



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería Mecánica Eléctrica

ARQUITECTURA BÁSICA DEL SWITCH FABRIC PARA MPLS (MULTI-PROTOCOL LABEL SWITCHING)

José David Crocker Fuentes

Asesorado por el Ing. MsEE. PhD. Enrique Edmundo Ruiz Carballo

Guatemala, Agosto de 2006

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**ARQUITECTURA BÁSICA DEL SWITCH FABRIC PARA MPLS
(MULTI-PROTOCOL LABEL SWITCHING)**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

JOSÉ DAVID CROCKER FUENTES

ASESORADO POR EL INGENIERO ENRIQUE EDMUNDO RUIZ CARBALLO

AL CONFERÍRSELE EL TÍTULO DE
INGENIERO EN ELECTRÓNICA

GUATEMALA, AGOSTO DE 2006

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Murphy Olympo Paiz Recinos
VOCAL I	Inga. Glenda Patricia García Soria
VOCAL II	Lic. Amahán Sánchez Álvarez
VOCAL III	Ing. Julio David Galicia Celada
VOCAL IV	Br. Kenneth Issur Estrada Ruiz
VOCAL V	Br. Elisa Yazminda Vides Leiva
SECRETARIA	Inga. Marcia Ivonne Véliz Vargas

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Murphy Olympo Paiz Recinos
EXAMINADOR	Ing. Enrique Edmundo Ruiz Carballo
EXAMINADOR	Ing. Guillermo Antonio Puente Romero
EXAMINADOR	Ing. Armando Alonzo Rivera Carrillo
SECRETARIA	Inga. Marcia Ivonne Véliz Vargas

HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

ARQUITECTURA BÁSICA DEL SWITCH FABRIC PARA MPLS (MULTI-PROTOCOL LABEL SWITCHING),

tema que me fuera asignado por la Dirección de la Escuela de Ingeniería de Mecánica Eléctrica, el 14 de febrero del 2006.

José David Crocker Fuentes

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA



FACULTAD DE INGENIERIA

Guatemala, 04 de Julio de 2006.

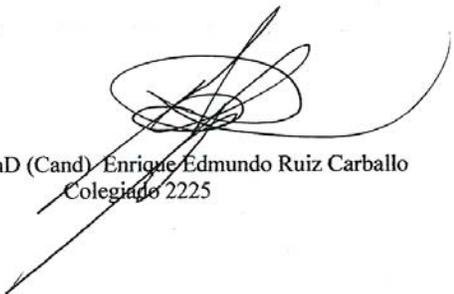
Ing. Julio Cesar Solares Peñate
Coordinador del Área de Electrónica
Escuela de Ingeniería Mecánica-Eléctrica
Facultad de Ingeniería
Universidad de San Carlos de Guatemala

Señor Coordinador:

Por este medio hago de su conocimiento de que he revisado el trabajo de graduación "**Arquitectura del Switch Fabric para MPLS (Multi-Protocol Label Switching)**", desarrollado por el estudiante **José David Crocker Fuentes**; con base a la revisión y corrección de dicho trabajo, considero que ha alcanzado los objetivos propuestos por lo cual el estudiante y mi persona nos hacemos responsables del contenido de este mismo.

Sin otro particular, me suscribo ante usted.

Atentamente,


Ing. MsEE. PhD (Cand) Enrique Edmundo Ruiz Carballo
Colegiado 2225

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA



Guatemala, 12 de julio 2006.

FACULTAD DE INGENIERIA

Señor Director
Ing. Mario Renato Escobedo Martínez
Escuela de Ingeniería Mecánica Eléctrica
Facultad de Ingeniería, USAC.

Señor Director:

Me permito dar aprobación al trabajo de Graduación titulado: **Arquitectura Básica del Switch Fabric para MPLS (Multi-Protocol Label Switching)**, desarrollado por el estudiante; José David Crocker Fuentes, por considerar que cumple con los requisitos establecidos para tal fin.

Sin otro particular, aprovecho la oportunidad para saludarle.

Atentamente,

ID Y ENSEÑAD A TODOS


Ing. Julio César Solares Peñate
Coordinador Área de Electrónica

JCSP/zo



UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA



FACULTAD DE INGENIERIA

El Director de la Escuela de Ingeniería Mecánica Eléctrica, después de conocer el dictamen del Asesor, con el Visto Bueno del Coordinador de Área, al trabajo de Graduación del estudiante; José David Crocker Fuentes titulado: **Arquitectura Básica del Switch-Fabric para MPLS (Multi-Protocol Label Switching)**, procede a la autorización del mismo.

A handwritten signature in black ink, appearing to read 'Mario Renato Escobedo Martínez', written over a circular stamp.

Ing. Mario Renato Escobedo Martínez

DIRECTOR



GUATEMALA, 18 DE JULIO 2,006.

ACTO QUE DEDICO A

Dios.

Mis Abuelos: Enma Sagastume de Crocker y Nicolás Fuentes.
Que en paz descansen.

AGRADECIMIENTOS A

Dios

Gracias, por darme la oportunidad de disfrutar la vida, por ser bueno y bendecirme tanto. No hubiera llegado al punto en donde estoy en mi vida, si no fuera por tu bendición.

Familia

A mis padres, Fredy y Delia: a quienes amo profundamente, por hacer de mí una persona recta y provechosa, por mostrarme de niño que el principio de la sabiduría es el temor a Jehová. Gracias por su ejemplo de lucha y esfuerzo, trabajo y determinación, lo que les ha permitido lograr cosas grandes, haciendo así que yo luche por alcanzar mayores metas. Gracias por apoyarme económicamente en mi formación como profesional. Y más que nada, gracias, por creer en mí. Los amo.

A mis abuelitos: Nicolás Fuentes y Enma Sagastume de Crocker, gracias por amarme tanto sólo por el simple hecho de ser yo, gracias por su apoyo, su guía, su amor, y su comprensión. Los veo en el cielo.

A mi hermana: gracias por estar conmigo siempre, por ser mi compañía, mi amiga, por verme como un ejemplo. Créeme que es muy halagador para mí que una persona tan especial, inteligente, tan buena y recta, me admire.

A mi tía Irma Fuentes Laparra: gracias por ser mi otra mamá. Te amo mucho.

A mi tía Helen Fuentes Laparra: gracias por interesarte en mi persona y estar conmigo cuando era un niño.

A la Familia Fuentes Laparra: gracias por quererme, apoyarme y consentirme tanto.

A mis tíos y primos en general: gracias por ser parte de mi vida.

Padrinos de Graduación

Al Dr. Enrique Ruiz Carballo; gracias por todo su apoyo en la elaboración de este trabajo de graduación, por ser uno de los mejores catedráticos de esta facultad, y por ser para mí un mentor digno de imitar.

Al Lic. Mario Bautista; tío, gracias por todo lo que me has apoyado siempre, por tus consejos, por tu amistad, por tu ejemplo y por consentirme tanto.

Profesores y Profesionales

Al Ing. Guillermo Puente, por su ejemplo tanto como profesional como empresario, gracias por su apoyo, y por el esfuerzo que pone para que sus alumnos aprendan al máximo.

A la Inga. Ingrid de Loukota, gracias por todo su apoyo, por sus consejos y por enseñarme que no hay que quejarse de la vida, sino que hay que enfrentarla y sacarle lo que más se pueda.

Al Ing. Juan David Alvarado: gracias por ayudarme a escoger el tema del presente trabajo de graduación, y por la información que me brindaste para el mismo.

Al Ing. Pedro Obregón Leche: gracias por todo tu apoyo durante mi carrera, por asesorarme siempre en qué era lo más conveniente hacer, y por explicarme tantas veces las cosas que no entendía.

Al Ing. Jesús Martínez: gracias por ser un ejemplo de excelencia, y por sus consejos.

Amigos

Patrick Allen, Claudia Castañeda, Carlos Menéndez, Luís Lira, Jaime García, Oscar Ponce, Fabio Steve Morales, Mario y Kelvin Silvestre, Eduardo Alvarado, Javier Alay, Jorge Monterroso, Ana Vásquez, Ricardo del Cid, Karla Morataya, Armando Gálvez, Luís Mejía, Carlos Rodríguez, Gerson Chamo, Luís García, Roland Márquez, Tatiana Vallejo, Mario Mérida, Diego de León, Andrés Reyes, Héctor Mendoza, Alfonso Alfaro, Edwin Morales, la Familia Allen y a la Familia Silvestre. Gracias por su apoyo y amistad.

Compañeros

Gracias por los desvelos y los buenos momentos que compartimos juntos en los proyectos; la verdad, sin ustedes no hubiera llegado hasta aquí.

Instituciones

Universidad de San Carlos de Guatemala. Facultad de Ingeniería. Escuela de Mecánica Eléctrica. Gracias por la buena educación que me brindaron, por el orgullo y el honor que me han dado de poder decir que aparte de ser Guatemalteco, también soy San Carlita.

Al Grupo Visión Estudiantil, por su apoyo y los grandes recuerdos que me dejaron de la Universidad.

A la Rama Estudiantil de la IEEE 2005, gracias por el apoyo y por todos los recuerdos de CONESIEE V.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
GLOSARIO	VII
RESUMEN	XI
OBJETIVOS	XIII
INTRODUCCIÓN	XV
1. MECANISMO DE SWITCHEO ESTÁTICO	1
1.1. Multiplexación	1
1.2. Mecanismos de switcheo	4
1.2.1. Circuito de switcheo	4
1.2.2. Circuito de switcheo multi-tasa.....	5
1.2.3. Circuitos de switcheo rápido.....	5
1.2.4. Switcheo de paquetes.....	6
1.2.5. Switcheo rápido de paquetes.....	8
1.3. Evolución del switch de paquetes	9
1.3.1. La primera arquitectura del switch.....	9
1.3.2. Switcheo de dirección descentralizada.....	9
1.3.3. El switch de rutas múltiples	11
1.4. Fundamentos de diseño del switch rápido de paquetes.....	11
1.5. Arquitectura del switch rápido de paquetes.....	14
1.5.1. Diagrama interno de switch de paquete rápido	14
1.5.2. Clasificación simple de diseños de switches	15
2. SWITCH FABRIC	17
2.1. Desempeño del switch fabric	19

2.2. Modelos de tráfico.....	21
2.3. El modelo de simulación	22
2.4. Switch fabric crossbar (de barra cruzada)	25
2.4.1. Rendimiento en estado de saturación	25
2.4.1.2. Media de retraso para tráfico en slots	27
2.4.2. Matrices de croconexión	29
2.4.2.1. Red Delta.....	29
2.4.2.2. Red Bannes.....	31
2.5. Tipos de switch fabric	34
2.5.1. Switch fabric de Space Domain	34
2.5.2. Switch fabric de tipo Time Domain	35
2.5.3. Switch fabric de memoria compartida.....	38
2.5.4. Switch fabric de crossbar con control central	39
2.5.5. Switch fabric de full matrix interconnect con control en las salidas.....	40
2.5.6. Switch fabric de full matrix interconnect con control en la entrada	41
2.5.7. Switch fabric de MIN con matriz de 8 X 8.....	41
2.5.8. Switch fabric de 16x16 tipo torus.....	42
2.5.9. Switch fabric de hiper cubo de 8x8.....	43
3. TENDENCIAS EN EL DISEÑO DE ROUTERS.....	45
3.1. Componentes de un Router.....	45
3.1.1. Puerto de entrada.....	46
3.1.2. El switching fabric:.....	47
3.1.3. Los puertos de salida	48
3.1.4. El procesador de routing.....	48
3.2. Asuntos de diseño	48
3.2.1. Backbone routers.....	49
3.2.2. Misión del router	51
3.2.3. Routers de acceso	53

3.3. Recientes avances y tendencias.....	54
3.3.1. Router de búsqueda de alta velocidad	54
3.3.1.1. Las técnicas orientadas en Hardware.....	56
3.3.1.2. Técnicas de Table Compaction	56
3.3.1.3. Técnica de Hashing (despedazar)	57
3.3.2. Avances en switch fabrics.....	58
3.3.3. Label Switch Router (LSR)	59
4. TECNOLOGÍA MPLS (MULTI-PROTOCOL LABEL SWITCHING)	61
4.1. El Camino hacia la convergencia de niveles: IP sobre ATM	61
4.2. Un paso más en la convergencia hacia IP: Conmutación IP	62
4.3. La Convergencia Real a MPLS.....	63
4.3.1. Ideas preconcebidas sobre El MPLS	63
4.3.2. Descripción funcional del MPLS	64
4.3.2.1. Funcionamiento del envío de paquetes en MPLS	64
4.3.2.2. Control de la información en MPLS.....	68
4.3.2.3. Funcionamiento global MPLS.....	72
4.4. Aplicaciones del MPLS.....	72
4.4.1. Ingeniería de tráfico.....	72
4.4.2. Clases de servicio (CoS).....	73
4.4.3. Redes privadas virtuales	74
CONCLUSIONES	77
RECOMENDACIONES	79
REFERENCIAS	81
BIBLIOGRAFÍA	83

ÍNDICE DE ILUSTRACIONES

FIGURAS

1	Multiplexación por división de tiempo	2
2	Ráfaga de información en comunicación estática	3
3	Arquitectura del switch de paquete con configuración en paralelo	10
4	Arquitectura del switch de paquete rápido	14
5	Topología del switch fabric	18
6	Rendimiento de un switch fabric crossbar en saturación	26
7	Media de retraso para switch de buffering de entrada para N grande	28
8	Efecto de introducir un by-pass de alineación de entrada, y buffering de salida a través de un switch fabric crossbar	29
9	Red delta de 64x6	30
10	Red benes de 64x64	32
11	Diagrama de un SF de tipo time domain	35
12	Diagrama de bloqueo de salida de un SF de tipo time domain	36
13	Diagrama de bloqueo interno de un SF de tipo time domain	36
14	Diagrama de un SF de memoria compartida	38
15	Diagrama de un SF de crossbar con control central	39
16	Diagrama de un SF de full matriz interconnect con control en las salidas	40
17	Diagrama de un SF de full matriz interconnect con control en la entrada	41
18	Diagrama de un SF de min con matriz de 8X8	41

19	Diagrama de un SF de 16X16 tipo torus	42
20	Diagrama de un SF de hiper cubo de 8X8	43
21	Composición interna del LSR	60
22	Ejemplo de red MPLS	65
23	Ejemplo de Forward Equivalent Class	66
24	Estructura de la cabecera en MPLS	68
25	Lambda Switch Capable	71
26	Red privada virtual, elaborada a base de la tecnología MPLS	75

TABLA

I	Parámetros de diseño del switch fabric	20
---	--	----

GLOSARIO

ATM	Modo de Transferencia Asíncrona. Fue un protocolo inicialmente diseñado para dar soporte a redes multi-servicio, puesto que permite integrar el tráfico procedente de distintas aplicaciones de una manera muy flexible.
BGP	Interdomain Routing Protocol, es el más comúnmente usado en Internet.
FEC	Es un conjunto de paquetes que se envían sobre el mismo camino a través de una red, aun cuando sus destinos finales sean diferentes.
Frame relay	Retransmisión de tramas. Protocolo conmutado estándar de la industria de la capa de enlace de datos, que maneja varios circuitos virtuales mediante una forma de encapsulación HDLC, entre dispositivos conectados.
IEEE	Instituto de Ingeniería Eléctrica y Electrónica. Organización profesional cuyas actividades incluyen el desarrollo de estándares de comunicaciones y redes. Los estándares de LAN de IEEE son los más importantes para las LAN de la actualidad.

IETF	Organización que se ocupa del diseño de la Internet. Sus siglas significan (Internet Engineering Task Force). Esta fuerza está compuesta por más de 80 grupos de trabajo responsables por el desarrollo de estándares de Internet. IETF opera bajo el auspicio de ISOC.
IGP	Protocolo de gateway interior. Protocolo de Internet que se utiliza para intercambiar información de enrutamiento dentro de un sistema autónomo.
JITTER	Fluctuaciones causadas por perturbaciones oscilatorias.
LAN	Red de Área Local. Red de datos de alta velocidad y bajo nivel de error que cubre un área geográfica relativamente pequeña.
LDP	Label Distribution Protocol
LSP	(Label Switched Paths) Conocidos como "camino virtuales" funcionalmente equivalentes a los PVCs de ATM y Frame Relay.
LSR	(Label Switching Router) Un LSR no es sino un router especializado en el envío de paquetes etiquetados por MPLS.
MESH	Tipo de configuración de la red del switch fabric.
PPP	Protocolo Punto a Punto. Suministra conexiones router a router y host a red a través de circuitos síncronos y asíncronos.
PSC	Capacidad de switcheo de paquetes.

PVC	Circuitos Virtuales Permanentes, que se establecen sobre la topología física de la red ATM.
QoS:	Calidad de Servicio. Es un set métrico usado para medir la calidad de disponibilidad de transmisión y de servicio de cualquier sistema dado de transmisión.
RSVP	Protocolo de reserva de recursos. Protocolo que hace posible la reserva de recursos a través de una red IP. Las aplicaciones que se ejecutan en los sistemas finales IP pueden usar RSVP para indicarle a los otros nodos la naturaleza de los flujos de paquetes que desean recibir. RSPV depende de IPv6. También denominado protocolo de configuración de reserva de recursos.
TDM	Multiplexación por división de tiempo.
TRUNKS	Tramas de red, en las redes LAN.
VC	Camino Virtual. Es un camino lógico ideado para asegurar comunicación fidedigna entre dos dispositivos en una red.
VCI	Identificador de un camino virtual.
VPI	Es el equivalente de un VCI en la tecnología ATM.
WRAPPING	Envolver paquetes de información

RESUMEN

Este trabajo estudia los diferentes mecanismos de switcheo, así como la implementación de éstos en los routers, y cómo es que aún hoy en día se utilizan en las tecnologías modernas.

Antes de entrar en detalles de cada uno de estos dispositivos, se hace mención a los dos tipos de comunicación sobre los cuales pueden operar. La comunicación síncrona y asíncrona. La primera opera con multiplexación TDM (Multiplexación por división de Tiempo), y la segunda que obedece a la multiplexación estática.

La función de los mecanismos de switcheo a grandes rasgos es la de unir una fuente con un destino, y éstos no son más que la conjunción de switch fabrics, puertos de entrada y de salida y una unidad central de proceso. Difieren uno de otro, más que todo en la forma en que la información es procesada en cada uno de ellos. En el capítulo 1 se habla de: el circuito de switcheo, el circuito de switcheo multi-tasa, el circuito de switcheo rápido, el switch de paquetes y el switch de paquetes rápidos, además de cómo son éstos diseñados.

El switch fabric, es una red de switch la cual puede ser configurada de distintas formas, y cómo estas tienen bastante ingerencia en la eficiencia de los mecanismos de switcheo. El capítulo 2 está dedicado a este dispositivo, se estudian tres configuraciones comunes, las cuales son Red Delta, Red Bannes y Red Crossbar. Y se observa detenidamente cómo varía la eficiencia del switch fabric al trabajar en cada una de éstas, cuando se ve sometido en los tres casos a las mismas condiciones. Como se verá, estas comparaciones se hacen a base de las simulaciones realizadas por el Dr. Newman [3].

El capítulo 3 se centra en el estudio de los routers, y se ve como éste está constituido por un switch de paquetes rápido, y se hace un estudio más a fondo de cada uno de sus componentes. Se mencionan los distintos routers que existen, así como las nuevas tendencias en el diseño de estos. Por último se habla del LSR (Label Switching Router), el cual es un router que trabaja a base de etiquetas y es un dispositivo de mucha importancia en la tecnología MPLS.

Por último, en el capítulo 4 se ve todo lo referente a la tecnología MPLS (Multi-Protocol Label Switching). Como por ejemplo, cuáles fueron sus razones de creación, el porqué se está emigrando a ella y es la base para el Internet del siglo XXI, y cuál es el papel de los LSR, en ella.

OBJETIVOS

❖ GENERAL

Generar información sobre nuevas tecnologías en enlaces de comunicación de alta capacidad.

❖ ESPECÍFICOS

1. Seleccionar apropiadamente los parámetros de diseños de los mecanismos de Switcheo, para caracterizar el efecto en su desempeño.
2. Inducir los conceptos básicos de Switch Fabric, y su ingerencia en la arquitectura de los Routers.
3. Conocer la labor que realizan los Label Switching Routers (LSR), en las redes de comunicación.
4. Mostrar a MPLS, como la arquitectura base de las nuevas aplicaciones de redes, para ofrecer diferentes niveles de servicio de Internet con mayor fiabilidad.

INTRODUCCIÓN

La demanda sostenida de nuevos servicios sofisticados, así como el aumento del uso del Internet, necesita de cambios fundamentales en la tecnología utilizada con relación a las prácticas desarrolladas a mitad de los años 90.

Actualmente existen nuevas tecnologías de transmisión usando fibra óptica, tales como, Dense Wavelength División Multiplexing (DWDM), que proporciona una alternativa al ATM para realizar múltiples servicios sobre circuitos individuales. Así mismo, una nueva generación de routers con funciones especializadas en el transporte de paquetes en el núcleo de las redes, los cuales son conocidos más comúnmente como LSR (Label switching Routers), los cuales son capaces de realizar el enrutamiento a base de etiquetas. Este nuevo tipo de routers está desplazando a los tradicionales routers de ATM.

Estos routers no sólo se utilizan para llevar paquetes a gran velocidad, sino que también se ocupan de asuntos de menor importancia, como programar soporte para los servicios diferenciales, tecnologías heterogéneas de enlace, y retraso en la compatibilidad con un gran rango de formatos de paquetes y protocolos de routeos.

En la construcción de los nuevos cimientos para el Internet del presente siglo, se considera fundamental, la reciente aparición de una nueva arquitectura de red conocida como Multi-protocol Label Switching (MPLS), la cual surgió en el seno de IETF (Internet Engineering Task Force).

Se persigue describir los más recientes adelantos de la tecnología de los MPLS, que consiste en la clara separación entre las funciones de routing (control de la

información sobre la topología y tráfico en la red) de las funciones de forwarding (envío de datos entre elementos de la red). Como preámbulo a lo anterior se hace un estudio, a los mecanismos de switcheo estático. Observando que el más robusto y eficiente es el switch de paquete rápido (fast packate switch); sabiendo que a base de este dispositivo de switcheo es que se elaboran los routers genéricos, y este fue el que dio la pauta para el desarrollo de estos tipos de dispositivos hasta llegar al anteriormente mencionado LSR.

Ente las partes del Switch de Paquete rápido encontramos al Switch fabric, el cual como su nombre lo deja ver, es un arreglo de switch, que se puede comparar a la tela, el cual puede trabajar de manera bidireccional. A través de él, es donde se realiza el intercambio de elementos, confiando en su habilidad, para el intercambio y control de la unidad central de proceso que también se encuentra conectada al switch.

Como las características y configuración de este dispositivo, son las que determinan el rendimiento de un switch de paquete rápido y por ende el de los routers; el capítulo 2, se dedica al estudio de éste.

Se hubiera deseado observar el rendimiento de estos dispositivos de una manera palpable, pero, por el hecho de no tener acceso a un router convencional, en el cual se podría observar el funcionamiento de un switch de paquete rápido, y menos a un LSR, a través del cual se podría estudiar más a fondo el funcionamiento de la tecnología de MPLS, se optó por revisar estudios anteriores del tema, buscando dar una exploración en profundidad de éste.

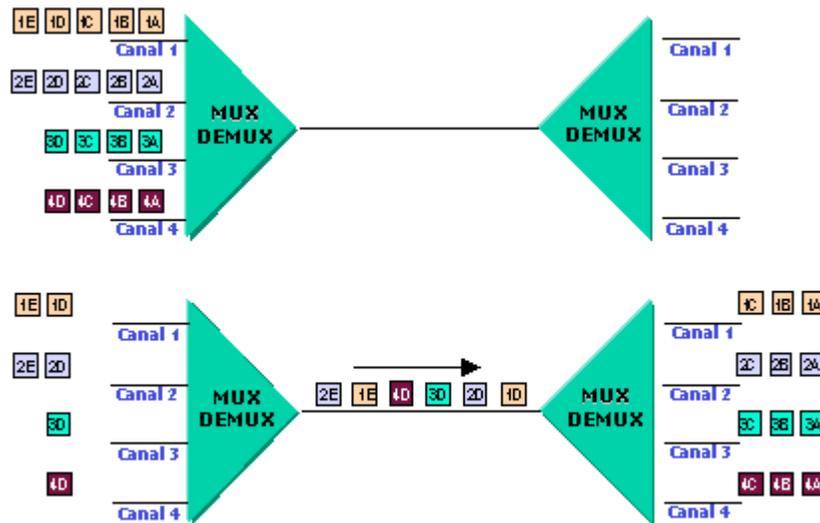
1. MECANISMO DE SWITCHEO ESTÁTICO

1.1. Multiplexación

Es la técnica por medio de la cual, es posible que la información de dos o más canales, viajen por un medio en común. Las redes telefónicas funcionan bajo este principio, pues a través de una sola línea de banda ancha, viajan varias conversaciones telefónicas. TDM es la cual es la multiplexación por división por tiempo y es una de las técnicas de multiplexación mas utilizadas. Su primera función al ser desarrollada fue eliminar el ruido y filtraje que tenia el método de FDM (Multiplexación por división de frecuencia). Pero se explota al máximo cuando se utiliza en comunicaciones digitales. En esta técnica, cada fuente tiene a su disposición toda la capacidad del canal por un determinado intervalo de tiempo. Aquí lo que se hace es que se acomoda, la información que lleva cada dispositivo conectado a un puerto, tomando en cuenta que el rango de bits para todos los dispositivos no debería exceder el rango de bits por segundo de la línea de salida. Lo anterior se logra utilizando normas de compresión.

En TDM se empaqueta un conjunto de información (tramas de bits) de diferentes fuentes en un solo canal de comunicación en tiempos (muy cortos) diferentes. En el otro extremo estas tramas son otra vez reensambladas y llevadas a su respectivo canal, ver figura 1. Los multiplexadores TDM como manejan tramas de bits son capaces además de comprimir la información al eliminar redundancias en los paquetes, muy útil en el caso de aplicaciones de voz.

Figura 1. Multiplexación por División de Tiempo.



Fuente: www.redesdecomunicacion.com/digitales/psk

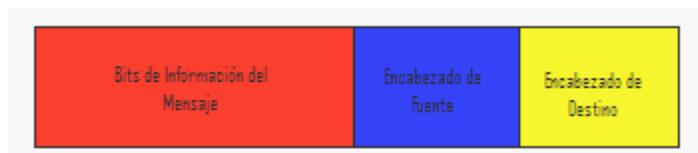
En la técnica de Multiplexación por división de tiempo, el canal cuenta con ancho de banda fijo y se trabaja a una tasa de bits determinada. Esta tiene una deficiencia en ancho de banda variable a la hora de trabajar con fuentes asíncronas. Una fuente asíncrona es aquella que su tasa de bits es ampliamente variable. Para solventar esta deficiencia se introdujo la Multiplexación estática.

En la Multiplexación estática, al igual que en TDM la información se transmite en ráfagas de información, cada fuente recibe gran capacidad de transmisión conforme esta lo necesite, estas ráfagas poseen distintas longitudes de información, y el ancho de banda ya no es fijo.

En la Multiplexación estática, una fuente asíncrona solo ocupa el canal, cuando esta tiene información para enviar, en cambio, en TDM cuando una fuente asíncrona esta en su periodo de tiempo de transmisión y esta no posee información para enviar, el canal permanece bloqueado para las otras fuentes.

En la comunicación estática, a diferencia de TDM, no se sabe en que momento transmitirá cada fuente, por lo que esta información adicional debe enviarse en cada ráfaga de información para identificar el canal, por lo tanto estas son mucho mas largas que las ráfagas de información en TDM, ver figura 2.

Figura 2. Ráfaga de Información en comunicación Estática.



En TDM convencional, es posible que se trabaje al 100% de la capacidad del medio indefinidamente, teniendo como resultado una estabilidad baja y un retraso que puede ser determinado. En cambio con Multiplexación estática para un periodo corto de operación no es posible que se exceda la máxima capacidad del canal pues esto produce un gran retraso, grandes errores, inestabilidad alta y pérdida de información. Por lo que en comunicación estática se debe trabajar al 80% de la capacidad del canal. A pesar de estas deficiencias la comunicación estática, es más eficiente al momento de trabajar con fuentes asíncronas de información. En esta el acceso al medio esta dividido en celdas, y cada celda esta ubicada para conducir el tráfico de una fuente asíncrona en demanda, cada celda cuenta con un encabezado para identificar la fuente.

1.2. Mecanismos de switcheo

La Multiplexación, permite el enlace de comunicación de varios canales, sobre un medio en común, conectando una fuente y un destino. Para que sea posible realizar esta comunicación entre la fuente y destino se ve involucrado un dispositivo de switcheo, el cual es el encargado de habilitar o bloquear determinada fuente de información. Dependiendo la técnica de Multiplexación, se utiliza un determinado dispositivo de switcheo, yendo desde el TDM convencional (la cual es una transmisión sincronía), hasta el switch de paquete que es el más comúnmente utilizado en el Multiplexación estática (la cual es una transmisión asíncrona).

1.2.1. Circuito de switcheo

Switcheo esta relacionado con la conexión. Este en si se basa en unir una fuente y un destino a través de un mecanismo de switcheo, puede ser unidireccional o bidireccional. Se pueden realizar conexiones punto a punto o multipunto: La conexión punto a punto conecta una fuente con un solo destino, y una conexión multipunto es aquella en la cual una solo fuente se conecta con muchos destinos. Y estas pueden estar en el dominio del tiempo o en el dominio del espacio. Cuando los canales llegan al interruptor por distintos medios de entrada, se dice la conexión se encuentra en el dominio del espacio y la conexión es ubicada en un camino físico a través de interruptor, y cuando los canales llegan en un TDM a un solo medio de entrada se dice que la conexión se encuentra en el dominio del tiempo, y la conexión se realiza por ranuras de tiempo constantes. En las conexiones en el dominio del tiempo es necesario contar con una unidad de memoria entre la trama con el fin de que esta cree un retardo del tamaño de la longitud de la trama de conexión.

La red no hace el intento de corregir errores, debido que el interruptor se presenta como invisible a la conexión. También es conveniente tener en cuenta que una vez que conexión es establecida esta no soportara degradación de retraso o de ancho de banda debido a recargo de información.

1.2.2. Circuito de switcheo multi-tasa

Este dispositivo de switcheo, puede trabajar con canales que tienen distintos anchos de banda, los cuales a su vez deben permanecer constantes. Por lo que la mayoría de veces trabajar con este tipo de interruptor puede ser muy complicado, en aspectos de diseño y control.

Su mayor desventaja es que este dispositivo no maneja las fuentes asíncronas de una manera más eficiente que un switch convencional. Debido a que no ofrece que el retraso entre ranuras de tiempo en la red se mantenga constante a la hora de sincronizar todas las tasas básicas de las que se encuentra formado un canal multi-tasa.

1.2.3. Circuitos de switcheo rápido

El circuito de switcheo rápido, se diseñó para trabajar con fuentes asíncronas. Siendo este un mecanismo de conexión con respuesta rápida, se puede establecer una conexión por cada ráfaga de información de una fuente asíncrona, bloquearla en el instante que esta termine de pasar, logrando de esta manera una mejor utilización del ancho de banda pues solo será utilizado por las fuentes en funcionamiento.

Un circuito virtual es una conexión lógica entre fuente y destino que desintegra el concepto de la conexión de ancho de banda. Al realizarse una llamada se establece un circuito virtual en un switcheo asíncrono, pero el ancho de banda de este solo es utilizado cuando aparece una ráfaga de información la cual se transmite a la misma tasa a la cual es recibida.

Al trabajar de esta manera se le introduce una memoria intermedia, la cual es utilizada cuando, una ráfaga de información llega y el ancho de banda no se encuentra disponible en ese momento, su función es retrasar la información hasta que el ancho de banda se vuelva disponible.

1.2.4. Switcheo de paquetes

En el switch de paquete rápido el ancho de banda, esta disponible para cada ráfaga de información, en vez de estar disponible para cada canal. Como su nombre lo indica este switch trabaja a base de paquetes, cada paquete esta constituido por el largo máximo de la ráfaga de información y los campos de control los cuales indican quien es la fuente y quien es el destino. Los requisitos del buffer del switch y los retrasos permitidos por paquete son los que limitan la longitud del paquete. Dichos paquetes se guardan dentro del switch y no son transmitidos hasta que son completamente recibidos.

Al trabajar con dichos interruptores la operación puede ser en dos modalidades: La conexión orientada y la conexiónes. En la conexión orientada todos lo circuitos forman parte de un circuito virtual, es decir trabajan sobre la misma ruta, y lo primero que se debe hacer para trabajar en este tipo de conexión es establecer el circuito virtual, con lo que se logra que el control de flujo pueda ser mas eficiente y selectivo, pues el procesamiento de paquetes subsiguientes puede ser simplificado. En la conexión conexiónes, los datagramas (los paquetes como son llamado aquí), viajan por distintas rutas, y no se puede asegurar que los paquetes arriben a su destino en el mismo orden en el que salieron. Este tipo de conexión es menos selectiva, pero menos vulnerable a problemas de nodo, por lo que es mucho más eficiente para el tráfico de patrones cambiantes, pero esto se da al precio de que se necesita mucho más procesamiento para cada paquete.

Entre la ventajas de utilizar el circuito de switcheo rápido, encontramos que es muy eficiente en a la hora de trabajar con fuentes asíncronas, no consume ancho de banda cuando la fuente no esta funcionando, puede soportar distintas demandas de ancho de banda, y soportan conexiones fuente y destino en diferentes tasas de datos. Pero por el gran procesamiento que llevan los paquetes en cada switch hace que la conexión padezca de grandes retrasos y alta variación de retraso, por lo que la capacidad máxima del switch de paquete es menor que la de los switchs convencionales.

1.2.5. Switcheo rápido de paquetes

El switch de paquete rápido trata de mantener las propiedades del switch de paquete, pero trata de corregir las deficiencias de retraso, logrando de esta manera que aumente la capacidad máxima switch. Hoy en día los crecientes aumentos en el desarrollo de transmisión por fibra óptica, ha hecho que los enlaces soporten mayores anchos de banda y sean más eficientes pues cuenta con tasas de errores bajas, debido a esto es que ya no es necesario el control de errores en cada switch, en toda la trayectoria del enlace, debido al retardo y a los requerimientos de los buffer. Los switch de paquete rápido trabajan en conexión orientada. Un circuito virtual es establecido a través de la red, y los encabezados cortos de paquete pueden usarse para distinguir entre cada uno de los circuitos virtuales multiplexados sobre un solo enlace, además también es posible realizarse por medio de tablas el enrutamiento de cada paquete.

Para el manejo de tráfico asíncrono eficiente, es necesario utilizar un switcheo estático, por lo que utilizamos el switch rápido y el switch rápido de paquete. Su diferencia se basa en que el switch de paquete rápido de paquetes requiere un encabezado por paquete, mientras que el circuito de switcheo rápido solo necesita un encabezado por cada ráfaga de información.

Se sabe por experiencia que en el switcheo estático, un sobre flujo ocurre cuando la información que entra en el switch es mayor a la capacidad de transmisión produciendo esta un retardo, y o pérdida de información. En el switch de paquete rápido estos efectos se pueden extender a un grupo determinado de llamadas, además de que el ancho de banda puede variar para fuentes individuales, por lo que presenta una mayor flexibilidad y desempeño en el manejo de tráfico de los datos en retransmisiones, usando un ancho de banda entero en vez de los canales de banda estrecha del circuito de switcheo rápido.

1.3. Evolución del switch de paquetes

1.3.1. La primera arquitectura del switch

El switch de paquete, fue diseñado con solo una unidad de proceso, para manejar las funciones de protocolo de conmutación de paquetes, enrutamiento. Esta unidad de proceso y la complejidad del protocolo utilizado son los entes que limitan el rendimiento de switch. Hoy en día el control de flujo, la detección y corrección de errores se hace a través del protocolo de bajo nivel, y esto es posible gracias a que se puede asignar una unidad de proceso por cada puerto del switch, pues cada uno de estos realiza estas operaciones en forma independiente y el procesador central se dedica únicamente a funciones de protocolo de nivel superior. Este cambio hubiera sido un éxito si no se hubiera dejado a la unidad de central de proceso (procesador central) como eje de en la interconexión de todos lo puertos de switch, pues se comportaba como un cuello de botella.

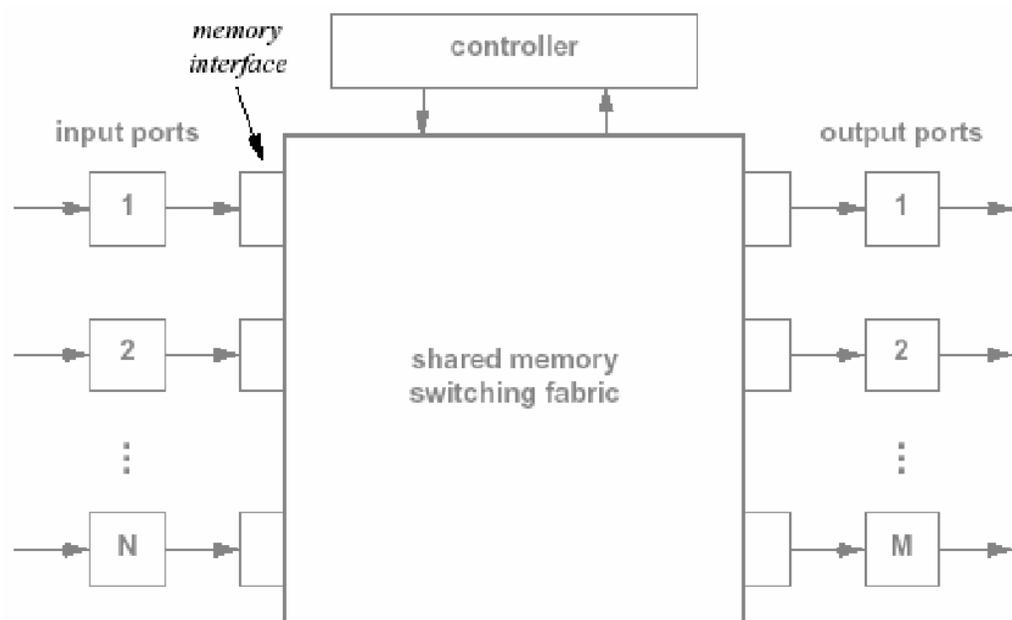
1.3.2. Switcheo de dirección descentralizada

Como se vio en la sección anterior, las deficiencias para alcanzar la máxima capacidad del switch son debido a la unidad central de proceso, por lo que se opto por aislar esta del camino de transmisión de los paquetes. Debido a lo anterior la unidad central de proceso tomo una función de supervisor, mientras las funciones de conexión de puertos son realizadas por un medio de interconexión insertado que trabaja en conjunto con los periféricos y los puertos inteligentes del switch. La tasa a la cual trabajan estos en conjunto con el ancho de banda son los que determinan el rendimiento del switch de paquetes.

El switch de paquetes, utilizaba una unidad de memoria para realizar las interconexiones de camino, y cada puerto podía acceder directamente a la unidad de memoria. Los diseños de este switch estaban basados en varias topologías, bus serial, bus en paralelo o de anillo. Como vemos la arquitectura de un switch de paquete se asemeja grandemente a la de una red de área local (LAN).

Una configuración en paralelo puede ser soportada por el switch de paquete y también utiliza una unidad de memoria compartida para la interconexión de los puertos con la unidad central de proceso como se aprecia en la figura 3.

Figura 3. Arquitectura de Switch de paquete con configuración en Paralelo



Fuente: Tilman Wolf de la Universidad Amherts de Massachussets.

1.3.3. El switch de rutas múltiples

Como hemos observado en la mayoría de diseños de switch, el ancho de banda toma un papel sumamente importante en el hecho de que este pueda trabajar a su capacidad máxima. Para poder solventar esta contrariedad se optó por desarrollar un switch que pudiera realizar la interconexión por múltiples rutas, logrando de esta manera que todas las demandas de comunicación en los puertos de switcheo sean respondidas. Usando la tecnología de rutas múltiples se puede obtener una mayor capacidad, pues el ancho de banda puede irse aumentando entre más se aumente el número de switch usados en el arreglo de las vías de interconexión y por ende este ya no es una limitación para la capacidad del switch. Esta mejora en capacidad nos brinda beneficios en el hecho de que un switch con esta característica no precisa de la velocidad, pero sí de alta tecnología que tienden a encarecer el dispositivo.

Esta arquitectura de rutas múltiples puede ser usada con el circuito de switcheo convencional, el switch de paquete.

1.4. Fundamentos de diseño del switch rápido de paquetes

El switch de paquete rápido, funciona bajo muchos conceptos de transferencia asíncrona que mencionamos en un inicio. En otras palabras su comportamiento es muy similar al de un dispositivo que funciona bajo Multiplexación estática. Aunque es conveniente mencionar que algunos casos los diseños de estos switch son solo capaces de manipular paquetes que tienen longitudes estrictamente fijas. Cuando se trabaja con exclusivamente con estos diseños, el ancho de banda, tanto en la línea de entrada como en la de salida es dividido en ranuras de tiempo, como en TDM, y ciertas ranuras llevan información y otras van vacías o son llenadas con ciertos patrones de sincronización.

Hablando de manera general del switch de paquete rápido se puede decir que esta compuesto por un conjunto de líneas de entradas y salida, cada una de ellas con sus respectivos puertos de entrada y salida los cuales se encuentran interconectados por un switch fabric. En el switch de paquete rápido la información es trasladada por paquetes o directamente por el switch fabric y esta transferencia es controlada por la unidad central de proceso que también se encuentra conectada al switch. En el switch de paquete rápido se realizan interconexiones de línea de una fuente a un destino de manera bidireccional.

Cuando se trabaja modo asíncrono, cada paquete de información contiene un encabezado o etiqueta que es el que determina el destino de dicho paquete. Esta etiqueta es asignada por la unidad central proceso, para poder establecer una conexión. Este antes de asignar una dirección a cada paquete, de una lista de direcciones, se cerciora que dicha dirección no se encuentre en uso. Cuando no es posible localizar la dirección, de la etiqueta, es probable que el número de circuitos virtuales dentro de la red sea limitado, se consume gran cantidad de tiempo buscando una etiqueta global, por lo que es requerido un soporte para las conexiones que atraviesan por varios switch de paquetes rápido de etiqueta diferente. Para poder dar esta clase de soporte, cuando se establece una conexión, se establece una tabla de consulta, que sirve para que el switch de paquete rápido, pueda asignarle a cada paquete una etiqueta de salida, en función de la etiqueta de campo de entrada y de la tabla.

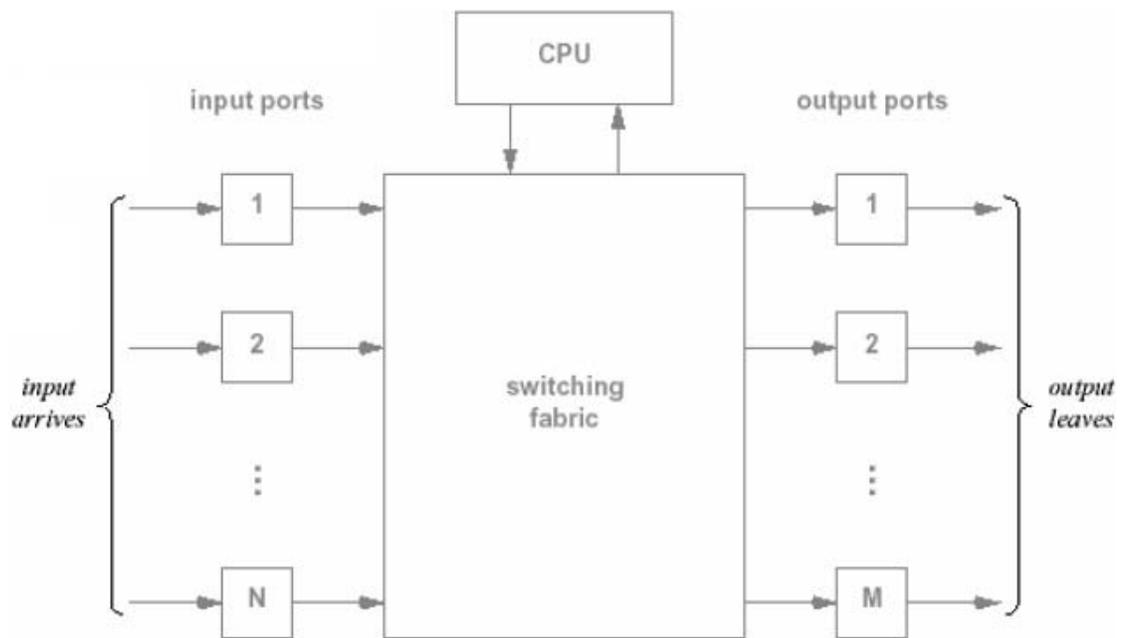
Cuando uno desea construir un switch de paquete rápido con alta capacidad, hay dos caminos a seguir, trabajando siempre el concepto de rutas múltiples que es requerido por este. El primer caso se realiza interconectando una gran cantidad de switch de paquetes rápido, donde cada uno de estos representa un dispositivo de switcheo, como si se estuviera trabajando con un switch de paquete rápido sencillo, este arreglo, debe contar con una unidad central de proceso, una tabla de consulta para las conexiones y se continúa con la manipulación de las etiquetas de los puertos de entrada como se mencionó anteriormente. Al trabajar de esta manera se tiene la deficiencia que se produce una réplica innecesaria de las funciones de control de los switches, pero nos brinda la ventaja que se posee una mejor flexibilidad a la hora de elegir la red de interconexión a utilizar. El segundo método consiste en implementar con switch simples un switch fabric como un switch multi-etapa, la ventaja que presenta este es que no ocurre el duplicado en las funciones de control de los switch. Los switch fabrics se forman con switch idénticos que poseen el mismo número de entradas como de salida, la cantidad de estos está definida por su grado de elemento al cuadrado. El grado de elementos es sumamente importante pues este es el que nos determina el número de interconexiones y etapas de switcheo requeridas para poder formar el switch de mayor capacidad.

Los paquetes en las redes de interconexión en multi-etapa pueden llegar al puerto correcto de salida, al seguir un algoritmo ya determinado. Este algoritmo se basa en que cada etiqueta específica contiene la información explícita, que dice a qué puerto de salida corresponde cada paquete de entrada, y esto se realiza a través de las tablas de consulta.

1.5. Arquitectura del switch rápido de paquetes

1.5.1. Diagrama interno de switch de paquete rápido

Figura 4. Arquitectura de Switch de paquete Rápido.



Fuente: Tilman Wolf de la Universidad Amherts de Massachussets.

Internamente un switch de paquete rápido esta compuesto de tres componentes básicos como podemos observar en la Fig. 4:

- ❖ Switch Fabric.
- ❖ Puertos de Entrada.
- ❖ Puertos de Salida.
- ❖ Controlador de Switch.

Switch fabric: Responsable de la transferencia de datos de las entradas para las salidas. Confía en encabezado interno del paquete para asignar una ruta para la salida correcta.

Puerto de entrada: Los puertos de entrada mantienen paquetes en espera para transmisión a través Switch Fabric. Los adaptadores de salida reciben paquetes del switch fabric y los sostienen (si necesario) antes de transmitirlos al siguiente enlace.

Controlador de switch: Este (cuando esta presente) Concede permisos de colocar información en los adaptadores de entrada para transmisión a través del switch fabric.

Puerto de salida: Su función es transmitir la información al destino correspondiente.

1.5.2. Clasificación simple de diseños de switches

El switch fabric y el buffer son los componentes indispensables en el diseño del switch y los switch de paquete rápido son clasificados según son posicionados estos dispositivos. El switch fabric puede ser con buffer o sin buffer y eso depende si la memoria interna se encuentra intermedia a los elementos de switcheo (con buffer) o si esta es externa al switch fabric (sin buffer). Cuando el switch trabaja con buffer, este puede ser buffer de entrada y de salida. Es buffer de entrada cuando este se encuentra antes que el switch fabric y buffer de salida cuando se encuentra después del switch fabric. El rendimiento de un switch con buffer de salida es del doble que el de un switch con buffer de entrada, pero este necesita de mayor hardware y mayores interconexiones para ser implementado.

Para mejorar el desempeño de los switch de buffer de entrada, con el fin que este se asemeje a un switch con un buffer de salida, se debe tomar muy en cuenta el tipo de switch fabric que se va a utilizar pues existen dos tipos fundamentales, con bloqueo o sin bloqueo. El switch fabric con bloqueo tiene menor desempeño, pero requiere menos interconexiones, este bloqueo se da cuando un paquete que se dirige a una salida que se encuentra en estado libre, y es detenido dentro del switch fabric por el trafico que hay en este.

En los switches con buffer de salida, cada buffer esta enfocado en un puerto de salida, pero todos los puertos de salida con los que cuenta el switch fabric pueden compartir una pila de buffers.

Por todo lo anterior podemos decir que los elementos de switcheo con buffer de entrada son mucho más sencillos de construir que los que tienen buffer de salida pero ofrecen un desempeño inferior. En general el desempeño de un diseño de un switch con buffer interno usando elementos de switcheo con buffer de entrada es similar a un diseño con buffer a la salida usando un **switch fabric** con bloqueo.

Como vemos el **switch fabric** toma una gran importancia en el diseño de los switch, por lo que lo definiremos como: Una interconexión de elementos conmutativos para facilitar un mecanismo particular de switcheo. El **switch Fabric** es caracterizado por los tipos e interconexión de sus elementos conmutativos. En el siguiente capitulo se tratara este dispositivo mas a fondo.

2. SWITCH FABRIC

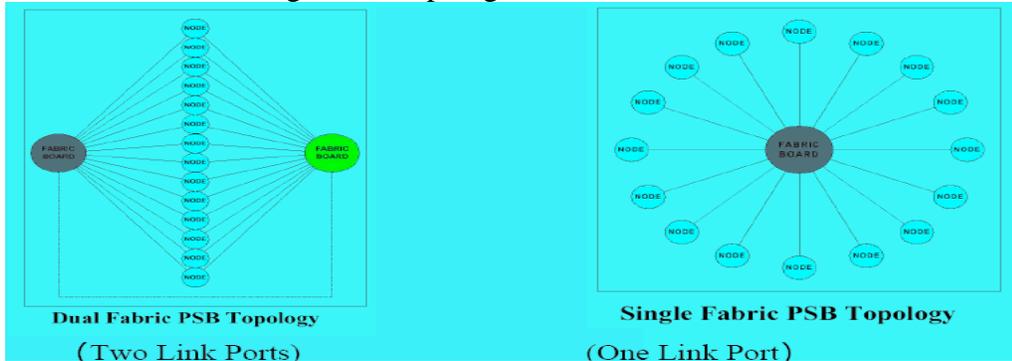
Como vimos en el capítulo 1, el switch fabric es de suma importancia, en los dispositivos de switcheo, y de él dependen en gran manera muchas de las características de estos, en el presente capítulo se hará un estudio más profundo de su funcionamiento.

El switch fabric es utilizado para suplementar, y eventualmente reemplazar, los buses tradicionales. Los switches fabrics son atractivos porque abstraen la transferencia de datos y se escapan de las complejidades del software de registro de dirección haciendo mapas.

También minimizan el número de switch de señales requerido, y eliminan problemas de desalineamiento de carga que limitan el desempeño de los buses paralelos.

Un Switch Fabric es una red interconectada de dispositivos conmutativos. La red contiene muchas terminales de entrada y de salida, y el switch fabric transporta los datos de las entradas para las salidas. Esto significa que estos han de viajar de una ranura a otra por medio de un dispositivo de alternación.

Figura 5. Topologías de Switch Fabric.



Fuentes: Bus Tronica an Elma Company, TreNew an Elma Company

Hay muchos tipos de topologías de Switch Fabric. Muchas de los nuevos estándares de la industria se basan en topologías centralizadas, dónde estos dispositivos activos de switcheo están ubicados en una especial tarjeta Fabric. Las otras tarjetas que se conectan a la tarjeta Fabric son llamados comúnmente tarjeta de nodos, que difiere de las conexiones de bus paralelo Fig. 5. Otros sistemas usan topologías de distribución, también conocido como las topologías de la malla, donde cada tarjeta está conectada con todas las otras tarjetas, y estos dispositivos activos de switcheo están ubicados en cada tarjeta.

El mecanismo de switcheo depende de la habilidad del switch fabric para dar a los controladores de la terminal de entrada información de una colisión entre un paquete que intenta entrar, y uno ya establecido. Esto es logrado estableciendo un camino inverso a través del switch fabric paralelamente al camino principal. Cada enlace en la red de la interconexión consta de dos caminos, un camino principal para transmitir los datos y un camino inverso para la señal de colisión. Estando interrumpido, el switch fabric devolverá una señal de colisión en el camino inverso y todos los mecanismos de switcheo del camino parcialmente establecido regresarán al estado desocupado tan pronto como el controlador de la terminal de entrada quite el paquete fracasado del switch fabric.

Tres clases de switch fabric de auto-ruteo se presentan merecedoras de investigación: No bloqueado, Nuevo arreglo no bloqueado, y bloqueado. Como se vio en el capítulo 1 un switch no bloqueado tienden a costar caro en términos de sus requisitos del hardware, pero tiene un desempeño mucho mayor que el del switch bloqueado. Por ello el switch fabric no bloqueado será utilizado como un estándar en contra el cual otros switch fabrics pueden ser comparados. Además de esto es importante mencionar que los switch fabric no importando si son bloqueados o no bloqueados también pueden clasificarse en Time Domain (cuando trabaja con TDM) y Space Domain (cuando trabaja en Multiplexación Estática).

2.1. Desempeño del switch fabric

Como mencionamos en el inciso anterior existen varias clases de switch fabrics, cada una con un rango de parámetros posibles de implementación, y estos requieren investigación para seleccionar el diseño preferido del switch.

La forma más fácil que se conoce para establecer el desempeño de una implementación particular del switch es medir el rendimiento promedio del switch normalizado cuándo esta saturado por tráfico que es una información aleatoria uniforme de destinos de paquete. Este es el rendimiento específico máximo del switch. Y nos da una medida útil de la capacidad del switch para la cual se pueden comparar los diferentes diseños del switch. La mayoría de veces es importante tomar en cuenta el desempeño de la media de retraso de paquetes desde la entrada hasta la salida del switch para fuentes idénticas de tráfico en cada uno de los puertos del switch. Para poder comparar el rendimiento en saturación y el desempeño medio de retraso se desarrollara un modelo de simulación de los diferentes switching fabrics utilizando los diversos resúmenes de parámetros de implementación que se describen en la tabla I.

Tabla I: Parámetros de diseño del Switch fabric.

PARAMETOS	RANGO
TAMAÑO DEL SWITCH FABRIC	Desde 2X2 hasta 4096x4096
Redes de Interconexión	Cossbar Delta Benes
GRADO DE ELEMENTOS DE DISEÑO	2, 4, 8, 10.
Múltiple	Buscando
Ruta	Desbordando
Algoritmos	Aleatorios
PLANOS MULTIPLES DE SWITCHEO	de 1 a 4
Controladores de Puertos	Regular By-Pass de la Cola de Entrada Doble buffer a la Salida De Lujo

Fuente: Fast Packet Switching for Integrated Services of Peter Newman

2.2. Modelos de tráfico

Se han desarrollados dos modelos de tráfico para la comparación de los diversos diseños del switch: Un modelo para la evaluación de rendimiento en saturación y una fuente de tráfico en los slots del switch para la comparación del retraso de los paquetes. En el modelo de saturación, se satura con un tráfico entrante cada puerto de entrada del switch con el único fin de que siempre este disponible un nuevo paquete en cada puerto de entrada al terminar la transmisión del anterior paquete. Para los switches que no utilizan un by-pass de alineación de entrada no es necesario modelar las alineaciones de entrada, mas que para poder suplir cada puerto de entrada con un paquete nuevo cada vez que completase una transmisión. Es modelada una alineación en cada puerto del switch para que esté siempre lleno, en los switch que en su entrada cuentan con by-pass de alineación. En switches de dimensiones 512x512 la alineación puede ser de unos 100 paquetes de largo mientras que para los paquetes usados para switches mayores es suficiente una alineación aproximadamente de 10 paquetes. Todos los paquetes deben ser del mismo largo y seguir una distribución aleatoria uniforme de destino, en tanto que como coladeros se utiliza a los puertos de salida.

Ningún buffer de switcheo de entrada puede ser efectivamente activado en saturación debido a que las alineaciones de entrada se llenarían permanentemente creando un retraso muy alto y altas pérdidas de paquetes, aunque en saturación es donde este nos da información acerca de su capacidad máxima. Para comparar el retraso medio a través de diferentes estructuras del switch, deben colocarse en cada puerto de entrada, fuentes de tráfico que produzcan una carga por debajo del rendimiento específico en saturación. Cada terminal de entrada recibe una corriente continua de timeslots, es decir de un solo paquete. La carga ofrecida por la fuente de tráfico puede ser una probabilidad aleatoria uniforme, pues cada timeslot puede estar vacío o lleno de un solo paquete, por lo que a todos los paquetes se les asigno una distribución aleatoria, para que trabaje con fuentes que poseen la misma distribución de carga para cada medida de retraso media.

2.3. El Modelo de simulación

Una estrategia exitosa de simulación depende de varias decisiones importantes del diseño. Aunque estas surgen dependiendo del problema que tiene que ser solucionado en cada simulación.

El cronometrar sistemas es un problema que potencialmente introduce un intercambio entre la validez, o sea la extensión para la cual los resultados de simulación son propios de medidas en un sistema real (el prototipo), por un lado vs. el tiempo de ejecución del simulador por otra parte. Decrementando el último aumenta el rango de validez del modelo, si bien al precio de un tiempo de ejecución más largo. Este dilema es no posee solución en el contexto secuencial, esto refleja el límite fundamental de la ejecución de instrucciones secuenciales en la máquina de Von Neumann. Una solución parcial para esta simulación secuencial, es la estrategia de evento conducido, simulación donde solo instantes en tiempo donde un evento potencial cambia de estado son considerados. Por lo tanto periodos de tiempo donde restos de estado no cambian son brincados y la resolución de tiempo se torna independiente del tiempo de ejecución. Esta estrategia intercambia código complejo para un desempeño en tiempo de ejecución, esto ciertamente es la única posibilidad en computadoras secuenciales. La estrategia de tiempo que fue escogida para el actuar en este tipo de experimentos es el método de conducción de tiempo [9].

Esto es debido al hecho que el hardware real opera en slots de tiempo, cada celda que arriba al switch fabric es alineada para el siguiente slot. Aunque un slot sea relativamente es el ancho de una unidad de tiempo (Una celda esta dividida en 8 slots), no es una influencia negativa de los resultados de simulación. El dilema básico del tiempo de resolución vrs. el tiempo de ejecución fue resuelto. Implementar el método de tiempo conducido es más simple que el de eventos de códigos conducidos [9].

Un modelo de simulación para un experimento particular es construido en dos niveles: El nivel de componente básico y el nivel estructural. Los componentes básicos son modelos del elemento básico de switcheo, es este caso el switch fabric, entradas y salidas formateados, colección de módulos estáticos, fuentes de error etc. Estos son combinados dentro un simulador global en el nivel estructural. Esta metodología permite la generación de simulaciones que son “validas por construcción” sobre la intersección de los rangos de valides de los componentes básicos de modulación [9].

El experimento que se presenta en esta sección pretende dar una vista general del tipo de trabajo que se realizo en el simulador y su asociación con el modelo analítico que puede ser usado. El modelo presenta infinitos buffers, celdas con probabilidad de perdida, la cual es computada como un porcentaje. Las unidades de tiempo y espacio son expresadas en slot, donde una celda ATM (modo de transferencia Asíncrona), sobre una Multiplexación estática, transmite 155.52 Mbit/s igual a 8 slots. Por lo tanto el espacio por slot es 68 bits y el tiempo de slot es $68 \text{ bits}/155.52 \text{ Mbit/s} \sim 0.437 \mu\text{s}$.

Como se menciona arriba, se trabajo sobre el modelo de evento conducido dentro del método de conducción de tiempo. Pues el Dr. Peter Newman, con el fin de reducir el tiempo de computadora requerido para realizar el modelo, hizo una simulación, bajo el algoritmo de flooding en las que cada componente fue simulado como un acontecimiento instantáneo, el cual es llevado por un camino completo en el cual puede ser rastreado a través de una red de interconexión (describe la operación conjunta de varios diseños de switch fabrics). En un principio todos los nodos de interconexión del de la red, fueron puestos en desocupado, y durante la transmisión del paquete el camino fue marcado como ocupado. Al iniciarse una nueva transmisión todos los caminos que llevan a la ruta deseada deben investigarse, con el fin de encontrar un camino que este desocupado, si se diera el caso de que ningún camino se encontrase en estado de desocupado, el intento se considera como un intento bloqueado, y un nuevo intento debe ser programado.

Cada paquete, se va transmitiendo de etapa por etapa, por lo que si un paquete falla en pasar a una nueva etapa podría bloquear su intento. Este efecto se puede simplificar, para sobre estimar el desempeño en saturación. Esta simplificación, la utilizo el Dr. Newman para mostrar que para redes de dos planos Delta, que son construidas con elementos de switcheo que van desde grado 8 al 16, el error introducido en la evaluación hecha en saturación es aproximadamente de 2%.

En este modelo de simulación, el Dr. Newman considero el hecho de que al terminar la transmisión, se debe poner el camino en estado de desocupado, considerando la transmisión del paquete como instantánea. En la realidad esto no ocurre de una manera instantánea por lo que debe de considerarse un retardo fijo de tráfico, por la respuesta de switcheo en cada etapa. Considerando lo anterior, se puede tomar un retraso para resolución de eje de tiempo de 1/100 de duración de paquete, esto debe hacerse siempre tomando en cuenta la velocidad del switch fabric.

El Dr. Newman, inicio la simulación en general, relacionando todos los paquetes durante un tiempo aleatorio, y dejo circular la información por la red durante el tiempo necesario para transmitir 200 paquetes, después de esto empezó las mediciones, pues el sistema ya se encontraba en estado estable. Por las redes se transmitieron 200,000 paquetes, trabajando siempre en estado de saturación, con lo cual observo una desviación estándar de 0.4% para las redes pequeñas y de 0.2% para las redes de 64*64 o mayores. Esta simulación incluía también los bits de label, tags o cualquier otro tipo de encabezado y el ancho de banda de puerto útil fue apenas menor que el especificado para el rendimiento en saturación.

En base a lo anteriormente dicho sobre el modelo de simulación se presenta, como están configuradas las matrices de crosconexión del fabric, su rendimiento al ser sometidas a la simulación anteriormente mencionada.

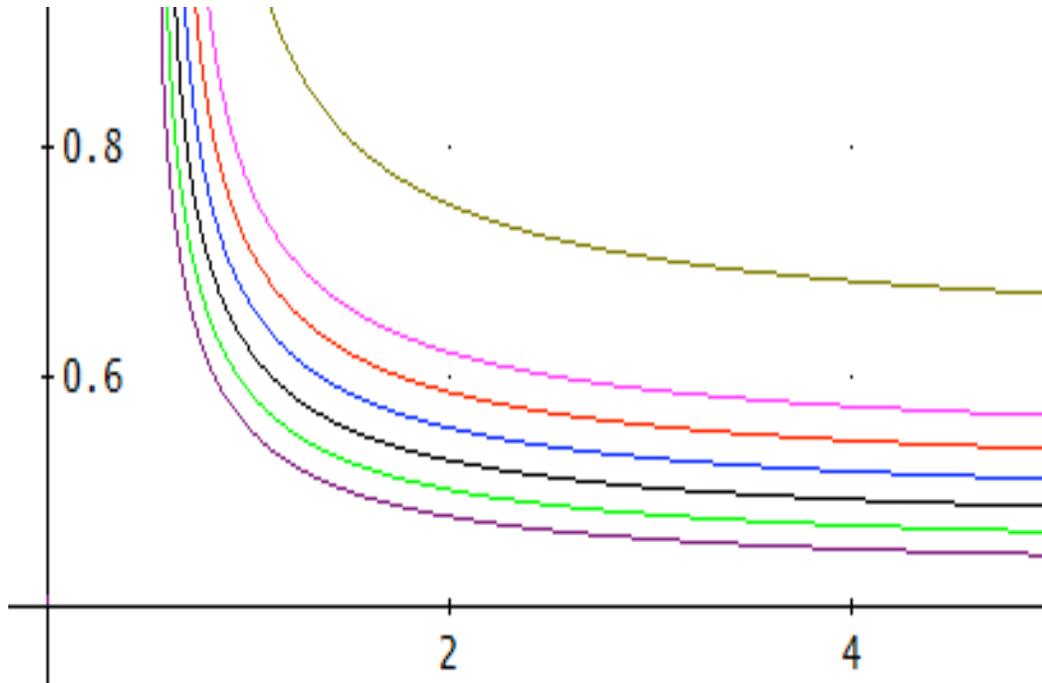
2.4. Switch fabric crossbar (de barra cruzada)

2.4.1. Rendimiento en estado de saturación

Este switch fabric es sin bloqueo con lo que ofrece un desempeño ideal en contra del cual otras redes de interconexión pueden ser comparadas. En las entradas de este switch, el bloqueo procede únicamente de la probabilidad de que fuentes múltiples intenten transmitir hacia el mismo destino al mismo tiempo. El rendimiento en los resultados de saturación de switch fabrics crossbar de hasta 512x512 de tamaño, es presentado en la fig. 6. bajo suposiciones diversas. El tamaño del interruptor ($N \times N$) es expresado como el $\log_2 N$ y las curvas son discretas, estando los puntos conectados únicamente por conveniencia visual. En la suposición sincrónica todos los paquetes son enviados al switch sincronizadamente y en una fase tal que la contención para todos los puertos de salida ocurre al mismo tiempo y esos paquetes son exitosamente transmitidos a través del switch fabric. Los paquetes que son bloqueados pueden ser descartados o pueden esperar exactamente un ciclo de duración del paquete para ser reenviados en el siguiente timeslot. En la suposición asincrónica hay relaciones de tiempo aleatorias entre todos los paquetes. Si un paquete es bloqueado espera hasta el próximo reintento y luego es independientemente de todos los otros paquetes para ser reenviado. El retraso de reintento es expresado como un porcentaje de la duración del paquete.

Según un análisis [1], se da una expresión para el rendimiento de un switch fabric crossbar en saturación bajo las suposiciones de que la llegada de los paquetes es sincronizada y para la que los paquetes bloqueados son descartados. La expresión es $1 - [1 - (1/N)]^N$. Tiene una asíntota de $(1 - e^{-1}) = 0.632$ para un N grande. Si los paquetes son reenviados en lugar de descartados, el switch fabric crossbar sincrónico, se vuelve mucho más difícil de analizar pero se tiene un resultado de $(2 - \sqrt{2}) = 0.586$ para la asíntota del rendimiento en saturación para una N grande.

Figura 6. Rendimiento de un switch fabric crossbar en saturación.



Fuente: Grafica de $1 - [1 - (1/N)]^N$ en Derive 6.1, para distintos valores de N.

El rendimiento en saturación para la operación asincrónica de un switch fabric crossbar en diversos valores de reintento de retraso son también dados en la fig. 6. La operación asincrónica con un retraso de reintento de cero es equivalente para la operación sincrónica, mientras la contención para los puertos de salida ocurre inmediatamente y estos se libran del paquete anterior. En la operación asincrónica el rendimiento en saturación se reduce tanto como el retraso de reintento se incrementa. Esto es debido al hecho de que mientras mayor sea el retraso de reintento, mayor tenderá a ser la probabilidad de que un puerto de salida gaste algún tiempo desocupado después de transmitir un paquete, y esto causará un retraso al puerto de entrada, cuando intente restablecer un paquete, [3].

Para el switch fabric crossbar sincrónico con paquetes bloqueados reintentados, el desempeño en saturación por tráfico con una distribución aleatoria uniforme de destino representa la probabilidad de que cualquier paquete será exitoso en cualquier intento de envío. Por lo tanto la operación de cualquier puerto de entrada del switch puede ser modelada como un servidor geométrico en el cual la probabilidad que un paquete requerirá j intentos de enviarse es $p(1-p)^{j-1}$. La media de retraso a través del switch fabric es, siendo p la carga de tráfico [3]:

$$\sum j * p(1-p)^{j-1} = 1/p$$

2.4.1.2. Media de retraso para tráfico en slots

Existen resultados analíticos disponibles de la media de retraso para tráfico en slots tanto para la entrada como la salida con buffer del switch fabric crossbar. Para una N grande, la ecuación para el switch con buffer de entrada es:

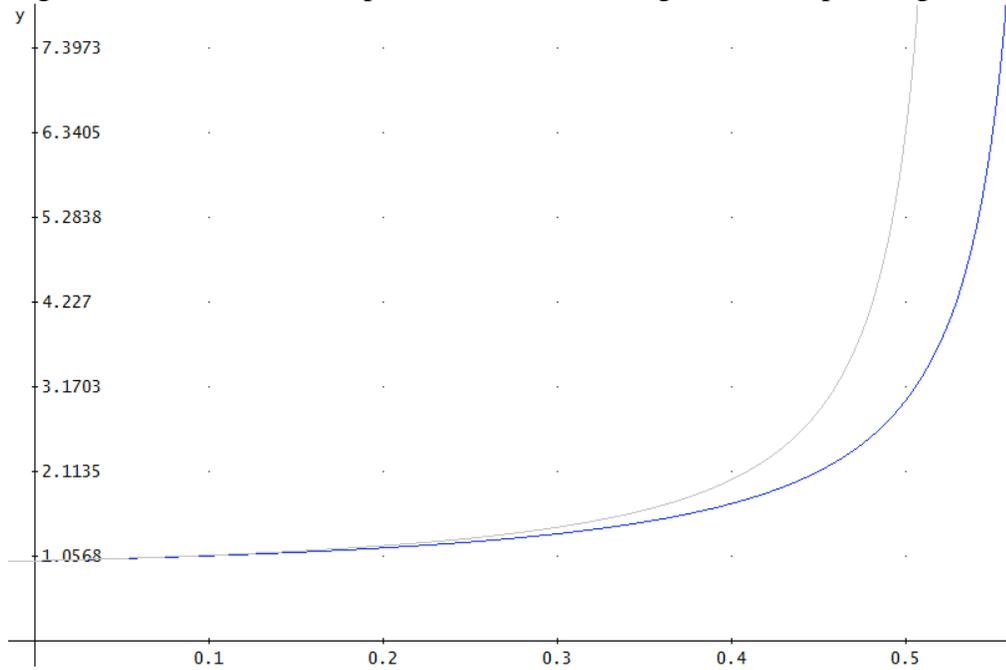
$$\frac{(2-p)(1-p)}{(2-\sqrt{2-p})(2+\sqrt{2-p})} \quad (2.1)$$

Y para el switch con buffer de salida es:

$$\frac{(N-1)}{N} \frac{p}{2(1-p)} + 1$$

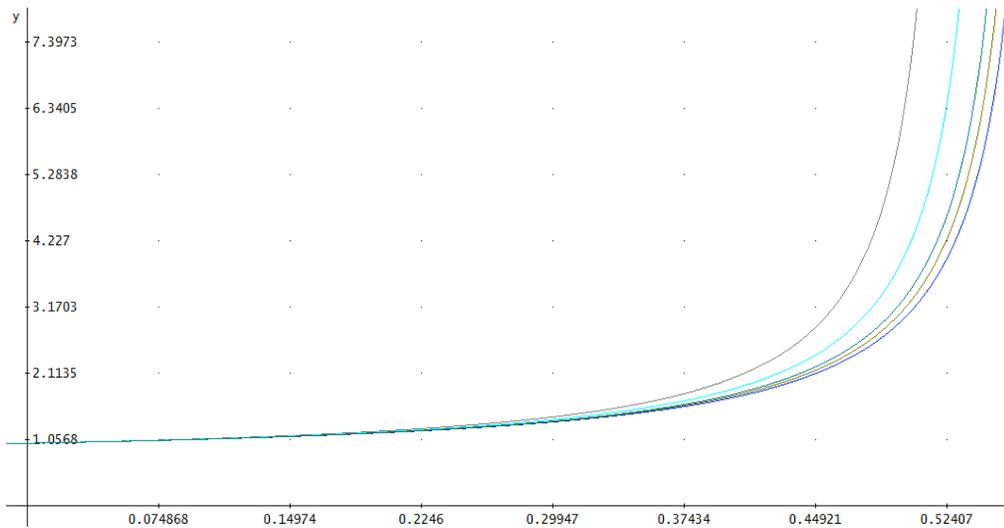
Donde la N es el tamaño del switch y p es la carga de tráfico (lo cuál es la probabilidad de que cualquier timeslot de entrada o de salida, respectivamente contenga un paquete).

Figura 7. Media de retraso para switch de buffering de entrada para N grande.



Fuente: Grafica de (2.1). en Derive 6.1

Figura 8. Efecto de introducir un by-pass de alineación de entrada, y buffering de salida a través de un switch fabric crossbar.



Fuente: Grafica de (2.1). en Derive 6.1 variando los valores de p (carga de tráfico).

2.4.2. Matrices de croconexión

2.4.2.1. Red delta

Se puede decir de la red Delta que ella da la apariencia de que es la red mas apropiada para trabajar en la clase de switch fabric bloqueado, con auto-ruteo mas controlado.

Podemos presentar un ejemplo de una red Delta de 64x64 construida con elementos de switcheo de grado 8, en la Fig. 9.

Figura 9. Red Delta de 64x64.

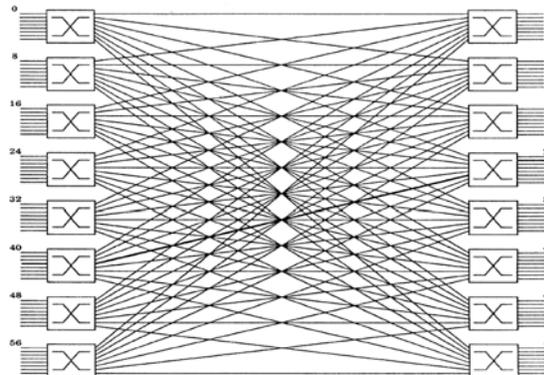


Figure 5.3: A 64×64 delta network of 8×8 switching elements.

Fuente: Dr. Peter Newman of Wolfson College, University of Cambridge

Como se puede apreciar en la figura las terminales que representan la salida del switch fabric están numerados de 0 a 63 siendo la terminal de grado mas alto asignada con el numero 0, y las terminales de salida de cada uno de los switch que forman el switch fabric, están numerados del 0 al 7, siendo de igual manera la terminal de puerto mas alto asignado con el numero 0. Para un switch de base 2, como el de la figura 2.5 dos dígitos son necesarios, el primer dígito el cual es el mas significativo nos dice cual es la salida del elemento de switcheo, y el dígito menos significativo, nos dice cual es la puerta de salida del switch fabric, y es de esta manera que se establece la ruta de interconexión. Es posible que hayan múltiples caminos entre un par de puerto de entrada y uno de salida. Para poder seleccionar el camino que nos brinde mas eficiencia se utilizan dos algoritmos: el de búsqueda y el flooding. En el de búsqueda el mecanismo del puerto de entrada es el encargado de controlar los intentos de comunicación a través de las vías hasta que una secuencia, logre hacerlo con éxito. En el método de flooding los paquetes de llegada, son enviados simultáneamente a través de todas las vías libres que llevan al destino, y los paquetes que colisionan con otro son removidos del switch fabric.

Siguiendo en la línea de la simulación, dentro del algoritmo de flooding mencionado con anterioridad, haciendo trabajar la red Delta de un solo plano en saturación con una operación asíncrona, con paquetes bloqueados de reintento con un retraso del 10% de la duración del paquete, se puede concluir que la red Delta posee un poder integrado dependiendo del grado de los elementos de switcheo, y cualquier entrada es conectada con cualquier salida por un camino único.

Cuando la red Delta trabaja bajo operación síncrona, tiene un mejor rendimiento de 8% cuando trabaja con paquetes bloqueado reintentado que cuando trabaja con paquetes bloqueados tirados, esta magnitud es muy similar o se asemeja a la red crossbar cuando trabaja bajo las mismas condiciones. Y como era de esperarse el rendimiento de una red Delta operando en saturación bajo la operación sincrónica es menor que cuando esta trabaja en operación asíncrona. Esto es debido a los intentos sin éxito de los paquetes que desean atravesar la red, haciendo que la red multi-etapa quede bloqueada, este efecto es directamente proporcional al número de etapas de interconexión que hay dentro de la red. Por su parte en la operación asíncrona, la duración del paquete, el tamaño del switch y el retraso son los que determinan la magnitud de dicho efecto.

2.4.2.2. Red bannes

La red bannes tiene una distribución de destino aleatoria pero ofrece un desarrollo para el tráfico aceptable. En la figura 10 se muestra una red bannes de 64x64 la cual cuenta con elementos de switcheo de grado 8.

Figura 10. Red Bennes de 64x64.

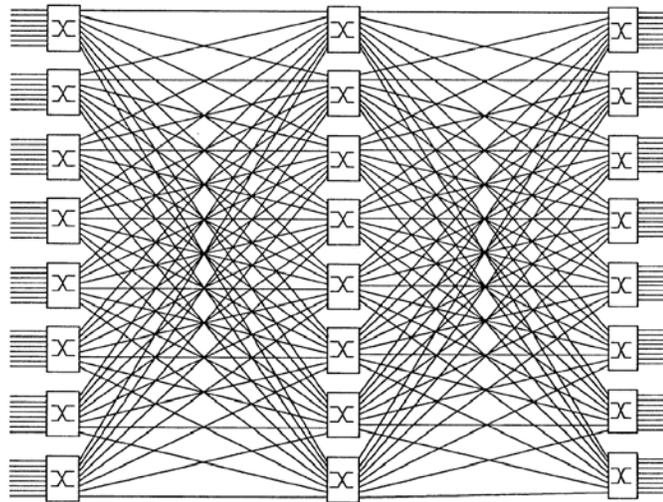


Figure 5.5: A 64×64 Beneš network of 8×8 switching elements.

Fuente: Dr. Peter Newman of Wolfson College, University of Cambridge

En la estructura de la red Bennes cada paquete se selecciona independiente, buscando que los paquetes incidentes no produzcan errores, cuando viajan por el mismo circuito virtual. Para seleccionar el camino más eficiente a través de la red, se han desarrollado tres algoritmos: El de búsqueda y el de flooding, los cuales, se comportan de la misma manera como lo hacen cuando se trabaja con la estructura de la red Delta con la diferencia que aquí se cuenta con muchos mas caminos. Y el algoritmo aleatorio, que es en el cual se selecciona un camino libre con la siguiente etapa al azar utilizando un switch, que cuenta con dispositivos de switcheo que pueden realizar esa operación.

Una investigación del rendimiento en saturación de los Bennes y las estructuras sub-equipadas Bennes revelan que los algoritmos de búsqueda producen un desempeño mucho más pobre que los algoritmos flooding y los aleatorios y son aun ligeramente inferiores al desempeño de la red delta equivalente de un solo plano. Esto se debe a que las estructuras Bennes ofrecen mucho más caminos que las Delta para ser buscados secuencialmente. También el mayor número de etapas de switch incrementa la posibilidad de bloqueo dentro de la switch fabric en cargas altas. El algoritmo aleatorio ofrece un mejor desempeño que el de la red delta equivalente de un solo plano porque siempre selecciona un camino a través de la distribución de etapas del switch fabric que esté libre. Como podría esperarse, el algoritmo de flooding ofrece el mejor desempeño de rendimiento el cual es muy similar al equivalente crossbar network. Para aplicaciones que requieren una duración corta de paquete, el algoritmo aleatorio de selección del camino es preferido. Para paquetes de duración más larga, el algoritmo de flooding ofrece mejor desempeño ya que la interferencia relativa entre los intentos múltiples de asignación de paquete es disminuido por el incremento en la duración del paquete.

Es difícil que una switch fabric Bennes sea seleccionada en vez de una red delta a causa de su desempeño en saturación. La diferencia en el desempeño entre las dos estructuras no es particularmente significativa. La estructura Bennes es de interés por su sensibilidad grandemente reducida para la distribución de destino del tráfico incidental cuando es comparada con la red delta.

2.5. Tipos de switch fabric

Como se vio al principio los switch fabric pueden trabajar, en el dominio del tiempo cuando este trabaja con una comunicación que esta basada en Multiplexación sincronía como TDM y en dominio del espacio cuando trabaja con Multiplexación asíncrona no importando si estos son bloqueados o no bloqueados, a continuación se presentan algunos ejemplos de ellos.

- ❖ Time domain.
- ❖ Space domain.

2.5.1. Switch fabric de space domain

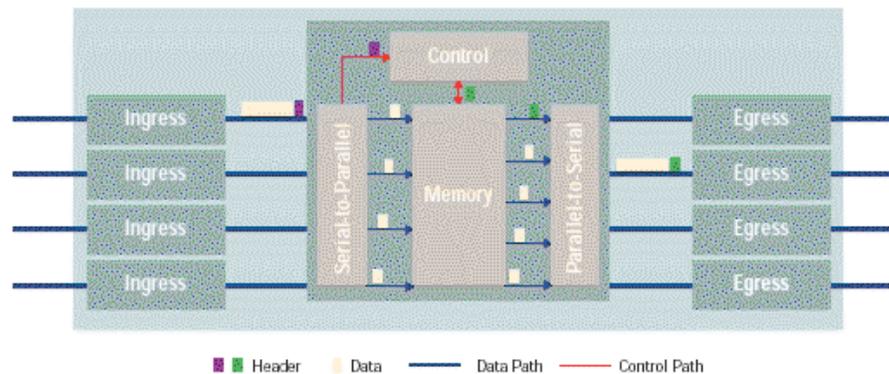
Según sus topologías existen de varias características:

- ❖ Debe existir uno o varios caminos entre el puerto de entrada y el de salida.
- ❖ Un camino puede ser dedicado por par, ó parcialmente compartido
- ❖ Un paquete puede pasar por varios segmentos (etapas) para llegar a su destino.
- ❖ La distancia a recorrer de un paquete es la cantidad de etapas que debe un paquete recorrer para llegar a su destino. Dependiendo de su característica puede ser variable o fijo.

De este tipo existen varias configuraciones como el Crossbar, MIN, Torus y el hiper cubo.

2.5.2. Switch fabric de tipo time domain

Figura 11. Diagrama de un SF de tipo Time Domain.



Fuente: TELLABS.

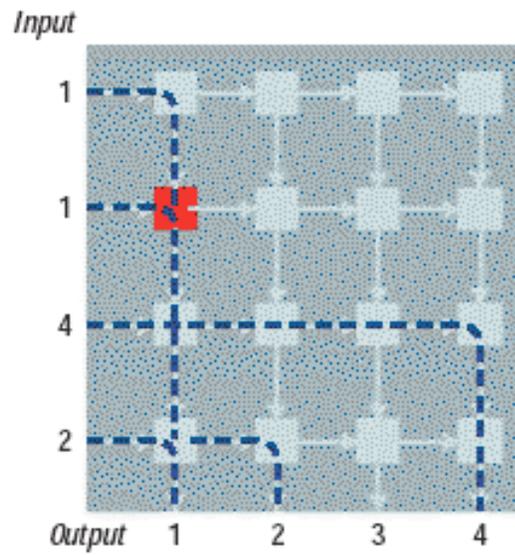
Las principales desventajas que deben mejorar para ser competitivos son:

- ❖ Tener potencia de control mayor, para control por flujo.
- ❖ El ancho de banda de la memoria tiene que ser dos veces el producto del número de puertos por la velocidad de línea del puerto.

Los esquemas de control de estas configuraciones son los que tienen que responder ante problemas como:

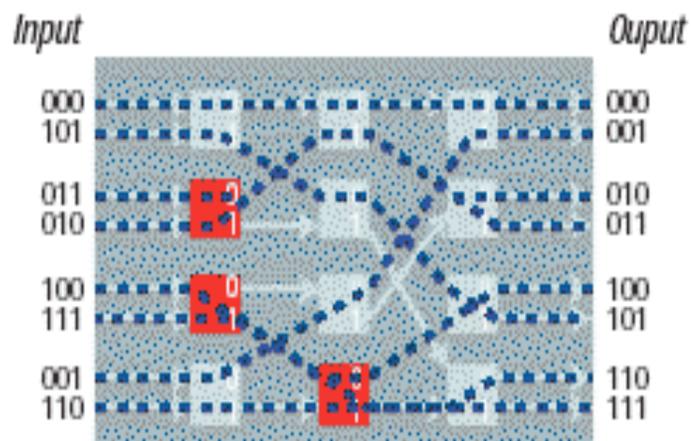
- ❖ Bloqueo de salida, bloqueo medio y bloqueo por HOLD.
- ❖ Este control puede ser centralizado o de manera distribuida.

Figura 12. Diagrama de bloqueo de salida de un SF de tipo Time Domain.



Fuente: TELLABS.

Figura 13. Diagrama de bloqueo de interno de un SF de tipo Time Domain



Fuente: TELLABS.

Para evaluar los SF se necesita verificar los siguientes puntos:

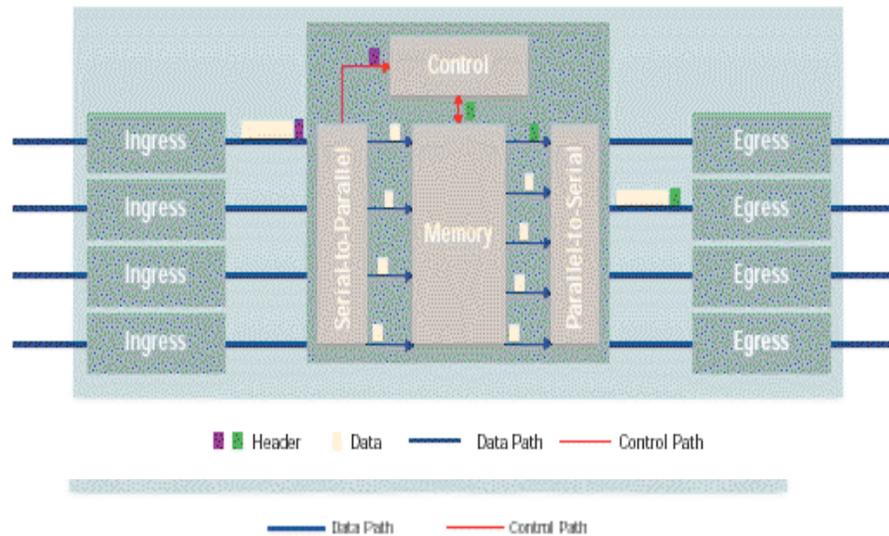
- ❖ Escalabilidad: Que se lleve de Gbps a Tbps con bajo costo inicial.
- ❖ Desempeño: Que lleve el paquete en bajo retardo y sin bloqueo.
- ❖ Diferenciación de servicio.
- ❖ Tolerancia a fallas.

Aparte de estos puntos deben medirse los puntos de arquitectura como:

- ❖ Etapa de conmutación: simple ó compuesta.
- ❖ Distancia puerto: fija o variable.
- ❖ Plano de control: centralizado o distribuido.
- ❖ Tipos de buffer: Dependiendo de su localización.

2.5.3. Switch fabric de memoria compartida

Figura 14. Diagrama de un SF de memoria compartida.

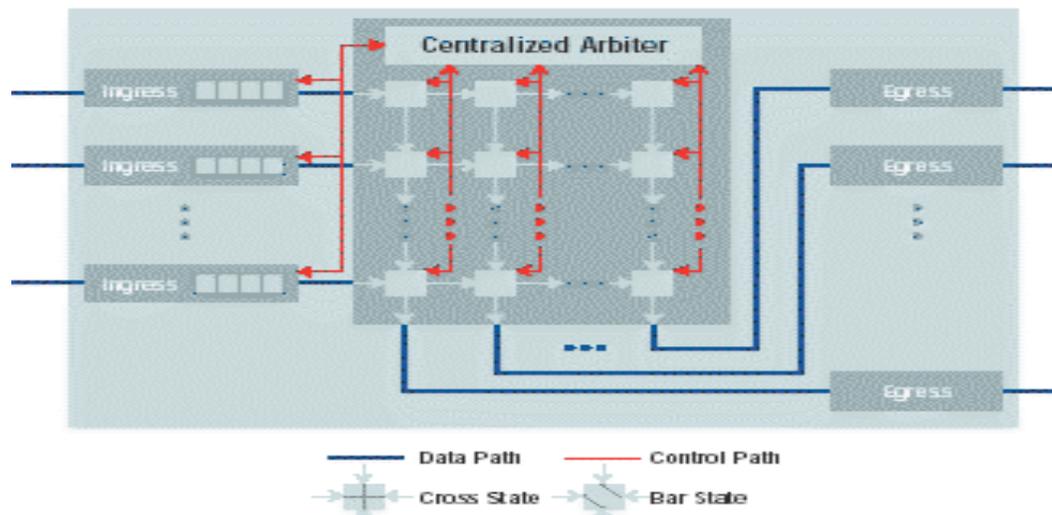


Fuente: TELLABS.

- ❖ STM-64 full duplex máximo.
- ❖ Control centralizado que se vuelve cuello de botella.
- ❖ Riesgo de paquetes perdidos por el diseño de los cambios de paralelo-serie.

2.5.4. Switch fabric de crossbar con control central

Figura 15. Diagrama de un SF de crossbar con control central.

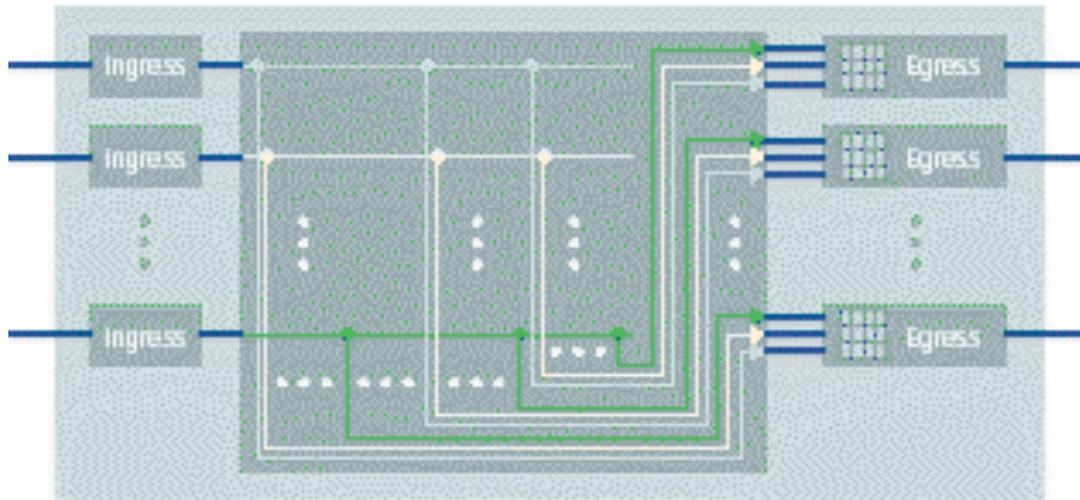


Fuente: TELLABS.

- ❖ N^2 elementos con N el número de pares entrada salida.
- ❖ No se puede evitar el bloqueo de salida por lo que se usan buffers.
- ❖ Adoptada para switches de TDM desde hace tiempo por su bajo retardo.
- ❖ Aun con su bajo retardo suele ser más eficiente cuando es adoptado para switches que trabajan con Multiplexación estática.

2.5.5. Switch fabric de full matrix interconnect con control en las salidas

Figura 16. Diagrama de un SF de full matrix interconnect con control en las salidas.

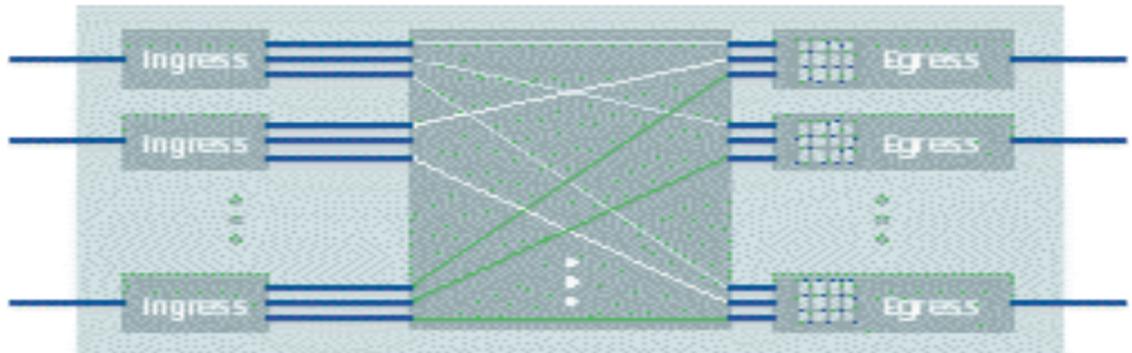


Fuente: TELLABS.

- ❖ Basado en un esquema de full mesh. Equivalente a un switch fabric de distancia fija puerto a puerto
- ❖ Crecimiento lineal de los buffers con respecto al número de puertos
- ❖ En la salida se debe enviar los paquetes en un broadcast de entrada.
- ❖ Se corre el riesgo de llegar a tener ráfagas completas en todas las entradas y para evitar esto se debe contar con un mecanismo de knock out.

2.5.6. Switch fabric de full matrix interconnect con control en la entrada

Figura 17. Diagrama de un SF de full matrix interconnect con control en la entrada.

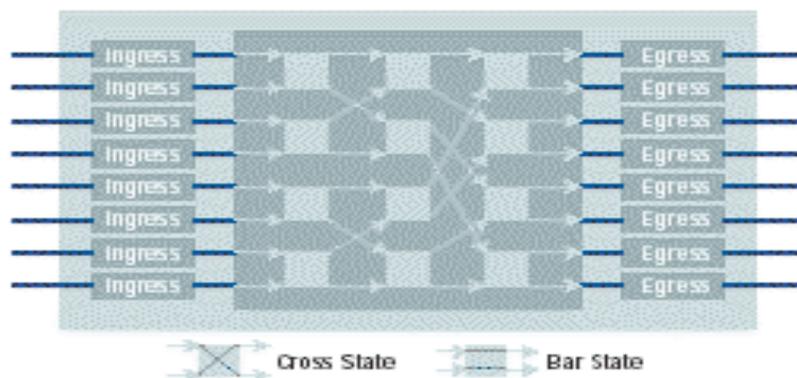


Fuente: TELLABS.

- ❖ Es similar al FMI de control en la salida, pero se debe hacer un mayor control, verificando la dirección de destino y haciendo un unicast.
- ❖ Esto dependiendo el diseño puede comprometer el retardo de los paquetes.

2.5.7. Switch fabric de MIN con matriz de 8 X 8

Figura 18. Diagrama de un SF de min con matriz de 8 X 8.

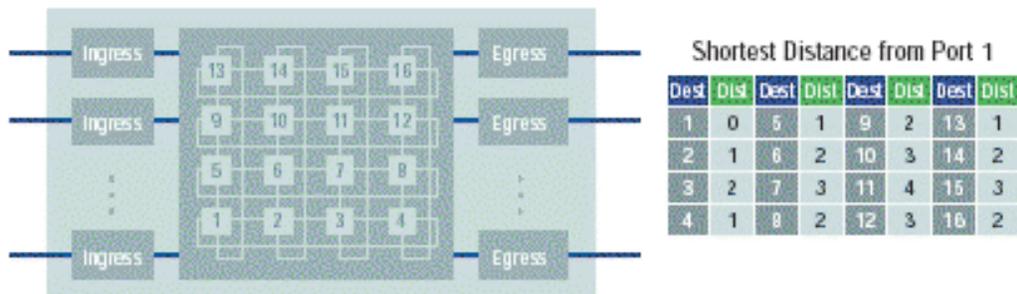


Fuente: TELLABS.

- ❖ Reducción a $\log_2(N)$ del número de etapas.
- ❖ Auto erutación.
- ❖ Produce bloqueo interno por compartir las rutas. Se debe usar un esquema de ordenamiento o buffers para evitarlo

2.5.8. Switch fabric de 16x16 tipo torus

Figura 19. Diagrama de un SF de 16X 16 tipo torus.

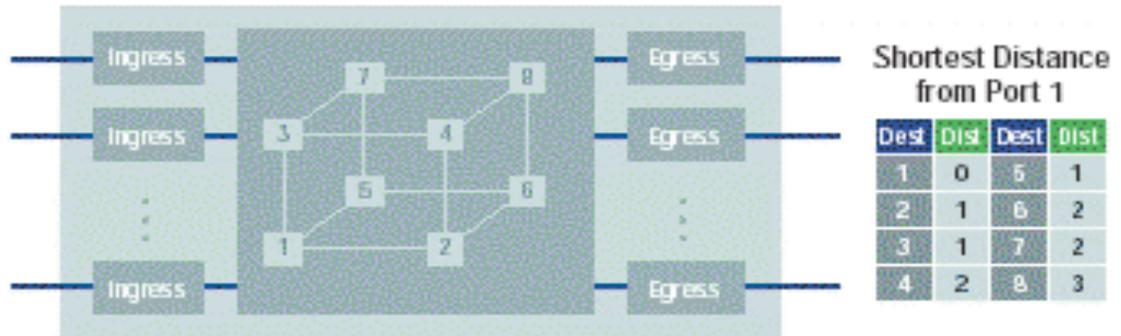


Fuente: TELLABS.

- ❖ Mesh bidimensional basado en un store-forward con un algoritmo de distance vector.
- ❖ Es un switch fabric de multi etapa y distancia variable.
- ❖ Puede llevarse a un mesh circular tridimensional.
- ❖ Puede ser impredecible su desempeño en cuanto a jitter y retardo.

2.5.9. Switch fabric de hiper cubo de 8x8

Figura 20. Diagrama de un SF de hiper cubo de 8X8.



Fuente: TELLABS.

- ❖ Usa topología genérica de N dimensiones para 2N puertos.
- ❖ Tiene una distancia mínima en relación al Torus, N siendo la máxima distancia
- ❖ Es un balance bueno entre número de puertos y corta distancia.

3. TENDENCIAS EN EL DISEÑO DE ROUTERS

Se ha hablado sobre los distintos dispositivos de switcheo, como es que constituidos y cual es su comportamiento al momento de trabajar, con una comunicación síncrona o asíncrona y cual es su eficiencia cuando es saturado por información. El Switch de paquete rápido es un dispositivo de switcheo, robusto que trabaja mejor para comunicación asíncrona, haciendo que cada paquete lleve dentro de su información, cual es su fuente y su destino trabajando sobre el concepto de conexión orientada, es decir la creación de circuitos virtuales, por donde debe viajar la información. La conexión también puede hacerse en función de tablas de enrutamiento. Como vimos entre sus dispositivos se encuentra el switch fabric, y en el capítulo 2 nos centramos especialmente en el funcionamiento de este. En este capítulo hablaremos del router, el cual es un dispositivo que trabaja en la sobre el nivel 3 si nos referimos al modelo OSI y su función es examinar la capa de red para determinar la red de destino por medio de una tabla de enrutamiento, que mantiene un registro de las rutas hacia los destinos de redes específicos, para poder llevar a cabo esta función veremos que dentro de el, cuenta con un mecanismo de switcheo que comúnmente es el switch de paquete rápido.

3.1. Componentes de un router

La arquitectura de un router genérico es prácticamente la arquitectura de un switch de paquete rápido. Como vimos en la sección 1.5.1, esta, está compuesta de 4 componentes básicos. Puerto de entrada, Puerto de salida, y el Switch Fabric del que tanto hemos hablado, y la Unidad Central de Proceso. A continuación trataremos cada una de ello por separado y de una manera más profunda.

3.1.1. Puerto de entrada

El puerto de entrada, es el medio por medio del cual entra la información al router, en forma de paquetes. Estos vienen en una tarjeta en la cual puede haber 4, 8 y 16 puertos. Estas tarjetas son llamadas, tarjetas alineadas.

Sus funciones principales son:

- ❖ Realiza las funciones, de capsulación y descapsulación de los paquetes. Trabajando dentro de la capa de enlace.
- ❖ Cuenta con una tablilla de reenvío con la cual es capaz de encontrar la ruta y el puerto de destino, lo hace por medio de un algoritmo de búsqueda, como los mencionados en el capítulo 2.
- ❖ Es posible que un puerto deba clasificar los paquetes, para servicios previamente definidos, con el fin de poder proveer garantías de QoS.
- ❖ Determinar el protocolo que necesite ejecutar un puerto para el enlace de datos (Serial Line Internet Protocol (SLPID) y Protocolo Punto a Punto (PPP) o Protocolo de transmisión de punto a punto (PPTP)) y realizar la búsqueda de ruta; cuando ya este hecha, envíese el paquete al puerto de salida usando la switching fabric. Si el router esta alineado a las entradas, varios puertos de entrada deben compartir la fabric.
- ❖ La función final de un puerto de entrada es regular el tráfico de paquetes de información para los protocolos para compartir este recurso común.

3.1.2. El switching fabric:

Existen muchas técnicas para implementar el switching fabric, dependiendo como este configurada su matriz de croconexión, sabemos que las mas comunes son la configuración en red Bennes, red Delta [2] y la más eficiente y mas utilizada por las tantas razones ya mencionadas en el capitulo 2 seria la red Crossbar. Un crossbar puede ser considerado como buses de $2N$ acoplados por $N*N$ puntos de intersección. Si un punto de intersección esta encendido, los datos en un bus de entrada se hacen disponibles a un bus de salida; de lo contrario, esto no sucede. Sin embargo, el encender y apagar los puntos de intersección por cada asignación de paquetes transferidos a través del crossbar esta a cargo de la unidad central de proceso, limitando éste la velocidad del crossbar switch fabric pues se comportaba como un cuello de botella. La solución a dicho problema fue el switcheo de dirección descentralizada como ya antes habíamos mencionado antes en el capitulo 1.

En un router de memoria compartida (los paquetes de entrada se guardan en un dispositivo de memoria), solo los apuntadores de paquetes son switchados, aumentado la capacidad de swicheo. La velocidad de acceso a la memoria limita la velocidad del switch. Los tiempos de acceso a la memoria disminuyen un 5%, mientras que la capacidad de la memoria se duplica anualmente.

También para conocer la eficiencia del router con que se esta trabajando es de suma importancia conocer si se esta trabajando con un switch fabric que posee buffer o sin buffer sabiendo que esto depende de cómo es la interacción del switch con la memoria compartida. Saber si es un buffer que encuentra a la entrada o a la salida y si este switch fabric es con bloqueo o sin bloqueo, estos conceptos ya fueron tratado en el capitulo 1.

3.1.3. Los puertos de salida

Antes de que sean transmitidos al enlace de salida guardan los paquetes. Para dar soporte a las garantías y prioridades implementan sofisticados algoritmos. Los puertos de salida, como los puertos de entrada, también necesitan darle sostenimiento a la encapsulación y decapsulación dentro del estrato de enlace de datos, además de una variedad de protocolos del nivel superior.

3.1.4. El procesador de routing

Computariza la tabla de reenvío, estableciendo protocolos de conmutación de paquetes para llevar a cabo el enrutamiento, y realiza el software para manejar y configurar la router. Ejecuta una serie de movimientos para cualquier paquete, cuyo destino, no puede ser encontrada en la tabla de reenvío en la tarjeta de línea.

3.2. Asuntos de diseño

Con la explicación de un router genérico en mano, estableceremos nuestra atención en aspectos de diseño para routers de backbone, de acceso y su misión.

3.2.1. Backbone routers

El Internet actualmente tiene unas pocas decenas de backbones, que sirven para una pequeña cantidad de miles de redes más pequeñas. Una de las ventajas de los routers de backbone es que su costo es compartido entre una gran base de clientes mediante la interconexión de redes empresariales. La fiabilidad y velocidad son los asuntos principales en el diseño de backbone routers, ya que el costo de enlaces de transmisión en un área amplia es tan alto actualmente que este se vuelve un asunto secundario.

La fiabilidad del hardware en backbone routers, se puede obtener utilizando algunas de las mismas técnicas usadas en los telephone switches.

Las piezas calientes, los suministros de fuerza duales y los caminos de duplicación de datos a través de los routers, son estándares en todos los routers de high_end, por lo cual pondremos mayor atención en el desarrollo de técnicas para obtener el routeo de alta velocidad.

El tiempo tomado para buscar una ruta en la tabla de reenvío, es donde se encuentra el cuello de botella principal en backbone IP routers. Un puerto de entrada al recibir un paquete, para indicar como llegar al puerto de destino del mismo, busca la dirección de destino en su tabla de reenvío, almacenando entradas de routeo como la dirección de red, la máscara y el puerto. El puerto se cicla conceptualmente mediante todas sus entradas de reenvío al recibir un paquete con dicha dirección. El router enmascara el paquete, con la máscara asignada a esa entrada, suma puertos para la asignación de puertos candidatos para destino, si se adjunta con la dirección correspondiente de la red. Decimos que trabaja con el prefijo de encuentro más largo, cuando el destino seleccionado es el puerto candidato correspondiente al enmascarado más largo.

Un ejemplo que se puede mencionar de lo discutido anteriormente podría ser un router que tiene tres diferentes entradas dentro de su tabla de routing, recibe un paquete con una dirección de destino que hace juego con las tres entradas, entonces, la asignación de candidatos de puerto de destino del paquete es (1, 3, 5). Sin embargo, el puerto 3, corresponde para la entrada de routeo con la máscara más larga, por consiguiente, el destino del paquete es puerto 3. Lo anterior ejemplifica las dos razones por las que es difícil buscar una ruta; ya que en primer lugar la tabla de routing puede tener miles de entradas, y en segundo lugar el paquete entrante puede hacer juego con múltiples entradas de routeo. Por lo anterior, se debe encontrar la entrada que hace juego más ampliamente.

Un caché de destinos frecuentemente visitados, es ineficaz, si los paquetes son pequeños o los paquetes son dirigidos para un gran número de destinos, aumentando el costo de búsqueda de ruta.

Para evaluar la situación de los backbones actualmente en uso, se debe considerar que el 40% aproximadamente de los paquetes tienen 40 bytes de largo (paquetes de reconocimiento TCP). Este porcentaje significa que en el presente los backbone routers en uso necesitan realizar un gran número de búsquedas de ruta a cada segundo.

Cada router tiene un cuello de botella adicional de desempeño, cuando se encuentran alineados al ingreso o a la salida, compartiendo el cuello de botella de búsqueda de ruta. Los switches alineados a las salidas deben ejecutar el swicheo Fabric a una velocidad mayor que la suma de los enlaces entrantes, lo cual puede ser solucionado construyendo redes de interconexión de tercera generación, pero deja el problema de almacenar paquetes en los buffers de salida.

Los tiempos de acceso DRAM o SRAM, delimitan la velocidad en que un router alineado a las salidas puede ser aplicado, limitando la tasa en la cual un buffer de salida puede ser accesado, sin embargo, colocando todo alineado en la entrada se puede sortear este problema, con este acercamiento, un controlador debe resolver lo contenido en el switching fabric y la línea de salida.

La estabilidad y la fiabilidad en la implementación de protocolos de routing, afectan críticamente la dimensionalidad del Internet. Se conocen redes estables en donde los routers utilizan diferentes versiones del mismo protocolo, o bien que corran enteramente otros protocolos. Existen problemas serios y casi no detectables cuando se registran mínimos cambios en la configuración de la red. La causa de un gran número de problemas de routing en el Internet, son la descripción de filtros de exportación (importación) que son rutas de un protocolo de portal de interior de acceso para BGP. Los pequeños problemas que no son resueltos en el Software cuando se implementa el protocolo o la mala configuración de routers, afectan a la caracterización del Internet debido a oscilaciones en el enrutamiento.

3.2.2. Misión del router

La meta principal es la de proveer conectividad para un gran número de puntos finales a menor costo. Además de dar soporte a las calidades de servicio, los routers son deseables ya que permiten al menos garantías QoS, para el tráfico confinado al área local.

La mayoría de redes para aplicaciones empresariales actualmente se construyen con segmentos Ethernet conectados por ejes o puentes, los cuales son baratos y de fácil instalación, y no necesitan ser configurados. Hay usualmente poco soporte para la diferenciación de servicio ya que no solo el desempeño de una red construida con ejes y puentes se degenera con el tamaño de la red. Contrariamente a una red construida con particiones de routers en las máquinas con dominios múltiples de colisión, escala de mejor manera con el tamaño de la red. La mayoría de routers permiten múltiples niveles de prioridades soportando alguna diferenciación de servicio. Tienden a ser más caras por puerto y necesitan ser configuradas antes de ser usadas ya que tienen varios requisitos adicionales de diseño por lo que los diseñadores deben solucionar las metas conflictivas de diseño debiendo proveer una riqueza de rasgos establecidos para cada puerto.

Las redes backbone tienden a soportar sólo el protocolo IP. Contrariamente las redes empresariales pueden transmitir una cantidad significativa de tráfico de multi-llamadas y transmisiones eficazmente, ocupándose del legado de tecnologías LAN, debiendo soportar múltiples protocolos, incluyendo IP, IPX y Vines, características como firewalls, filtros de tráfico, las extensivas políticas administrativas y de seguridad, y los LAN virtuales; conectan un puñado de trunks y proveen un gran número de puertos.

El desafío es construir routers empresariales que posean un gran número de puertos, que sean fáciles de configurar y soportar QoS y que tengan un bajo costo por puerto.

3.2.3. Routers de acceso

Las redes de acceso han sido poco más que un grupo de módems agregados a concentradores de terminales, prestando servicio a un gran número de dial up de velocidad lenta el cual se está alterando. En primer lugar, el acceso a las redes usan una colección variada de tecnologías, tales como: módems de alta velocidad, ADSL, y cable módems. En segundo lugar, para acceder a Internet desde el hogar usando líneas telefónicas, ha aumentado la carga en la red telefónica, creando problemas para switches de teléfono debido a la conexión larga que se soporta para conexiones de dial up de Internet, existiendo una presión considerable en las redes de acceso de la red telefónica subadyacente, tratando de bordear el switch de voz para las llamadas con datos. En tercer lugar, los routers de acceso proveen no solo un SLIP o una conexión PPP, sino que también protocolos de red privada virtual como PPTP e IPSec, [8], siendo necesario que estos protocolos puedan ser aplicados en cada puerto del router. Finalmente, se fomentará un incremento en la carga de las routers de acceso, mediante tecnologías como circuito de suscriptor digital asimétrico (ADSL) incrementando el ancho de banda disponible en cada casa.

Por estas cuatro tendencias, los routers de acceso tratarán de bordear el switch de voz de ser posible, pronto necesitarán soportar un gran número de puertos heterogéneos, potencialmente de alta velocidad y una gran variedad de protocolos funcionando en cada puerto.

3.3. Recientes avances y tendencias

Por la creciente demanda a nivel mundial del uso de la Internet, los diseñadores han puesto más empeño en el desarrollo de routers que soporten todas las demandas puestas por los usuarios, presentando a continuación ciertos avances de las nuevas tendencias en el diseño de routers.

3.3.1. Router de búsqueda de alta velocidad

La necesidad para computar la unión prefijada más larga para cada paquete entrante, vimos que era el mayor cuello de botella en las backbone routers.

La velocidad de un algoritmo de búsqueda de la ruta, esta determinada, por el número de accesos de memoria que requiere encontrar la entrada que hace juego con la memoria y la velocidad de la memoria.

Una segunda consideración en diseñar tablas de reenvío de datos, es el tiempo que toma actualizar la tabla. Estudio recientes muestran que una tabla de routing se altera relativamente despacio, actualizándose solo alrededor de una vez cada dos minutos, permitiéndonos usar estructuras más complicadas de datos cuyo objetivo es optimizar la búsqueda de la ruta a expensas del tiempo que toma actualizar la tabla de routing.

El estándar de las estructuras de datos para guardar rutas es un árbol, donde cada camino (trunk) que recorre de la raíz a las hojas se convierte en una entrada en la tabla de reenvío, en la que el prefijo de la dimensión más larga para combinar es el camino más largo, haciendo juego con la dirección de destino de un paquete entrante. El algoritmo basado en árboles empieza en la raíz del árbol y hace juego en forma recursiva con los vástagos del nodo en uso y con unos pocos bits de la dirección de destino, deteniéndose si no encuentra ninguna combinación. Para encontrar la combinación prefijada más larga ocurre el peor caso pues su tiempo es proporcional al largo de la dirección de destino. La idea en un algoritmo basado en árboles es que la mayoría de nodos precisan almacenamiento para unos pocos vástagos en lugar de todas las posibilidades, haciendo uso en forma escasa y poco escogida de memorias a expensas de tener que buscar más memorias. El algoritmo del comúnmente usado árbol-Patricia regresa para encontrar la combinación más larga, conduciendo a obtener un pobre desempeño en el peor de los casos. La decisión equivocada de diseño, está precisamente en la caída de los precios de la memoria.

El desempeño de algoritmos de búsqueda de ruta puede mejorarse de varias maneras. Clasificándose estas técnicas de la siguiente manera:

3.3.1.1. Las técnicas orientadas en hardware

Las soluciones orientadas en hardware se basan en memorias caches y memorias direccionales en contenido (CAMs). Estas técnicas tienen un escalamiento pobre con la dimensión de la tabla de routing, y no pueden servir para routers backbone que soportan tablas grandes de routing. Para reducir drásticamente el tiempo de acceso de memoria, en hardware esencialmente se combinan lógica y memoria en un solo dispositivo. Este acercamiento de la “memoria inteligente”, es muy utilizado en conjunción con las técnicas de software. Incrementar la cantidad de memoria usada para guardar la tabla routing, es una segunda solución orientada al hardware. Como los costos de memoria caen paulatinamente con el paso del tiempo, este acercamiento podría ser mejor aún para routers empresariales, sin embargo, un problema sutil se da con este acercamiento, es que la tabla se vuelve muy trabajosa de actualizar, ya que cambiar una sola entrada de reenvío puede causar varias miles de posiciones de memoria para estar actualizado.

3.3.1.2. Técnicas de table compaction

Para construir una estructura complicada pero compacta de datos para la tabla de reenvío, las técnicas de table compaction, saca provecho de la escasa distribución de entradas de reenvío en el espacio de todas las direcciones posibles de la red. La tabla se guarda luego en el caché primario de un procesador, permitiendo la búsqueda de la ruta a una velocidad de gigabit.

3.3.1.3. Técnica de hashing (despedazar)

Las técnicas basadas en despedazar han sido propuestas también para búsquedas de ruta, ya que para determinar la combinación de prefijo más larga, hay necesidad de limitar el uso de pedazos. Para encontrar la combinación más larga, no sabemos que prefijo usar, dada una dirección de destino. Para solucionar este problema debemos probar diferentes enmascarados, seleccionando el que tenga el enmascarado más largo.

La elección de máscaras puede hacerse mediante la primera parte de los pocos bits de la dirección utilizada para encontrar una lista de prefijos largos o en forma iterativa o jerárquica, sin embargo, ninguna de estas soluciones escala bien con la dimensión de destino.

El algoritmo hace una búsqueda binaria en los largos del prefijo, en lugar de ingenuamente ir en busca de un pedazo exitoso a partir del prefijo posible más largo, requiriendo que las tablas de pedazos lleven dentro de si rotuladores que, apunten hacia la tabla de pedazos del largo más pequeño correcta para la búsqueda.

Para una entrada particular de reenvío, el camino de búsqueda, se guarda compactamente en forma de una cuerda lo cual disminuye los requisitos de almacenamiento para rotuladores. Al pre-computar las tablas de pedazos que mantienen firmemente todas las entradas de reenvío asociadas con cada prefijo de 16 bits, pueden reducir el número de accesos de memoria a un promedio de dos por búsqueda, mutando la tabla de pedazos al vuelo.

3.3.2. Avances en switch fabrics

Los switch fabrics son usualmente implementadas como un larguero, comparten memoria o autobús. A mediados de los 80s pocos fueron los diseños de switch fabric que se construyeron a pesar de que se realizó un montón de investigación, porque los avances en las velocidades de bus los hacían innecesarios, incluyendo los bien conocidos de la familia fabrics Delta, junto con otros como los Bennes, siendo revividos en su mayor parte para los grandes switches de ATM, en los 90s. Con la disminución de la demanda para ATM, las routers IP están siendo construidas por la segmentación de wrapping (empacamiento de paquetes) y los módulos de reagrupación alrededor de los núcleos de switch fabric ATM. En estos routers los circuitos virtuales son establecidos desde un puerto a todos los demás puertos. Las celdas de ATM, son reensambladas al puerto de salida antes de la transmisión.

Para superar los problemas asociados con el switching de paquetes de tamaño variable, usar un núcleo de ATM permite a la router soportar diferentes corrientes de QoS en la switching fabric. Estos diseños, sin embargo, heredan algunas de las desventajas del ATM. Primeramente, en la mayoría de los casos los switches ATM, no tienen buen soporte (actualmente el multicast no es ampliamente soportado en backbone). Los esfuerzos como el IP Multicast Initiative (IPMI) han variado esto, haciendo que soporte del multicast sea un rasgo importante para backbone routers. Estos problemas generales reducen la eficiencia de la switching fabric. En segundo lugar, un problema más sutil se origina porque los algoritmos de regulación de tráfico IP son usualmente especificados en términos de paquetes en vez de términos de celdas.

Contar con switch fabrics basadas en celdas, implementando semántica como las requeridas por filtros compartidos en RSVP, es algo difícil de conseguir.

Para solventar problemas como este dados en las redes de comunicación basadas en IP, se han desarrollado, un tipo de routers modernos conocidos como LSR.

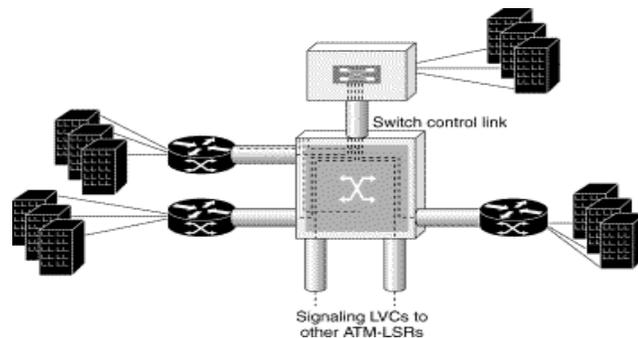
3.3.3. Label switch router (LSR)

Los LSR son un tipo especial de routers que trabajan dentro de un MPLS (Multi-Protocolo Label Switching), la cual es una nueva arquitectura de Red, base para el trabajo del Internet en el futuro. Esta propuesta se estandarizo a principios del 2001, en el seno de IETF (Internet Engineering Task Force). Los LSR son capaces de realizar el enrutamiento de paquetes determinado básicamente por la etiqueta que es asignada por el MPLS. Esta etiqueta sirve para identificar si el paquete pertenece a determinado ruta o no. Para establecer la relación entre las rutas de los paquetes y la etiqueta de los mismos, se utilizan los LSR que se encuentran en los extremos, que también son llamados LERs (Label Edge Routers). Cada uno de los LSR de la red se encarga de analizar la etiqueta de los paquetes recibidos para determinar el enlace de salida por el que deben ser retransmitidos. Debido a que al salir del grupo de LSRs la información debe de llevar una nueva etiqueta, estos también se encargan de establecerla. Toda la información referente a la ruta, la etiqueta de entrada y la etiqueta de salida es almacenada en unas tablas de conmutación que generalmente se encuentran en todos los nodos del circuito y deben de ser inicializadas antes de enviar cualquier paquete por el medio.

Como se vio anteriormente, a cada paquete se le asigna un camino especificado entre los LSR de bordes. A este camino se le llama LSP por sus siglas en inglés (Label Switching Path). Que no es más que un camino virtual por el cual debe de viajar la información.

Hay varias fases dentro del arreglo de los LSRs. En la primera fase, es donde se establecen los LSPs por los que van a ser transmitidos los paquetes. La forma en que lo hacen es que primero envían un mensaje hacia el LSR de destino y este al contestar el mensaje, sigue el camino inverso por el cual el paquete va a ser enviado desde el primer LER. Cada uno de los nodos de este camino, empezando del último, van enviando el mensaje y al mismo tiempo guardando la etiqueta de envío de los paquetes hasta que llegan al primer nodo.

Figura 21. Composición Interna del LSR.



Fuente: Cisco Systems

En la figura 21. Se muestra un arreglo de LSRs con sus respectivas conexiones. Como puede observarse, este arreglo cuenta con un switch control link, el cual es el corazón del circuito. Este switch no es más que un switch de paquete rápido; por la forma de funcionamiento del LSR. Esto se hace obvio cuando se observa, la descripción de este dispositivo de switcheo y como se acopla a el funcionamiento de la arquitectura MPLS. Por lo que se puede decir que este moderno router esta desarrollado bajo la arquitectura de un router tradicional en principio. Su diferencia más significativa es que el router label switching puede trabajar en el nivel 2 y nivel 3 del modelo OSI, pero de esto hablaremos en el capítulo 4, donde se busca profundizar más en la tecnología MPLS.

4. TECNOLOGÍA MPLS (MULTI-PROTOCOL LABEL SWITCHING)

4.1. El camino hacia la convergencia de niveles: IP sobre ATM

A mediados de los 90's se prefería la utilización de la ATM en comparación con la IP. Debido a que esta ATM fue diseñada para su utilización en equipos de bajo costo, a diferencia de la IP que posee la posibilidad de enviar paquetes de longitud variable y con un encabezado más largo. Sin embargo, en los últimos tiempos, los avances tecnológicos en el tema del enrutamiento, le han dado una mayor oportunidad al IP de sobresalir sobre la ATM.

Si se utilizan algoritmos de búsqueda rápida de enrutamiento, los routers pueden hacer match de longitud larga de una manera muy eficiente, es decir tan rápida y barata como si se utilizaran VCIs. Además, si la información se divide en paquetes de IP con un tamaño constante en los puertos de entrada y salida, en los routers se pueden utilizar los switch fabric que son muy similares a los de ATM.

Gracias a los avances de diseño del IP, se han reducido los costos de los algoritmos de planificación lo suficiente, para que los gastos fijos para planificar el tamaño variable de los paquetes se haya reducido. Con esto se prevee que en un futuro muy cercano el costo por megabit de switcheo por segundo del ancho de banda del IP sea ligeramente mas baja que el de la ATM, lo que provocará que la utilización de la ATM quede relegada únicamente a papeles muy limitados en las redes, principalmente como una tecnología de enlace de datos para la telefonía y para los portadores de ADSL.

4.2. Un paso más en la convergencia hacia IP: Conmutación IP

Debido a las ventajas que se proveen con la utilización del IP, se ha desencadenado una convergencia hacia esta de todas las aplicaciones existentes. Esas técnicas se conocen como "conmutación IP" (IP switching) o "conmutación multinivel" (multilayer switching). Estos sistemas tienen el problema de la falta de interoperatividad, debido a que usan diferentes tecnologías para combinar la conmutación de nivel 2 con el enrutamiento de IP de nivel 3. A continuación se describirán brevemente los fundamentos de las IP y multilayer switching, los cuales se basan principalmente en dos componentes básicos:

- ❖ A estos sistemas les era muy difícil realizar por separado funciones de envío (forwarding) y funciones de control (enrutamiento).
- ❖ Creencias e ideas preconcebidas en el contexto de intercambio de las etiquetas para la transmisión de los datos.

La componente de control utiliza protocolos estándar de enrutamiento para el intercambio de información con los otros routers para la construcción y el mantenimiento de las tablas de enrutamiento. Al llegar los paquetes al router, la componente de envío busca en la tabla de envío, para tomar la decisión de la ruta de cada paquete.

Es importante que todos los componentes de la red posean cierto nivel de visión. Los algoritmos de intercambio de etiqueta deben estar hechos de manera que permitan una conectividad entre los extremos de una red. Sin ellos no tendría conectividad.

4.3. La convergencia real a MPLS

4.3.1. Ideas preconcebidas sobre el MPLS

Debido a la interoperatividad entre productos de diferentes fabricantes, se ha buscado establecer un estándar y durante el tiempo en que se ha desarrollado dicho estándar, se han creado algunas ideas un tanto equivocadas sobre el avance, aplicabilidad y objetivos de los MPLS. A continuación se habla de los objetivos de la elaboración del estándar:

Por el crecimiento constante de la Internet, lo primero que se buscaba de la tecnología MPLS era su capacidad de crecer con esta, trabajando siempre bajo los lineamientos del Modelo de Servicios Integrados del IETF. Por lo cual MPLS tendría que ser capaz de soportar y actuar sobre cualquier tecnología de transporte. El envío de todos paquetes debería poderse realizar en unicast y multicast, siendo desde luego compatible con todos los requerimientos de las redes que trabajan sobre IP.

Si se observa con detenimiento el funcionamiento de los sistemas actuales de hosts finales, se concluye que estos no han implementado MPLS, ya que necesitan enviar los paquetes a un dispositivo de red de nivel 3 para que este pueda identificar la cabecera del paquete para después tomar las decisiones correspondientes sobre el destino final del mismo. Es en esta etapa es donde puede decidirse si se envía el paquete por enrutamiento convencional o si se le asigna una etiqueta para enviarlo por un LSP.

Las etiquetas MPLS únicamente tienen alcance local. De la misma manera, el último LSR de un LSP utiliza enrutamiento de nivel 3 para entregar los paquetes a su destino después que se ha eliminado la etiqueta.

4.3.2. Descripción funcional del MPLS

Con esta arquitectura se pretende, que la administración de tráfico del nivel dos y la dimensionabilidad y la capacidad de acceder a cambios del nivel 3, trabajen en conjunto.

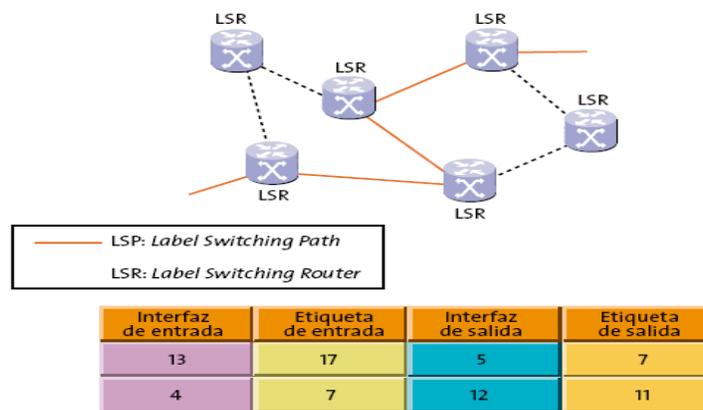
Se introduce una etiqueta entre el nivel 2 y el nivel 3. Con esto este protocolo es capaz de realizar la conmutación de paquetes IP. Lo que se busca es una administración de los recursos, en función de la habilidad de transmisión de extremo a extremo con la que la red cuenta, haciendo que la información viaje en datagramas a través de circuitos virtuales que se establecen entre dichos extremos de red conocidos en esta como LSP. Esta etiqueta evita procesos a nivel de red como algoritmos de encaminamiento necesarios para la conmutación de los datagramas de IP [4].

4.3.2.1. Funcionamiento del envío de paquetes en MPLS

Los MPLS se basan principalmente en la asignación y el intercambio de etiquetas para el establecimiento de las rutas LSP por la red. Existen dos tipos básicos de LSP, los simples y los duplex. Estos últimos requieren la utilización de dos LSPs, uno en cada dirección. Los LSP se crean a base de la concatenación de uno o más saltos en los que se van intercambiando las etiquetas de los paquetes, y estos se van enviando de un Label switching router (LSR) a otro utilizando como medio de MPLS.

Así como las soluciones de conmutación multinivel, los MPLS separan las componentes funcionales de control y de envío. La forma de implementar el envío es mediante el intercambio de las etiquetas de los paquetes en los LSPs de MPLS, RSVP o LDP (un nuevo estándar de señalización), ver fig. 22.

Figura 22. Ejemplo de Red MPLS.

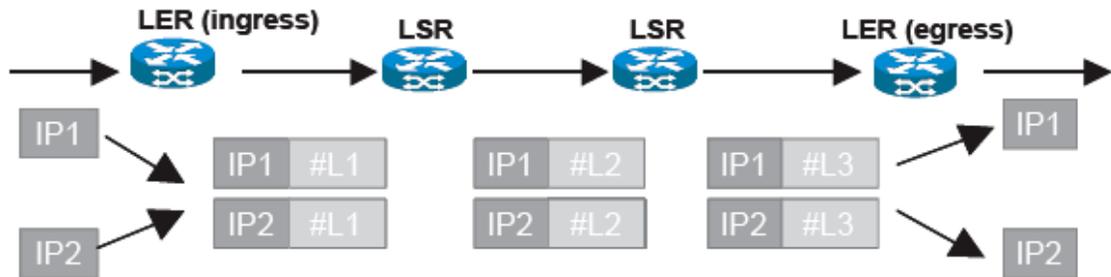


Fuente: /www.tdx.cesca.es/TESIS_UPC/AVAILABLE/TDX-0323104-125205//Chapter1.pdf

Debido a la utilización del procedimiento de intercambio de etiquetas MPLS, no hay necesidad de administrar dos arquitecturas diferentes para la transformación de las direcciones IP y las tablas de enrutamiento, el papel de la ATM queda reducido únicamente al transporte de datos a base de celdas.

Un camino LSP es un circuito que siguen todos los paquetes asignados a la misma FEC. Se necesitan dos LSR en un LSP, los cuales están en el exterior del dominio MPLS. Al primero se le denomina de entrada o de cabecera y al último de salida o de cola, también existe el LSRs que se encuentra internamente en el dominio de MPLS.

Figura 23. Ejemplo de Forward Equivalent Class



Fuente: /www.tdx.cesca.es/TESIS_UPC/AVAILABLE/TDX-0323104-125205//Chapter1.pdf

FEC es un grupo de paquetes de IP que se reenvió sobre el mismo LSP y tratado de la misma manera como si fuera un solo paquete con una etiqueta por un LSR aun si los paquetes difieren en su información del encabezado de nivel de red, ver fig. 23. [4]

Se podría decir que un LSR es un tipo de router especializado en el envío de los paquetes que son etiquetados por los MPLS.

En cada router, a partir de toda la información de enrutamiento se elaboran tablas de información, la cual brinda elementos de control. De igual manera, un LSR trabaja sobre este tipo de tablas envió a través de intercambio de etiquetas en la tecnología MPLS.

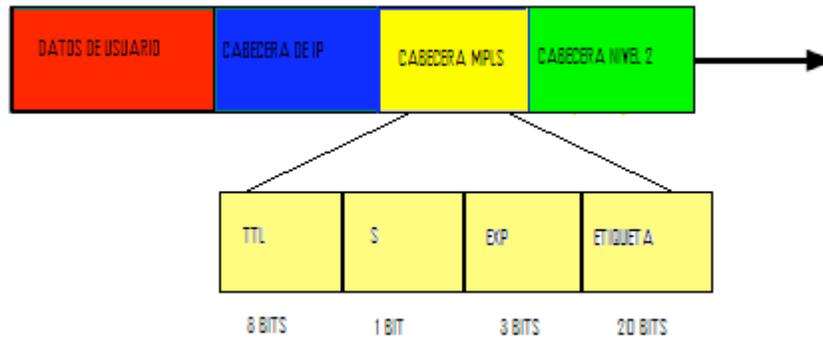
Un LSR funciona a base de intercambiar etiquetas según una tabla de envío. Esta tabla se construye a partir de la información de encaminamiento que proporciona la componente de control. Las tablas se encuentran elaborados con pares de etiquetas, y su función es la de acompañar a los paquetes que llegan por la interfaz con una etiqueta análoga. Existe una etiqueta de entrada de cola y una de salida en la cabecera.

Los paquetes son clasificados para poder asignárseles una cabecera en el LSR, pues a su entrada los paquetes son recibidos sin etiqueta. Es dentro de este donde las etiquetas son asignadas y este procedimiento como ya sabemos se hace a través de las tablas de encaminamiento, y cada paquete es asignado a una FEC. Después de esto el paquete es reenviado a otro LSR dentro de un LSP. En la tecnología MPLS los LSR se basan en el algoritmo de sustitución de etiquetas, pues son las etiquetas las que son estudiadas a través de las tablas de alternación de etiquetas. Cuando la información sale de la red MPLS la etiqueta que el paquete lleva es consultada en la tabla y es quitada, entonces el paquete es reenviado por un router convencional. Se puede notar que en la operación de MPLS la cabecera IP es prácticamente ignorada, pues se encuentra enmascarada, y toma de nuevo su valor cuando el paquete es transmitido fuera de la red MPLS.

MPLS cuenta con cabeceras en donde se insertan las etiquetas, entre los niveles 2 y 3, si hablamos de niveles desde el punto de vista del modelo OSI, sabemos que el nivel 2 es la capa de enlaces de datos la cual se ocupa del direccionamiento físico de datos, ofreciendo un tránsito confiable de un enlace, y el nivel 3 que es la capa de red la cual se centra en la conectividad y en la selección de rutas entre dos puntos que pueden estar en redes distintas. Podemos decir, que el nivel 3 agrega un encabezado de red y el nivel 2 agrega un encabezado de trama a un paquete de datos [5]. Como el nivel 3 ya posee un campo para etiquetas, estos son los que utilizamos, en cambio en el nivel 2 no es posible trabajar con campos que soporten etiquetas por los que en la tecnología MPLS, se creó una cabecera de 4 octetos, la cual va en medio de la cabecera del nivel 2 y la del nivel 3, la cual cuenta con un campo determinada exclusivamente para la etiqueta.

A continuación se presentará como esta estructurada la cabecera en MPLS a través de una gráfica. Fig. 24, en donde se puede apreciar que por medio de esta se puede hacer una cohesión entre tecnologías de transporte. Esto da la facilidad de extinción de las redes a los que proveedores de IP.

Figura. 24. Estructura de la Cabecera MPLS



4.3.2.2. Control de la información en MPLS

El manejo de la información en MPLS esta basada en dos aspectos fundamentales que trataremos a continuación.

¿Cuál es la forma de generación de de las tablas de envío, las cuales son lo cimientos de los LSPs?

¿Como las etiquetas influyen en la distribución de los datos a los LSRs?

Para poder responder a la primer pregunta, se debe tomar en cuenta la información que conocemos de la red: Por ejemplo su topología, estereotipo de tráfico, y las características que puede poseer dicho enlace. Toda esta información es de suma importancia en la tecnología MPLS pues son la base para la creación de los LSPs. Como vimos anteriormente la cabecera IP cuando el paquete viaja por una red MPLS, no es relevante pues la información viaja en función de la etiqueta. Lo que se hace es que por cada ruta de IP se crea una ruta de etiquetas, con el fin de unir las entradas con salidas en las tablas de los LSRs. Tales tablas son creadas a base de la información de conducción de un protocolo interno ya predeterminado es el que se encarga de pasar la información.

Para responder la segunda pregunta, centrémonos en el concepto de señalización, como es que se asigna una etiqueta entre LSR. Es muy importante mencionar que la tecnología MPLS no posee un protocolo de etiquetas único. Podemos mencionar el protocolo RSVP y el protocolo LDP, y el protocolo BGP.

- ❖ LDP (Label Distribution Protocol). Es específico de MPLS. Su extensión CR-LDP (Constrained Routing LDP) tiene en cuenta los requisitos del flujo y las condiciones de la red para establecer el LSP por el que debe ser transportado [6].
- ❖ RSVP (Resource reSerVation Protocol) con extensiones, Se amplía el protocolo de reserva RSVP, originalmente diseñado para realizar demandas de calidad de servicio extremo a extremo [6].
- ❖ BGP (Border Gateway Protocol) con transporte de etiquetas. Las redes privadas virtuales y la interconexión entre redes de proveedores de servicio se perfilan como las principales aplicaciones de este protocolo, que distribuirá etiquetas de segundo nivel de jerarquía dentro de los núcleos de red [6].

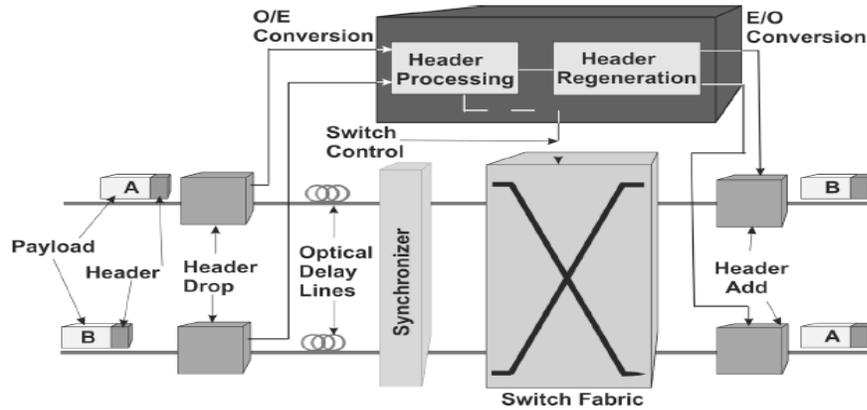
Para complementar lo anterior hay que hablar de la nueva tendencia. El GMPLS, acrónimo del término MPLS generalizado, es una propuesta de extensión del estándar MPLS. Su objetivo es integrar en un mismo plano de control la red IP y los conmutadores ópticos, de forma que el operador vea el encaminamiento óptico como una funcionalidad más de los routers IP, [4].

Podemos decir que la etiqueta de los LSRs puede ser de cuatro tipos principales. A continuación hablaremos de estos.

Más exactamente las interfaces de los LSRs, se pueden subdividir en las clases siguientes:

- ❖ Existe un tipo de LSRs, el cual su funcionamiento se basa en reconocer los límites del paquete por cada celda, en este tipo el destino a cual los datos son reenviados, depende de la cabecera que posee cada paquete por celda. A esto dispositivos se les denomina PSC (Packet switching capable). Como ejemplos de estos podemos citar, los routers que según la cabecera “Shim” transmiten los datos, y los LSRs de ATM, los cuales transmiten los datos según los VPI/VCI que son los identificadores virtuales de ATM.
- ❖ Existe otro dispositivo de LSRs, el cual se basa en el envío de los datos por ranuras, en un ciclo constante de tiempo, trabaja sobre comunicación síncrona haciendo uso de la Multiplexación de TDM, por lo que se le denomina (Time-División Múltiple Capable). Un ejemplo de estas interfaces es una interfase en un Cross-Connect de SONET/SDH.
- ❖ Se cuenta con un dispositivo LSRs, en el cual la longitud de onda con la cual los datos arriban a este, determina la forma en que estos son retransmitidos. Este tipo de LSRs son denominados LSC (Lambda Switch Capable). Como el más común de estos podemos mencionar la interfase de Cross-Connect Óptico, la cual funciona para distintas longitudes de onda independientes, ver fig. 25.

Figura 25. Lambda Switch Capable



Fuente: PhD Diptish Dey, Twente University **Press**

- ❖ Existe también una LSRs, en la cual de acuerdo a la posición que los datos poseen en el espacio físico del mundo real, determina la manera en que los datos son transmitidos. Este tipo de LSRs son denominados FSC (Fiber-Switch Capable). Un ejemplo de este tipo de tecnología podría ser el Cross-Connect Óptico el cual cuenta con la capacidad de operar con una o múltiples fibras.

En los LSPs (Label Switched Path) anidados, se trabaja sobre una jerarquía de envío, que va de las interfaces PSC, las cuales siguen a las interfaces TDM, las cuales siguen a las interfaces LSC, las cuales siguen a las interfaces FSC, cuando es visto de la de menor jerarquía a la de mayor jerarquía. Vemos que FSC es la de mayor jerarquía y PSC la de menor. Por lo anterior dentro de un camino LSP que empieza y termina en una interfase PSC, puede residir en un LSP que empieza y termina en una interfase TDM. Por turno este LSP puede residir en un LSP que empieza y termina en una interfase LSC, que por turno puede residir en un LSP que empieza y termina en una interfase FSC [7].

4.3.2.3. Funcionamiento global MPLS

En la tecnología MPLS, es mas simple la creación de caminos virtuales, los cuales se adaptan mas fácilmente, su importancia se basa en que los paquetes IP siempre se mantienen visibles y esto es una de las diferencias principales que posee con las redes convencionales. La función principal de MPLS es unir dos routers de dos redes distintas que se encuentran a larga distancia, y lo hacen a través de las ya mencionados LSPs o caminos virtuales de MPLS de un solo salto.

4.4. Aplicaciones del MPLS

4.4.1. Ingeniería de tráfico

Al seleccionar flujos por medio del algoritmo IGP para trasladar enlaces que son congestionados a los que se encuentran mas descongestionados, sin que importe que no sea la ruta mas corta. Estamos de hablando de ingeniería de trafico, la cual su función principal consiste en adaptar estos flujos a las propiedades de la red. Lo que se busca es que no haya suprautilizacion de recursos en un área de la red, mientras en otra haya subutilización, evitando de esta manera calentamiento y los problemáticos cuellos de botella.

- ❖ Se puede especificar exactamente el camino físico de un LSP para establecer rutas.
- ❖ Para evitar el problema de sobrecarga el en los enlaces y los cuellos de botella, se posible conseguir estadísticas de cuanto es el uso que tiene cada uno de los LSP. Estas estadísticas tienden a tener gran importancia a la hora del diseño de la red.

- ❖ Es posible asignar rutas para servicio, las cuales son posibles dejar en modo restringido.

Independientemente que se trabaje sobre la tecnología ATM, MPLS puede trabajar con ingeniería de tráfico sobre una red IP.

4.4.2. Clases de servicio (CoS)

La técnica que consiste en enmarcar los paquetes que son transmitidos a través de la red se conoce como QoS, es por medio de esta que es posible la diferenciación de servicios tales como, el WWW, el correo electrónico, y la transferencia de ficheros. Como la tecnología MPLS fue diseñada pensada con el fin de trabajar con servicios que se encontraran diferenciados, se adapta muy bien a esta técnica, pues cuenta con un campo llamado EXP, con el cual sobre un determinado LSP sea posible la propagación de la clase de servicio CoS.

Diferentes tipos de tráfico son transportados en MPLS:

- ❖ Dependiendo de la información en los bits en el campo EXP, en los LSR le son adjudicadas colas al tráfico que fluye por un determinado LSP.
- ❖ Es posible que existan LSP de distinta prioridad de tráfico. Podemos mencionar el de máxima prioridad, el de prioridad media, y el de tráfico best-effort, los cuales funcionan sobre niveles de servicio que poseen distinto precio para el usuario. Estos niveles son el primero, preferente y turista. Cada LSP trabaja a distinto ancho de banda, y son varios de estos los que se encuentran entre un par de LSRs.

4.4.3. Redes privadas virtuales

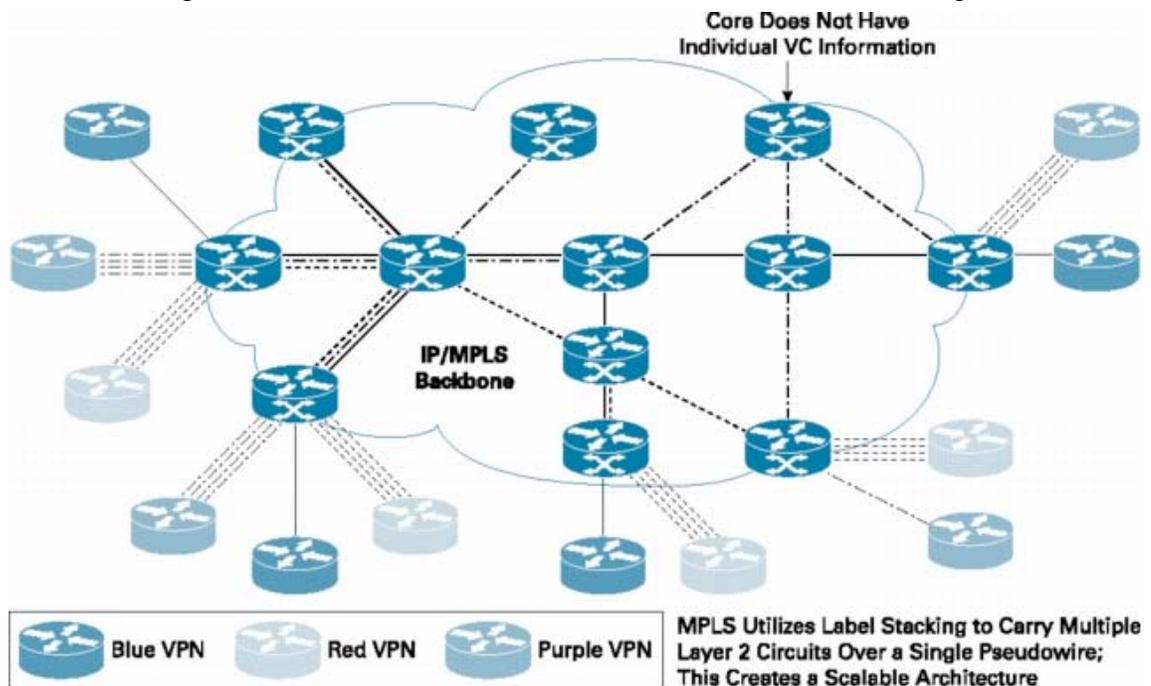
Como se sabe existe una gran cantidad de empresas, que poseen mas de una sucursal, las cuales se encuentran muy lejos unas con otras, y se aprecia más en el caso de las grandes transnacionales en que estas se encuentran en distintos países. La mayoría de estas empresas necesita poder tener conectadas todas sus sucursales, pero esto no es posible realizarse con una red tradicional, y una red WAM (Wide Área Network), significaría un problema económico por su gran costo. Por esta razón seideo, la manera de poder comunicar estas sucursales a través de las redes públicas, pero de una manera que fuese segura pues los datos que se transfieren son de mucho valor para la empresa, y el que estos fueran interceptados por ajenos representaría un problema mayor.

Una red privada virtual VPN, es un medio de interconexión por medio del cual, estas distintas sucursales se pueden conectar a través de una red de servicios publico, solventando el problema de la seguridad de los datos. Una VPN trabaja normalmente sobre el Internet, estableciendo enlaces seguros. Normalmente la VPN se trabaja sobre el soporte de la red ATM a través de los PVC. El principio es que dos puntos se enlazan a través de un IP y un esquema de encriptación de la información para la seguridad de la misma el cual es realizado por unos firewalls. Al final el comportamiento de una VPN es como el de una red tradicional, pero funciona para componentes remotos con funcionamiento transparente y seguro.

A pesar de que las VPN de túneles de IP sobre los PVCs, son funcionales, trabajarlos sobre la tecnología MPLS es muy eficiente y a continuación veremos el porque:

En las VPNs basadas en IP, lo que pasa es que su topología que es a base de túneles esta sobrepuesta a la topología existente del proveedor y hay una conexión entre todos los extremos de la VPN, mientras que en la tecnología MPLS, acopla la topología de la VPN a la red del proveedor, El acople que se da en MPLS permite que todos los usuarios se conecten a la VPN por medio de una nube común, la cual es implementada por medio de los caminos virtuales LSPs, trabajando como ya hemos mencionado tantas veces por medio de etiquetas. En este el encabezado de IP no posee ninguna importancia y no se ve en el proceso de encaminamiento de la información, el IP solo toma su valor al trabajar en el encaminamiento de una red tradicional, por lo que no habría ningún inconveniente en que este sea examinado por técnicas de QoS. Todo esto con el fin de aplicar a la VPN, técnicas de ingeniería de trafico, y creas CoS.

Figura 26. Red Privada Virtual, elaborada a base de la tecnología MPLS.



Fuente: www.Cisco.com

CONCLUSIONES

1. Un router es un dispositivo que trabaja sobre el nivel tres si nos referimos al modelo OSI, y su función es examinar la capa de red, con el fin de determinar el destino de los datos. Este trabaja bajo una comunicación estática con el fin de ser más eficiente a la hora de trabajar con fuentes asíncronas. Su arquitectura interna está basada en el switch de paquete rápido. La conexión es a base de configuración orientada y entre los componentes se encuentra el switch fabric, el cual influye grandemente en el desempeño del router. Se puede decir que la capacidad máxima se da cuando se utiliza un switch fabric con buffer de salida, sin bloqueo y configurado como una red crossbar, y de ahí la importancia de este dispositivo en el diseño de los routers.
2. Un LSR (Label Switching Router) es una nueva tecnología de router, al igual que el router genérico cuenta en su arquitectura interna con un switch de paquete rápido. Se diferencia del router convencional en que éste trabaja en el nivel dos y nivel tres del modelo OSI. Lo que éste hace es introducir una etiqueta entre el nivel dos y tres combinando eficazmente las funciones de control del routing con la simplicidad y rapidez de la conmutación. La etiqueta es asignada según los parámetros de la arquitectura MPLS, lo que le brinda la capacidad de realizar conmutación y enrutamiento de paquetes IP. Cada uno de los LSR de la red se encarga de analizar la etiqueta de los paquetes recibidos para determinar el enlace de salida por el que debe ser retransmitido.

3. Se puede finalmente decir, que MPLS es una tecnología de red que se basa en la asignación e intercambio de etiquetas, para el establecimiento de las rutas virtuales LSP a través de la red, haciendo uso de los LSR. Su bondad está en que en ésta no se necesita la administración de dos arquitecturas, para poder transformar direcciones IP y tablas de enrutamiento. Haciendo de esta manera a la tecnología ATM a un lado y consolidándose ésta como la base del Internet del siglo XXI.

RECOMENDACIONES

1. En este trabajo se desarrollaron conceptos sobre nuevas tecnologías como MPLS en las redes, así como de los dispositivos que intervienen en ésta. Por lo que es importante para La Escuela de Ingeniería Mecánica Eléctrica, promover investigaciones futuras, con temas acordes a éste.
2. Es de suma importancia que, la Escuela de Ingeniería Mecánica Eléctrica, cuente con un equipo de vanguardia para que la investigación se pueda llevar a niveles más elevados, pudiendo obtener de ella mayores resultados, pues esta falta de recursos es muchas veces una limitación para el estudiante.
3. Sería interesante que el gobierno, en conjunto con todas las universidades del país, crearan un centro de investigación nacional, donde se llevaran a cabo proyectos de investigación, en el que trabajaran profesionales de las distintas ramas. Esto haría que Guatemala empezara a ser un país productor en vez de consumista, además que se promovería la exportación y por ende el desarrollo.

REFERENCIAS

- [1] J H Patel. *Performance of processor-memory interconnections for multiprocessors*. IEEE Trans. Computers, **C-30** (10), 771-780, Oct. 1981.
- [2] S. Keshav, *An Engineering Approach to Computer Networking*, Reading, MA: Addison-Wesley, 1997.
- [3] Peter Newman, *Fast Packet Switching for Integrated Services*, Wolfson College, A dissertation submitted for the degree of Doctor of Philosophy. December 1988.
- [4] Sociedad de la Información. **La Red de Tránsito**. Publicaciones, www.telefonica.es, 2003.
- [5] Vito Amato. **Guía del Primer Año**. Academia de networking de Cisco Systems, 1999.
- [6] **MULTI-PROTOCOL LABEL SWITCHING**,
www.tdx.cesca.es/TESIS_UPC/AVAILABLE/TDX-0323104-125205//Chapter1.pdf.
- [7] L. Berger, *Generalized Multi-Protocol Label Switching (GMPLS) Descripción Funcional de la Señalización*. Network Working Group, Enero 2003.
- [8] **S. Keshav and Rosen Sharma**. Issues and Trends in Router Design. Cornell University, Mayo 1998.

- [9] Liu, Erik Dirkx. *Parallel Simulation of ATM Switches*. Vrije Universiteit Brussel, dept INFO Pleinlaan 2, 1997.

BIBLIOGRAFÍA

1. Kevin Deierling, Méllanos, *Advanced System Architectures Drive Choice of Switch Fabric Solution*, COTS JOURNAL, abril 2004.
2. S. Keshav, *An Engineering Approach to Computer Networking*, Reading, MA: Addison-Wesley, 1997.
3. Dpto. Ingeniería de Sistemas Telemáticos. **Aplicaciones Distribuidas Avanzadas**, Universidad Politécnica de Madrid. Curso de Doctorado 2004-2005.
4. **Comunicaciones de Telefónica I+D**, Telefónica Investigación y Desarrollo, Número 36, junio 2005.
5. Cisco Systems, *Documentation Product*, Copyright © 2004.
6. Tilman Wolf, *ECE 697J ECE 697J – Advanced Topics Advanced Topics in Computer Networks in Computer Networks*. Switching Fabrics, University of Massachusetts Amherst, febrero 2003.
7. Newman. Peter. *Fast Packet Switching for Integrated Services*. Wolfson College, A dissertation submitted for the degree of Doctor of Philosophy, December 1988.
8. Firoiu, Victor. *Feedback Output Queuing A Novel Architecture for Efficient Switching Systems*. Performance Engineering Center / Advanced Technology, Nortel Networks, August 21, 2002.

9. Vito Amato. **Guía del Primer Año**. Academia de networking de Cisco Systems, 1999.
10. Elio Salvadori, *International Doctorate School in Information and Communication Technologies*. Traffic Engineering in Dynamic Optical Networks, DIT - University of Trento. february 2006.
11. *Label Switch Router*. the free encyclopedia, Wikipedia, modified 9 February 2006.
12. Sociedad de la Información. **La Red de Tránsito**. Publicaciones, www.telefonica.es, 2003.
13. Kranti Deshpande, Sheetal Nalole, Shukra Shah , Prof. Rajesh Ingle *Department of Computer Engineering*, Ashutosh Kelkar, Hemant P. Kelkar; *Modeling of Standalone and Networked ATM switches (N-AtmSim) using Discrete Event Simulation with TCP/IP Sockets*; Pune Institute of Computer Technology (PICT), Department of Computer Science, AT College, Pune 411043, India.
14. Cisco Systems, **Módulo de red de enlace troncal digital T1/E1 de paquetes de voz**, Hoja de Datos, Copyright © 2001.
15. *MULTI-PROTOCOL LABEL SWITCHING*, www.tdx.cesca.es/TESIS_UPC/AVAILABLE/TDX-0323104-125205//Chapter1.pdf.
16. Bus Tronica an Elma Company, TreNew an Elma Company; *Next-Gen Backplane Architectures*.
17. Duato, José. Dept. of Computer Engineering (DISCA); *On the Design of High-Speed Switch Fabrics*; Universidad Politécnica de Valencia, Spain.

18. Liu, Erik Dirkx. *Parallel Simulation of ATM Switches*. Vrije Universiteit Brussel, dept INFO Pleinlaan 2, 1997.
19. Tenenbaum. Andrew S. **Redes de Computadoras**. Prentice Hall.
20. JabberES, **Redes Privadas Virtuales**, Actualización 19 de octubre de 1999.
21. de Miguel, Tomás P. **Redes Privadas Virtuales y MPLS**, DIT Univ. Politécnica Madrid, marzo del 2003.
22. Wernicke, John. *Simulative Analysis of QoS in Avionics Networks for Reliably Low Latency*, Journal of Undergraduate Research, February 2006.
23. Hybricon. *The Switch Fabric are Coming, and They Have Already Arrived*, Technology Focus, September 2002.
24. Regula, Jack. *Switch Fabric Architecture*. PLX TECHNOLOGY, the I/O Interconnect Solution, January 2001.
25. Saleem Bhatti, *Switches*, GMT, Actualizado en 1995.
26. Diptish Dey, *Theory Towards an all-optical WDM slotted-ring MAN support for optical multicasting*, PhD thesis, University of Twente, 2003
27. Barberá, José. **Una arquitectura de Backbone para la Internet del siglo XXI**. Red IRIS, Actualización octubre 2003.