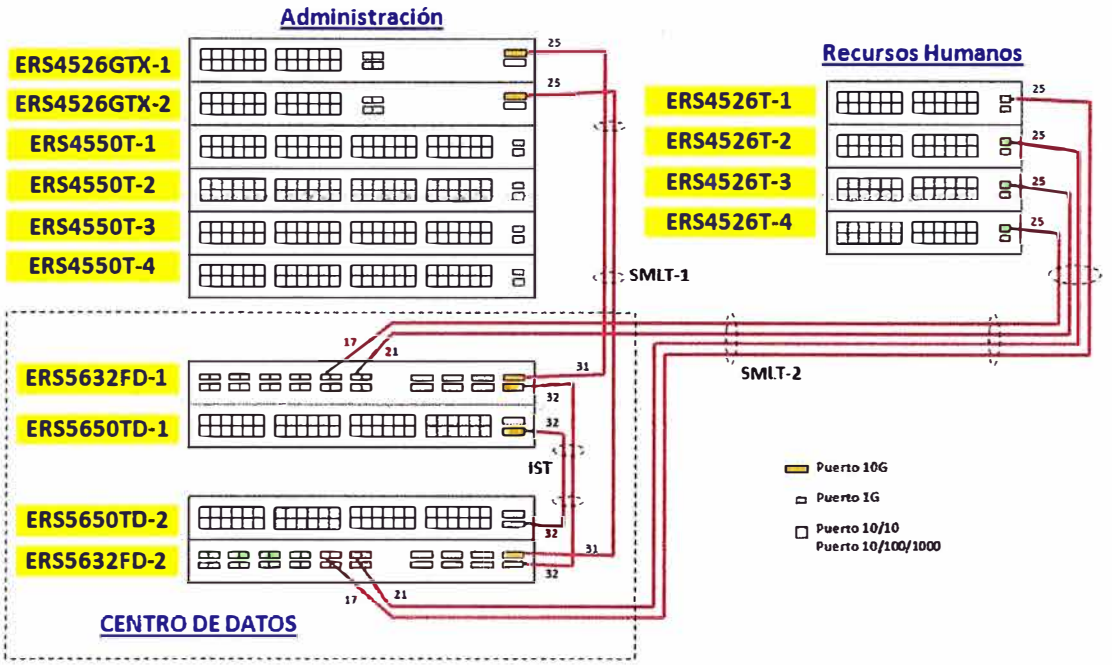


Fig. 4.3 Interconexión a nivel de puertos

En la figura 4.3 anterior aparece la numeración de los puertos por cada equipo, la cual sigue el orden siguiente:

- El puerto 1 de un equipo corresponde al puerto ubicado en la parte superior izquierda de la vista frontal (como se ve en la figura). El puerto 2 corresponde al puerto inmediatamente inferior y el puerto 3 vuelve a ser el puerto superior ubicado a la derecha del puerto 1. De esta forma, todos los puertos en la fila superior corresponden a puertos impares y los de la fila inferior a puertos pares.
- El conmutador ubicado en la parte superior del apilamiento corresponde a la unidad 1, siendo la unidad 2 el equipo inmediatamente inferior y así sucesivamente.
- La nomenclatura para designar un puerto del apilamiento es del tipo "unidad/puerto". Por ejemplo, el puerto 2/4 corresponde al puerto 4 de la unidad 2 dentro de un apilamiento.

En la figura 4.4 siguiente puede apreciarse el esquema lógico correspondiente a la figura anterior, en el lado izquierdo mostrando los enlaces SMLT y la conectividad entre los conmutadores del núcleo.

En el lado derecho se muestra como se extiende por todo el bloque básico la VLAN 2 (datos) que es la única que vamos a configurar a modo de ejemplo. El procedimiento para agregar más VLANs según se necesite en un caso particular, es totalmente replicable al que se va a mostrar para la VLAN 2.

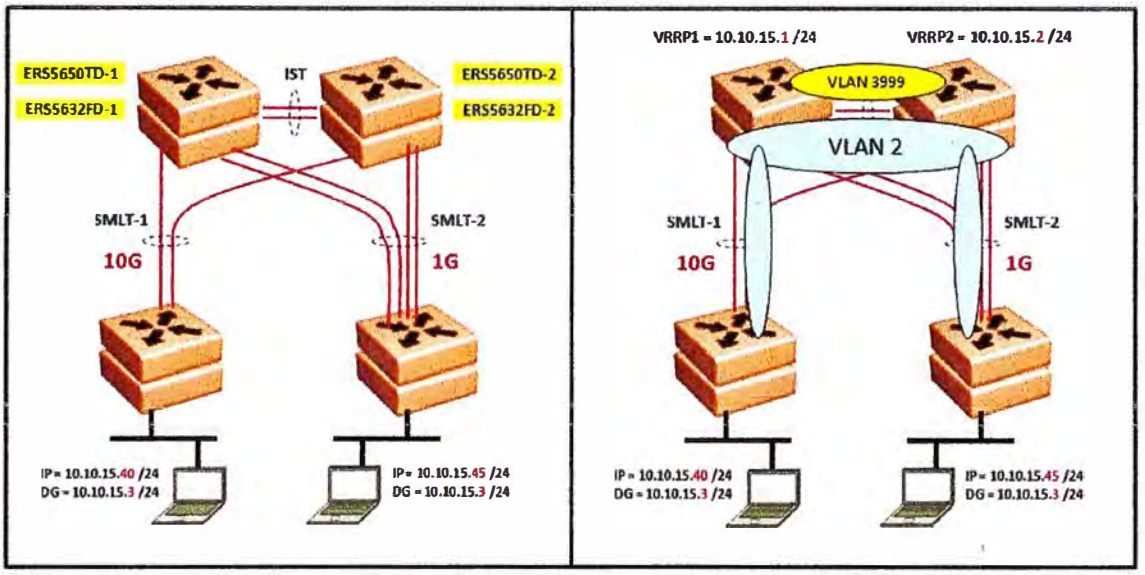


Fig. 4.4 Esquema Lógico de VLANs

4.9.1 Conmutadores del Núcleo

TABLA N° 4.5 Configuración Apilamiento ERS5632FD / ERS 5650TD

Paso	Acción	Modo	Comando de Configuración	Descripción
1	Crear VLAN para IST y VLAN de acceso	Global	**vlan create** *3999* **name** *grupo* **type port**	Crear VLAN 3999 del tipo PUERTO con nombre "grupo" para la troncal IST.
		Global	**vlan create** *2* **name** *grupo2* **type port**	Crear VLAN 2 del tipo PUERTO con nombre "grupo2" para la primera VLAN de acceso.
		Global	**vlan ports** *1/17, 1/21, 1/31, 1/32, 2/32* **tagging tagAll filter-untagged-frame enable**	Configurar los puertos indicados para marcar con 802.1Q todas las tramas de salida (tagAll) y para que filtren las tramas que no llegan marcadas.
		Global	**vlan members remove** *1 1/17, 1/21, 1/31, 1/32, 2/32*	Remover los puertos indicados de la VLAN por defecto (VLAN 1).
		Global	**vlan members add** *3999 1/32, 2/32*	Añadir los puertos indicados a la VLAN 3999
		Global	**vlan members add** *2 1/17, 1/21, 1/31, 1/32, 2/32*	Añadir los puertos indicados a la VLAN 2
2	Configurar VLACP	Global	**vlacp macaddress** *180.c200.f*	Configurar la dirección MAC utilizada para VLACP.
		Global	**vlacp enable**	Habilitar VLACP a nivel global.
3	Crear IST	Global	**mlt** *1* **name** *grupo* **enable member** *1/32, 2/32* **learning disable**	Crear y habilitar grupo MLT 1 con nombre "grupo". Agregar los puertos miembros indicados al grupo MLT. Deshabilitar STP en el grupo MLT.
		Global	**ip routing**	Habilitar el enrutamiento IP.
		Global	**interface vlan** *3999*	Ingresar a la VLAN asociada a la troncal IST.

		Interface	**ip address** *1.1.1.x 255.255.255.252*	Asignar una dirección IP local distinta a la VLAN 3999 en cada apilamiento.
		Interface	**exit**	Salir del modo de configuración.
		Global	**interface mlt** *1*	Ingresar a la interface lógica MLT 1
		Interface	**ist enable peer-ip** *1.1.1.z* **vlan** *3999*	Habilitar el protocolo IST apuntando a la dirección IP remota del otro apilamiento. Asociar la VLAN que tiene la dirección IP del apilamiento local.
		Interface	**exit**	Salir del modo de configuración.
		Global	**interface fastEthernet ALL**	Ingresa al modo de configuración de las interfaces LAN que se indican explícitamente.
		Interface	**vlacp port** *1/32, 2/32* **enable**	Configurar VLACP en los puertos IST con parámetros por defecto.
		Interface	**exit**	Salir del modo de configuración.
4	Configurar SMLT para un grupo	Global	**mlt** *2* **name** *grupo2* **enable member** *1/17,1/21* **learning disable**	Crear y habilitar grupo MLT 2 con nombre "grupo2", agrega los puertos miembros indicados y deshabilita STP en este grupo MLT.
		Global	**interface mlt** *2*	Ingresar a la interface lógica MLT 2
		Interface	**smlt** *2*	Crea un grupo SMLT 2 a partir de un grupo MLT. Se escoge el mismo número de identificación para ambos grupos, por simplicidad. Este número de identificación de grupo debe sincronizar con el otro conmutador a través de la troncal IST.
		Interface	**exit**	Salir del modo de configuración.
5	Configurar SMLT para un puerto	Global	**interface fastEthernet ALL**	Ingresa al modo de configuración de las interfaces LAN que se indican explícitamente.
		Interface	**smlt port** *1/31 33*	Crea directamente un puerto SMLT (el único en este caso) para otro grupo SMLT y le asocia un número de identificación 33. Este número de identificación debe sincronizar con el otro conmutador a través de la troncal IST.
		Interface	**exit**	Salir del modo de configuración.
6	Configurar VLACP hacia los Accesos	Global	**interface fastEthernet ALL**	Ingresa al modo de configuración de las interfaces LAN que se indican explícitamente.
		Interface	**vlacp port** *1/17, 1/21, 1/31* **timeout short**	Configurar VLACP en los todos los puertos SMLT con un valor de temporizador corto.
		Interface	**vlacp port** *1/17, 1/21, 1/31* **enable**	Habilitar VLACP en los puertos SMLT.
		Interface	**exit**	Salir del modo de configuración.
7	Configurar IP en la VLAN de Acceso. Habilitar OSPF en el Core solamente.	Global	**interface vlan** *2*	Ingresar a la VLAN asociada al acceso.
		Interface	**ip address** *10.10.15.x 255.255.255.0*	Asignar una dirección IP local distinta en cada equipo del núcleo, para la VLAN de acceso.
		Interface	**exit**	Salir del modo de configuración.
8	Habilitar VRRP	Global	**router vrrp enable**	Habilita el protocolo VRRP a nivel global.
		Global	**interface vlan** *2*	Ingresa a la configuración de la interfase VLAN.
		Interface	**ip vrrp address** *2 10.10.15.3*	Crea un grupo VRRP y le asigna un número 2 de identificación. Asigna una dirección IP virtual para el enrutador virtual.
		Interface	**ip vrrp** *2* **backup-master enable**	Habilita la funcionalidad Backup-Master para el grupo VRRP indicado.
		Interface	**ip vrrp** *2* **holddown-timer** *90*	Establece el temporizador en 90 segundos.

		Interface	**ip vrrp** *2* **priority** *200*	Establece la prioridad del conmutador VRRP Maestro. El otro conmutador debe dejarse en su valor por defecto (100).
		Interface	**ip vrrp** *2* **fast-adv enable**	Habilita la funcionalidad "Fast Advertisement" para que el tiempo de anuncio sea de 1 segundo.
		Interface	**ip vrrp** *2* **enable**	Habilita el grupo VRRP 2 indicado.
		Interface	**exit**	Salir del modo de configuración.
9	Habilitar DHCP (Opcional)	Global	**ip dhcp-relay fwd-path** *10.10.15.x 172.30.30.1* **mode dhcp**	Habilita la funcionalidad "DHCP Relay" con el agente 10.10.15.x, para que reenvíe hacia el servidor DHCP 172.30.30.1 (es un ejemplo) los paquetes del protocolo DHCP.
10	Verificación	Admin	**show mlt**	Muestra los grupos MLT creados.Para los conmutadores de núcleo, debe haber un grupo para IST y otro grupo para SMLT.
		Admin	**show vlan**	Muestra las VLAN creadas. Para los conmutadores de núcleo, debe haber una VLAN para IST, otra VLAN para el acceso y la VLAN por defecto.
		Admin	**show vlan interface info** *1/17, 1/21, 1/31, 1/32, 2/32*	Verifica que los puertos indicados están configurados para 802.1Q. Verifica el filtrado de tramas en los puertos indicados. El PVID de todos estos puertos es 1.
		Admin	**show vlan interface vids** *1/17, 1/21, 1/31, 1/32, 2/32*	Verifica a que VLANs pertenecen los puertos indicados.
		Admin	**show smlt**	Muestra el estado de los grupos y puertos SMLT.
		Admin	**show vlacp**	Muestra el estado del protocolo VLACP a nivel global.
		Admin	**show vlacp interface** *1/17, 1/21, 1/31, 1/32, 2/32*	Muestra el estado del protocolo VLACP en las interfaces principales.
		Admin	**show ip vrrp interface verbose vrid** 2	Muestra el estado de un grupo VRRP

4.9.2 Conmutadores de Acceso

TABLA N° 4.6 Configuración Apilamiento ERS4526T

Paso	Acción	Modo	Comando de Configuración	Descripción
1	Crear VLANs para IST y Acceso (una sola)	Global	**vlan create** *2* **name** *core* **type port**	Crear VLAN 2 del tipo PUERTO con nombre "core". Esta VLAN corresponde a los enlaces de subida hacia el núcleo.
		Global	**vlan ports** *1/25, 2/25, 3/25, 4/25* **tagging tagAll filter-untagged-frame enable**	Configurar los puertos indicados para marcar con 802.1Q todas las tramas de salida (tagAll) y para que filtren las tramas que no llegan marcadas.
		Global	**vlan members remove** *1 1/1-1/26, 2/1-2/26, 3/1-3/26, 4/1-4/26*	Remover los puertos indicados de la VLAN por defecto (VLAN 1).
		Global	**vlan members add** *2 1/1-1/26, 2/1-2/26, 3/1-3/26, 4/1-4/26*	Añadir los puertos indicados a la VLAN 2
2	Configurar VLACP	Global	**vlacp macaddress** *180.c200.f*	Configurar la dirección MAC utilizada para VLACP.
		Global	**vlacp enable**	Habilitar VLACP a nivel global.
		Global	**interface fastEthernet ALL**	Ingresa al modo de configuración de las interfaces LAN que se indican explícitamente.

		Interface	**vlacp port** *1/25, 2/25, 3/25, 4/25* **timeout short**	Configurar VLACP en los puertos SMLT que van al núcleo con un temporizador corto.
		Interface	**vlacp port** *1/25, 2/25, 3/25, 4/25* **enable**	Habilitar VLACP en los puertos SMLT.
		Interface	**exit**	Salir del modo de configuración.
3	Configurar MLT	Global	**mlt** *1* **name** *core* **member** *1/25, 2/25, 3/25, 4/25* **learning disable**	Crear grupo MLT 1 con nombre "core". Agregar los miembros indicados al grupo MLT. Deshabilitar STP en el grupo MLT.
		Global	**mlt** *1* **enable**	Habilitar el grupo MLT 1.
4	Habilitar FastStart y Rate-Limit	Global	**interface fastEthernet** *1/1-1/24, 1/26, 2/1-2/24, 2/26, 3/1-3/24, 3/26, 4/1-4/24, 4/26*	Ingresa al modo de configuración de las interfaces LAN indicadas explícitamente.
		Interface	**spanning-tree learning fast**	Habilita STP con la funcionalidad "FastStart".
		Interface	**rate-limit both** *10*	Establece el ratio máximo para tráfico de difusión por puerto. El ejemplo muestra el valor máximo disponible (10% del tráfico del puerto).
		Interface	**exit**	Salir del modo de configuración.
5	Verificación	Admin	**show mlt**	Muestra los grupos MLT creados.
		Admin	**show vlan**	Muestra las VLAN creadas..
		Admin	**show vlan interface info** *1/25, 2/25, 3/25, 4/25*	Verifica que los puertos indicados están configurados para 802.1Q. Verifica el filtrado de tramas en los puertos indicados.
		Admin	**show vlan interface vids** *1/25, 2/25, 3/25, 4/25*	Verifica a que VLANs pertenecen los puertos indicados.
		Admin	**show vlacp**	Muestra el estado del protocolo VLACP a nivel global.
		Admin	**show vlacp interface** *1/25, 2/25, 3/25, 4/25*	Muestra el estado del protocolo VLACP en las interfaces principales.

TABLA N° 4.7 Configuración Apilamiento ERS4526GTX

Paso	Acción	Modo	Comando de Configuración	Descripción
1	Crear VLANs para IST y Acceso (una sola)	Global	**vlan create** *2* **name** *core* **type port**	Crear VLAN 2 del tipo PUERTO con nombre "core". Esta VLAN corresponde a los enlaces de subida hacia el núcleo.
		Global	**vlan ports** *1/25, 2/25* **tagging tagAll filter-untagged-frame enable**	Configurar los puertos indicados para marcar con 802.1Q todas las tramas de salida (tagAll) y para que filtren las tramas que no llegan marcadas.
		Global	**vlan members remove** *1 1/1-1/26, 2/1-2/26, 3/1-3/50, 4/1-4/50, 5/1-5/50, 6/1-6/50*	Remover los puertos indicados de la VLAN por defecto (VLAN 1).
		Global	**vlan members add** *2 1/1-1/26, 2/1-2/26, 3/1-3/50, 4/1-4/50, 5/1-5/50, 6/1-6/50*	Añadir los puertos indicados a la VLAN 2
2	Configurar VLACP	Global	**vlacp macaddress** *180.c200.f*	Configurar la dirección MAC utilizada para VLACP.
		Global	**vlacp enable**	Habilitar VLACP a nivel global.
		Global	**interface fastEthernet ALL**	Ingresa al modo de configuración de las interfaces LAN que se indican explícitamente.
		Interface	**vlacp port** *1/25, 2/25* **timeout short**	Configurar VLACP en los puertos SMLT que van al núcleo con un temporizador corto.
		Interface	**vlacp port** *1/25, 2/25* **enable**	Habilitar VLACP en los puertos SMLT.
		Interface	**exit**	Salir del modo de configuración.
3	Configurar MLT	Global	**mlt** *1* **name** *core* **member** *1/25, 2/25* **learning disable**	Crear grupo MLT 1 con nombre "core". Agregar los miembros indicados al grupo MLT. Deshabilitar STP en el grupo MLT.
		Global	**mlt** *1* **enable**	Habilitar el grupo MLT 1.

4	Habilitar FastStart y Rate-Limit	Global	**interface fastEthernet** *1/1-1/24, 1/26, 2/1-2/24, 2/26, 3/1-3/50, 4/1-4/50, 5/1-5/50, 6/1-6/50*	Ingresa al modo de configuración de las interfaces LAN indicadas explícitamente.
		Interface	**spanning-tree learning fast**	Habilita STP con la funcionalidad "FastStart".
		Interface	**rate-limit both** *10*	Establece el ratio máximo para tráfico de difusión por puerto. El ejemplo muestra el valor máximo disponible (10% del tráfico del puerto).
		Interface	**exit**	Salir del modo de configuración.
5	Verificación	Admin	**show mlt**	Muestra los grupos MLT creados.
		Admin	**show vlan**	Muestra las VLAN creadas.
		Admin	**show vlan interface info** *1/25, 2/25*	Verifica que los puertos indicados están configurados para 802.1Q. Verifica el filtrado de tramas en los puertos indicados.
		Admin	**show vlan interface vids** *1/25, 2/25*	Verifica a que VLANs pertenecen los puertos indicados.
		Admin	**show vlacp**	Muestra el estado del protocolo VLACP a nivel global.
		Admin	**show vlacp interface** *1/25, 2/25*	Muestra el estado del protocolo VLACP en las interfaces principales.

CAPÍTULO V
COSTOS Y PRESUPUESTOS

En este capítulo voy a presentar los costos asociados al equipamiento que fue presentado en el diseño del capítulo anterior. Esto nos permitirá tener una idea de la inversión total que satisface los requerimientos para la infraestructura de red de campus revisada anteriormente, considerando como requisito principal la alta disponibilidad que se logra incorporando la tecnología de virtualización SMLT.

Los costos que considero a continuación (Tablas N° 5.1 al 5.10) corresponden a los precios de lista del fabricante Avaya (antes Nortel) para inicios de este año 2010, lo cual representa una muy buena estimación de presupuesto para los costos reales que se obtendrían en el mercado local, considerando el proceso de importación.

5.1 Costos Detallados del Equipamiento

TABLA N° 5.1 Centro de Datos

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 5632FD				
AL1001E15-E5	Ethernet Routing Switch 5632FD with 24 SFP ports, 8 XFP ports, 300W AC PS, 1.5 foot Stacking Cable., and Base Software License Kit (See Note 1). [EUED RoHS 5/6 compliant]. NA Power Cord	2	10,198	20,396
AL1905E03-E5	Ethernet Routing Switch 5600 redundant 300W AC power supply. For use in the ERS5698TFD, 5650TD, and 5632FD. [EUED RoHS 5/6 compliant]. NA Power Cord	2	253	506
AL1016001	Ethernet Routing Switch 5000 series Advanced License Kit, for 1 switch or stack. Enabled features: SMLT, OSPF, ECMP, VRRP, and IPFIX. (one license required per stack or standalone unit).	2	2,072	4,144
AVAYA ERS 5650TD				
AL1001E14-E5	Ethernet Routing Switch 5650TD with 48 10/100/1000 ports, 2 XFP ports, 300W AC PS, 1.5 foot Stacking Cable., and Base Software License Kit (See Note 1). [EUED RoHS 5/6 compliant]. NA Power Cord	2	4,078	8,156
AL1905E03-E5	Ethernet Routing Switch 5600 redundant 300W AC power supply. For use in the ERS5698TFD, 5650TD, and 5632FD. [EUED RoHS 5/6 compliant]. NA Power Cord	2	253	506
Interfaces y Cables				
AL2018027-E6	Ethernet Routing Switch 5600 family 5-meter stack cable [EUED RoHS 6/6 compliant]	2	269	538
AA1403005-E5	1-port 10GBase-SR XFP. Supports high modal bandwidth MMF (i.e. 50um, 2000MHz*km) for interconnects up to 300m. Core 62.5um fiber also supported. Please refer to documentation for fiber loss budgets.	6	1,060	6,360
AA1419048-E6	1-port 1000Base-SX Small Form Factor Pluggable (SFP) Gigabit Ethernet Trancsceiver, connector type: LC. Digital Diagnostic Monitoring Interface.	24	207	4,968
			Sub-Total (US$)	45,574

TABLA N° 5.2 Administración

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4526GTX				
AL4500E06-E6	Ethernet Routing Switch 4526GTX with 24 10/100/1000 BaseTX ports and 4 shared SFP ports plus 2 10Gig XFP slots, HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	2	2,548	5,096
AVAYA ERS 4550T / 4550T-PWR				
AL4500E02-E6	Ethernet Routing Switch 4550T with 48 10/100 BaseTX ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	2	1,171	2,342
AL4500E12-E6	Ethernet Routing Switch 4550T-PWR with 48 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	2	2,038	4,076
Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as return cable for resiliency). [RoHS compliant].	1	151	151
AA1403005-E5	1-port 10GBase-SR XFP. Supports high modal bandwidth MMF (i.e. 50um, 2000MHz*km) for interconnects up to 300m. Core 62.5um fiber also supported. Please refer to documentation for fiber loss budgets.	2	1,060	2,120
			Sub-Total (US$)	13,785

TABLA N° 5.3 Operaciones

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4526GTX				
AL4500E06-E6	Ethernet Routing Switch 4526GTX with 24 10/100/1000 BaseTX ports and 4 shared SFP ports plus 2 10Gig XFP slots, HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	2	2,548	5,096
AVAYA ERS 4526T / 4550T-PWR				
AL4500E03-E6	Ethernet Routing Switch 4526T with 24 10/100 BaseTX ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	1,171	1,171
AL4500E12-E6	Ethernet Routing Switch 4550T-PWR with 48 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	2,038	2,038
Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as return cable for resiliency). [RoHS compliant].	1	151	151
AA1403005-E5	1-port 10GBase-SR XFP. Supports high modal bandwidth MMF (i.e. 50um, 2000MHz*km) for interconnects up to 300m. Core 62.5um fiber also supported. Please refer to documentation for fiber loss budgets.	2	1,060	2,120
			Sub-Total (US$)	10,576

TABLA N° 5.4 Distribución

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4548GT				
AL4500E04-E6	Ethernet Routing Switch 4548GT with 48 10/100/1000 BaseTX ports and 4 shared SFP ports, plus HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	3,211	3,211
AVAYA ERS 4550T-PWR				
AL4500E12-E6	Ethernet Routing Switch 4550T-PWR with 48 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	2,038	2,038
Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as retum cable for resiliency). [RoHS compliant].	1	151	151
AA1419048-E6	1-port 1000Base-SX Small Form Factor Pluggable (SFP) Gigabit Ethernet Transceiver, connector type: LC. Digital Diagnostic Monitoring Interface.	4	207	828
			Sub-Total (US$)	6,228

TABLA N° 5.5 Logística

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4550T / 4550T-PWR				
AL4500E02-E6	Ethernet Routing Switch 4550T with 48 10/100 BaseTX ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	1,171	1,171
AL4500E12-E6	Ethernet Routing Switch 4550T-PWR with 48 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	2,038	2,038
Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as retum cable for resiliency). [RoHS compliant].	1	151	151
AA1419048-E6	1-port 1000Base-SX Small Form Factor Pluggable (SFP) Gigabit Ethernet Transceiver, connector type: LC. Digital Diagnostic Monitoring Interface.	4	207	828
			Sub-Total (US$)	4,188

TABLA N° 5.6 Ventas

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4548GT				
AL4500E04-E6	Ethernet Routing Switch 4548GT with 48 10/100/1000 BaseTX ports and 4 shared SFP ports, plus HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	3,211	3,211
AVAYA ERS 4550T-PWR				
AL4500E12-E6	Ethernet Routing Switch 4550T-PWR with 48 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	2,038	2,038

Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as return cable for resiliency). [RoHS compliant].	1	151	151
AA1419048-E6	1-port 1000Base-SX Small Form Factor Pluggable (SFP) Gigabit Ethernet Transceiver, connector type: LC. Digital Diagnostic Monitoring Interface.	4	207	828
			Sub-Total (US$)	6,228

TABLA N° 5.7 Laboratorio

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4524GT				
AL4500E05-E6	Ethernet Routing Switch 4524GT with 24 10/100/1000 BaseTX ports and 4 shared SFP ports (inc 100FX SFP support), HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	1,528	1,528
AVAYA ERS 4526T / 4550T-PWR				
AL4500E03-E6	Ethernet Routing Switch 4526T with 24 10/100 BaseTX ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	1,171	1,171
AL4500E12-E6	Ethernet Routing Switch 4550T-PWR with 48 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	2,038	2,038
Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as return cable for resiliency). [RoHS compliant].	1	151	151
AA1419048-E6	1-port 1000Base-SX Small Form Factor Pluggable (SFP) Gigabit Ethernet Transceiver, connector type: LC. Digital Diagnostic Monitoring Interface.	4	207	828
			Sub-Total (US$)	5,716

TABLA N° 5.8 Ingeniería

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4548GT				
AL4500E04-E6	Ethernet Routing Switch 4548GT with 48 10/100/1000 BaseTX ports and 4 shared SFP ports, plus HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	3,211	3,211
AVAYA ERS 4526T-PWR				
AL4500E13-E6	Ethernet Routing Switch 4526T-PWR with 24 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	1,528	1,528
Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as return cable for resiliency). [RoHS compliant].	1	151	151
AA1419048-E6	1-port 1000Base-SX Small Form Factor Pluggable (SFP) Gigabit Ethernet Transceiver, connector type: LC. Digital Diagnostic Monitoring Interface.	2	207	414
			Sub-Total (US$)	5,304

TABLA N° 5.9 Finanzas

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4548GT				
AL4500E04-E6	Ethernet Routing Switch 4548GT with 48 10/100/1000 BaseTX ports and 4 shared SFP ports, plus HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	3,211	3,211
AVAYA ERS 4526T-PWR				
AL4500E13-E6	Ethernet Routing Switch 4526T-PWR with 24 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	1,528	1,528
Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as return cable for resiliency). [RoHS compliant].	1	151	151
AA1419048-E6	1-port 1000Base-SX Small Form Factor Pluggable (SFP) Gigabit Ethernet Transceiver, connector type: LC. Digital Diagnostic Monitoring Interface.	2	207	414
			Sub-Total (US$)	5,304

TABLA N° 5.10 Recursos Humanos

Código	Descripción	Cant	Precio Unit (US$)	Precio Total (US$)
AVAYA ERS 4526T / 4526T-PWR				
AL4500E03-E6	Ethernet Routing Switch 4526T with 24 10/100 BaseTX ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS slot. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	3	1,171	3,513
AL4500E13-E6	Ethernet Routing Switch 4526T-PWR with 24 10/100 802.3af PoE ports plus 2 combo 10/100/1000 SFP ports, HiStack ports and RPS connector. Inc. Base Software License & 46cm stack cable. [RoHS compliant] (N America power cord)	1	1,528	1,528
Interfaces y Cables				
AL4518002-E6	4500-SSC HiStack Stacking Cable 1.5m (5ft) for Ethernet Routing Switch 4500 series (spare or for use as return cable for resiliency). [RoHS compliant].	1	151	151
AA1419048-E6	1-port 1000Base-SX Small Form Factor Pluggable (SFP) Gigabit Ethernet Transceiver, connector type: LC. Digital Diagnostic Monitoring Interface.	4	207	828
			Sub-Total (US$)	6,020

5.2 Resumen de Costos

A continuación se presenta la tabla N° 5.11 con un resumen de los costos detallados y mostrados en la sección anterior.

Todos los costos indicados están en dólares americanos y no incluyen el IGV. Obsérvese que como es de esperarse, el costo más alto corresponde al núcleo de la red donde se encuentran ubicados los 2 conmutadores agrupadores SMLT.

Los costos de los equipos de la capa de acceso son menores en proporción pero se encuentran dentro del promedio del mercado, dependiendo si los puertos son del tipo 10/100 0 10/100/1000.

Tabla N° 5.11 Resumen de Costos

Área de la Empresa	Sub-Totales (US$)
Centro de Datos	45,574.00
Administración	13,785.00
Operaciones	10,576.00
Distribución	6,228.00
Logística	4,188.00
Ventas	6,228.00
Laboratorio	5,716.00
Ingeniería	5,304.00
Finanzas	5,304.00
Recursos Humanos	6,020.00
Total Final (US$)	**108,923.00**

5.3 Análisis Comparativo en el Mercado

De la lista anterior, se deduce que los costos asociados a los conmutadores de acceso se encuentran dentro del promedio con otras marcas para este tipo de equipos considerando las mismas características y tipos de puertos. Esto tiene sentido porque en ellos solo se utiliza tecnología estándar.

Si resulta muy resaltante que el costo del núcleo (ubicado en el Centro de Datos) completamente redundante (incluyendo fuentes de poder redundantes en todos los conmutadores) y basado en la tecnología SMLT resulte significativamente inferior en costos a tecnología similar en el mercado que brinda los mismos niveles de resiliencia.

Por ejemplo, el fabricante Cisco ofrece la tecnología VSS (Virtual Switching System, en inglés) que permite a dos conmutadores Catalyst 6500 (tipo chasis) operar virtualmente como si fuera uno solo para lograr un alto grado de resiliencia. VSS exige que estos conmutadores operen con procesadores de última generación, lo cual implica una inversión muy superior al monto de todo el equipamiento estimado en la sección anterior (para la misma cantidad y tipo de puertos) y esto solo por un equipo Catalyst 6500. Esto demuestra lo altamente competitivo en costos que resulta el diseño SMLT planteado.

CONCLUSIONES Y RECOMENDACIONES

1. En el diseño de redes de campus, la consideración de alta disponibilidad es normalmente un factor que se deja de lado por los altos costos y la complejidad asociada.
2. SMLT es una tecnología de virtualización que permite extender el concepto de agrupamiento de enlaces, logrando que dos conmutadores del núcleo operen como si fuera uno solo, evitando el uso de la tecnología RSTP la cual se emplea en el diseño tradicional de redes de campus.
3. En el capítulo IV, se presento un ejemplo de diseño utilizando SMLT para una red de 1,000 usuarios, lo cual incluyo el detalle hasta el nivel de las líneas de comando por equipo, lo que permitió apreciar la simplicidad de la aplicación de esta tecnología y su escalabilidad con el creciente número de VLANs de acceso que se requieren en la actualidad.
4. SMLT no solo supera en simplicidad y escalabilidad a RSTP, sino que ofrece una real característica de resiliencia en el diseño de campus, ofreciendo un tiempo de recuperación frente a fallas, menor a 1 segundo, lo cual es imperceptible para las aplicaciones más exigentes de los usuarios finales en la actualidad.
5. En el capítulo "Costos y Presupuestos" se presento un detalle de los costos asociados al equipamiento utilizado en el ejemplo de diseño, donde se demostró la alta competitividad en este aspecto de los equipos que usan esta tecnología.
6. Se recomienda el uso de la tecnología SMLT explicada en este informe, la cual brinda un alto nivel de resiliencia y simplicidad en el diseño de redes de campus.

ANEXO A
Hojas de Datos de los Equipos

NORTEL

Hot Sheet

Nortel Ethernet Routing Switch 4500

The Nortel Ethernet Routing Switch 4500 product family is a stackable system that provides the high-performance, convergence-ready, secure and resilient Ethernet switching connectivity required by today's application- and competition-driven enterprise networks. The Ethernet Routing Switch (ERS) 4500 delivers 10/100 and 10/100/1000 switching with Power over Ethernet and 10 Gigabit uplink models for simplified and flexible network deployments, driving lower Total Cost of Ownership.

The Ethernet Routing Switch 4500 is available in ten models, as detailed below. All models include built-in HiStack stacking ports that can deliver up to 320 Gbps stacking performance, plus redundant power support.

Ethernet Routing Switch 4500 Series

Nortel Ethernet Routing Switch 4500 highlights

- Simplified converged deployments through PoE, Advanced QoS and IP Handset port auto-configuration
- High-density desktop connectivity supporting up to 400 user ports
- Resilience through fail-safe stacking, distributed trunking and power redundancy
- Flexible mix-and-match stacking capabilities to best meet customers requirements

Ordering information

Order Code	Description
AL4500701-E6	ERS 4526FX with 24 100BaseFX ports plus 2 combo 10/100/1000/SFP ports
AL4500703-E6	ERS 4526T with 24 10/100 BaseTX ports plus 2 combo 10/100/1000/SFP ports
AL4500713-E6	ERS 4526T-PWR with 24 10/100 802.3af PoE ports plus 2 combo 10/100/1000/SFP ports
AL4500702-E6	ERS 4550T with 48 10/100 BaseTX ports plus 2 combo 10/100/1000/SFP ports
AL4500712-E6	ERS 4550T-PWR with 48 10/100 802.3af PoE ports plus 2 combo 10/100/1000/SFP ports
AL4500705-E6	ERS 4524GT with 24 10/100/1000 BaseTX ports and 4 shared SFP ports
AL4500706-E6	ERS 4526GTX with 24 10/100/1000 BaseTX ports and 4 shared SFP ports, plus 2 XGE XFP slots
AL4500716-E6	ERS 4526GTX-PWR with 24 10/100/1000 802.3af PoE ports and 4 shared SFP ports, plus 2 XGE XFP slots
AL4500704-E6	ERS 4548GT with 48 10/100/1000 BaseTX ports and 4 shared SFP ports
AL4500714-E6	ERS 4548GT-PWR with 48 10/100/1000 802.3af PoE ports and 4 shared SFP ports

Each unit includes base software license kit and 46cm stack cable.
? Signifies the power cord requirement. SFPs, RPS and spare stacking cables sold separately.

* The 100FX and T1 SFPs are supported on ERS 4526FX, 4526T, 4526T-PWR, 4550T, 4550T-PWR, and 4524GT.

Technical specifications

- 100FX ports: 24 MTRJ ports per switch
- 10/100 Ethernet ports: 24/48 per switch
- 10/100/1000 Ethernet ports: 24/48 per switch
- SFP Gigabit ports: 2-4 per switch
- SFP support: SX, LX, XD, ZX, CWDM, 100FX*, T1*
- Resilient Stacking: up to 8 units /480 ports per stack
- Stacking ports: 2 built-in HiStack ports per switch
- Total stacking capacity: 320 Gbps
- Total packet throughput: 6.6 – 138 Mpps
- Individual switch capacity: 48.8 – 184 Gbps
- Concurrent VLANs: 256
- Maximum MAC addresses: 8,000
- Jumbo Frame Support on all Gigabit and 10 Gigabit ports

Standards compliance

- IEEE 802.3 10BASE-T Ethernet
- IEEE 802.3u 100BASE-TX Fast Ethernet
- IEEE ANSI 802.3 Auto-negotiation
- IEEE 802.3z Gigabit Ethernet
- IEEE 802.3x Flow Control
- IEEE 802.1Q VLANs
- IEEE 802.1p Priority Queues
- IEEE 802.1D Spanning Tree
- IEEE 802.1w Rapid Spanning Tree

- IEEE 802.1s Multiple Spanning Tree Groups
- IEEE 802.3AB Link Layer Discovery Protocol
- IEEE 802.3ad Link Aggregation
- IEEE 802.1X Ethernet Authentication Protocol
- RFC 783 Trivial File Transfer Protocol (TFTP)
- RFC 791/950 Internet Protocol (IP)
- RFC 792 Internet Control Message Protocol (ICMP)
- RFC 826 Address Resolution Protocol (ARP)
- RFC 854 Telnet Server and Client
- RFC 951 / 1542 BOOTP
- RFC 1112 Internet Group Management Protocol v1
- RFC 1215 SNMP Traps Definition
- RFC 1271 / 1757 / 2819 RMON
- RFC 1361 / 1769 Simple Network Time Protocol (SNTP)
- RFC 1493 Bridge MIB
- RFC 1573 / 2863 Interface MIB
- RFC 1643 / 2665 Ethernet MIB
- RFC 1905 / 3416 SNMP
- RFC 1906 / 3417 SNMP Transport Mappings
- RFC 1907 / 3418 SNMP MIB
- RFC 1945 HTTP v1.0
- RFC 2011 SNMP v2 MIB for IP
- RFC 2012 SNMP v2 MIB for TCP
- RFC 2013 SNMP v2 MIB for UDP
- RFC 2138 RADIUS
- RFC 2236 Internet Group Management Protocol v2
- RFC 2474 Differentiated Services Support
- RFC 2570 / 3410 SNMPv3
- RFC 2571 / 3411 SNMP Frameworks
- RFC 2572 / 3412 SNMP Message Processing
- RFC 2573 / 3413 SNMPv3 Applications
- RFC 2574 / 3414 SNMPv3 USM
- RFC 2575 / 3415 SNMPv3 VACM
- RFC 2576 / 3584 Co-existence of SNMP v1/v2/v3
- RFC 2660 HTTPS (Secure Web Server)
- RFC 2665 Ethernet MIB
- RFC 2863 Interfaces Group MIB
- RFC 2674 Q-Bridge MIB
- RFC 2737 Entity MIBv2
- RFC 2819 RMON MIB

Additional features

- Customizable Auto-negotiation Advertisements (CANA)
- Multi-Link Trunking (6 groups of up to 4 links)
- Distributed Multi-Link Trunking (DMLT)
- Distributed Link Aggregation Groups (802.3ad LAG)
- Virtual Link Aggregation Control Protocol (VLACP)
- Nortel Multiple Spanning Tree groups
- Single IP address for stack management
- Resilient fail-safe stacking
- Automatic Unit Replacement (Configuration and Software)
- DSCP-based Recognition, Marking and Recolouring
- Ingress and Egress Port Mirroring
- Broadcast and Multicast Rate limiting per port
- ASCII Configuration File
- Web, NNCLI, IDM
- SSHv2 and SNMPv3 secure management support
- Automatic Detection Automatic Configuration (ADAC)
- 802.1X Single Host Single Authentication
- 802.1X Multiple Host Multiple Authentication
- 802.1X Guest VLAN
- 802.1X Single Host Multiple Authentication
- 802.1X Non-EAP (NEAP) access
- Nortel Secure Network Access (NSNA) support
- BPDU Filtering
- Stack Monitoring
- Backup and Restore switch software and configuration files
- USB Interface for configuration and software storage

Power over Ethernet specifications

- 802.3af compliant with Power classification support
- Signal pair power delivery
- Maximum 15.4 watts per port
- Maximum DTE Power AC 320 watts
- Maximum DTE Power AC + RPS 740 watts

Electrical specification

- Power supply: AC 100-240V, 50-60Hz
- Input current at 110v: 1.3A - 7.1A
- Input current at 220v: 0.7A - 3.6A
- Max power consumption: 150W - 470W
- Actual power consumption: 55W - 470W

Dimensions

- Width: 438.2mm (17.25 in)
- Height: 1RU 43.7mm (1.72 in)
- Depth: 368.3mm (14.5 in)
- Weight: 5kg (11lb) - 6.4kg (14lb)

Environmental specifications

- Operating temperature: 0 to 50 degrees C
- Storage temperature: -25 to 55 degrees C
- Relative humidity: 5% to 85% non-condensing
- Peak noise level: 37.2 - 43.6 dBA
- Thermal rating: 188 - 788 BTU/hr
- Calculated MTBF: 214,542 - 312,001 hrs

Safety Agency Approvals

- IEC 60950 International CB Certification
- EN 60950 European Certification
- UL60950 US certification
- CSA22.2, #60950 Canadian Certification
- NOM Mexican Certification

Electromagnetic Emissions and Immunity

- CISPR22, Class A/CISPR24 International
- EN55022, Class A/EN55024 European
- FCC, Part 15, Class A US Certification
- ICES-003, Class A Canadian Certification
- AN/NZS 3548 Australian/NZ Certification
- BSMI - Taiwan - CNS 13438, Class A
- MIC - Korea - MIC, No 2001-116
- VCCI Class A Japanese Certification

- IEEE 802.1s Multiple Spanning Tree Groups
- IEEE 802.3AB Link Layer Discovery Protocol
- IEEE 802.3ad Link Aggregation
- IEEE 802.1X Ethernet Authentication Protocol
- RFC 783 Trivial File Transfer Protocol (TFTP)
- RFC 791/950 Internet Protocol (IP)
- RFC 792 Internet Control Message Protocol (ICMP)
- RFC 826 Address Resolution Protocol (ARP)
- RFC 854 Telnet Server and Client
- RFC 951 / 1542 BOOTP
- RFC 1112 Internet Group Management Protocol v1
- RFC 1215 SNMP Traps Definition
- RFC 1271 / 1757 / 2819 RMON
- RFC 1361 / 1769 Simple Network Time Protocol (SNTP)
- RFC 1493 Bridge MIB
- RFC 1573 / 2863 Interface MIB
- RFC 1643 / 2665 Ethernet MIB
- RFC 1905 / 3416 SNMP
- RFC 1906 / 3417 SNMP Transport Mappings
- RFC 1907 / 3418 SNMP MIB
- RFC 1945 HTTP v1.0
- RFC 2011 SNMP v2 MIB for IP
- RFC 2012 SNMP v2 MIB for TCP
- RFC 2013 SNMP v2 MIB for UDP
- RFC 2138 RADIUS
- RFC 2236 Internet Group Management Protocol v2
- RFC 2474 Differentiated Services Support
- RFC 2570 / 3410 SNMPv3
- RFC 2571 / 3411 SNMP Frameworks
- RFC 2572 / 3412 SNMP Message Processing
- RFC 2573 / 3413 SNMPv3 Applications
- RFC 2574 / 3414 SNMPv3 USM
- RFC 2575 / 3415 SNMPv3 VACM
- RFC 2576 / 3584 Co-existence of SNMP v1/v2/v3
- RFC 2660 HTTPS (Secure Web Server)
- RFC 2665 Ethernet MIB
- RFC 2863 Interfaces Group MIB
- RFC 2674 Q-Bridge MIB
- RFC 2737 Entity MIBv2
- RFC 2819 RMON MIB

Additional features

- Customizable Auto-negotiation Advertisements (CANA)
- Multi-Link Trunking (6 groups of up to 4 links)
- Distributed Multi-Link Trunking (DMLT)
- Distributed Link Aggregation Groups (802.3ad LAG)
- Virtual Link Aggregation Control Protocol (VLACP)
- Nortel Multiple Spanning Tree groups
- Single IP address for stack management
- Resilient fail-safe stacking
- Automatic Unit Replacement (Configuration and Software)
- DSCP-based Recognition, Marking and Recolouring
- Ingress and Egress Port Mirroring
- Broadcast and Multicast Rate limiting per port
- ASCII Configuration File
- Web, NNCLI, JDM
- SSHv2 and SNMPv3 secure management support
- Automatic Detection Automatic Configuration (ADAC)
- 802.1X Single Host Single Authentication
- 802.1X Multiple Host Multiple Authentication
- 802.1X Guest VLAN
- 802.1X Single Host Multiple Authentication
- 802.1X Non-EAP (NEAP) access
- Nortel Secure Network Access (NSNA) support
- BPDU Filtering
- Stack Monitoring
- Backup and Restore switch software and configuration files
- USB Interface for configuration and software storage

Power over Ethernet specifications

- 802.3af compliant with Power classification support
- Signal pair power delivery
- Maximum 15.4 watts per port
- Maximum DTE Power AC 320 watts
- Maximum DTE Power AC + RPS 740 watts

Electrical specification

- Power supply: AC 100-240V, 50-60Hz
- Input current at 110v: 1.3A - 7.1A
- Input current at 220v: 0.7A - 3.6A
- Max power consumption: 150W - 470W
- Actual power consumption: 55W - 470W

Dimensions

- Width: 438.2mm (17.25 in)
- Height: 1RU 43.7mm (1.72 in)
- Depth: 368.3mm (14.5 in)
- Weight: 5kg (11lb) - 6.4kg (14lb)

Environmental specifications

- Operating temperature: 0 to 50 degrees C
- Storage temperature: -25 to 55 degrees C
- Relative humidity: 5% to 85% non-condensing
- Peak noise level: 37.2 - 43.6 dBA
- Thermal rating: 188 - 788 BTU/hr
- Calculated MTBF: 214,542 - 312,001 hrs

Safety Agency Approvals

- IEC 60950 International CB Certification
- EN 60950 European Certification
- UL60950 US certification
- CSA22.2, #60950 Canadian Certification
- NOM Mexican Certification

Electromagnetic Emissions and Immunity

- CISPR22, Class A/CISPR24 International
- EN55022, Class A/EN55024 European
- FCC, Part 15, Class A US Certification
- ICES-003, Class A Canadian Certification
- AN/NZS 3548 Australian/NZ Certification
- BSMI - Taiwan - CNS 13438, Class A
- MIC - Korea - MIC, No. 2001-116
- VCCI Class A Japanese Certification

NORTEL

Product Brief

Nortel Ethernet Routing Switch 5600 Series

Resilient, high-performance switching for the edge, aggregation or the network core

The Nortel Ethernet Routing Switch (ERS) 5600 Series is a set of high-performance stackable LAN switches providing the resiliency, security and convergence readiness required for today's high-end wiring closets, high-performance data centers and small/medium core environments.

Nortel's industry-leading resilient stacking technology provides high availability for delay-sensitive and business-critical voice and data. Up to eight Ethernet Routing Switch 5600 units can be stacked to form a single manageable and resilient stacked solution of up to 400 ports, with continuous uptime even if a switch within the stack should fail.

Scalable, resilient and flexible Ethernet switching

Recognizing that networking requirements vary from business to business with differing needs at the edge, core and distribution layer, Nortel offers five different models within the Ethernet Routing Switch 5600 family. This provides enterprises the flexibility to choose the model that best fits their networking requirements (refer to Table 1).

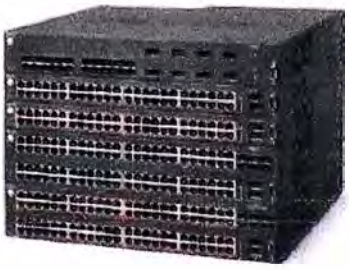
The Ethernet Routing Switch 5600 further introduces a per-unit stacking capacity of 144Gbps per switch and a total stacking backplane of 1.152Tbps for a stack of eight units. This is unprecedented in the industry and ensures optimal performance for delay-sensitive application traffic traversing the stack.

Ethernet Routing Switch 5600 models come with two in-built auto-sensing stacking ports for simple, quick and cost-effective stacking. This design frees up Ethernet Routing Switch 5600 uplink ports for dedicated connectivity to the backbone. The Ethernet Routing Switch 5600's non-blocking architecture further ensures full bandwidth across stack, uplink and user ports.

Ethernet Routing Switch 5600 switches can be stacked with one another or with any Ethernet Routing Switch 5500 unit. When connected to an adjacent

Table 1. Nortel Ethernet Routing Switch 5600 Series

Model	Specification
ERS 5650TD	48 10/100/1000 TX ports, plus 2 10GbE XFP ports
ERS 5650TD-PWR	48 10/100/1000 TX and 802.3af PoE ports, plus 2 10 GbE XFP ports
ERS 5698TFD	96 10/100/1000 TX ports including 6 Combo SFP ports, and 2 10GbE XFP ports (2RU High)
ERS 5698TFD-PWR	96 10/100/1000 TX and 802.3af PoE ports including 6 Combo SFP ports, and 2 10GbE XFP ports (2RU High)
ERS 5632FD	24 SFP ports, plus 8 10GbE XFP ports (1.5RU High)



Ethernet Routing Switch 5600 Series

Nortel's innovative, resilient stackable architecture

Nortel has continued to evolve and perfect its enterprise stackable portfolio into a truly resilient, high-performance stacking solution. Nortel's Ethernet Routing Switch 2500, 4500, 5500 and 5600 families all incorporate key capabilities designed for scalable stack performance and ease of management, including:

- **Ease of growth/management:** Simply cable a new unit into the stack, extend the configuration and more capacity can be added to the stack. The stack acts as a single entity that can be managed via a single IP address.
- **Scalable performance:** The non-blocking design of Nortel's stackable switches along with their high-speed stack connections (up to 144Gbps) ensure that the stack scales proportionally as each new unit is added.
- **Optimal path forwarding:** A bi-directional multi-path forwarding algorithm between switches ensures that the shortest, most optimal path is chosen for each data flow. This design offers a clear performance advantage over logical ring or Token approaches used by other vendors.
- **No single point of failure:** Stack operation is unaffected by the failure of any unit in the stack. Each unit's independent switch fabric along with redundant power ensures continuous operation of the stack. Individual stack units can further be replaced without bringing down the stack (Auto-Unit Replacement).
- **Split Multi-Link Trunking (SMLT):** Nortel's SMLT technology can further be incorporated across stacks to enhance overall resiliency of the stack solution providing load balancing and split second failover.

ERS 5500, the ERS 5600 automatically adjusts to the ERS 5500's stack bandwidth of 80 Gbps. This enhanced flexibility provides the opportunity to mix and match port configurations based on operations needs, while preserving existing ERS 5500 investment.

Ethernet Routing Switch 5600 deployment scenarios

Ethernet Routing Switch 5600 is a flexible solution that can be deployed in a variety of enterprise environments. These include:

High-density wiring closet

With its non-blocking design, high-density GbE and integrated 10GbE ports, the Ethernet Routing Switch 5600 is perfectly suited as a wiring closet solution for high-availability LAN edge user connectivity. Up to eight ERS 5600 switches can be combined into a single stack with each unit providing at least two 10GbE XFP ports for high-performance uplinks to the core or aggregation layers in the network. ERS 5698 models further provide high-density economical GbE connectivity with the added flexibility of up ro 6 SFP ports per switch for gigabit fiber connections. With the ability to support Power over Ethernet across all its ports, the ERS 5600 is an economical and flexible wiring closet/edge solution in support of desktop and powered devices.

High-performance data center edge

The high-performance, low-latency and high-availability requirements of the data center make the Ethernet Routing Switch 5600 a perfect solution for the data center edge. By utilizing the ERS 5600's innovative horizontal stacking capability — a switch at the top of each server rack, stacked horizontally — the ERS 5600 can support high-density server connections. Up to 384 GbE ports and multiple 10GbE uplinks can be deployed in this configuration (see Figure 1).

Nortel's Switch Cluster solution (SMLT) can further enhance overall resiliency through addition of a second horizontally-stacked "top of rack" switch. Connected servers can then be "dual homed" to separate Ethernet Routing Switch 5600 stacks, which in turn appear as a single, logical switch to the network. This enables active-active connections, load balancing and sub-second failover across the stacks. These features make the ERS 5600 a truly cost-effective data center solution that combines always-on resiliency with full non-blocking capabilities.

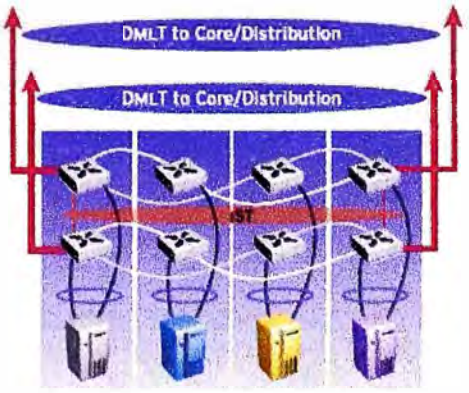


Figure 1. Ethernet Routing Switch 5600 in horizontal stacking/data center application

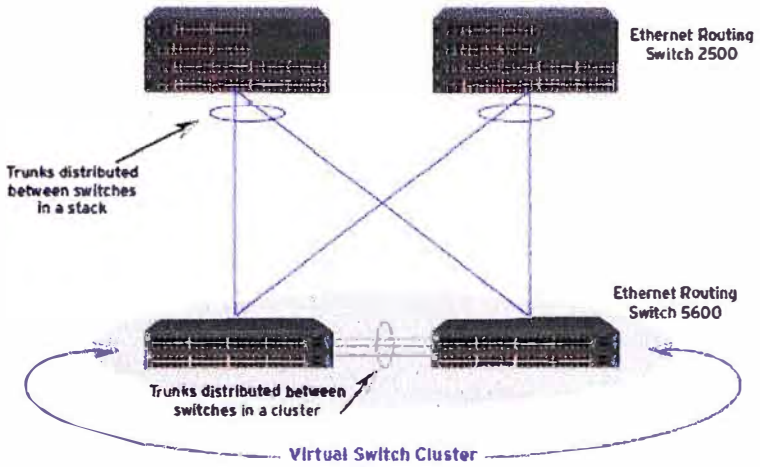


Figure 2. Ethernet Routing Switch 5600 in small core application using SMLT with ERS 2500s at the edge for redundancy and efficiency

Highly scalable small to medium core
The Ethernet Routing Switch 5600 can also serve as a high-performing, feature-rich core solution. With its high-density 10GbE and SFP ports (up to 64 10GbE and 192 SFP ports per stack), the Ethernet Routing Switch 5632FD is particularly well-suited to core applications. Coupled with dynamic routing protocol (RIP/OSPF) support and in-unit redundant power supplies, the ERS 5600 can provide a low-cost, resilient small core solution.

Ethernet Routing Switch 5600s can further act as a virtual switch cluster to other ERS 2500, 4500 or even third-party access switches (see Figure 2) using Nortel's switch clustering (SMLT) technology. ERS 5600s in this configuration can provide up to 800 ports and over 2.3 Terabits of performance — all while enabling full use of all switches and links across the network.

Convergence-ready capabilities

Power over Ethernet
The ERS 5650TD-PWR and 5698TFD-PWR models provide standards-based Power over Ethernet (PoE) and are designed to power devices such as IP phones, wireless access points, network cameras, security and lighting devices, and access control devices such as badge readers.

Full interoperability with all standards-based equipment means the switch has the flexibility to power multiple vendors' devices. Ethernet Routing Switch 5600 switches can supply up to 15.4 Watts per port, as defined in the IEEE 802.3af standard.

Link Layer Discovery Protocol and Auto Detection/Auto Configuration
The Ethernet Routing Switch 5600 series combines the functionality of the Link Layer Discovery Protocol (LLDP) and Auto Detection/Auto Configuration (ADAC) to detect and configure Nortel IP Phones. The implementation includes a Location Type-Length Value (TLV) which provides a mechanism to enable Emergency Location Services for IP phones (for example, Enhanced 911).

Auto-discovery is extended with LLDP-MED (Media Endpoint Discovery) to automatically perform power and VLAN assignments for enabled devices.

Nortel further extends the use of this protocol by coupling it with the ability to automatically configure QoS settings and VLAN membership when a Nortel IP handset is detected. This ensures fast and easy IP handset deployments across the network.

Resilient operations — redundant power and auto-unit replacement

Field-replaceable power supplies
Ethernet Routing Switch 5600 models support integrated AC or DC power supplies in the rear of each unit for improved redundancy and uptime. 5632FD and 5650TD models support two and ERS 5698TFD models support three integrated field-replaceable supplies. This power supply design not only offers N+1 power redundancy, but also provides savings in valuable rack space and reduction in overall system, servicing and sparing costs.

Nortel's industry-leading resilient stacking technology provides high availability for delay-sensitive and business-critical voice and data.

Automatic Unit Replacement
Ethernet Routing Switch 5600 units support Automatic Unit Replacement (AUR) functionality to allow network managers to quickly and easily replace failed units in the stack. When a unit is replaced, the configuration and software image are automatically synchronized with that of the failed unit. This minimizes manual intervention as well as any interruption to existing stack users.

Ethernet Routing Switch 5600 models further support the ability to load a software image into the base unit of the stack and have it automatically propagated to other switches in the stack. This further simplifies stack management.

Security for safeguarding the network

Protecting the network against both external and increasingly prevalent internal attacks is a critical part of every IT manager's job.

The ability to do this requires simple to manage, yet intelligent security solutions that not only look at the identity of the person logging in, but also at the device connecting into the network.

This network access control approach ensures that only authorized individuals and properly scanned/secured devices are allowed onto the network.

Nortel's Secure Network Access
Secure Network Access (SNA) — Nortel's endpoint security and policy compliance solution — inspects, assesses, ensures compliance to policy, and remediates at the network endpoint source, prior to network access.

Nortel Secure Network Access dramatically simplifies the complexity of enterprise network access architectures with a solution that assures endpoint security compliance. Nortel SNA provides this security through seamless device quarantine and containment, remediation and repair for LAN users and remote users, with both fixed and mobile connectivity devices.

Network access control through 802.1X-based authentication
The Ethernet Routing Switch 5600 supports the IEEE 802.1X-based security feature — EAP. Based on the IEEE 802.1X standard, EAP limits access to the network based on user credentials. A user is required to "login" to the network using a username/password. The user database is maintained on the authentication server, for example RADIUS (not the switch).

The Ethernet Routing Switch features MAC address-based security, which allows authentication of all access devices. Network access is granted or denied via proper MAC address identification. The ERS 5600 also supports Guest VLAN.

Additionally, for Voice over IP (VoIP) deployments where the PC connects to the network through a VoIP handset, the PC and the IP device are authenticated individually.

External attack protection
The Ethernet Routing Switch 5600 also supports a set of security features designed to protect against external attack. These include Dynamic Host Control Protocol (DHCP) snooping, Dynamic Address Resolution Protocol inspection, IP Source Guard and IGMP snooping. Together, these protect against snooping and man-in-the middle attacks that might compromise the network.

Table 2. Nortel Ethernet Routing Switch 5500 and 5600 Series model comparisons

Features	ERS 5510	ERS 5520-PWR	ERS 5530	ERS 5650	ERS 5698	ERS 5632
10/100/1000 ports	24 / 48	24 / 48	24 (12 shared)	48	96 (6 shared)	–
GbE SFP ports	2	4	12	–	6	24
10Gbps XFP ports	–	–	2	2	2	8
Power over Ethernet	No	Yes	No	Yes (PWR models)	Yes (PWR models)	No
Resilient fail-safe stacking	Yes	Yes	Yes	Yes /o	Yes	Yes
Stack capacity	640Gbps	640Gbps	640Gbps	1.152Tbps[1]	1.152Tbps[1]	1.152Tbps[1]
Number of switches supported by an RPS15 module	Up to 4	1	1	n/a	n/a	n/a
Number of AC or DC modular supplies	n/a	n/a	n/a	Up to 2	Up to 3	Up to 2
Typical deployment	Desktop Connectivity & Server Aggregation	Desktop Connectivity, PoE for Convergence Devices	Small Core & Server Aggregation	Desktop Connectivity (optional PoE) & Server Aggregation	Desktop Connectivity (optional PoE)	Server Aggregation & Small/Medium Core

[1] Auto-sensing stack compatibility with ERS 5500.

Simplified management

A stack of Ethernet Routing Switch 5600 switches can be managed as a single entity through one IP address to simplify network management.

The Ethernet Routing Switch 5600 offers the highest level of security with features including Secure Shell (SSH), IP Manager Lists, RADIUS (Remote Authentication Dial-In User Services) and TACACS+ Authentication, Simple Network Management Protocol (SNMPv3), and Nortel's Secure Network Access.

Dual image support

The Ethernet Routing Switch 5600 supports dual agent images. This allows the administrator to download a newer image to the switch/stack, and specify when the unit should reboot to take the new image. Once the new image is installed and running, the administrator has the ability to "roll back" to the previous image.

IPv6 management

The Ethernet Routing Switch 5600 supports either IPv4 or IPv6-based management. Simply select whether the switch/stack should be managed via IPv4 or IPv6, enter in the switch/stack IP address in the correct format, and the switch/stack is fully manageable.

Many-to-many port mirroring

The Ethernet Routing Switch 5600 introduces support for many-to-many port mirroring to address more complex, converged networks. This feature provides the ability to have multiple instances (up to 4) of many-to-one port mirroring, and flows to be captured simultaneously to traffic analyzers, call recorders, IDS/IPS devices, etc.

Advanced management features

BootP and TFTP support allows centralized Switch IP Address assignment, software upgrades and SNMP agent updates over the network. The RADIUS-based security feature uses the RADIUS protocol to authenticate local console and TELNET logins.

Lifetime warranty

All Nortel Ethernet Routing Switches offer a lifetime warranty. A Nortel product hardware warranty is supported for as long as the original end user continues to own or use the product, including fan and power supply. In the event of a discontinuance of product manufacture, Nortel warranty support is available for up to five (5) years after discontinuance.

ANEXO B
Estudio de Rendimiento de la Tecnología SMLT

TEST REPORT

Tolly.

#210118

March 2010

Commissioned by Avaya Inc.

Testing conducted September 2005

Avaya Ethernet Routing Switch 5000 Series

Competitive Performance Evaluation versus Cisco Catalyst 3750G and HP ProCurve 3400cl

Premise

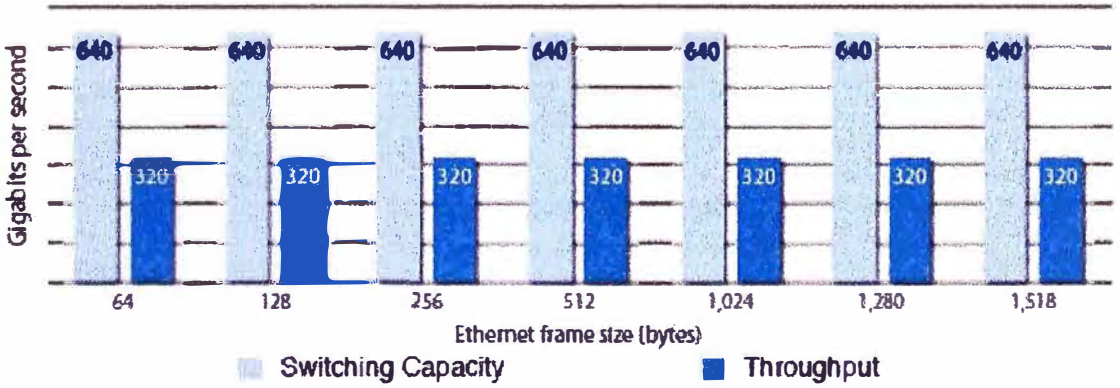
When considering the purchase of stackable switches, network managers need to know the performance characteristics of the available products. Buyers need to know the bidirectional performance characteristic in the multiple switch stack configuration, plus what impact, if any, a device outage will have on the overall performance of the switching stack.

Avaya commissioned The Tolly Group to evaluate the Layer 2 switching performance, resiliency and ease of use delivered by the company's stackable Ethernet Routing Switch 5000 series of switches. The Avaya Ethernet Routing Switch 5000 series of stackable switches tested include 24- and 48-port versions of 5510, 5520 and 5530 models – single rack-unit stackable Gigabit Ethernet (GbE) Layer3 routing switches designed to provide high-density GbE desktop connectivity to mid and large enterprise customers' wiring closets.

Test Highlights

1. Delivers superior stacking performance of up to 640 Gbps of switching capacity in an eight-unit stack of Avaya 5500 switches
2. Achieves line-rate performance of 202 Gbps frame-forwarding in an eight-unit stack, while Cisco and HP switches support only 25.7 Gbps and 114.7 Gbps respectively
3. Demonstrates 36% to 44% less average latency, when compared to Cisco and HP devices tested
4. Recovers from link and switch outages almost 10X faster using Avaya's SMLT implementation than the RSTP implementation in the Cisco Catalyst and HP ProCurve solutions tested
5. Offers the lowest cost per megabit of throughput among the switches tested at just below $90 versus almost $100 HP and over $300 for Cisco

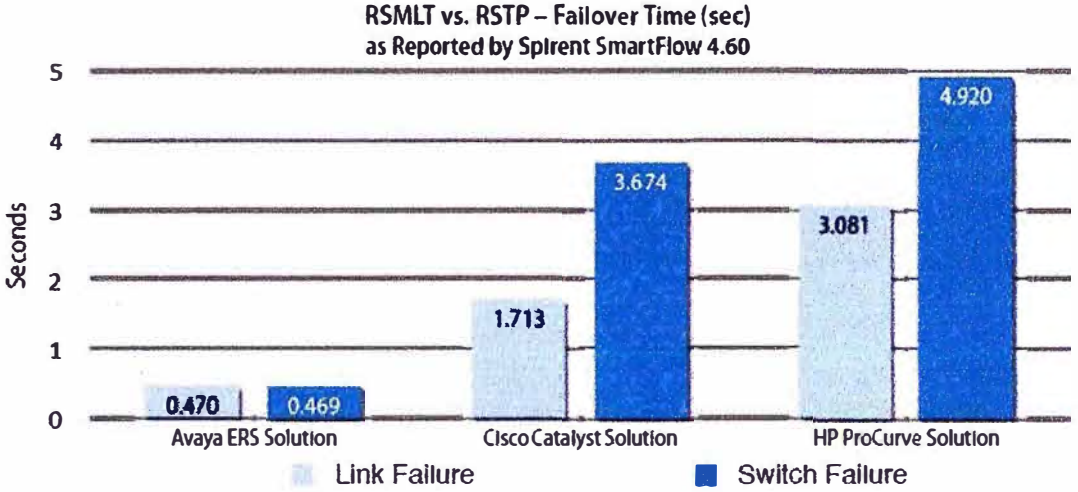
Aggregate Layer 2 Throughput of Ethernet Routing Switch 5510-48T in an Eight-switch Stack Configuration (320 GbE ports) as Reported by SmartBits SmartFlow 4.60



Note: A custom port-pairing scheme was used. See the "Test Configuration and Methodology" section

Source: The Tolly Group, September 2005

Figure 1



Avaya Ethernet Routing Switch solution: Consists of two Ethernet Routing Switch 5530 access switches and two Ethernet Routing 8600 core switches using SMLT for resiliency.
Cisco Catalyst Switch solution: Consists of two Catalyst 3750G access switches and two Catalyst 6500 core switches using RSTP for resiliency. HP **ProCurve Switch solution:** Consists of two ProCurve 3400cl access switches and two ProCurve 9300 core switches using RSTP for resiliency.

Source: The Tolly Group, September 2005 Figure 5

average cut-through latency. The Avaya Ethernet Routing Switch 5500 demonstrated at least 1.5X less average cut-through latency for all frame sizes tested. In terms of standard deviation of latency, Avaya demonstrated almost 7X less standard deviation compared to the HP ProCurve 3400cl for 64-byte frames, while the Avaya switch demonstrated 1.4X less standard deviation with 1,518-byte frames. The latency results show that the Avaya Ethernet Routing Switch 5500 offered superior performance in terms of latency and standard deviation of latency compared to the HP ProCurve 3400cl and Cisco Catalyst 3750G.

Stack Resiliency: Single Unit Failure

The resiliency of an eight-unit Avaya switch stack was compared to the Cisco Catalyst 3750G and the HP ProCurve 3400cl switches in a similar resilient stacking network configuration. The 202 GbE switch ports in the stack were distributed as 162 GbE ports in VLAN 1 and 40 GbE ports inVLAN 2, and the corresponding input traffic consisted of 64-byte frames at approximately 240 million fps into VLAN 1 and 46 million fps into VLAN 2. Switch failure was introduced in VLAN 2. Test results show that Avaya exhibited 2X to 10X the maximum frame forwarding rate compared to similarly configured Cisco and HP stacks. Avaya achieved a frame forwarding rate as high as 240 million fps, while HP achieved 110 million fps and Cisco fared the worst achieving a frame forwarding rate of just 24 million fps. (See Figure 3.)

Standalone Switch Performance

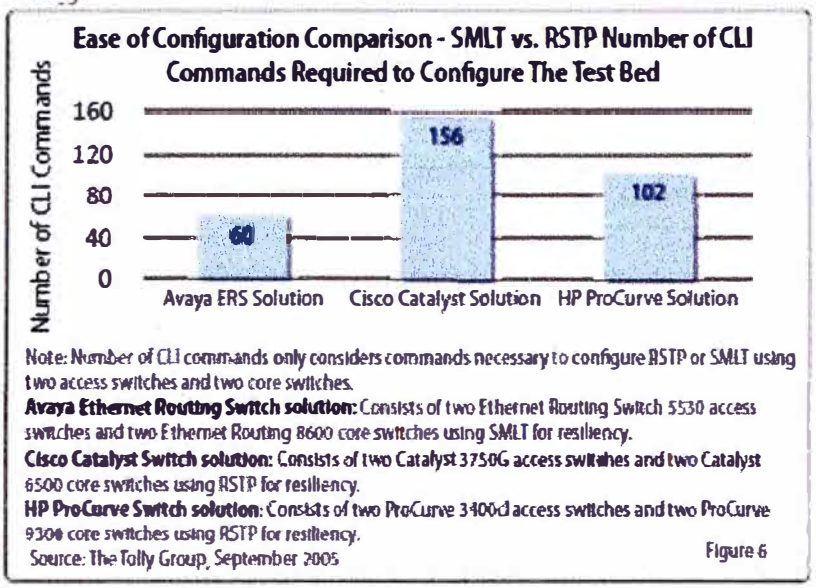
The Avaya Ethernet Routing Switch 5510-48T was compared with Cisco Catalyst 3750G-48TS and HP ProCurve 3400cl-48G switches in terms of Layer 2 zero-loss throughput, average cut-through latency and standard deviation of latency.

Results show that both Avaya and HP switches achieved 100% of the maximum theoretical throughput while handling Layer 2 test traffic consisting of 64-, 512- and 1,518-byte frames transmitted across 48 ports in a port-to-port configuration. Cisco could only achieve throughput of 55% of the theoretical maximum for 64-byte frames, 61% for 512- byte frames and 62% for 1,518-byte frames.

The Tolly Group also used the throughput results to calculate a cost-per-megabit of throughput for the three switches. This is done by dividing the switch price by the zero-loss throughput achieved. The Avaya switch offered the lowest cost/MB of throughput at just under $90, while Cisco offered the highest with a cost/MB of throughput exceeding $300. (See Figure 4.)

RSTP vs SMLT Performance

Tolly Group engineers tested the failover times of the Rapid Spanning Tree Protocol (RSTP) and Avaya's Split Multi-Link Trunking (SMLT) technologies in the event of a link failure and a switch failure. Avaya's solution



Note: Number of CLI commands only considers commands necessary to configure RSTP or SMLT using two access switches and two core switches.

Avaya Ethernet Routing Switch solution: Consists of two Ethernet Routing Switch 5530 access switches and two Ethernet Routing 8600 core switches using SMLT for resiliency.

Cisco Catalyst Switch solution: Consists of two Catalyst 3750G access switches and two Catalyst 6500 core switches using RSTP for resiliency.

HP ProCurve Switch solution: Consists of two ProCurve 3400cl access switches and two ProCurve 9300 core switches using RSTP for resiliency.

Source: The Tolly Group, September 2005

Figure 6

consisted of Ethernet Routing Switch 5530 and 8600 switches implementing SMLT, while Cisco's solution of Catalyst 3750G and 6513 switches, and HP's solution using ProCurve 3400cl and 9403 switches, both implemented RSTP.

Tests show that Avaya's Ethernet Routing Switch 8600 and 5530 solution using SMLT demonstrated the fastest network failover time in the event of a link or switch failure. In the event of a link failure, Avaya's solution using SMLT failed-over in 0.5 seconds while Cisco's solution took 1.7 seconds and HP's solution took 3.1 seconds. (See Figure 5.)

In the event of a switch failure, Avaya's solution using SMLT again failed-over in 0.5 seconds, while Cisco's solution using RSTP failed-over in 3.7 seconds, and HP's solution using RSTP failed- over in 4.9 seconds.

This shows that even with the same network topology, the Avaya SMLT implementation achieved significantly faster fail-over times compared to the RSTP implementations offered by Cisco and HP.

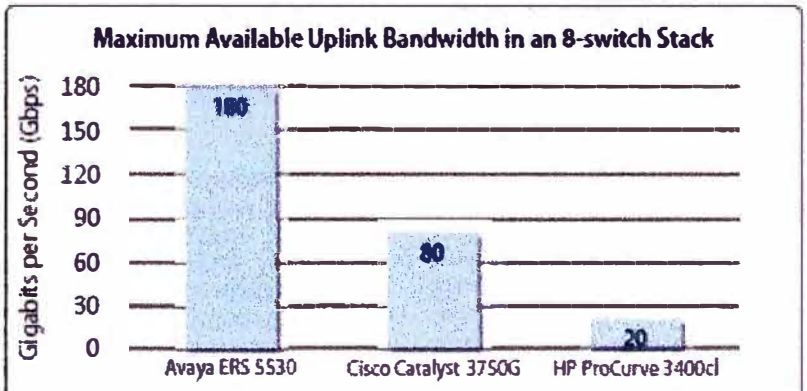
Ease of Use - Number of Commands to Configure RSTP vs. SMLT

The engineers counted the number of CLI commands required to configure the switches in the test bed for SMLT versus RSTP. The result showed that Avaya's test bed consisting of two Ethernet Routing Switch 5530 access switches and two Ethernet Routing Switch 8600 core switches required a total of 60 commands to configure SMLT. (See Figure 6.) In comparison, HP's test bed consisting of two ProCurve 3400cl access switches and two ProCurve 9300 core switches needed 102 commands to configure RSTP. Cisco's test bed consisting of two Catalyst 3750G access switches and two Catalyst 6500 core switches needed 156 commands to configure RSTP.

This shows that Avaya's SMLT implementation requires less number of CLI commands to configure the test bed compared to HP and Cisco's implementation of RSTP.

Highest Available Uplink Bandwidth

Avaya Ethernet Routing Switch 5530 and Cisco Catalyst 3750G switches support dedicated uplink connections in addition to dedicated stacking connections, where as the HP ProCurve 3400cl does not have dedicated stacking ports. This necessitates using an optional 10-GbE module with two

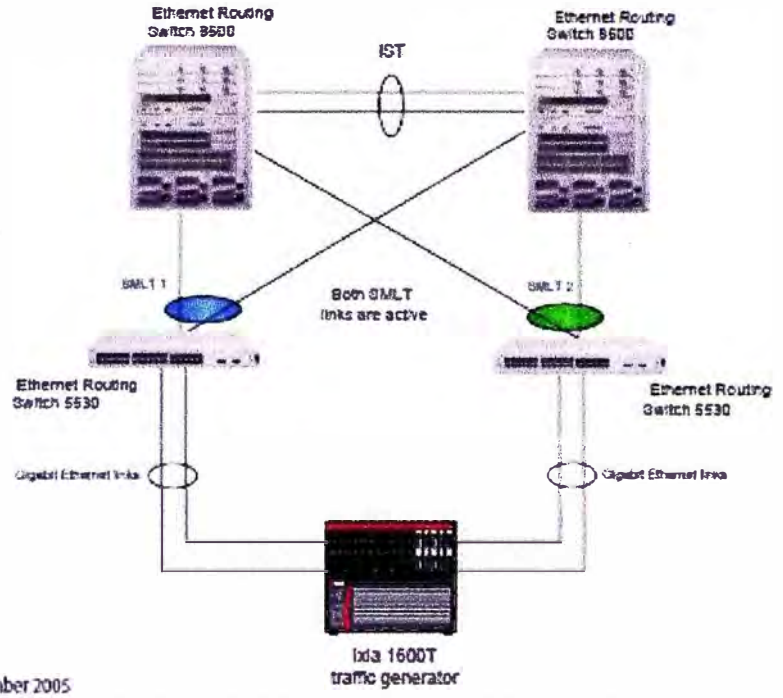


Each Avaya Ethernet Routing Switch 5530 switch has two 10-GbE ports for uplink.
Each Cisco Catalyst 3750G switch has one 10-GbE port for uplink
Each HP ProCurve switch has two 10-GbE ports on the optional 10-GbE module, but only two 10-GbE ports available for uplink in an 8- switch stack, as the rest of the 10-GbE ports are used for stacking purposes due to lack of dedicated stacking ports.

Source: The Tolly Group, September 2005

Figure 7

Test Bed for RSTP vs. SMLT Failover Time Comparison Showing Avaya Ethernet Routing Switch Solution Using SMLT



Source: Tolly, September 2005

Figure 8

Systems Under Test

Vendor	Product	Software
Avaya	ERS 5510-48T	4.2.0.004
	ERS 5510-24T	
	ERS 5520-48T-PWR	
	ERS 5530-24TFD	
	ERS 8600	4.0.1.0
Cisco Systems	Catalyst 3750G-48TS	12.2 (25) SEB1
	Catalyst 3750G-48PS	
	Catalyst 3750G-24TS	
	Catalyst 3750G-16TD-S	
	Catalyst 6500	12.2 (18) SXD5
HP ProCurve	ProCurve 3400cl-48G	M.08.66
	ProCurve 3400cl-24G	
	ProCurve 9304M	07.8.00aT53

Source: Tolly, September 2005

Figure 9

10-GbE ports for stacking connections on the HP ProCurve 3400cl. In an eight switch stack configuration, the Ethernet Routing Switch 5530 solution has 16

10-GbE links available for uplink connections, the Cisco Catalyst 3750G solution has eight 10-GbE links available for uplink, while the HP ProCurve 3400cl solution only has two 10-GbE links available for uplink. This means that in an eight switch stack, the Avaya switch has 160 Gbps of maximum available uplink bandwidth compared to 80 Gbps for the Cisco and 20 Gbps for the HP devices. (See Figure 7.) This shows that Avaya's Ethernet Routing Switch 5530 solution offers the highest uplink bandwidth among the devices tested.

Test Configuration and Methodology

For performance tests, The Tolly Group tested Avaya Ethernet Routing Switch 5000 series stackable switches (models 5510, 5520 and 5530) against Cisco Catalyst 3750G series switches and HP ProCurve 3400cl series switches. According to Avaya, all the switches were tested with production software generally available to the customer base. (See Figure 9.)

For evaluating the Layer 2 stack performance of the DUTs in an eight- unit high stack, engineers tested the frame forwarding rate, average cut-through

BIBLIOGRAFIA

1. "Nortel Ethernet Routing Switch Solution" – Nortel Press, 2008
2. "Nortel Ethernet Routing Switch 5000 Series Configuration" – Nortel Networks, 2008
3. "Building Cisco Multilayer Switched Network" – Cisco Press, 2000
4. "CCNP BCMSN Official Exam Certification Guide, 4th Edition" – Cisco Press, 2007
5. "Designing Cisco Network Service Architectures, 2nd Edition" – Cisco Press, 2009
6. "Understanding Rapid Spanning Tree Protocol (802.1w)" – Articulo Técnico
http://www.cisco.com/en/US/tech/tk389/tk621/technologies_white_paper09186a0080094cfa.shtml
7. "Spanning Tree Protocol Tutorial" – Articulo Técnico
http://tutorials.beginners.co.uk/spanning-tree-protocol-stp-.htm
8. "VRRP Tutorial" – Artículo Técnico
http://www.estoile.com/links/vrrp.htm
9. "Gigabit Campus Network Design" – Artículo Técnico

http://www.cisco.com/warp/public/cc/so/neso/lnso/cpso/gcnd_wp.pdf

10. SMLT (Split Multilink Trunking) – Internet Draft - July 07, 2008
http://tools.ietf.org/id/draft-lapuh-network-smlt-08.txt