

UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA



**DESCRIPCIÓN DE LAS TECNOLOGIAS DE REDES DE
ALMACENAMIENTO (SAN) Y SU APLICACIÓN EN LOS
SISTEMAS COMPUTACIONALES**

INFORME DE SUFICIENCIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ELECTRÓNICO

PRESENTADO POR:

LUIS MIGUEL SAKIHAMA MIYASHIRO

**PROMOCIÓN
2002- II**

**LIMA – PERÚ
2007**

**DESCRIPCIÓN DE LAS TECNOLOGÍAS DE REDES DE
ALMACENAMIENTO (SAN) Y SU APLICACIÓN EN LOS SISTEMAS
COMPUTACIONALES**

***Dedico este trabajo
A mi madre,
por su apoyo constante e incondicional
A Roxana,
por su sorprendente capacidad de amar***

SUMARIO

El presente trabajo pretende, a través de su desarrollo, explorar las tecnologías de redes de almacenamiento utilizadas en la actualidad y sus aplicaciones en las empresas de pequeña, mediana y gran envergadura.

En el capítulo I se plantea claramente el objetivo del informe, así como también se exponen las razones por las cuales este tipo de redes son cada vez más imprescindibles dentro de una empresa.

En el capítulo II se revisa la historia de las redes de almacenamiento desde sus inicios, hasta las diversas opciones disponibles hoy en día.

En el capítulo III, el más extenso, se exploran en detalle las tecnologías de redes de almacenamiento, revisando sus características, protocolos, interfases y modos de operación y, en la medida de lo posible, los estándares que los formalizan.

Finalmente en el capítulo IV se describen algunas aplicaciones principales de las tecnologías desarrolladas en los capítulos anteriores.

INDICE

PROLOGO	1
CAPITULO I	
INTRODUCCION	2
1.1. Objetivos del informe	2
CAPITULO II	
HISTORIA DE LA CONECTIVIDAD DE LOS SISTEMAS DE ALMACENAMIENTO	3
2.1. Los inicios	3
2.2. La conexión directa a servidores	4
2.3. Las redes de almacenamiento (Storage Area Network – SAN)	6
CAPITULO III	
DESCRIPCION DE LAS TECNOLOGIAS DE REDES DE ALMACENAMIENTO (SAN)	8
3.1. La tecnología SCSI	9
3.1.1. Definición de la tecnología SCSI	9
3.1.2. Los estándares SCSI	9
3.1.3. El modelo de arquitectura SCSI	12
3.1.4. La capa física del modelo SCSI	15
3.1.5. La capa de protocolos de transporte del modelo SCSI	16
3.1.6. La capa de aplicación del modelo SCSI	17
3.2. La interfaz SCSI paralela	19
3.3. La evolución de SCSI: Fibre Channel	21
3.3.1. La pila de protocolos Fibre Channel	21
3.3.2. Topologías Fibre Channel	30

3.3.3. Productos Fibre Channel	36
3.4. Fibre Channel sobre un tunel IP (iFCP)	45
3.5. Internet SCSI (iSCSI)	47
3.6. Conexión de Fibra para entornos Mainframe (FICON)	50
3.7. Almacenamiento en red (NAS)	50
3.8. Tendencias del mercado de almacenamiento actual	51
3.9. Resumen	52
CAPITULO IV	
APLICACIONES DE LAS REDES DE ALMACENAMIENTO EN SISTEMAS COMPUTACIONALES	53
4.1. Consolidación de un sistema de almacenamiento local	53
4.2. Replicación de un sistema de almacenamiento	57
4.3. LAN-free Backups	60
CONCLUSIONES	63
GLOSARIO	64
BIBLIOGRAFÍA	65

PRÓLOGO

En la actualidad si una empresa o institución se ve afectado por una interrupción o la degradación en los procesos de acceso a la información, ya sea desde agentes internos o externos de la institución, es muy probable que exista una pérdida económica considerable. Por ejemplo si debido a alguna falla en los sistemas de un banco sus clientes no pueden realizar pagos con sus tarjetas de crédito por varias horas, esta interrupción del servicio representa una pérdida que puede llegar a muchos cientos de miles o inclusive millones de dólares dependiendo de la severidad del problema. Si gran parte de la información de este banco llegara a perderse, sin duda representaría la quiebra total.

Con este sencillo caso es posible demostrar la tremenda importancia que tiene salvaguardar la accesibilidad y la integridad de la información y es por esta razón que existe una industria creciente en tecnologías de redes de almacenamiento y los sistemas de almacenamiento en sí.

En este documento nos centraremos solamente en las tecnologías de redes de almacenamiento y sus aplicaciones en la industria.

CAPITULO I

INTRODUCCION

1.1. Objetivos del informe

A partir de los años 90 el paradigma de la industria de almacenamiento ha cambiado radicalmente conjuntamente con la cantidad de datos manejados por los sistemas de cómputo. Por ejemplo en 1980 la cantidad aproximada de información que circulaba en la red como páginas web estáticas bordeaba algunos GB (GB=Gigabytes que equivale a 1000'000'000) a nivel mundial, mientras que para el año 2000 se estimaba en 170 TB (TB=Terabytes que equivale a 1000'000'000'000 bytes) en el 2002, ¡y solo en páginas estáticas! Para el 2003 se estimó que el crecimiento de información por año es de 5 EB (EX=Exabytes que equivale a 1000'000'000'000'000 bytes), es decir que cada año se crea esa cantidad de información en todo el mundo, cifra que aumenta año tras año. Este crecimiento explosivo en la cantidad de información almacenada ha revolucionado la forma cómo almacenar y acceder a la información desde los sistemas de procesamiento.

Este crecimiento no es casual; estamos dentro de un periodo en donde la información es el activo más importante para la economía, lo que algunos estudiosos llaman la “sociedad del conocimiento” o “sociedad de la información” por lo cual es indudable que, día tras día, desarrollar técnicas, redes y sistemas que permitan un acceso rápido y continuo a la información almacenada cobra cada vez mayor importancia. Sin embargo el concepto de redes de almacenamiento es generalmente desconocido y poco difundido. Este informe busca cubrir esa brecha y describir las distintas tecnologías para redes de almacenamiento disponibles en la actualidad y aquellas en desarrollo y sus aplicaciones en los sistemas computacionales, así como también las tendencias que rigen el mercado actual en donde los conceptos claves son la integración de tecnologías y la consolidación del hardware. Asimismo se realizarán comparaciones entre las distintas tecnologías en donde se resaltarán las ventajas y desventajas de cada una de ellas.

CAPITULO II

HISTORIA DE LA CONECTIVIDAD DE LOS SISTEMAS DE ALMACENAMIENTO

El concepto de redes de almacenamiento es relativamente nuevo, tanto es así que la primera referencia al término "Storage Area Network (SAN)" se realizó en 1999, hace solamente 7 años atrás.

El nacimiento de esta nueva tecnología parte de la necesidad cada vez más urgente de conectar muchos host o servidores a dispositivos de almacenamiento externo en disco y/o en cinta magnética lo cual se realizaba en forma uno-a-uno y no permitía compartir recursos y aumentaba la complejidad y el costo de los sistemas.

Durante el presente capítulo recorreremos la historia de las redes de almacenamiento desde los inicios en donde simplemente no eran necesarias hasta la actualidad en donde son indispensables en cualquier mediana-gran empresa.

2.1. Los inicios

El hombre a través de los tiempos siempre ha deseado dejar una huella de su paso por el mundo o simplemente almacenar en algún medio no volátil su conocimiento y sus vivencias, y lo ha logrado a través de medios tan disímiles como la piedra, el suelo, el papiro, el papel, etc.

Sin embargo si hablamos de almacenamiento de datos de computadores o de algún dispositivo mecánico-eléctrico propiamente dichos nos podemos remontar a 1884 en donde el ingeniero norteamericano Herman Hollerith utilizó las tarjetas perforadas para almacenar datos estadísticos. Las tarjetas perforadas fueron utilizadas también por las primeras computadoras en los años 40's y se mantuvieron en vigencia hasta la mitad de los 70's.

Mas adelante, en la búsqueda de una mayor flexibilidad y una capacidad de almacenamiento superior, IBM lanzó al mercado las cintas magnéticas que en un primer momento podían almacenar la información de hasta 10,000 tarjetas perforadas. Fue un éxito total y fueron el medio de almacenamiento más utilizado hasta la mitad de los años 80's.

Algunos años después, IBM nuevamente revolucionó el mercado con el lanzamiento del computador 305 RAMAC (Random Access Method of Accounting and Control) que contenía un sistema de almacenamiento en base a discos y podría almacenar hasta 5 millones de caracteres (4.4 MB).

2.2. La conexión directa a servidores

Desde la aparición de las cintas y discos magnéticos, estos crecieron en forma ininterrumpida, su capacidad y complejidad se fue incrementando año tras año, por lo cual se optó por separarlos del sistema de procesamiento para abstraer la complejidad del sistema de almacenamiento y poder además compartirlos, lo cual resultaba la solución de un problema y la aparición de otro: ¿cómo conectar el sistema de procesamiento con el sistema de almacenamiento?

La interfaz SCSI para servidores y workstations

En principio el problema se abordó utilizando una interfaz basada en el sistema CHS (Cilindro, Cabeza y Sector) para el direccionamiento de la información en los discos, sin embargo en 1979 Shugart Associates, un fabricante de discos, inventó una interfaz que soportaba un direccionamiento lógico de bloques en lugar del sistema que se venía utilizando y que además proporcionaba una interfaz paralela de 8 bits en lugar de las señales analógicas serie que utilizaban por entonces las controladoras, así como una serie de comandos genéricos en sustitución de las líneas de control que acompañaban a las líneas de datos. La interfaz fue denominada SASI (Shugart Associates Systems Interface), y su especificación incluía algunos comandos de 6 Bytes y una interfaz de terminación sencilla.

A finales de 1981, Shugart y NCR, un fabricante de ordenadores, presentaron la especificación al comité ANSI bajo el nombre de SCSI ("Small Computer System Interface") En los años siguientes se fueron agregando capacidades y funcionalidades

adicionales, incluyendo la capacidad de los dispositivos de competir por el uso del bus (arbitraje); desconectar y reconectar temporalmente del bus durante la ejecución de comandos, y tener más de un adaptador host en un bus SCSI.

En 1984 se presentó al ANSI (American National Standards Institute) un borrador con la propuesta del estándar cuando ya existían en el mercado muchos de los nuevos dispositivos. El estándar fue aprobado en 1986 como SCSI-1 que básicamente describía un bus de 8 bits que operaba a 2 MB/s y soportaba hasta 7 dispositivos dentro de un mismo bus.

La aparición de esta interfaz es sin duda uno de los hitos más importantes en la historia de las redes de almacenamiento y constituye la base de las nuevas tecnologías de almacenamiento. Incluso con más de 20 años de antigüedad esta interfaz se mantiene todavía vigente para algunos dispositivos de almacenamiento y se actualiza continuamente. En la próxima sección explicaremos en forma extensa las características físicas y lógicas de esta interfaz, ya que si bien algunas veces no es considerada como una red sino como un bus, será de mucha utilidad para entender las nuevas tecnologías.

A pesar de todas las ventajas brindadas por la interfaz SCSI, sus limitaciones se fueron haciendo evidentes cuando empezó a aumentar la cantidad de host o servidores que necesitaban conectarse a un dispositivo de almacenamiento externo. La interfaz SCSI, al tener una naturaleza de Bus, limita el número de dispositivos a conectar a un mismo bus y por lo tanto la única solución viable ante este problema era comprar dispositivos de almacenamiento adicionales aunque se desperdicia capacidad de almacenamiento en el actual dado que ya no podían conectarse más dispositivos, lo cual creaba asimismo un problema de eficiencia de uso del almacenamiento.

La conexión para supercomputadoras basada en el protocolo HIPPI

HIPPI (High Performance Parallel Interface) es una interfaz que fue utilizada para conectar dispositivos de almacenamiento a las supercomputadoras entre los años 1980's y 1990's.

El primer estándar HIPPI consistía en un cable de pares trenzados de 50 hilos por donde discurría la información a una velocidad de 800 Mbps o 100 MB/s. Luego este estándar

sufrió una actualización en donde se incluía un modo que alcanzaba velocidades de hasta 1600 Mbps o 200 MB/s sobre un cable de fibra óptica.

Esta interfaz tuvo una vida corta debido a la fuerte competencia contra tecnologías más masivas como SCSI y Fibre Channel que superaron su velocidad de transmisión y prestaciones sin embargo cabe resaltar que al momento de su lanzamiento otras tecnologías como SONET no superaban los 155 Mbps.

Un enfoque diferente: La interfaz inteligente IPI

IPI (Intelligent Peripheral Interface) fue un intento por centralizar la inteligencia del almacenamiento dentro de la interfaz, en contraste con otras tecnologías como SCSI en donde la interfaz solamente se utilizaba para mover la información entre el host y el sistema de almacenamiento; esta diferencia teóricamente lograba una mejor coordinación entre el host y los distintos sistemas de almacenamiento conectados a él. Sin embargo en la práctica muchos usuarios solamente tenían un solo sistema de almacenamiento por lo cual no presentaba ventaja alguna sobre las otras tecnologías sino por el contrario problemas. Debido a esta situación esta tecnología fue descontinuada en la segunda mitad de los años 1990's.

2.3. Las redes de almacenamiento (Storage Area Network – SAN)

Cuando hablamos de redes SAN generalmente se relacionan a la tecnología Fibre Channel la cual es la que tiene mayor presencia a nivel mundial sin embargo cabe resaltar que existen otras tecnologías como iSCSI o FICON que comparten esta definición. Según el SNIA (Storage Networking Industry Association) la definición formal de un SAN es la siguiente: "Una red cuyo propósito primario es la transferencia de datos entre los sistemas informáticos y los elementos de almacenamiento y entre elementos de almacenamiento. SAN abreviado. Una SAN consiste en una infraestructura de comunicación, que proporciona conexiones físicas, y una capa de administración, que organiza las conexiones, los elementos de almacenamiento, y los sistemas informáticos de modo que la transferencia de datos sea segura y robusta. El término SAN se identifica generalmente (pero no necesariamente) con servicios orientados a bloques de I/O más que con servicios de acceso a archivo. Un sistema de almacenamiento que consiste en elementos de almacenamiento, dispositivos de almacenamiento, sistemas informáticos, y/o aplicaciones, más todo el software de control, comunicándose sobre una red."

En resumen una SAN es básicamente una red utilizada para comunicar los hosts con los dispositivos de almacenamiento a nivel de bloques (como en la tecnología SCSI) o a nivel de archivos; y justamente esa es su gran ventaja ya que mantiene lo bueno del mundo SCSI (acceso eficiente a nivel de bloques) y lo combina con un direccionamiento de red que le proporciona escalabilidad, consolidación de recursos y flexibilidad.

En la siguiente sección del presente capítulo exploraremos al detalle las tecnologías vigentes que encajan dentro de este concepto para más adelante descubrir la utilidad de estas tecnologías en el mundo real mediante las múltiples aplicaciones principalmente implementadas en el mundo empresarial.

CAPITULO III

DESCRIPCION DE LAS TECNOLOGIAS DE REDES DE ALMACENAMIENTO (SAN)

Muchos factores han influido en la aparición de las tecnologías de redes de almacenamiento pero podemos mencionar a dos como los principales:

La evolución tecnológica

- Las necesidades del mercado

Es así que actualmente conviven tecnologías relativamente antiguas como SCSI con tecnologías muy recientes como FCIP (Fibre Channel sobre IP) dado que sus aplicaciones específicas no son las mismas y por lo tanto son complementarias, en otras palabras estas tecnologías pueden coexistir porque cubren nichos de mercado diferentes.

Un ejemplo muy claro de esto es la evolución de la tecnología SCSI llamada iSCSI o SCSI sobre IP que consiste básicamente en comandos SCSI encapsulados en paquetes IP; esta tecnología cumple la misma función que el protocolo Fibre Channel pero como habíamos comentado anteriormente cubren necesidades diferentes: iSCSI actualmente puede correr sobre redes Gigabit ya instaladas (bajo costo) a 1 Gbps como máximo (bajo rendimiento) mientras que Fibre Channel corre sobre redes propias dedicadas (alto costo) a velocidades de 4 Gbps (alto rendimiento). La primera tecnología puede encajar para una empresa pequeña, sin embargo no cubriría las expectativas de una gran empresa.

A través del presente capítulo exploraremos las características de funcionamiento de cada una de estas tecnologías, sus ventajas y desventajas, y su posición en el mercado actual.

3.1. La tecnología SCSI

En el capítulo anterior expliqué en forma rápida los inicios de la interfaz SCSI y su utilización en la industria, lo que no mencioné es que la interfaz SCSI paralela hace referencia solo a una parte del universo de estándares que componen la tecnología SCSI la cual ha sido extendido a lo largo del tiempo y proporciona las bases para otras nuevas tecnologías de almacenamiento.

3.1.1. Definición de la tecnología SCSI

Existen muchas definiciones para SCSI, sin embargo considero la dictada por el SNIA como la más exacta: “SCSI es una colección de estándares ANSI y propuestas de estándares que definen buses de I/O (Input/Output) diseñados principalmente para conectar subsistemas de almacenamiento o dispositivos a un host o servidores a través de los HBA (Host Bus Adapter o Adaptadores de bus). Originalmente creado para su uso con computadoras pequeñas, SCSI ha sido extendido para cubrir una gran cantidad de necesidades de cómputo y podría afirmarse que es el bus de I/O mas implementado en la actualidad”

3.1.2. Los estándares SCSI

Durante la evolución de la tecnología SCSI aparecieron muchos estándares algunos de los cuales fueron designados con diferentes nombres por distintos fabricantes lo cual crea confusión en lo referido a los estándares. Para complicar aún más el tema, como parte de la evolución el estándar SCSI se desglosa en tres partes por lo cual cada una genera una serie de estándares y actualizaciones independientes.

En el siguiente cuadro se muestra un resumen de la evolución de la tecnología SCSI en general y de la interfaz SCSI paralela en particular; la exploración dentro de cada uno de estos estándares sería muy extensa y escapa al alcance del presente informe.

Tabla 3.1.2.1.- Evolución de la interfaz SCSI paralela

Estándar	Comentario
SASI	1979 Shugart Associates Systems Interface. En realidad no fue un estándar sino la solución técnica de un fabricante.

<p>SCSI-1</p>	<p>El primer estándar SCSI se publicó en 1986. Opera sobre un bus asíncrono de 8 bits a unos 2 MB/s, sobre cables de 50 conductores, denominado cable A, con una longitud máxima de 6 metros. Este cable debía estar dotado de un terminador pasivo (resistencias) de 132 ohmios; admite 7 dispositivos (además del adaptador).</p> <p>Tenía el problema de que muchos de sus comandos eran optativos, por lo que había gran posibilidad que un dispositivo concreto no respondiese exactamente como se esperaba. Para subsanar el problema, un grupo de fabricantes se agruparon para adoptar un conjunto de 18 comandos básicos, que se denominó CCS y que fuese admitido por todos los periféricos. Este conjunto de comandos fue la base de lo que luego sería el estándar SCSI-2.</p>
<p>SCSI-2</p>	<p>Aprobado en 1989, es una ampliación de la especificación SCSI-1 que introdujo una mayor velocidad operativa y algunas otras mejoras.</p> <p>Opera sobre un bus síncrono a 5 y 10 MB/s, esta última denominada Fast-SCSI-2, sobre un bus de 8 bits, aunque existe una versión de 16 bits denominada Wide-SCSI-2. Esta última admite un máximo de 15 dispositivos (además del adaptador) en el mismo bus. Las versiones de 8 y 16 bits utilizan cables de 50 y 68 conductores respectivamente, este último denominado cable P.</p> <p>Esta norma define un terminador de cable activo (regulado por voltaje) que mejora la integridad de la transmisión; mejora la normalización de los comandos, e introduce la posibilidad de enviar al dispositivo una cola de comandos que este ejecutará de la forma que estime más conveniente. Esta habilidad es especialmente útil en Sistemas Operativos multiproceso en los que diversas tareas pueden enviar diversos comandos a la misma unidad.</p>
<p>SCSI-3</p>	<p>En este punto el estándar se desglosa en una serie de normas independientes que comprenden tres áreas: Conjunto de órdenes (comandos); protocolos de transporte y capa física. En consecuencia, a partir de aquí no puede hablarse propiamente de SCSI-4, SCSI-5, etc. porque cada especificación sigue un desarrollo independiente.</p> <p>Nota: Esta arquitectura es conocida como SAM-2 ("SCSI Architecture</p>

	<p>Model"), en contraposición a SAM (o SAM-1) que se refiere a las normas anteriores (todo-en-uno).</p> <p>Las normas relativas al conjunto de órdenes se dividen en dos grupos; las que son comunes a todos los dispositivos ("Primary Command Set") SPC, SPC-2, SPC-3, y las que son específicas de un determinado tipo de dispositivo (discos, escaneres, cintas, etc) que tienen sus propias siglas:</p> <ul style="list-style-type: none"> • Block commands (por ejemplo, discos); SBC, SBC-2 • Reduced Block Commands (discos); RBC • Stream Commands (dispositivos de cinta); SSC, SSC-2, SSC-3 • Media Changer Commands (jukeboxes); SMC, SMC-2 • Multi Media Commands (DVDs); MMC, MMC2, MMC-3, Etc. • Enclosure Services; SES, SES-2. • Controller Commands (p.e. dispositivos RAID); SCC, SCC-2. • Object-based Storage Devices; OSD. • Management Server Commands; MSC. <p>Las especificaciones relativas a los protocolos de transporte adoptan sus propias siglas, como por ejemplo; FCP ("Fibre Channel Protocol"); SSP ("Serial SCSI Protocol"); SMP ("Serial Management Protocol") y STP ("Serial ATA Tunneling Protocol").</p> <p>Finalmente, las especificaciones relativas a la capa física tienen también sus propias denominaciones: SPI ("SCSI Parallel Interface") SPI-2 SPI-3, SPI-4; SAS ("Serial Attached SCSI"), etc.</p> <p>A partir de aquí se introdujeron mejoras que extendían la funcionalidad de SCSI-2 manteniendo compatibilidad descendente con las antiguas versiones. Simultáneamente la maquinaria de márketing introducía nombres que hicieron más o menos fortuna, como Ultra-X SCSI y Fast-Y SCSI.</p>
Fast-Wide SCSI	Aparecida en 1992 incorpora los estándares SPI/SIP. Consigue tasas de transferencia de 20 MB/s.

Ultra-SCSI	Aparecida en 1995, proporciona una amplitud de banda doble, acercándola al siguiente nivel de rendimiento, pero manteniendo los cables, conexiones y alojamientos existentes Permite velocidades de transferencia de 20 MB/s con 8 bits y de 40 MB/s con 16 bits.
Ultra-2	Aparecida en 1997. Sigue la norma SPI-2 ("SCSI Parallel Interface") y alcanza 80 MB/s. Utiliza la señal por diferencial de bajo voltaje LVD ("Low Voltage Differential") antes comentada.
Ultra-3	Aparece en 1999 y consigue velocidades de 160 MB/s con la norma SPI-3. Incluye un sistema CRC ("Cyclic Redundancy Chec") de detección y corrección de errores. También denominado Ultra160.
Ultra320	Aparece en el 2002 y puede decirse que es el estándar Ultra160 con el doble de velocidad (320 MB/s) y se encuentra basada en la norma SPI-4.

3.1.3. El modelo de arquitectura SCSI

El modelo de arquitectura SCSI (SAM por sus siglas en inglés) es una vista abstracta y de alto nivel de la manera de cómo se comunican los dispositivos SCSI. Su objetivo es mostrar como los diferentes estándares SCSI se encuentran interrelacionados.

En la Tabla 3.1.2.1 observamos un punto de quiebre en el estándar SCSI-3. A partir de ese punto la arquitectura de la tecnología SCSI se divide en tres capas o niveles:

- Capa de aplicación (SAL por sus siglas en inglés)
- Capa de protocolos de transporte (STPL por sus siglas en inglés)
- Capa de interconexión o capa física (SIL por sus siglas en inglés)

En base a este modelo se han creado y actualizado muchos estándares de manera totalmente independientes y se ha extendido el uso de estos estándares SCSI a nuevas tecnologías de almacenamiento como se muestra en la Figura 3.1.3.1.

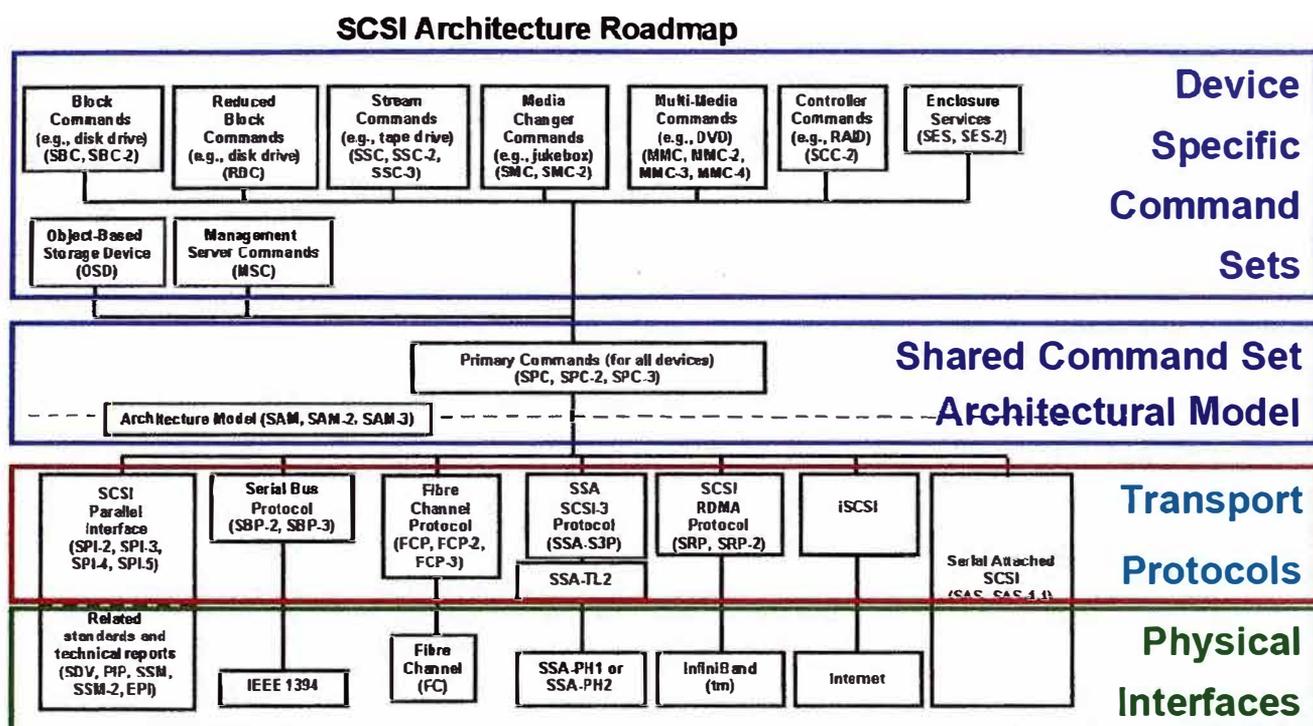


Figura 3.1.3.1.- Modelo de arquitectura SCSI

El modelo de arquitectura SCSI se describe en términos de objetos, capas de protocolos, e interfases de servicio entre los objetos. La descripción de cada uno de estos elementos escapa al alcance del presente informe, sin embargo explicaré los más importantes para que brinden los conceptos necesarios para entender los temas subsiguientes:

El modelo de servicios distribuidos SCSI

Las interfases de servicio entre los objetos distribuidos son representadas por el modelo cliente-servidor mostrado en la Figura 3.1.3.2. Las líneas horizontales punteadas con flechas denotan una simple transacción de solicitud-respuesta tal y como se da entre un cliente y un servidor. Las líneas sólidas con flechas indican el camino real a través del subsistema de transporte. En este modelo, cada cliente o servidor puede correr una transacción en forma concurrente con otros servidores o clientes.

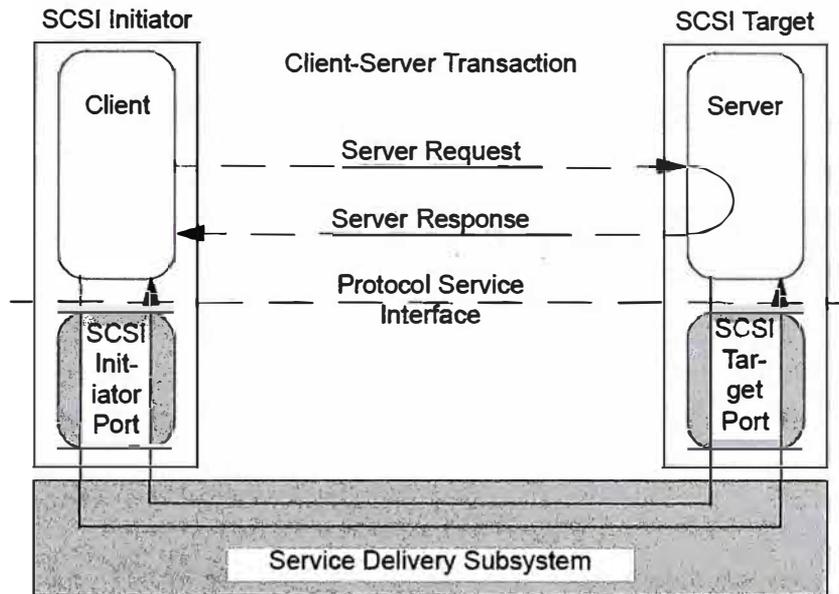


Figura 3.1.3.2.- Modelo cliente-servidor

Una transacción del tipo cliente-servidor es representada por una subrutina con entradas proporcionadas por el que inicia la llamada (p.e. el cliente). La subrutina es procesada por el servidor y devuelve una salida y un estado de la subrutina. Un cliente envía solicitudes a un servidor remoto vía el puerto iniciador SCSI (SCSI initiator port) vía el subsistema de transporte de servicios (Service Delivery Subsystem) y recibe una respuesta de estado completado o una notificación de error.

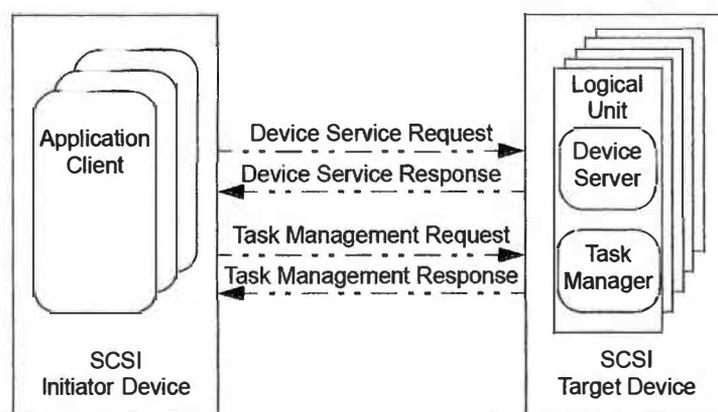


Figura 3.1.3.3.- Modelo cliente-servidor

En las siguientes secciones describiré cada una de estas capas poniendo especial énfasis en la capa física y de transporte del estándar SCSI-3 relacionada a la interfaz SCSI paralela (SPI) objeto de estudio del presente capítulo.

3.1.4. La capa física del modelo SCSI

La capa física o también llamada capa de interconexión consiste en un conjunto de estándares que definen los mecanismos de comunicación empleados por los protocolos de transporte SCSI. Estos estándares describen los requerimientos eléctricos y de señalización esenciales para que los dispositivos SCSI puedan interoperar sobre una conexión dada.

Los estándares que forman parte de la capa física del modelo de arquitectura SCSI se muestran en la tabla 3.1.4.1.

Tabla 3.1.4.1.- Estándares que componen la capa física del modelo SCSI

Nombre del estándar	Abrev	Norma técnica
Fibre Channel Arbitrated Loop – 2	FC-AL-2	[ISO/IEC 14165-122] [ANSI NCITS.332-1999] [ANSI NCITS.332-1999/AM1]
Fibre Channel Physical Interfaces	FC-PI	[ISO/IEC 14165-115] [ANSI INCITS.352-2002]
Fibre Channel Physical Interfaces – 2	FC-PI-2	[T11/1506-D]
Fibre Channel Framing and Signaling Interface	FC-FS	[ISO/IEC 14165-251]
High Performance Serial Bus High Performance Serial Bus (supplement to ANSI/IEEE 1394-1995)		[ANSI INCITS.373-2003] [ANSI IEEE 1394-1995] [ANSI IEEE 1394a-2000]
SCSI Parallel Interface – 2	SPI-2	[ISO/IEC 14776-112] [ANSI X3.302-1999]
SCSI Parallel Interface – 3	SPI-3	[ISO/IEC 14776-113] [ANSI NCITS.336-2000]
SCSI Parallel Interface – 4	SPI-4	[ISO/IEC 14776-114] [ANSI INCITS.362-2002]
SCSI Parallel Interface – 5	SPI-5	[ISO/IEC 14776-115] [ANSI INCITS.367:2003]
Serial Storage Architecture Physical Layer 1	SSA-PH	[ANSI X3.293-1996]
Serial Storage Architecture Physical Layer 2	SSA- PH-2	[ANSI NCITS.307-1998]

Serial Attached SCSI	SAS	[ISO/IEC 14776-150] [ANSI INCITS.376:2003]
Serial Attached SCSI - 1.1	SAS-1.1	[ISO/IEC 14776-151] [T10/1601-D]

En la tabla anterior podemos observar los estándares que norman la conexión y señalización de distintas tecnologías como Fibre Channel, la interfaz SCSI paralela o el Serial Attached SCSI (SAS). Mas adelante abordaré en mayor detalle algunas de éstas.

3.1.5. La capa de protocolos de transporte del modelo SCSI

La capa de protocolos de transporte SCSI define los requerimientos para intercambiar información para que distintos dispositivos SCSI sean capaces de comunicarse entre sí.

Los estándares que forman parte de la capa de protocolos de transporte del modelo SCSI se muestran en la tabla 3.1.5.1.

Tabla 3.1.5.1.- Estándares que componen la capa de transporte del modelo SCSI

Nombre del estándar	Abrev	Norma técnica
Automation/Drive Interface - Transport Protocol	ADT	[ISO/IEC 14776-191] [T10/1557-D]
Serial Storage Architecture Transport Layer 1	SSA-TL-1	[ANSI X3.295-1996]
Serial Storage Architecture Transport Layer 2	SSA-TL-2	[ANSI NCITS.308-1998]
SCSI-3 Fibre Channel Protocol	FCP	[ISO/IEC 14776-221] [ANSI X3.269-1996]
SCSI Fibre Channel Protocol – 2	FCP-2	[ISO/IEC 14776-222] [ANSI NCITS.350-2003]
SCSI Fibre Channel Protocol – 3	FCP-3	[ISO/IEC 14776-223] [T10/1560-D]
Serial Bus Protocol – 2	SBP-2	[ISO/IEC 14776-232] [ANSI NCITS.325-1999]
Serial Bus Protocol – 3	SBP-3	[ISO/IEC 14776-233] [T10/1467-D]
Serial Storage Architecture SCSI-3 Protocol	SSA-S3P	[ANSI NCITS.309-1998]

SCSI RDMA Protocol	SRP	[ISO/IEC 14776-241] [T10/1415-D]
SCSI RDMA Protocol – 2	SRP-2	[ISO/IEC 14776-242] [T10/1524-D]

3.1.6. La capa de aplicación del modelo SCSI

Los estándares que componen esta capa se encuentra dividida en dos grupos:

- Conjunto de comandos compartidos
- Conjunto de comandos para un dispositivo específico

El conjunto de comandos compartidos define un modelo común a todos los tipos de dispositivos SCSI (subsistema de discos, librerías de backup, unidades de CD-ROM, etc). Este estándar especifica los comandos requeridos y el comportamiento que es común a todos los dispositivos SCSI, sin tener en cuenta el tipo de dispositivo, e indica los requerimientos a ser cumplidos por un dispositivo SCSI Initiator cuando envíe comandos cualquier dispositivo SCSI target.

El conjunto de comandos para un dispositivo específico define un modelo particular para cada tipo de dispositivo SCSI. Estos estándares especifican los comandos requeridos y el comportamiento específico para un tipo de dispositivo dado e indica los requerimientos a cumplir por un dispositivo SCSI Initiator cuando envíe comandos a un tipo de dispositivo específico.

Tabla 3.1.6.1.- Estándares que componen la capa de aplicación del modelo SCSI

Nombre del estándar	Abrev	Norma técnica
SCSI-3 Primary Commands	SPC	[ANSI X3.301-1997]
SCSI Primary Commands – 2	SPC-2	[ISO/IEC 14776-452] [ANSI NCITS.351-2001]
SCSI Primary Commands – 3	SPC-3	[ISO/IEC 14776-453] [T10/1416-D]

Tabla 3.1.7.1.- Estándares que componen la capa de aplicación del modelo SCSI

Nombre del estándar	Abrev	Norma técnica
SCSI-3 Block Commands	SBC	[ISO/IEC 14776-321] [ANSI NCITS.306-1998]
SCSI Block Commands – 2	SBC-2	[ISO/IEC 14776-322] [T10/1417-D]
SCSI-3 Stream Commands	SSC	[ISO/IEC 14776-331] [ANSI NCITS.335-2000]
SCSI Stream Commands – 2	SSC-2	[ISO/IEC 14776-332] [ANSI INCITS.380-2003]
SCSI Stream Commands – 3	SSC-3	[ISO/IEC 14776-333] [T10/1611-D]
SCSI-3 Medium Changer Commands	SMC	[ISO/IEC 14776-351] [ANSI NCITS.314-1998]
SCSI Media Changer Commands - 2	SMC-2	[ISO/IEC 14776-352] [T10/1383-D]
SCSI-3 Multimedia Command Set	MMC	[ANSI X3.304-1997]
SCSI Multimedia Command Set - 2	MMC-2	[ISO/IEC 14776-362] [ANSI NCITS.333-2000]
SCSI Multimedia Command Set - 3	MMC-3	[ISO/IEC 14776-363] [ANSI INCITS.360-2002]
SCSI Multimedia Command Set - 4	MMC-4	[ISO/IEC 14776-364] [T10/1545-D]
SCSI Multimedia Command Set - 5	MMC-5	[ISO/IEC 14776-365] [T10/1675-D]
SCSI Controller Commands - 2	SCC-2	[ISO/IEC 14776-342] [ANSI NCITS.318-1998]
SCSI Reduced Block Commands	RBC	[ISO/IEC 14776-326] [ANSI NCITS.330-2000]
SCSI-3 Enclosure Services Commands	SES	[ISO/IEC 14776-371] [ANSI NCITS.305-1998]
SCSI Enclosure Services Commands - 2	SES-2	[ISO/IEC 14776-372] [T10/1559-D]
SCSI Specification for Optical Card Reader/Writer	OCRW	[ISO/IEC 14776-381]

Object-based Storage Device Commands	OSD	[ISO/IEC 14776-391] [T10/1355-D]
SCSI Management Server Commands	MSC	[ISO/IEC 14776-511] [T10/1528-D]
Automation/Drive Interface - Commands	ADC	[ISO/IEC 14776-356] [T10/1558-D]

3.2. La interfaz SCSI paralela

Luego de revisar el modelo de la arquitectura SCSI haré referencia rápidamente a la interfaz SCSI paralela, el cual aunque muchas veces no sea considerada como una tecnología de redes de almacenamiento es todavía muy utilizada para conectar servidores con dispositivos de almacenamiento y respaldo. La interfaz SCSI paralela sigue las normas SPI (SPI, SPI-2, SPI-3, SPI-4, SPI-5) que contienen una serie de lineamientos para las conexiones entre dos dispositivos a través de una interfase paralela.

Hasta hace unos cuantos años el método más común de conectar un disco y un dispositivo de cinta a servidores de aplicaciones o de archivos era vía el tradicional cable de bus SCSI.

La interfase física original SCSI consistía en un cable paralelo con 8 líneas para datos y otras más para control. Transmitir 8 bits de datos durante cada ciclo del reloj proporciona un ancho de banda relativamente elevado, sin embargo existen cuestiones eléctricas que restringen la distancia total permitida por la mayoría de implementaciones SCSI a 15~25 metros como máximo. La arquitectura del bus paralelo SCSI ha evolucionado con el tiempo, con ancho de bandas cada vez más elevados debido a ruta de datos más anchas (16 y 32 líneas de datos) y relojes con ciclos más rápidos.

Una de las dificultades presentada por la arquitectura de bus paralelo es el fenómeno conocido como skew o desviación. Si los 8 o 16 bits de datos son enviados en forma simultánea en paralelo, una pequeña diferencia en el retardo de propagación entre las líneas de datos puede ocurrir, y de esta manera no todos los bits llegarán a su destino al mismo tiempo. El skew o desviación se define como la diferencia de tiempo de llegada para cada bit que comprende la palabra de datos. Como se puede observar en la Figura 3.2.1, mientras mayor sea la diferencia en el retardo de propagación mayor debe ser la

ventana para asegurarse que todos los bits de datos han sido capturados. Cuando la longitud del cable se hace mayor el retardo de propagación también crece, esto limita la longitud del cable SCSI paralelo.

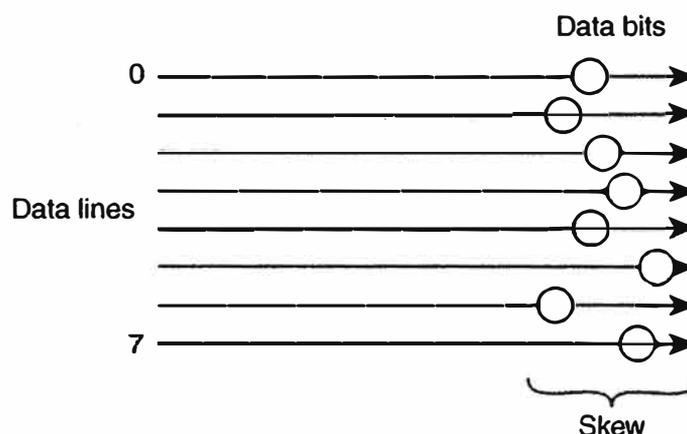


Figura 3.2.1. – Desviación o Skew en una transmisión paralela de bits

Adicionalmente, la interfaz SCSI paralela requiere de una terminación de los puertos no utilizados ya que los dispositivos SCSI se conectan juntos en cadena (daisy-chain en inglés), los dispositivos del final de la cadena deben cerrar correctamente el circuito para evitar interferencia por reflexión de la señal. Por todo lo expuesto es muy importante tener buenos cableados y terminadores para evitar corrupción de datos o fallas en las transacciones.

Podemos encontrar la interfaz SCSI paralela al interior de algunos servidores que la utilizan para conectar sus discos duros internos, sin embargo esta configuración es impráctica para la mayoría de las grandes empresas porque limita el crecimiento de la capacidad de almacenamiento a la cantidad máxima de discos que caben dentro de un servidor (normalmente no más de 10). Cuando se desea disponer de cientos de Gigabytes o Terabytes la mejor opción es la conexión de subsistemas de discos externos y para ello también se puede utilizar la interfase SCSI paralela, sin embargo el crecimiento explosivo de la cantidad de información ha relegado la utilización de esta tecnología por dos razones principales:

- El tamaño limitado del cable SCSI
- La máxima cantidad de dispositivos conectados a un bus SCSI es también limitada

Para superar estas debilidades la ANSI ha desarrollado una serie de estándares para introducir una nueva interfaz SCSI serial (SAS por sus siglas en inglés) que no compite con las tecnologías de redes de almacenamiento pero que permite la conexión directa de

subsistemas de almacenamiento a un servidor. Otro esfuerzo por superar las limitaciones de la interfaz SCSI paralela se encuentra representado por la tecnología Fibre Channel que permite conectar una mayor cantidad de dispositivos a uno o varios subsistemas de almacenamiento basado en redes de Fibra Optica y dispositivos de direccionamiento y ruteo.

3.3. La evolución de SCSI: Fibre Channel

Cómo ya había adelantado en el capítulo anterior Fibre Channel es una tecnología que supera las limitaciones de la interfaz SCSI paralela pero va más allá. Fue la primera tecnología Gigabit en ser exitosamente comercial y hoy por hoy es la tecnología más utilizada en redes de almacenamiento.

A través de las siguientes secciones iremos descubriendo el marco conceptual de la tecnología Fibre Channel que incluye el protocolo propiamente dicho, las topologías de conectividad y los productos que componen una red basada en esta tecnología.

3.3.1. La pila de protocolos Fibre Channel

La tecnología Fibre Channel posee una arquitectura orientada a redes y al igual que el modelo OSI posee niveles o capas en donde cada una tiene una finalidad distinta. Estas capas o niveles han sido definidos por el comité NCITS/ANSI T11X3 y básicamente consisten en una capa de nivel físico, de codificación, de control de transporte y de interfase con otros protocolo de niveles superiores los cuales entregan información al enlace Fibre Channel para que se encargue de transportarla. En resumen Fibre Channel es un protocolo que cabe en las dos primeras capas del modelo OSI (Físico y Enlace) y transporta información de otros protocolos de capas superiores como por ejemplo TCP/IP o Serial SCSI.

Los estándares que definen el protocolo Fibre Channel hacen mención a cuatro capas o niveles numerados desde FC-0 para la capa inferior hasta FC-4 para la capa superior.

Nivel	Función
FC-4	Nivel de interfase para protocolos de nivel superior
FC-3	Nivel de servicios comunes (todavía no implementado)
FC-2	Nivel de control de transporte

FC-1	Nivel de codificación y control de enlace
FC-0	Nivel de interfase física

Nivel FC-0: La capa física

El nivel FC-0 o nivel de interfase física como lo dice su nombre define las características de los elementos que componen la interfaz física recomendados para el protocolo Fibre Channel.

Para el enlace físico se puede utilizar cable de fibra óptica o de cobre (de ahí el nombre de Fibre Channel y no Fiber Channel). Generalmente se prefiere el cable de fibra por su inmunidad al ruido electromagnético que podría tener efectos catastróficos a las velocidades que maneja esta tecnología. Dentro de los cables de fibra existen varios tipos de los cuales cada uno es elegido según en base a la distancia del enlace físico. Los cables dentro del estándar Fibre Channel, junto con las distancias soportadas se muestran en la tabla 3.1.1.1.

Tabla 3.1.1.1.- Tipos de cables contemplados por el estándar Fibre Channel y las distancias soportadas

Tipo de cable	Transmisor	Velocidad	Distancia
Cobre	ECL/PECL	200 MB/s	0m -10m
		100 MB/s	0m – 30m
9 um Fibra Óptica Monomodo	1550nm Longwave Laser	200 MB/s	2m - > 50km
		100 MB/s	2m - > 50km
	1300nm Longwave Laser	400 MB/s	2m – 2km
		200 MB/s	2m – 2km
		100 MB/s	2m – 10km
			2m – 2km
50um Fibra Óptica Multimodo	850nm Shortwave Laser	400 MB/s	0.5m – 175m
		200 MB/s	0.5m – 300m
		100 MB/s	0.5m – 500m
62.5um Fibra Óptica Multimodo		400 MB/s	0.5m – 70m
		200 MB/s	0.5m – 150m
		100 MB/s	0.5m – 300m

Para su conexión con los dispositivos, el cable de cobre puede utilizar un conector estándar (p.e. DB-9) mientras que para los cables de fibra existen dos tipos de conectores: SC (Conector Estándar por sus siglas en inglés) muy utilizado es enlaces de 1 Gbps o los SFF (Conector de Formato Pequeño por sus siglas en inglés) que fue introducido en el estándar de 2 Gbps y se ha mantenido vigente en la actualización a 4 Gbps por su menor tamaño en comparación con el tipo SC (ver figura 3.1.1.1)

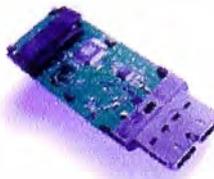


Figura 3.1.1.1.- Conectores tipo SC (el más grande) y LC

Adicionalmente a los cables y conectores se encuentran los elementos activos del enlace físico Fibre Channel que son los Transmisores y los Adaptadores de Interfase. Los transmisores son los encargados de transformar las señales eléctricas provenientes de los circuitos electrónicos de los dispositivos Fibre Channel en señales ópticas o eléctricas en formato serial a los niveles y velocidades establecidas en el protocolo. Actualmente la mayoría de productos Fibre Channel en el mercado utilizan los transmisores del tipo SFP (Transmisor de Formato Pequeño Removible por sus siglas en inglés) el cual acepta conectores del tipo SFF. Anteriormente a estos estuvieron en vigencia los transmisores del tipo GLM o GBIC, el primero contenía un convertidor paralelo/serial pero no era removible en caliente por lo que dificultaba el mantenimiento, mientras que el segundo ya era removible en caliente pero sucumbió ante la mayor eficiencia de espacio de los SFPs.



DB-9 GLM



Optical GML



Figura 3.1.1.2.- GLM para interfaz eléctrica y óptica

Figura 3.1.1.3.- Transmisores de tipo SFP

En los casos donde se requiere conectar un dispositivo Fibre Channel con interfase de cobre a otros dispositivos con interfase de fibra se deben utilizar los Adaptadores de Interfase (MIA por sus siglas en inglés) los cuales convierten las señales eléctricas en señales ópticas y viceversa.

Todos los elementos mencionados deben estar cuidadosamente diseñados y en buen estado para asegurar la transmisión a las velocidades definidas por el estándar ya el BER debe mantenerse por debajo de 10^{-12} . Cualquier defecto o mala utilización de estos elementos podría producir efectos indeseados (p.e. jitter) que generen errores en la información que circula por el enlace eléctrico

Nivel FC-1: El nivel de codificación

El Nivel FC-1 o nivel de codificación de datos es el encargado de codificar los bytes de información que recibe de la capa superior en una corriente de bits con propiedades que permita su fácil recuperación en el receptor con la menor cantidad de errores. Esto es necesario porque frente a una transmisión de muchos bits iguales (p.e. muchos 1's) sin codificación podría interpretarse como un patrón sin cambio o llevado al mundo electrónico como un voltaje sostenido en donde no existe ningún cambio; cuando esta secuencia es leída por el receptor, éste no podrá distinguir la cantidad de bits y perderá la sincronización y la señal de reloj.

Para superar este inconveniente Fibre Channel utiliza un sistema de codificación del tipo 8b/10b en donde cada 8 bits se convierten en 10 bits para ser transmitidos de tal manera de mantener constante el balance entre números de 1's y 0's dentro de la corriente de transmisión; esto lo logra monitoreando la disparidad del carácter precedente. Si un carácter con disparidad positiva es transmitido entonces el siguiente carácter a ser transmitido debe tener disparidad negativa. Este efecto se logra tomando los 8 bits provenientes del nivel FC-2 y una variable de control para que el nivel FC-1 los codifique utilizando el algoritmo 8b/10b.

FC-2 bit notation:	7	6	5	4	3	2	1	0	Control Variable
FC-1 unencoded bit notation:	H	G	F	E	D	C	B	A	Z

Figura 3.1.1.4.- Etiqueta de los bits

En la Figura 3.1.1.4 a cada carácter de transmisión se le asigna un nombre según la siguiente convención: Zxx.y en donde:

- Z es la variable de control del byte de información no codificado FC-1. El valor de Z es utilizado para indicar si el carácter de transmisión es un carácter de datos (Z=D) o un carácter especial (Z=K)
- xx es el valor decimal del número compuesto por los bits E, D, C, B y A de un byte de información no codificado FC-1.
y es el valor decimal del número binario compuesto por los bits H, G y F del byte de información no codificado FC-1

De esta manera se crean las palabras de transmisión en donde a través de una tabla se mapean dos caracteres de transmisión de 10 bits uno con disparidad positiva y otro con disparidad negativa y donde solo uno de éstos será transmitido dependiendo del estado de la disparidad actual de la transmisión (si la disparidad actual es negativa se transmite el carácter de disparidad positiva o neutral y si la paridad actual es positiva se transmite el carácter de disparidad negativa o neutral, de esta manera se mantiene balanceada la disparidad actual). Adicionalmente este tipo de codificación permite identificar algún error durante la transmisión por violación de código o por violación de disparidad.

Se incluye también dentro de este nivel la codificación de un grupo de comandos denominados "ordered sets". Estas palabras de transmisión son compuestas por 4 bytes codificados o 40 bits, en donde el primer byte es un carácter especial llamado K28.5 y los otros 3 bytes restantes definen la función del ese ordered set específico.

Los ordered sets se pueden clasificar en tres categorías generales: inicio de trama, fin de trama, y señales primitivas.

Nivel FC-2: Protocolo de tramas

Para poder mover información desde un dispositivo Fibre Channel hacia otro dispositivo, los bloques de datos entregados por los niveles superiores deben ser organizados en paquetes de datos para que puedan ser transportados. En la nomenclatura de Fibre Channel estos paquetes de datos son llamados tramas.

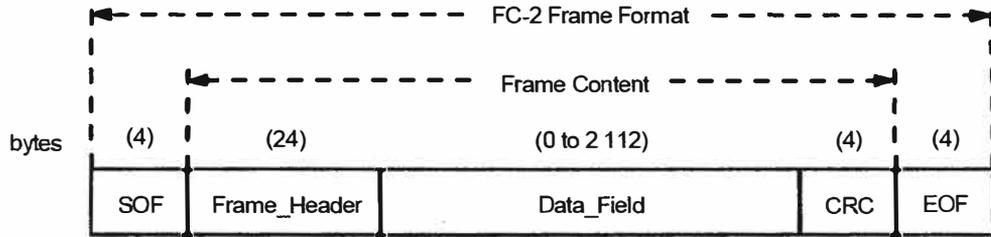


Figura 3.1.1.5.- Estructura de la trama FC-2

Como se muestra en la figura 3.1.1.5 las tramas siempre llevan por delante un ordered set llamado SOF (Star of Frame por sus siglas en inglés o Inicio de Trama en español) que indica el inicio de una trama FC-2 y se utiliza además para definir la clase de servicio utilizada y para especificar si la trama es la primera en una serie o es simplemente una en una serie de tramas relacionadas. La longitud del SOF es de 4 bytes.

Seguido del SOF aparece una cabecera de trama de 24 bytes (Frame Header en inglés) que contiene las direcciones de origen y de destino así como campos de control que indican el tipo de la trama (de datos o de control) y su posición dentro de una serie de tramas.

Bits Word	31 .. 24	23 .. 16	15 .. 08	07 .. 00
0	R_CTL	D_ID		
1	CS_CTL/Priority	S_ID		
2	TYPE	F_CTL		
3	SEQ_ID	DF_CTL	SEQ_CNT	
4	OX_ID		RX_ID	
5	Parameter			

Figura 3.1.1.6.- Estructura de la cabecera de la trama FC-2

A continuación un resumen del significado de cada uno de estos campos de la cabecera:

R_CTL.- Campo de 1 byte que contiene bits de ruteo e información para categorizar la función de la trama.

D_ID, S_ID.- Campos de tres bytes cada uno que contienen la dirección del puerto destino (D_ID) y origen (S_ID) respectivamente.

- **CS_CTL/Priority.**- Campo que contiene parámetros de clase de servicio o de prioridad según el valor contenido en el bit 17 del campo FC_CTL
- **Type.**- El protocolo asociado con el payload o parte de usuario (por ejemplo SCSI-3)

Frame Control (F_CTL).- Bits que identifican la transferencia de la secuencia; el comienzo, el medio o el final de la secuencia; y el final de la conexión.

Data Field Control (DF_CTL).- Indica la presencia de cabecera opcional.

- **Sequence ID (SEQ_ID).**- Identificador numérico único entre dos puertos.
- **Sequence Count (SEQ_CNT).**- Contador de trama o identificador de trama de 16 bits.
- **Originator Exchange ID (OX_ID).**- Identificador de intercambio que utiliza el nodo que origina el mismo

Responder Exchange ID (RX_ID).- Similar al OX_ID pero insertado por el que responde al intercambio.

Parameter.- Su contenido puede variar dependiendo del tipo de trama. Normalmente es utilizado como relleno para la sección de datos de usuario o payload

Luego de la cabecera continúa los datos en sí que pueden ser desde 0 hasta 2112 bytes. Esta longitud variable permite a Fibre Channel cumplir con los requerimientos de varias aplicaciones con un buen balance entre la cabecera y los datos en sí. Debido a que la trama FC-2 está basada en múltiplos de palabras de transmisión de 4 bytes, la data de usuario debe ser rellena con la cantidad de bits necesarios para convertirse en múltiplo de 4.

A continuación se introduce un campo de detección de errores o CRC (cyclic redundancy check) con el objeto de asegurar la integridad de la data dentro de la trama y se genera de la misma forma que el FCS (Frame Check Sequence) de FDDI.

Luego de que la trama ha sido ensamblada, las tramas son entregadas a través de un protocolo jerárquico de sequences (secuencias) y exchanges (intercambios). Una secuencia puede incluir una o mas tramas relacionadas (p.e. un bloque extenso de datos que se escribe a un dispositivo Fibre Channel y se divide en varias tramas para su transmisión), en cambio un intercambio puede incluir una o más secuencias no relacionadas. Dos dispositivos Fibre Channel pueden tener varios intercambios establecidos al mismo tiempo cada uno con sus propios IDs de intercambio e IDs de

secuencia en forma separada. Esta estructura simple permite maximizar la utilización del enlace entre los dispositivos Fibre Channel con un mínimo de cabeceras.

Para minimizar la longitud de la cabecera en las tramas FC-2, Fibre Channel limita la recuperación de errores a nivel de secuencia, es decir que cuando se detecta una trama con error se descarta toda la secuencia pues a velocidades de Gigabit es más rápido descartar toda una secuencia que implementar la lógica necesaria para gestionar la recuperación de una trama individual.

En este nivel también se realiza un control de flujo para manejar la congestión que podría presentarse en las comunicaciones entre dispositivos Fibre Channel. El control de flujo se basa en un sistema de créditos donde cada crédito representa la facultad de aceptar una trama adicional. Mientras que el receptor no envíe créditos al emisor, ninguna trama podrá ser transmitida. En la práctica este esquema de créditos se encuentra basado en el número de buffers (memoria) que un nodo Fibre Channel reserva para almacenar la data recibida. Estos buffers reciben una nueva información cada vez que el nivel FC-1 se la entrega y son liberados cada vez que el nivel FC-2 identifica las tramas individuales para la reconstrucción del bloque de datos origen.

Adicionalmente en este nivel se manejan las clases de servicio, las cuales permiten entregar QoS (Quality of Service) a las conexiones entre dos o más dispositivos, en base al tipo de interacción que tengan entre ellos. Actualmente existen 5 clases de servicio definidas por el estándar, sin embargo solamente las tres primeras se encuentran implementadas en los productos del mercado. A continuación se muestra una breve descripción de cada una de estas clases:

- Clase 1 – Servicio de conexión dedicada.
 - o Orientado a conexión
 - o Confirmación de entrega

- Clase 2 – Servicio de conexión multiplexada
 - o No orientado a conexión
 - o Confirmación de entrega

- Clase 3 – Servicio de datagramas
 - o No orientado a conexión

- Sin confirmación de entrega
- Clase 4 – Servicio fraccional. Utiliza una fracción del ancho de banda del enlace entre dos puertos de comunicación
 - Orientado a conexión
- Clase 6 – Servicio de conexión dedicado unidireccional. Proporciona conexiones unidireccionales dedicadas.
 - Orientado a conexión

Nivel FC-3: Servicios comunes

El nivel FC-3 define las funciones que afectan a todos los puertos de un mismo nodo o dispositivo, es decir existe un solo nivel FC-3 por cada nodo y proporciona las funciones que habilitan funcionalidades avanzadas que será utilizadas por todos los puertos de un solo nodo. Algunas de las funciones soportadas por este nivel son:

- **Striping**, el cual es utilizado para multiplicar el ancho de banda de un enlace Fibre Channel Múltiples N_Ports son conectados en paralelo para transmitir una unidad de información a través de múltiples enlaces.
- **Hunt groups**, los cuales son grupos de N_Ports asociados a un solo nodo y en donde cualquiera de esos puertos puede transportar la información para ese nodo. Esto disminuye la latencia que se genera al esperar que un solo puerto se encuentre disponible.
- **Multicasting**, esta funcionalidad puede ser comparada con un mensaje broadcast, es decir permite que una sola transmisión pueda ser recibida por múltiples nodos.

Nivel FC-4: Mapeo de aplicaciones

La capa más alta contiene un grupo de reglas definidas para poder mover información desde las interfases de los protocolos superiores hacia los niveles inferiores de Fibre Channel.

Para enviar información, el nivel FC-4 toma una unidad de información desde un protocolo de nivel superior (esto es el campo de datos en la trama FC-2 y es la data en sí que será transmitida. También denominado como Information Unit (IU)) y lo transforma en

secuencias para los niveles FC-3 y FC-2. Para decepcionar la información, el nivel FC-4 toma una secuencia desde los niveles FC-3 y FC-2 y los transforma en un IU para el protocolo de nivel superior.

3.3.2. Topologías Fibre Channel

En la tecnología Fibre Channel existen tres topologías bien definidas para interconectar los dispositivos Fibre Channel: Point-to-Point, Arbitrared Loop, y Fabric.

En una topología Point-to-Point se conectan solamente dos dispositivos y una vez que se ha establecido la conexión entre ellos asumen que disponen de todo el ancho de banda para el enlace. En una topología Arbitrared Loop la operación es distinta ya que la comunicación es similar a una red Token Ring en donde los dispositivos se comunican a través de un medio compartido y por lo tanto solo puede existir una sola comunicación simultánea dentro del anillo.

Por último la topología Fabric permite que existan múltiples conexiones punto a punto concurrentes a través de conmutaciones a nivel de enlace lo cual hace a ésta la topología más utilizada actualmente.

A continuación describiré brevemente cada una de estas topologías, sus características y sus modos de operación.

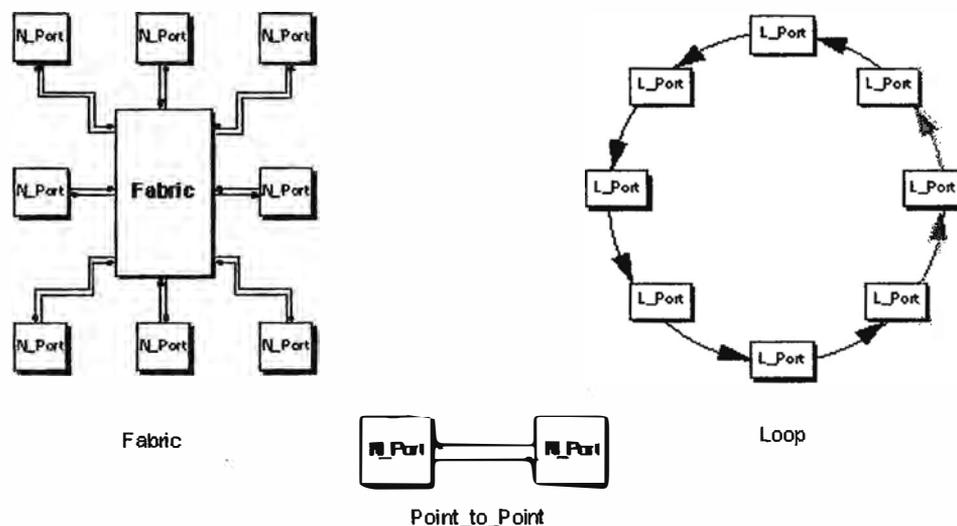


Figura 3.3.2.1.- Topologías Fibre Channel

Topología Point-to-Point

La topología Point-to-Point (Punto a punto) consiste en una simple conexión directa entre dos N_Ports de tal manera que el punto de transmisión de uno se conecte con el punto de recepción del otro y viceversa.

Antes que la información pueda ser transmitida de un dispositivo hacia el otro, ambos N_Ports deben realizar un N_Port login con el objeto de asignar direcciones o IDs a los N_Port. Luego de esto la conexión es mantenida y utilizada como una conexión dedicada por las aplicaciones. Hoy en día los N_Port permiten manejar anchos de banda de 200, 400 o 800 MB/s Full Duplex dependiendo de la generación de los dispositivos; para esta topología el 100% de este ancho de banda estaría dedicado a una sola conexión.

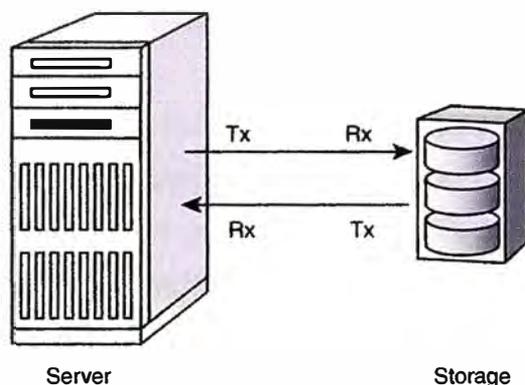


Figura 3.3.2.2.- Aplicación de la topología Point-to-Point

Para que la conexión tenga sentido, siempre debe existir un dispositivo Initiator que inicie la comunicación y un dispositivo Target que atiende los requerimientos del Initiator.

Topología Arbitrated Loop

Comparado con las configuraciones point-to-point, la topología Arbitrated Loop (Lazo arbitrado) permite realizar la interconexión de más de dos dispositivos Fibre Channel a través de la creación de un lazo o anillo.

En un principio el lazo se creaba conectando varios dispositivos Fibre Channel en la modalidad de "daisy chain" que consiste en que el puerto de transmisión de un dispositivo (NL_Port) se conecta al puerto de recepción del dispositivo que sigue en el anillo

(NL_Port) y así sucesivamente para luego cerrar el anillo conectando el último dispositivo con el primero.

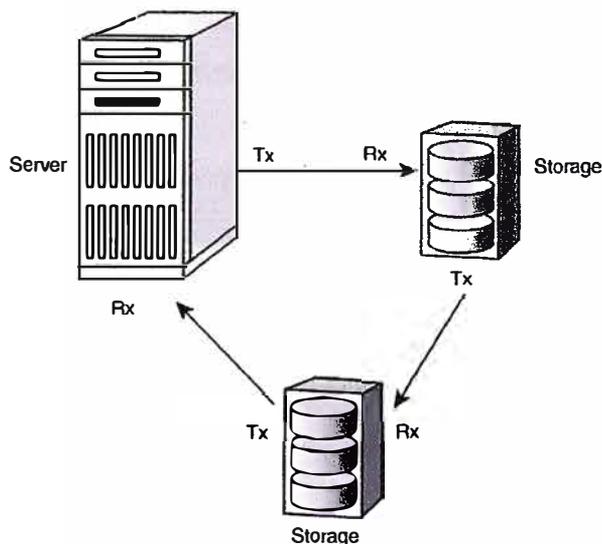


Figura 3.3.2.3.- Conectividad en la topología Arbitrared Loop

Esta topología recibe la denominación de lazo arbitrado porque tiene un esquema de arbitraje que permite decidir que dispositivo toma el control del lazo y luego lo cede hacia otro dispositivo. Asimismo posee un esquema de asignación de direcciones embebido que permite priorizar e identificar a cada dispositivo dentro del loop.

A pesar de sus ventajas con respecto a la conexión punto a punto esta topología presenta serios problemas en los casos que uno de los elementos del lazo falle, por lo cual se introdujeron los hubs que brindan la misma conectividad pero utilizando un solo puerto por dispositivo y con un circuito de bypass que sacan fuera de línea a un elemento del Loop en caso de falla del mismo sin interrumpir la operatividad de todo el loop. Por otra parte debido a que solamente pueden comunicarse dos dispositivos en el loop en un momento dado, el ancho de banda se comparte del loop para todos los dispositivos.

No entraré en mayor detalle sobre esta topología ya que en la actualidad casi a desaparecido y ha sido reemplazo por otra topología que brinda mayores beneficios por un costo similar: la topología Switch Fabric.

Topología Switch Fabric

Se denomina Fibre Channel Fabric o simplemente Fabric al conjunto de uno o mas switches interconectados entre si que permiten a los dispositivos Fibre Channel

conectados a la misma establecer conexiones punto a punto con todo el ancho de banda soportado entre los dos dispositivos (en la actualidad puede ser 200, 400 o 800 MB/s Full Duplex). De esta manera en una Fabric de 8 dispositivos que soporten 800 MB/s se puede lograr un ancho de banda total de 6400 MB/s.

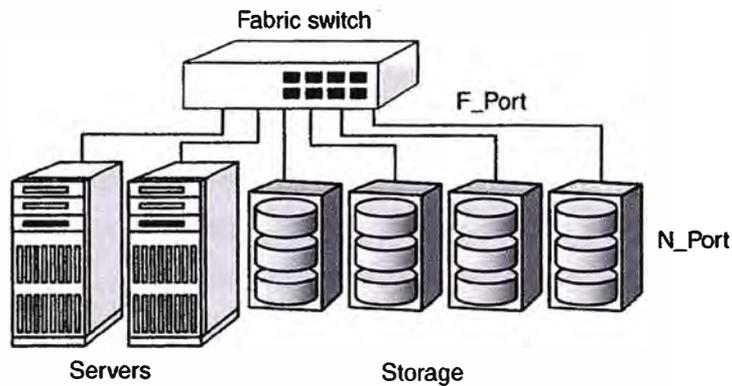


Figura 3.3.2.4.- Conectividad en la topología Fabric

El mecanismo de conmutación normalmente utilizado en las Fabric es llamado “cut-through” que funciona de la siguiente manera: cuando una trama entra desde un dispositivo conectado a la Fabric (N_Port), la lógica del tipo cut-through analiza solamente la dirección de destino a nivel de enlace FC-2 (D_ID, ver la figura 3.1.1.6) y toma la decisión de ruteo solamente en base a esta dirección de 24 bits. Este esquema permite aumentar el rendimiento en la Fabric al reducir el tiempo para la toma de decisión del ruteo ya que para ello solamente tiene que revisar el campo D_ID que se encuentra entre los primeros 4-bytes, los bytes que llegan a continuación no son necesarios para el ruteo.

El número de buffers para las tramas que se encuentran en cada F_Port (cada puerto de un switch fabric que se conecta a un Initiator o a un Target toman el nombre de F_Port) es un factor importante ya que durante periodos de mucha carga es posible que un puerto que recibe tramas de mas de un puerto no pueda procesar toda esa información y se vea obligado a descartar las tramas; en este caso cuanto mayor sea el número de buffers en ese puerto mayor será la cantidad de tramas que pueden ser encoladas y por lo tanto el rendimiento se ve afectado positivamente.

Volviendo al tema del direccionamiento, el estándar FC-SW divide las direcciones AL_PA de 3 bytes en tres campos de 1 byte cada uno. El byte más significativo se define el dominio, el segundo byte más significativo define el área, y el último byte define el puerto. El campo del dominio consta de 8 bits que proporcionan 239 valores o mejor dicho 239

dominios para una misma Fabric; es decir es posible interconectar hasta 239 switches Fibre Channel y asignarle a cada uno una diferente. El campo de área que soporta hasta 256 direcciones permite identificar los FL_Ports individuales que soportan loops, o también puede identificar un grupo de F_Ports. Un ejemplo de lo anterior es una tarjeta multipuerto dentro de un switch, los puertos de esta tarjeta se encontraran dentro de una sola área. Finalmente el campo puerto proporciona 256 direcciones para identificar individualmente los N_Ports conectados a cada área. Esta división de las direcciones de 24-bits en tres partes: dominio, área y puerto proporciona un rendimiento de conmutación superior porque la decisión de ruteo puede darse analizando la menor cantidad de bytes necesarios; por ejemplo si una trama tiene una dirección cuyo dominio es diferente al dispositivo que la origina el switch Fibre Channel sabe que debe rutear esa trama hacia otro switch sin necesidad de analizar los campos de área o puerto.

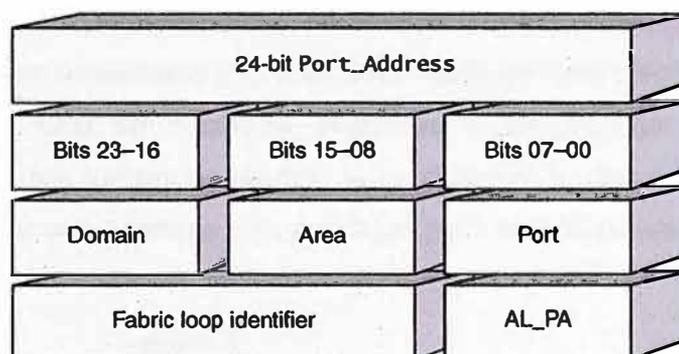


Figura 3.3.2.5.- División de la dirección de 24 bits (D_ID)

Basado en este direccionamiento varios switches Fibre Channel pueden conectarse unos con otros en forma malla para obtener redundancia de rutas desde uno hacia otro dispositivo.

Al momento que un dispositivo se conecta a la Fabric o red de switches, éste debe solicitar una dirección AL_PA a la Fabric quien le asigna una para que pueda comunicarse con los demás dispositivos ya conectados a la Fabric, a este proceso se le denomina Fabric Login (FLOGIN). Por otra parte cada N_Port que se conecta a una Fabric posee información propia que es de utilidad para el switch Fibre Channel y los demás dispositivos Fibre Channel que se conectan a la misma Fabric como por ejemplo una dirección física de 64 bytes llamada World Wide Port Name (WWPN) y otra dirección llamada World Wide Node Name (WWNN) que identifican a cada puerto y dispositivo en la industria Fibre Channel, similar a la dirección MAC en redes LAN; esta información es solicitada por la Fabric hacia el nuevo dispositivo que se conecta en un proceso llamado

Port Login (PLOGIN). La información recolectada se almacena en un Name Server contenido dentro de cada switch Fibre Channel de la Fabric para luego realizar un mapeo del AL_PA con la WWPN para realizar las operaciones de zonificación, el cual es una operación fundamental en las Fabric de hoy y que se explica a continuación.

Dado que en una misma Fabric pueden estar conectados muchos dispositivos y en la mayoría de casos se desea una comunicación limitada entre cierto número de ellos se hace necesario aislar cada grupo de dispositivos que deben conectarse entre sí de todos los demás. Por ejemplo sea una Fabric con dos servidores con sistema operativo Windows y Unix respectivamente (se supone que cada uno tiene su tarjeta Fibre Channel para conectarse a la Fabric) y que existen dos sistemas de almacenamiento, uno que contiene la data del servidor Windows y el otro contiene la información del servidor Unix, por lo tanto es obvio que no se desea que el servidor Windows acceda a la data Unix y viceversa para ello es indispensable crear dos zonas alrededor de cada par de dispositivos que deben conectarse y que aíslan a cualquier otro dispositivo conectado a la misma Fabric; con esto se consigue seguridad y se protege la integridad de la información. Existen dos formas de realizar la zonificación: a nivel de puertos o a nivel de WWN, cada cual tiene sus ventajas y desventajas pero ambos cumplen la misma función.

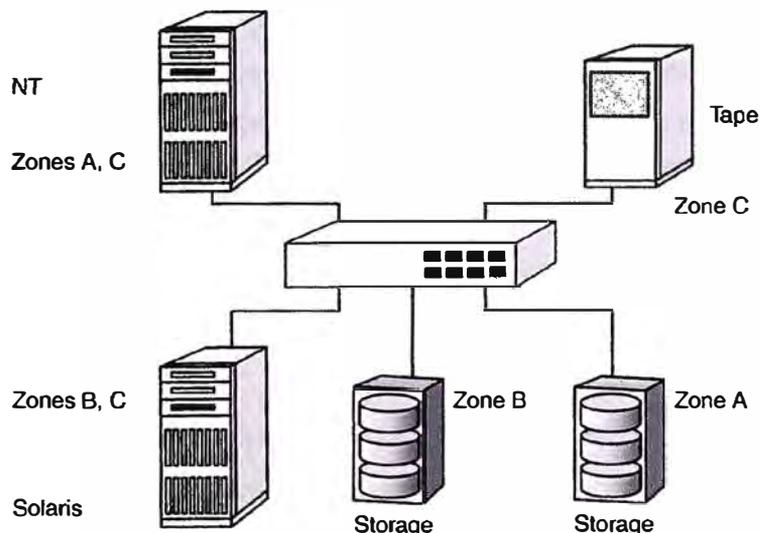


Figura 3.3.2.5.- Ejemplo de zonificación en una Fabric

Estas características de la topología Switch Fabric presentan muchas ventajas respecto a las dos topologías anteriores y es por ello que actualmente es la más utilizada.

3.3.3. Productos Fibre Channel

Las redes de almacenamiento Fibre Channel están compuestas de una serie de elementos de hardware y software de distintas marcas pero con funcionalidades similares.

En esta sección describiré los productos más importantes dentro de una red Fibre Channel prestando atención especial a aquellos que se encuentran vigentes en la actualidad.

Transceivers

El transceiver es como su nombre lo indica una combinación entre un transmisor (Transmitter) y un receptor (Receiver) y se utiliza como parte de muchos dispositivos Fibre Channel y tiene como misión transmitir y recibir señales ópticas o eléctricas hacia o desde otro dispositivo Fibre Channel a través de un medio de transmisión adecuado. Hoy en día las redes de almacenamiento Fibre Channel se basan en su gran mayoría en transceivers y medios de transmisión ópticos.

Existen muchas marcas y modelos de transceivers pero actualmente la mayoría de ellos son del tipo SFP y alcanzan velocidades de 4 Gbps aunque con estos mismos transceivers es posible conectarse a otros con velocidades menores de 1 y 2 Gbps. Adicionalmente es posible encontrar dos tipos de transceivers en el mercado, los del tipo ShortWave que emiten rayos de luz de 800nm a través de un LED y cubren distancias de hasta 500 metros con fibras ópticas multimodo y los del tipo LongWave que emiten rayos de luz de 1550nm a través de un Laser y permiten enlaces del orden de centenas de kilómetros utilizando fibras ópticas monomodo.

Estos elementos cumplen funciones del Nivel FC-0 y ya fueron descritos en el presente informe en la sección que trata sobre la pila de protocolos Fibre Channel.

Host Bus Adapters (HBA)

Los Host Bus Adaptadores Fibre Channel (Adaptadores Fibre Channel en español) son elementos que proporcionan la interfase entre el bus interno de un servidor o estación de trabajo con la red de almacenamiento externo. La interfase entre las HBA Fibre Channel y los servidores se realiza normalmente a través de un bus PCI, PCI-X o PCI-Express. Los

fabricantes de HBAs también proporcionan los drivers para que éstas puedan interactuar con los distintos Sistemas Operativos de la actualidad y con los protocolos de nivel superior.



Figura 3.3.3.1.- Adaptador Fibre Channel con SFP de 2 Gbps

La gran mayoría de HBAs tienen un solo transceiver para conectarse a la red SAN, sin embargo también existen HBAs de 2 y 4 puertos, estos dispositivos multipuertos aparecen como un solo nodo Fibre Channel que contiene 2 o más N_Ports, cada uno con su propia dirección de 24 bits (World Wide Port Name).

Tal y como se muestra en la figura 3.3.3.1 una HBA tiene incluida todos los niveles Fibre Channel desde FC-0 hasta FC-4. Al nivel FC-0 la HBA implementa funciones de transmisión y recepción para conectarse físicamente al medio de transmisión mayormente a través de un SFP de fibra óptica. Asimismo en este nivel se implementan otras funciones como los circuitos de recuperación de datos y de reloj, serialización y deserialización de la data, etc.

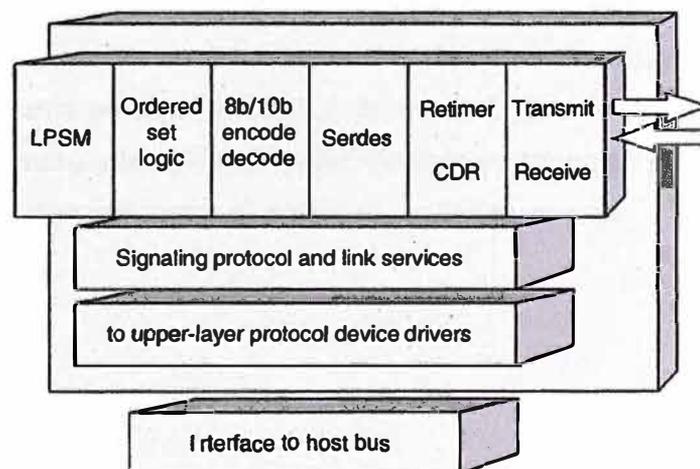


Figura 3.3.3.2.- Diagrama funcional de un Adaptador Fibre Channel

De la misma manera los requerimientos del protocolo FC-1 se cumplen a través de una lógica implementada dentro de la misma HBA que se encarga de realizar la codificación y decodificación 8b/10b para la data que se transmite y se recibe respectivamente. También podemos encontrar funciones del nivel FC-2 como la generación de tramas a partir del flujo de datos que proviene de los niveles superiores. El nivel FC-4 cumple sus funciones a través de los drivers del tipo SCSI-3 proporcionados por el fabricante de las HBAs para cada sistema operativo que son los responsables de mapear los recursos Fibre Channel hacia la combinación SCSI bus/target/LUN la cual es requerida por el sistema operativo para poder utilizar estos recursos Fibre Channel como discos de almacenamiento de información.

Almacenamiento externo con arreglos RAID

En la actualidad los equipos de almacenamiento externo en disco más utilizados son aquellos que utilizan arreglos RAID implementados con discos y puertos Fibre Channel. Los arreglos RAID se forman de un conjunto de discos independientes que se agrupan para formar un arreglo y en donde la información proveniente de algún otro dispositivo (p.e. un servidor) se almacena. La información se distribuye entre los discos de un arreglo siguiendo un algoritmo que define el tipo de RAID y que le brinda cierto tipo de propiedades al arreglo por ejemplo la redundancia de la información almacenada o el rendimiento en la lectura y/o escritura de cierto tipo de información. En la figura 3.3.3.2 se pueden apreciar dos tipos de arreglos RAID, el de la izquierda de tipo 1+0 en donde la información se escribe en forma intercalada en uno y otro drive y además se duplica lo cual crea un nivel de protección contra la falla de alguno de los discos, el arreglo de la derecha es del tipo 5 en donde los bloques de datos se distribuyen a través de varios discos y adicionalmente se crea un bloque de paridad basado en un algoritmo conocido de tal manera que frente a la caída de algún disco la información puede ser recuperada a partir de los otros discos aplicando el algoritmo en forma inversa.

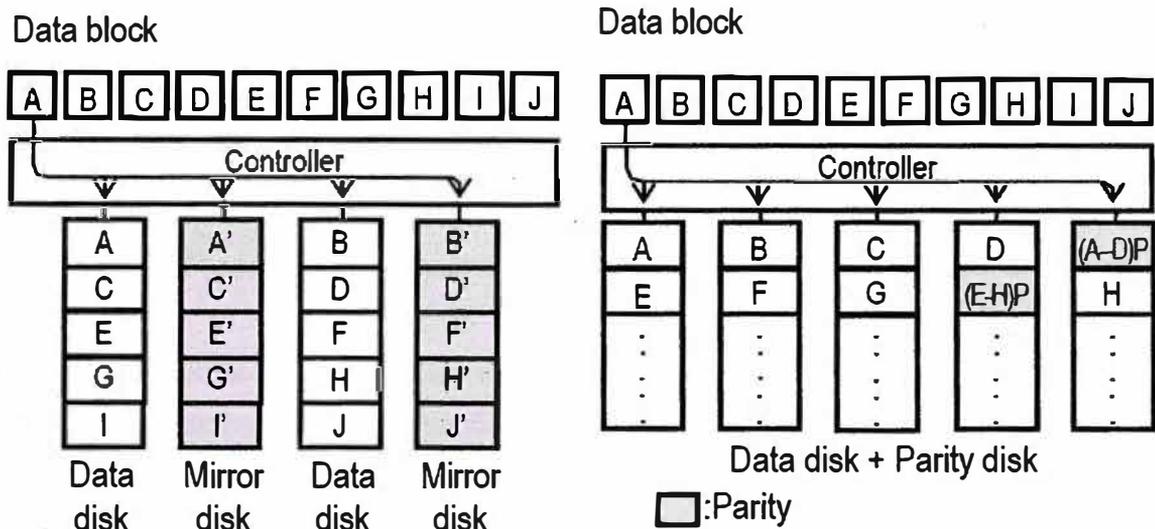


Figura 3.3.3.3.- Distribución de los bloques de datos en un arreglo RAID del tipo 1+0 y del tipo 5

Este tipo de equipos consta pues de discos de tecnología Fibre Channel o SCSI en el caso de los más antiguos con los cuales se forman arreglos RAID. Asimismo tienen una parte inteligente llamada controlador Fibre Channel o controlador RAID que por un lado se conectan a otros dispositivos a través de uno o varios puertos Fibre Channel (N_Port o NL_Port) y por otro lado se conectan a los discos y administran los arreglos formados con estos discos.

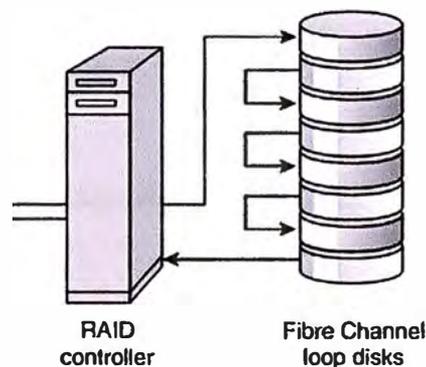


Figura 3.3.3.4.- Esquema de conectividad de un RAID controller

Switches Fibre Channel

Actualmente los switches Fibre Channel actualmente son la forma de conexión entre dos o más dispositivos, dejando de lado tecnologías antiguas como los Hubs que permitía una conexión del tipo Arbitrared Loop.



Figura 3.3.3.5.- Switch Fibre Channel de 32 puertos de 4 Gbps

La tecnología actual permite fabricar switches con puertos de hasta 4 Gbps de throughput Full Duplex, motores de conmutación de alta velocidad y ofrecen servicios de Fabric Login, Zoning, Servidor de nombres, monitoreo vía SNMP, entre otros. El tipo de implementación de estos servicios depende del fabricante que además puede incluir otros servicios de valor agregado.

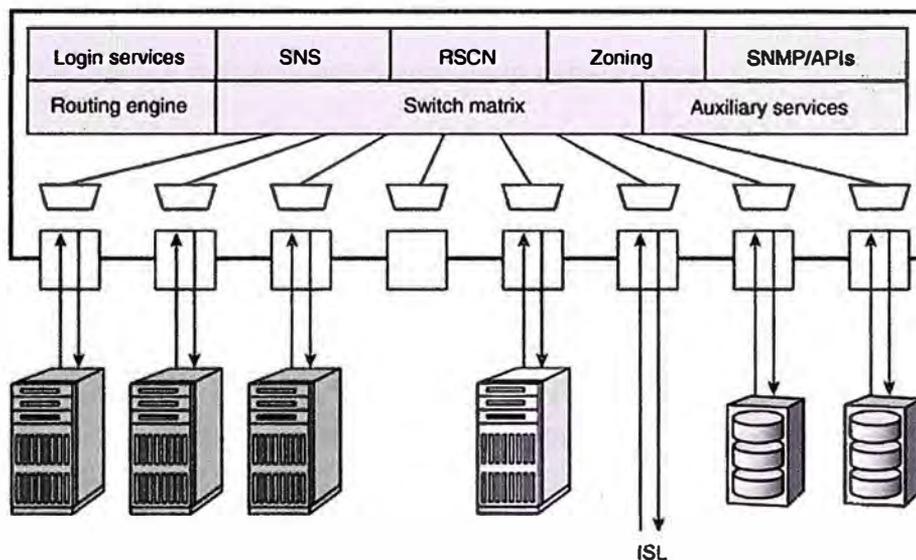


Figura 3.3.3.6.- Diagrama funcional de un switch Fibre Channel

Los switches Fibre Channel normalmente soportan hasta 32 puertos en sus modelos orientados a empresas pequeñas o medianas y hasta 64 puertos en sus modelos para grandes empresas. Sin embargo para Fabricas muy grandes generalmente se prefieren los Directores Fibre Channel que proporcionan una gran cantidad de puertos en una sola caja que además contiene características que aseguran una alta disponibilidad y generalmente son utilizados como la carretera de interconexión entre muchos switches más pequeños.

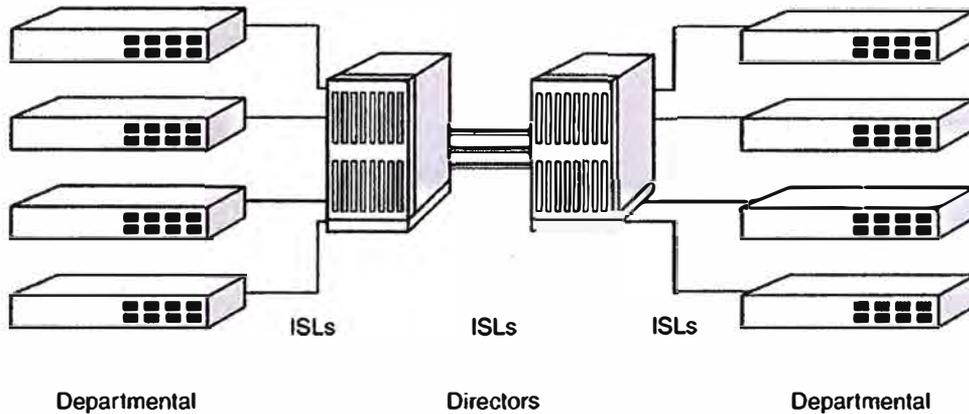


Figura 3.3.3.7.- Diagrama funcional de un switch Fibre Channel

Del gráfico anterior podemos apreciar que es posible crear Fabricas con una gran cantidad de switches y aún más dispositivos conectados. Dentro de una misma fabric las bases de datos de servidores de nombres de cada switch se mezclan de tal manera que todos tengan la misma información en sus servidores de nombre y por lo tanto la zonas creadas puedan incluir a cualquier dispositivo dentro de la misma fabric.

Por otra parte los puertos de los switches Fibre Channel pueden ser configurados de diferentes formas dando una gran flexibilidad de conexión a estos dispositivos. En el estándar Fibre Channel se definen los siguientes tipos de puertos:

- **F_Port.-** Se utiliza para conectar nodos (N_ports)
- **FL_Port.-** Se utiliza para conectar loops Fibre Channel (NL_ports)
- **E_Port.-** Se utiliza para la conexión entre switches también llamados ISL (Inter Switch Link)
- **G_Port.-** Este tipo de puerto puede actuar como un E_Port o como un F_port. Un puerto se encuentra definido como G_Port luego de que se conecta pero todavía no ha recibido respuesta al proceso de inicialización del loop o todavía no ha completado.
- **U_Port.-** Tipo de puerto universal, más genérico que el G_Port. Este puerto puede operar como un F_Port, FL_Port o como un E_Port. Un puerto se encuentra definido como U_Port cuando todavía no se encuentra conectado o todavía no ha asumido una función específica dentro de la Fabric.

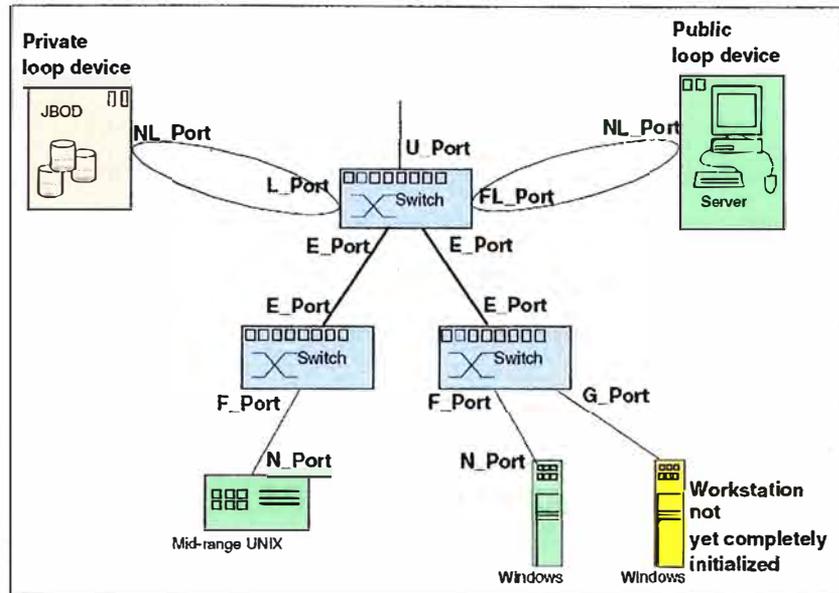


Figura 3.3.3.8.- Tipo de puertos en un switch Fibre Channel

Por último cabe resaltar que aunque cada fabricante implementa los servicios de una manera distinta normalmente la administración y el monitoreo de estos switches se realiza fuera de banda a través de un puerto ethernet vía telnet, web o SNMP.

Convertidores Fibre Channel-a-SCSI

Si bien en la actualidad la mayor cantidad de productos en la industria de almacenamiento externo y respaldo en cinta es de tecnología Fibre Channel todavía existe una cantidad considerable (y para algunos tipos de dispositivos como las librerías de backup continúan produciendo) de dispositivos con interfases SCSI paralelo. La necesidad de integrar los dispositivos SCSI dentro de una red Fibre Channel impulsó la aparición de convertidores Fibre Channel-a-SCSI que cumplan esta función.



Figura 3.3.3.9.- Convertidor Fibre Channel-SCSI con una interfaz Fibre Channel y dos interfases SCSI paralelo

Normalmente estos dispositivos poseen una o dos interfases Fibre Channel y de dos a cuatro interfases SCSI. Dado que Fibre Channel es una forma de transporte de

comandos SCSI este dispositivo simplemente realizar la conversión de un tipo de SCSI (Fibre Channel) hacia otro (SCSI paralelo)

La aplicación más común de este equipo es cuando se requiere conectar un dispositivo de cinta con interfaz SCSI paralelo a una red Fibre Channel para optimizar la utilización del dispositivo de cinta sin necesidad de comprar uno nuevo con interfaz Fibre Channel. Esto se logra utilizando el convertidor Fibre Channel-a-SCSI.

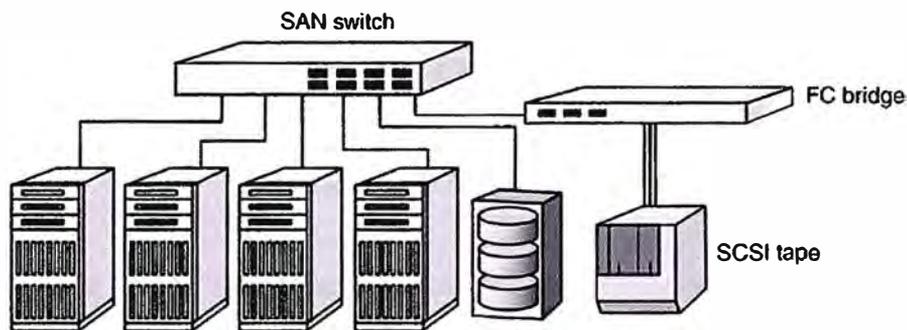


Figura 3.3.3.10.- Un dispositivo de cinta con interfaz SCSI paralelo conectado a un switch Fibre Channel a través de un convertidor

Extensores Fibre Channel

En las grandes empresas que tienen múltiples locales o sucursales distribuidos en sitios geográficamente apartados algunas veces se desea consolidar sus recursos de almacenamiento y de respaldo en cinta en un solo sitio con lo cual se centraliza la administración de los equipos, y se optimiza la eficiencia de recursos al ser compartidos por una gran cantidad de sucursales.

Sin embargo los switches y los componentes de una Fabric han sido diseñados para conectarse dentro de un mismo datacenter a distancias cortas (generalmente 500 metros) por lo cual se hace necesario incluir equipos que puedan extender la longitud de una Fabric a cientos o miles de kilómetros. Entre los dispositivos más utilizados se encuentran los multiplexores DWDM y los routers Fibre Channel-IP. A continuación una breve reseña sobre cada uno de estos equipos:

La tecnología Dense Wave Division Multiplexing (DWDM) permite el transporte de varios flujos de data sobre un mismo enlace óptico. La luz que atraviesa un cable óptico está compuesta por varios modos o longitudes de onda. Si un flujo o canal de datos está

asociado a una longitud de onda específico en el punto de transmisión, éste puede ser separado de otros flujos de datos en el receptor tal y como se muestra en la figura 3.3.3.11. Con la tecnología actual pueden transportarse hasta 64 canales de datos sobre una misma fibra, cada canal con su propia longitud de onda.

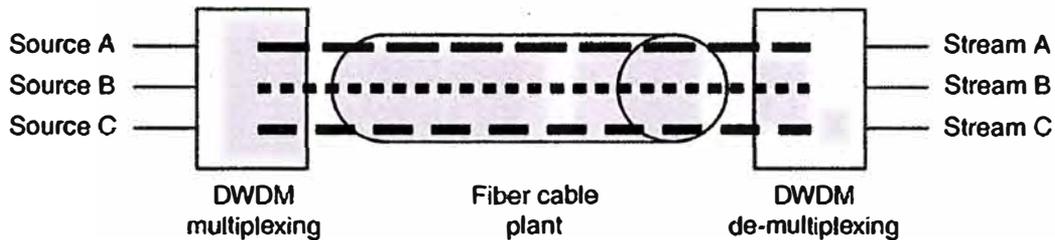


Figura 3.3.3.11.- DWDM transporta el flujo de datos de fuentes separadas en longitudes de onda separadas

DWDM requiere el uso de una fibra óptica dedicada también conocida como fibra oscura, este requerimiento incrementa los costos de esta tecnología pero proporciona un ancho de banda dedicado y siempre disponible para aplicaciones muy críticas.

Para lograr extender una Fabric simplemente hay que conectar uno o varios switches a un Multiplexor DWDM en uno de los sitios y conectar otros switches a otro multiplexor DWDM en el otro sitio, considerando que existe un enlace de fibra oscura entre ambos multiplexores.

Por otra parte también es posible extender una Fabric a través de una red IP utilizando equipos con tecnología FCIP los cuales al conectarse entre dos sites apartados a través de una red IP crean un tunel por donde se enviará la información Fibre Channel. Estos equipos fragmentan las tramas Fibre Channel y las encapsulan en datagramas IP que son enviados por la red IP los cuales al ser recibidos por el otro equipos FCIP son procesados para reconstruir nuevamente la trama Fibre Channel origen. Para acceder a la red IP pueden contar con una interfaz ethernet para conectarse a un router de borde o con una interfase para conectarse directamente a la WAN.

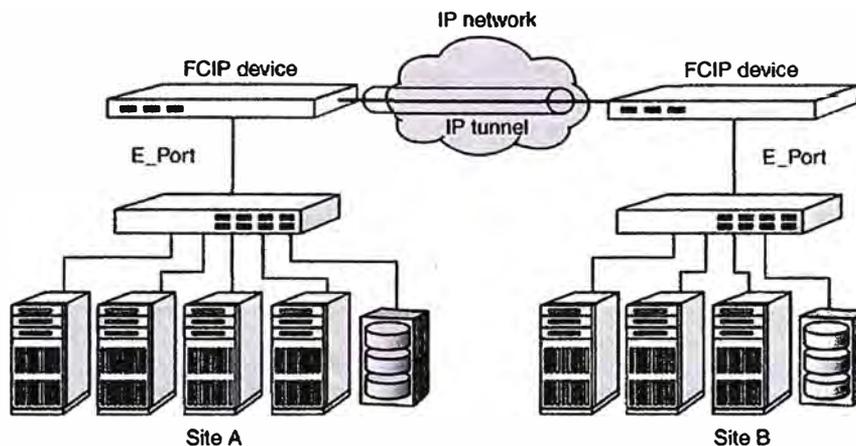


Figura 3.3.3.12.- Equipos FCIP para extender la distancia de una Fabric

Existen otras tecnologías que utilizan la red IP para conectar dispositivos de almacenamiento que serán discutidas en las siguientes secciones.

3.4. Fibre Channel sobre un tunel IP (iFCP)

El protocolo iFCP (Internet Fibre Channel Protocol) es un protocolo de gateway a gateway que se utiliza para proporcionar los servicios de una Fabric Fibre Channel a dispositivos Fibre Channel finales sobre una red TCP/IP. En otras palabras este protocolo permite conectar varias Fabric Fibre Channel aisladas a través del protocolo TCP/IP. De esta manera es posible aislar las fallas en cada una de las fabrics manteniendo la conectividad entre ellas.

Según se muestra en la figura 3.4.1 un gateway iFCP tiene una personalidad dual ya que un N_Port conectado a este gateway es tratado por un F_Port estándar. Esto implica que un gateway iFCP puede soportar servicios de una fabric Fibre Channel como por ejemplo Fabric Login, Servidor de nombres, entre otros. El gateway iFCP es el encargado de la comunicación entre el dominio Fibre Channel y el dominio IP realizando la traducción entre ambos lenguajes. Por ejemplo en la capa iFCP (iFCP layer) la dirección de 24 bits de un dispositivo Fibre Channel es mapeado a una única dirección IP de 32 bits, proporcionando de esta manera una dirección IP a cada dispositivo Fibre Channel a diferencia del protocolo FCIP que solamente asigna direcciones IP a los dispositivos que se conectan a la red IP. Adicionalmente el nivel FC-2 del protocolo Fibre Channel es sustituido por la pila de protocolos TCP/IP para una transmisión confiable vía una red IP.

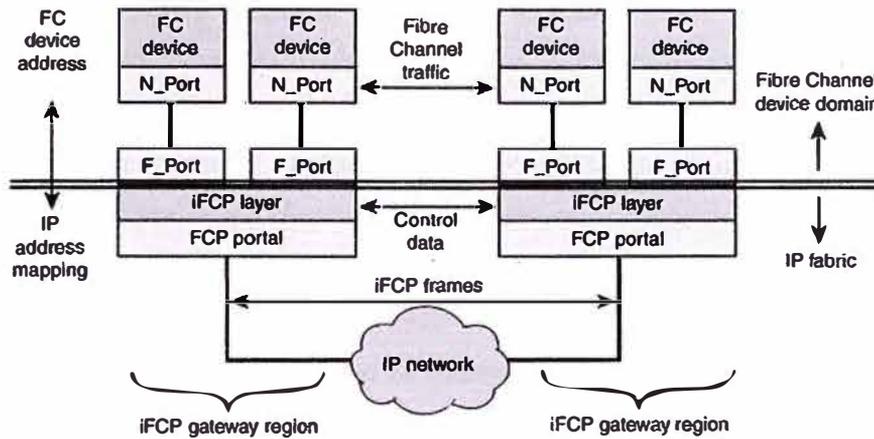


Figura 3.4.1.- Una red SAN IP basada en gateways iFCP

Dado que cada dispositivo Fibre Channel posee una dirección IP propia estos dispositivos se pueden comunicar entre todos sin restricciones tal y como se muestra en la figura 3.4.2. Este tipo de comunicación proporciona flexibilidad en el crecimiento de las redes de almacenamiento medianas y grandes.

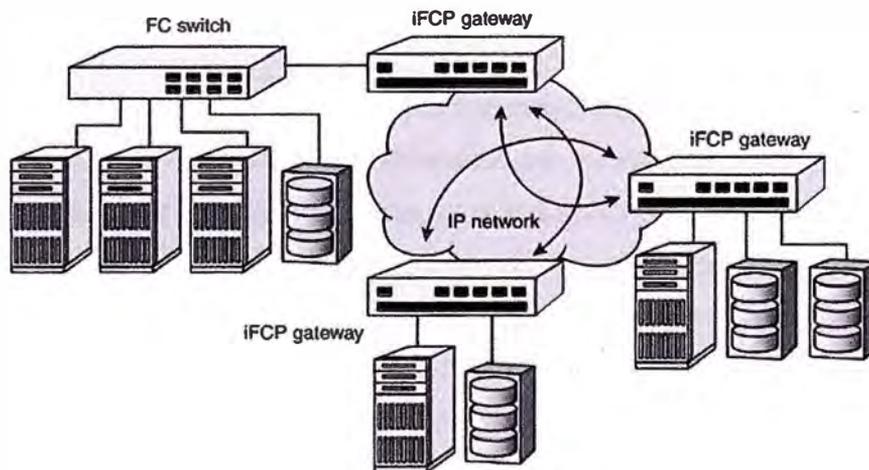


Figura 3.4.2.- Comunicaciones desde y hacia cualquier dispositivo Fibre Channel vía una red IP

A diferencia de la tecnología FCIP que solamente realiza un túnel vía una red IP la tecnología iFCP interviene activamente entre el transporte del tráfico de tramas entre N_Ports. El gateway iFCP debe manejar las transacciones Fibre Channel constantemente e interceptar los requerimientos de Port Login cuando se cruzan los dominios de Fibre Channel e IP. Cuando un E_Port se conecta a un switch Fibre Channel, el gateway iFCP debe emular una fabric terminando la conexión del E_Port en el site local. Por ejemplo en la red mostrada en la figura 3.4.2, el switch Fibre Channel vería al gateway iFCP al que se

encuentra conectado como un punto terminal en la Fabric local; todos los dispositivos en los sitios remotos de la red aparecerán como conectados localmente al gateway iFCP. De esta manera el gateway iFCP se presenta así mismo como un gran switch Fibre Channel virtual con una gran cantidad de dispositivos conectados directamente a ese switch aunque en realidad todos estos dispositivos pueden estar dispersos en distintos sitios.

La complejidad de iFCP comparado con FCIP ha impulsado a los vendedores de switches Fibre Channel a seleccionar FCIP para la extensión de la SAN, sin embargo iFCP tiene muchas ventajas que por el momento no son muy notorias pero que se podrán apreciar cuando las SAN empiecen a migrar a redes IP.

3.5. Internet SCSI (iSCSI)

El protocolo iSCSI a diferencia de los otros protocolos SAN IP vistos anteriormente reemplaza totalmente la Fabric y los dispositivos Fibre Channel con equipos IP nativos. Los Initiators y los Targets son iSCSI y el ruteo en la red es totalmente IP mientras que el transporte se encuentra a cargo de Ethernet.

Sin embargo ¿cómo es posible pasar de una red con inteligencia con Fibre Channel a una red de tipo IP? Durante las secciones anteriores se ha visto que una red Fibre Channel brinda cierta inteligencia a todo el sistema ya que asigna direcciones automáticamente a los dispositivos que se conectan, ejecuta logins, descubre dispositivos, y envía notificación de cambios. Toda esta inteligencia es necesaria ya que los targets son nodos pasivos en la red. Por otra parte en una red IP son los nodos extremos los que asumen la mayor parte de la inteligencia dejando tareas básicas para la red en sí ya que las redes IP no han sido diseñadas para conectar dispositivos de almacenamiento. Ya que la red no puede asumir la inteligencia necesaria entonces tiene que hacerlo el dispositivo iSCSI en sí, sin embargo todavía quedan algunas tareas restantes que deben ser asumidas por equipos adicionales conectados a la red.

La función principal del protocolo iSCSI es encapsular los comandos, los estados y la data SCSI proveniente desde los dispositivos de almacenamiento y respaldo o desde los sistemas operativos en forma de un bloque de descriptor de comandos (CDB por sus siglas en inglés). Este bloque es insertado en una unidad de data de protocolo iSCSI (iSCSI PDU) y este PDU es transportado por el protocolo TCP/IP.

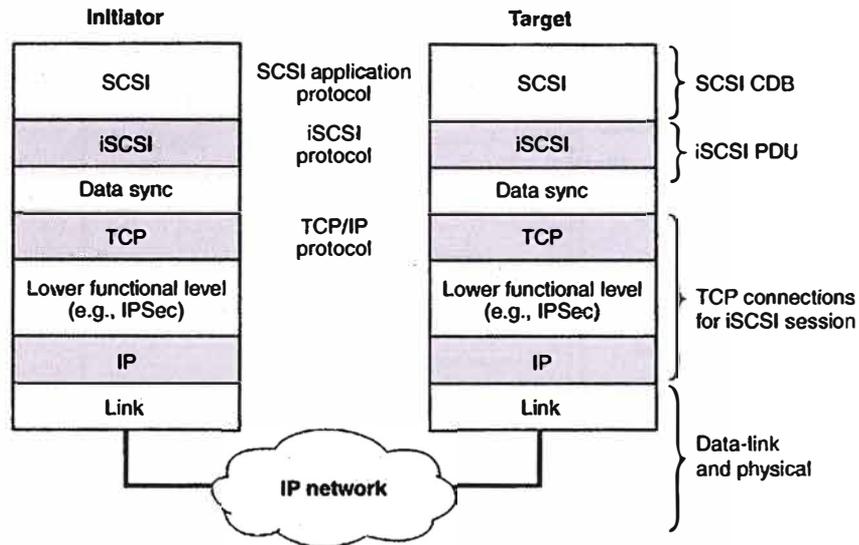


Figura 3.5.1.- La pila de protocolos iSCSI

Se pueden establecer una o más conexiones TCP para soportar las sesiones iSCSI entre un initiator y un target. Una mayor cantidad de sesiones TCP permite aumentar la cantidad de transacciones optimizando la utilización del enlace.

Para propósitos de ruteo de la data desde un initiator hasta un target se utiliza cada interface iSCSI posee una dirección IP. Sin embargo si el dispositivo iSCSI se mueve a otra red es posible que su dirección cambie, por lo cual el protocolo implementa una dirección adicional muy aparte de la dirección IP (similar a la WWN en Fibre Channel) única por dispositivo.

Asimismo los dispositivos iSCSI poseen un mecanismo de login sin embargo a diferencia de Fibre Channel el login no se realiza con la red sino con otro dispositivo iSCSI. El proceso de login tiene como objetivo negociar cualquier parámetro variable entre un initiator y un target y adicionalmente pueden invocar una rutina de seguridad para autenticar las conexiones. Cabe resaltar que la dirección IP del dispositivo se obtiene antes del proceso del login.

Luego que el proceso de login se encuentre completado, la sesión iSCSI entra en un modo de transacciones SCSI normales. Si se han establecido múltiples conexiones para una sesión iSCSI, sobre cada conexión fluirán comandos y respuestas SCSI relacionados y cada conexión es independiente de la otra.

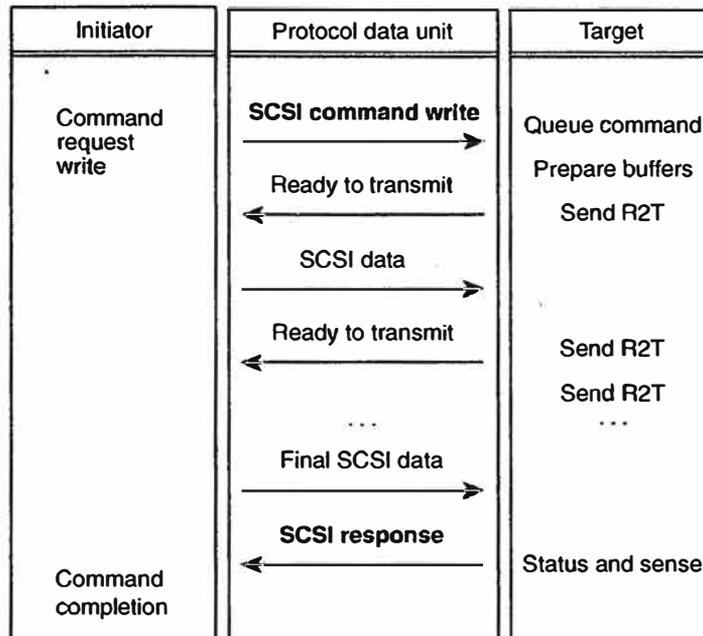


Figura 3.5.2.- Ejecución de una operación de escritura iSCSI

En la figura 3.5.2 se observa que el Initiator envía una petición de escritura y luego inicia el envío de la información mientras que el Target envía PDUs del tipo R2T (Ready to Transmit) que tienen por función controlar el flujo entre el Target y el Initiator al nivel de la capa SCSI. Estos PDU R2T son enviados por el Target cada vez que su buffer se encuentra disponible para recibir más data.

Las sesiones iSCSI y sus conexiones TCP/IP asociadas normalmente se mantienen abiertas, esperando por comandos SCSI adicionales desde la aplicación de las capas superiores a SCSI. Por ejemplo, un Initiator típicamente tiene discos asignados en la SAN y raramente rompe la conexión con éstos a menos que el Initiator o el Target sean reiniciados. Sin embargo en un entorno iSCSI el Initiator puede requerir más conexiones TCP para algunas transacciones que para otras, por lo tanto puede ser recomendable cerrar una sesión con todas sus conexiones TCP/IP relacionadas con la finalidad de brindar más recursos para la sesión iSCSI que los necesita.

Un tema central en el desarrollo de iSCSI es la seguridad de la información almacenada que fluye sobre redes IP. Este tema se está enfocando a través de la implementación de una serie de especificaciones agrupadas bajo el nombre de IP Security o IPsec. La arquitectura IPsec está definida en el RFC 2401 de la IETF (Internet Engineering Task Force). En este documento se definen los dos componentes principales de IPsec:

- Autenticación de la identidad de los puntos de comunicación en la red.

Encriptación de la data

Ambos componentes dependen de llaves privadas y públicas que se utilizan en los algoritmos de encriptación en el emisor y el receptor.

3.6. Conexión de Fibra para entornos Mainframe (FICON)

FICON (Fiber Connectivity) es un protocolo de almacenamiento que encaja en la pila de protocolos Fibre Channel en el nivel FC-4. Es utilizado por los Mainframes IBM y por dispositivos periféricos que se conectan a estos como por ejemplo arreglos de discos, lectores de cintas, etc. FICON es el reemplazo natural de las interfases ESCON utilizadas con anterioridad y básicamente mapea el protocolo ESCON sobre Fibre Channel de manera que lo utilice como medio de transporte, con ello consigue superar la velocidad de transferencia de 17 MB/s para ESCON consiguiendo ratios de 400 MB/s con FICON en la actualidad.

3.7. Almacenamiento en red (NAS)

Generalmente la tecnología NAS (Network Attached Storage) no es considerada como una SAN, sin embargo he decidido incluir una breve reseña de esta tecnología debido a su gran popularidad en el mercado y a que muchas veces compite y algunas otras se complementa con las demás tecnologías mencionadas anteriormente.

NAS utiliza un protocolo basado en archivos mientras que las otras tecnologías de redes de almacenamiento utilizan protocolo basados en bloques como por ejemplo el protocolo SCSI. En otras palabras un dispositivo NAS no es más que un disco (o un arreglo de discos) conectado a una red IP con su propia dirección IP y su propio sistema operativo que ofrece espacio de almacenamiento a los servidores que se conecten a él en la forma de carpetas compartidas o unidades de red (generalmente para almacenar archivos planos aunque también pueden almacenar información algunas Bases de Datos) a diferencia de un arreglo de discos SAN en donde los servidores que se conecten a él tendrán a su disposición recursos de almacenamiento reconocidos como si fuera discos físicos internos.

Entre los protocolos más conocidos utilizados por NAS podemos mencionar los siguientes:

- Network File System (NFS).- Utilizado en sistemas UNIX
- Server Message Block (SMB) / Common Internet File System (CIFS).- Utilizado en sistemas Windows.

3.8. Tendencias del mercado de almacenamiento actual

Aunque actualmente la tecnología Fibre Channel es la más utilizada a nivel global existen otros mercados que se encuentran en crecimiento constante como iSCSI o NAS.

Actualmente muchos clientes descartan la opción de utilizar la tecnología iSCSI debido a algunas limitaciones como el ancho de banda (1 Gbps como máximo) siendo ampliamente superado por las conexiones Fibre Channel (4 Gbps como máximo), o la seguridad de la información. Sin embargo este panorama podría cambiar radicalmente con la consolidación de la nueva evolución de las redes Ethernet: 10 Gigabit Ethernet. Con este aumento de la velocidad x10, y el surgimiento de nuevos protocolos de encriptación para iSCSI, se espera un crecimiento explosivo de esta tecnología debido a un factor determinante: el costo.

En cuanto a la tecnología NAS, ha experimentado un crecimiento sostenible en estos últimos años. Solo como ejemplo, el mayor fabricante de sistemas NAS en el mundo, Network Appliance ha mantenido un crecimiento de alrededor del 30% en los últimos años, cifra superior al ratio de crecimiento de la industria de storage. La ventaja de NAS frente a las otras tecnologías es que, debido a su naturaleza, permite una mayor integración con las aplicaciones y los Sistemas Operativos, entregando valor agregado al almacenamiento, es decir nuevas funcionalidades que, actualmente, son imposibles de realizar a través de una red SAN convencional. Sin embargo NAS, más que un competidor frente a las redes SAN es mayormente un complemento y se espera que se vea afectado positivamente por la consolidación de 10 Gigabit Ethernet.

¿Y que sucederá con Fibre Channel? Para Fibre Channel el futuro también es alentador. El roadmap proporcionado por la Fibre Channel Industry Association (FCIA) muestra una evolución en el throughput, disponible para el mercado para el próximo año y futuras evoluciones. Asimismo se esperan nuevas clases de servicio para el nivel FC-2 y funcionalidades mejoradas en la seguridad de la información transportada.

Tabla 3.9.1.- Resumen del Roadmap de la tecnología Fibre Channel

Nombre del producto	Throughput (MBps)	Ratio de Línea (Gbaud)	Especificación T11 Completada (Año)	Disponible en el mercado (Año)
1GFC	200	1.065	1996	1997
2GFC	400	2.125	2000	2001
4GFC	800	4.25	2003	2005
8GFC	1,600	8.5	2006	2008
16GFC	3200	17	2009	2011
32GFC	6400	34	2012	Según demanda
64GFC	12800	68	2016	Según demanda
128GFC	25600	136	2020	Según demanda

3.9. Resumen

En el presente capítulo hemos revisado la evolución de las diferentes tecnologías de redes de almacenamiento, y sus mecanismos de operación, los cuales difieren entre sí, pero que finalmente resuelven un mismo problema: el acceso compartido de distintos sistemas hacia un mismo repositorio de datos.

La introducción de la tecnología SCSI a finales de los años 70' significó el inicio de la evolución de la forma como los sistemas computacionales acceden a la información almacenada en discos magnéticos. Aún en la actualidad el protocolo SCSI se mantiene dentro de la mayoría de tecnologías de redes de almacenamiento como protocolo de aplicación, mientras evolucionan los niveles inferiores de infraestructura física, enlace y red. Tal es así que hoy en día es posible transportar el protocolo SCSI sobre redes IP convencionales o sobre una red dedicada de Fibra Canal. Cada tecnología posee sus ventajas y desventajas, y deben ser seleccionadas cuidadosamente en busca de satisfacer las necesidades de los usuarios.

Finalmente el horizonte de estas tecnologías es promisorio, especialmente para aquellas basadas cuyo transporte se encuentra basado en redes Ethernet debido a la aparición de la tecnología 10 Gigabit Ethernet. Sin embargo es necesario considerar que la evolución a nivel de Fibre Channel no se detiene, y la industria ya espera con ansias el lanzamiento al mercado de los productos basados en tecnología de 8 Gbps.

CAPITULO IV

APLICACIONES DE LAS REDES DE ALMACENAMIENTO EN SISTEMAS COMPUTACIONALES

Las redes de almacenamiento poseen múltiples aplicaciones en los sistemas computacionales, pueden utilizarse como medio de acceso desde servidores hacia sistemas de almacenamiento en disco magnético o sistemas de respaldo en cintas, pueden también utilizarse como medio de conexión entre sistemas de almacenamiento ubicados en locales geográficamente separados.

Esta flexibilidad de las redes de almacenamiento permite a las empresas consolidar el almacenamiento de decenas, cientos o miles de servidores, en un cantidad muchísimo menor de unidades de almacenamiento, permitiendo de esta manera una utilización más eficiente de los recursos, incrementando la disponibilidad de la información y reduciendo los costos de administración de los recursos.

En este capítulo revisaremos en forma general algunas de las aplicaciones más extendidas de estas redes en entornos corporativos, especialmente las medianas y grandes empresas. La forma de presentar cada aplicación será a través de la exposición del problema, las causas y la solución basada en la aplicación de las redes y dispositivos de almacenamiento y respaldo.

4.1. Consolidación de un sistema de almacenamiento local

Problema:

La empresa X1 se encuentra en expansión, y conjuntamente con el negocio, se han incrementado la cantidad de información a almacenar. Actualmente cuenta con un parque de 30 servidores cada uno con discos duros internos que varían desde discos duros individuales de 80GB hasta arreglos RAID internos que llegan hasta los 600GB. Asimismo la empresa X1, luego de realizado un inventario de sus servidores, a detectado que en la mayoría de servidores, solamente se utiliza menos del 60% de la capacidad total del

almacenamiento interno, calculando en total un desperdicio de recursos de almacenamiento del 50%.

Por otra parte, en los últimos dos años han sufrido la falla de los discos internos de 4 servidores causando pérdidas incalculables de información.

Solución:

El problema de la empresa X1 se origina debido a que cuenta con un número considerable de servidores y la información de éstos se encuentra distribuida en diversos dispositivos de almacenamiento interno; el espacio restante en cada uno de éstos dispositivos individuales no puede ser utilizado por otro servidor que tal vez si lo necesite causando un uso ineficiente de ese recurso. Asimismo debido a que resulta muy costoso adquirir controladores RAID internos la mayoría de servidores no tienen un mecanismo de protección contra fallas de disco, por lo tanto ante la falla del mismo la información se pierde, o requiere un tiempo muy elevado para su recuperación desde respaldos a cintas.

La solución consiste en implementar un sistema de almacenamiento consolidado que pueda conectarse a muchos servidores en forma simultánea, de tal manera que dichos servidores compartan un mismo recurso de almacenamiento según sus propias necesidades.

En el mercado actual existen varios fabricantes de sistemas de almacenamiento centralizado como el comentado en el párrafo anterior que, según las necesidades del cliente, pueden escalar a cientos de TeraBytes y manejar complejos algoritmos RAID para asegurar la disponibilidad de la información. Específicamente, para sistemas de crecimiento medio (denominados Mid-Range en el mercado de almacenamiento), podemos mencionar los elementos constitutivos más importantes:

- **Controladoras de discos.**- Las controladoras de discos constituyen la inteligencia del equipo, contienen procesadores, buses internos, memorias caché, y puertos a donde se conectarán las HBAs y/o los switches Fibre Channel. Normalmente se instalan dos controladoras redundante por sistema para asegurar que ante la falla de una de ellas, la otra asuma sus funciones, de tal manera que el acceso a la información se mantiene constante.

- **Bandeja base.-** La bandeja base es el cerebro del equipo y contiene los elementos físicos para que los diversos módulos (como las controladoras y los discos duros) puedan interconectarse entre sí.
- **Bandejas adicionales (o de expansión).-** Estas bandejas se apilan con la bandeja base para poder incrementar el número de discos duros del sistema. De esta manera el sistema crece a medida que crecen las necesidades del cliente.
- **Discos duros.-** En forma son similares a los discos duros de una PC o servidor, sin embargo su tecnología es totalmente diferente, ya que se conectan, entre ellos y hacia la controladora mediante un Loop Fibre Channel.
- **Fuentes de energía.-** Las fuentes de energía reciben la energía eléctrica AC convencional y la convierten a los niveles de DC requeridos por el sistema. Al igual que las controladoras vienen en configuración redundante (dos por sistema).

En la figura 4.1.1 y 4.1.2 podemos apreciar una vista frontal y posterior de un sistema de almacenamiento como el descrito con anterioridad.

De esta manera el almacenamiento de cada disco duro que se encuentra dentro del sistema puede ser agrupado para luego ser distribuido entre los distintos servidores que se conecten al sistema. En la figura 4.4.3 se puede apreciar un bosquejo de la solución para la empresa X1.

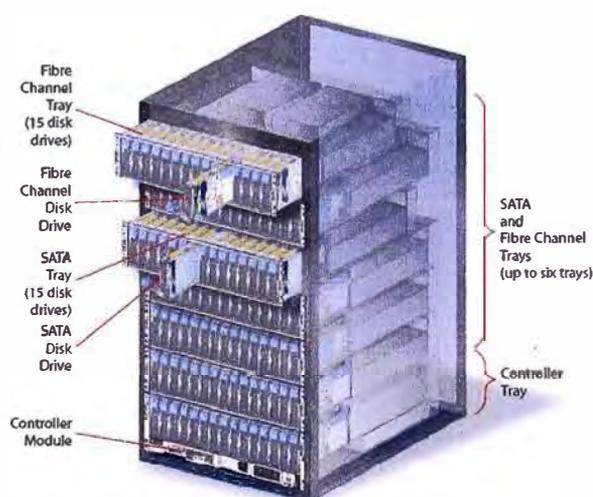


Figura 4.1.1.- Sistema de almacenamiento centralizado (Vista Frontal)

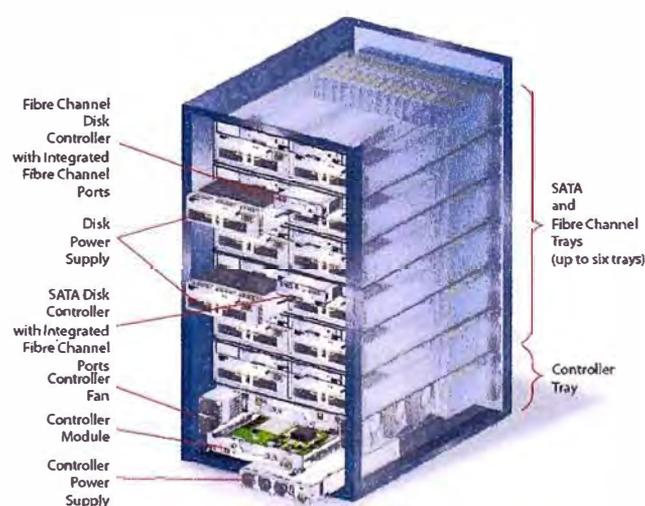


Figura 4.1.2.- Sistema de almacenamiento centralizado (Vista Posterior)

Solución de Almacenamiento Centralizado para 30 servidores basada en tecnología Fibre Channel

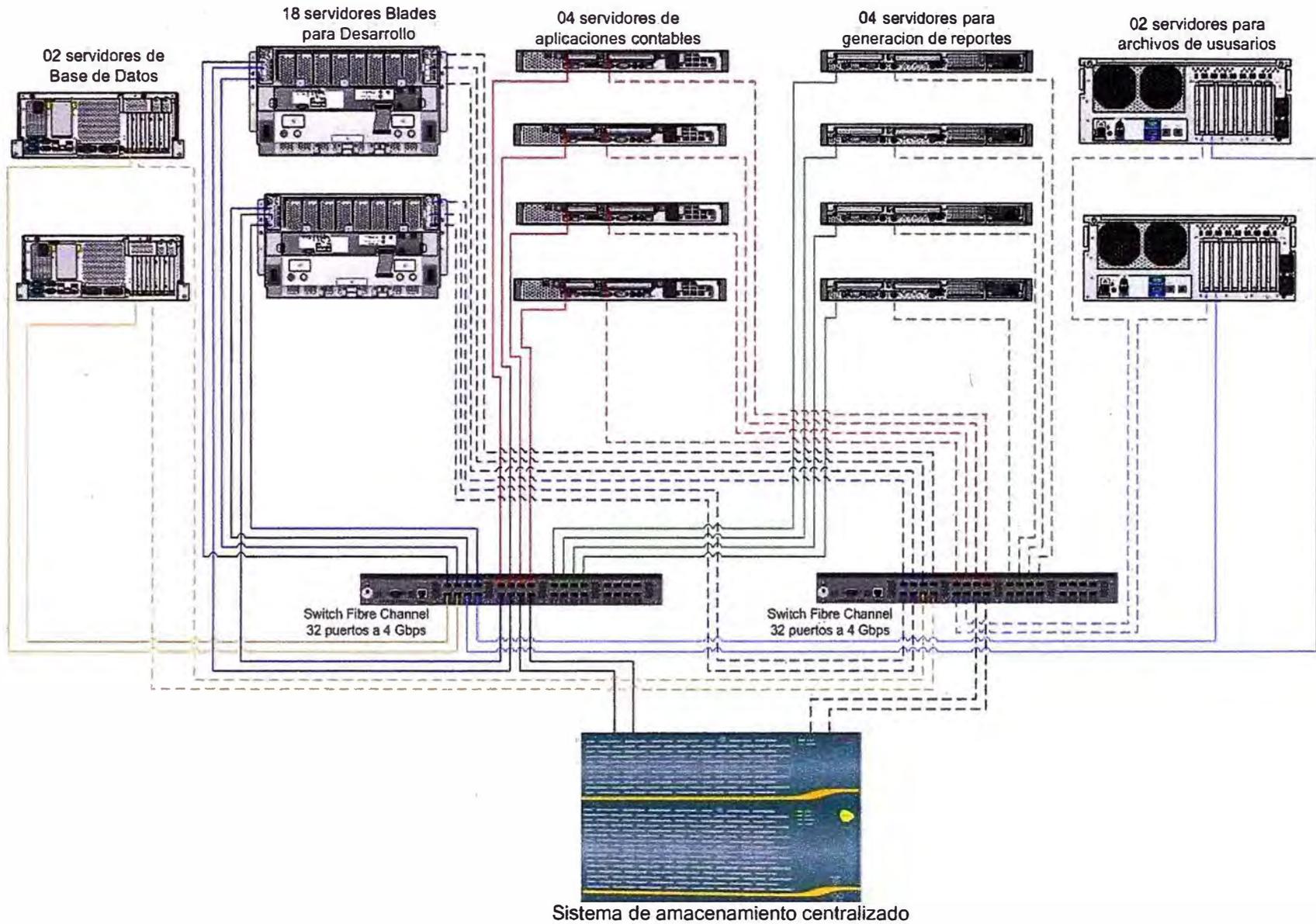


Figura 4.1.3.- Solución de almacenamiento centralizado para la EmpresaX1

4.2. Replicación de un sistema de almacenamiento

Problema:

La empresa X2 es una empresa financiera con buena presencia en el mercado y cuenta con varios miles de clientes, y está lanzando al mercado nuevos productos que, según sus propias estimaciones, incrementará su base de clientes en 70%. Dado que es una empresa consolidada cuenta con un sistema de almacenamiento centralizado con una capacidad de almacenamiento de 15 TBs. La empresa X2 optó por un sistema de almacenamiento centralizado debido a la tremenda importancia que tiene su información; una pérdida significativa de información significaría una pérdida de millones de dólares.

Asimismo esta empresa, diariamente respalda toda su valiosa información a cintas magnéticas para protegerla contra cualquier evento indeseado. Sin embargo, la cantidad de información es tal, que los respaldos toman aproximadamente de 14 a 18 horas en completarse. Frente a algún desastre que destruya totalmente su sistema de almacenamiento centralizado o todo el Centro de Cómputo, tomará varias semanas o meses adquirir un nuevo sistema de almacenamiento similar al anterior y varias horas o días adicionales recuperar toda la información guardada en cintas magnéticas. La empresa X2 tiene conocimiento que en el sector financiero existen regulaciones que obligan a las empresas a implementar la infraestructura necesaria para que la información crítica del negocio pueda ser recuperada en unas cuantas horas frente a desastres de todo el Centro de Cómputo.

Solución

Existen varias alternativas de solución para este problema, sin embargo dada la importancia de la información manejada por esta empresa y las regulaciones del gobierno, la solución óptima es implementar un centro de cómputo alternativo de contingencia que le permita reiniciar sus operaciones en pocas horas luego de un desastre sobre su centro de cómputo actual.

Este centro de cómputo alternativo estaría ubicado a una distancia prudencial del centro de cómputo actual (mayor a 30 km) y debe contener al menos un sistema de almacenamiento y algunos servidores similares a los que posee actualmente. Asimismo debe existir un enlace de fibra óptica dedicado (fibra oscura) entre ambos sites, de tal manera que utilizando un multiplexor óptico se logre extender la red LAN y la red SAN al local remoto.

La extensión de la red SAN se logra instalando switches Fibre Channel en el centro alternativo que se conectan, vía el multiplexor, a los switches Fibre Channel del centro principal a través del protocolo Fibre Channel.

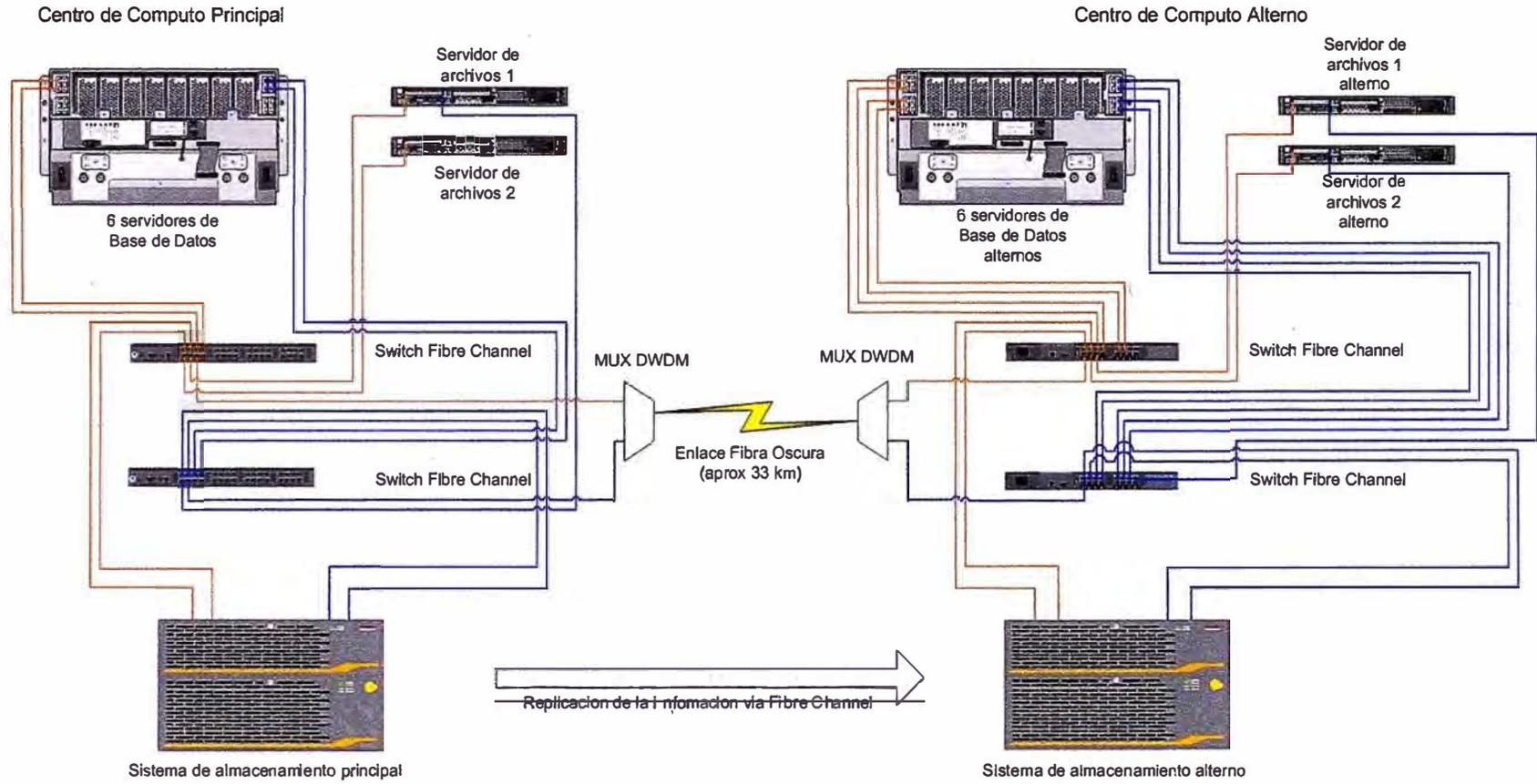
Una vez que la red SAN ha sido extendida, es posible conectar los dos sistemas de almacenamientos (el principal y el alternativo) vía canales Fibre Channel, con el objetivo que toda la información (o al menos la más crítica) almacenada en el sistema de almacenamiento principal sea replicada en forma continua al sistema de almacenamiento alternativo, de tal manera que frente a una caída del centro de cómputo principal las operaciones del negocio pueden seguir corriendo sobre la infraestructura y la información replicada en el centro de cómputo alternativo. Todo el mecanismo de replicación es controlado por el software embebido en los sistemas de almacenamiento.

Existen configuraciones aún más complicadas para brindar protección frente a desastres de gran magnitud, para ello muchas de las grandes corporaciones multinacionales replican la información de sus sistemas de almacenamiento a centros de cómputos ubicados en otros países o inclusive continentes vía enlace IP dedicados utilizando el protocolo FCIP.

En la figura 4.2.1 se presenta un esquema de la solución descrita.

Figura 4.2.1.- Solución de almacenamiento replicado para la EmpresaX2

Solucion de almacenamiento replicado en un Centro de Computo Alterno



4.3. LAN-free Backups

Problema:

La empresa X3 es una empresa minera que ha sufrido un incremento considerable de personal debido a un crecimiento inesperado de la demanda del mineral que explotan. Debido a ello, la cantidad de información generada por sus trabajadores se ha incrementado en un 60% por lo cual gastaron gran parte del presupuesto asignado al área de sistemas en adquirir un nuevo sistema de almacenamiento y una red SAN para soportar este crecimiento de información.

Por otra parte, el área de sistemas debe realizar copias de respaldo de toda su información relevante (“backups” en inglés) diariamente entre las 21:00 horas hasta las 07:00 horas del día siguiente (este lapso de 10 horas es denominado ventana de backup), con el objeto de mantener un nivel de seguridad contra pérdidas de información. La restricción del horario para realizar los respaldos responde al mecanismo de operación de los mismos. La empresa X3 tiene implementada una arquitectura estándar de respaldos basada en una librería de cintas de tecnología SCSI de 4 años de antigüedad, conectada a un servidor de respaldos, el cual a su vez tiene instalado un software de respaldo que, a partir de las 21:00 horas de cada día inicia en forma automática los trabajos de respaldos que consisten, en extraer la información relevante de todos los servidores de la empresa X3 a través de la red LAN de 100 Mbps para luego grabar esa información en las unidades de cinta de la librería robótica conectada al servidor vía un enlace SCSI. En la figura 4.3.1 puede apreciarse la arquitectura de una solución de respaldo centralizada basada en tecnología SCSI.

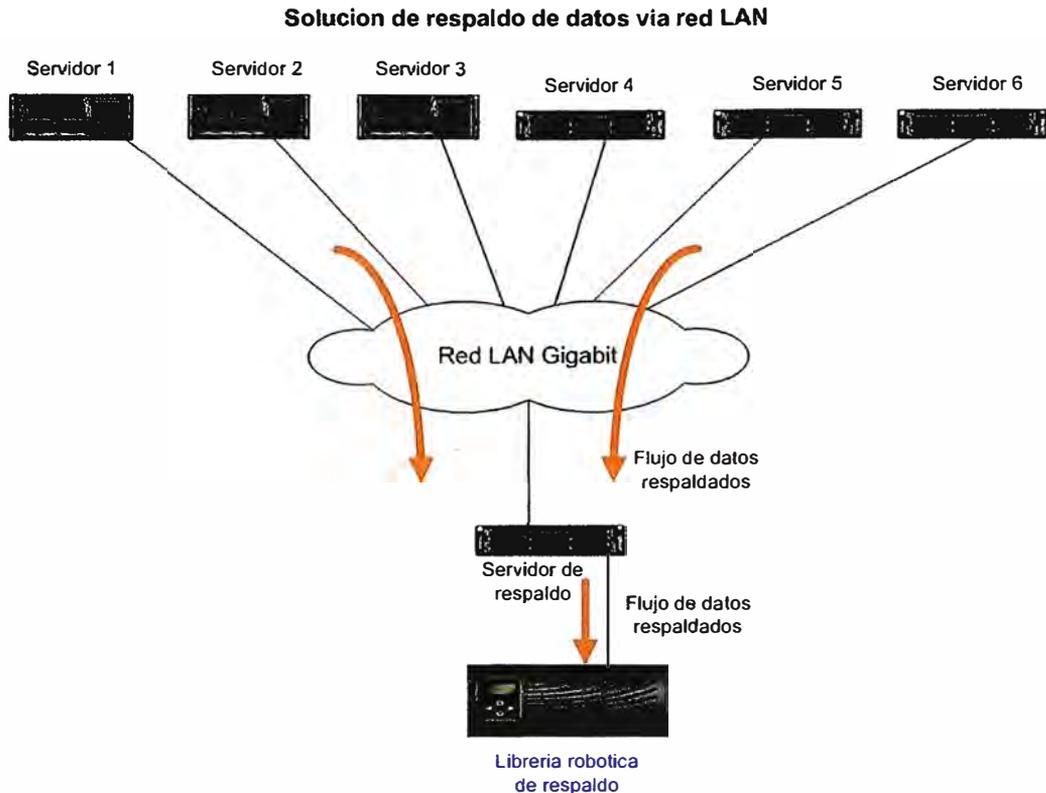


Figura 4.3.1.- Solución de respaldo de datos vía red LAN

Debido al crecimiento inesperado de la cantidad de información a respaldar, la empresa X3 se han encontrado con un nuevo problema: la cantidad de tiempo para realizar sus respaldos (“backups” en inglés) está excediendo las 10 horas de ventana de backups y está afectando los sistemas productivos, debido a las tareas de respaldo consumen recursos del servidor respaldado y de toda la red LAN.

Pensar en un centro alternativo de respaldo es económicamente inviable debido a que el área de sistemas de la empresa X3 ha consumido gran parte de su presupuesto en la adquisición del nuevo sistema de almacenamiento con su respectiva red SAN, por lo cual deben buscar una solución mucho más económica.

Solución

Como en los casos anteriores, existen diversas alternativas de solución para este problema, sin embargo dadas las condiciones planteadas, una alternativa adecuada sería utilizar la arquitectura de LAN-free backups (traducido al español: respaldos sin usar la LAN).

Esta arquitectura permitiría reducir los tiempos de respaldos a la empresa X3 debido a que los respaldos ya no se realizarían por una red LAN de 100 Mbps sino a través de una red SAN de 4 Gbps. Para ello la empresa X3 solamente tendrá que adquirir una librería de cintas robóticas basada en tecnología Fibre Channel 4 Gbps y conectarla al switch SAN que adquirió con su sistema de almacenamiento, de tal manera que todos los servidores conectados a la SAN realicen sus respaldos directamente a la librería robótica sin necesidad de pasar a través del servidor de respaldo. En la figura 4.3.2 puede apreciarse la arquitectura de una solución de respaldo centralizada basada en tecnología Fibre Channel.

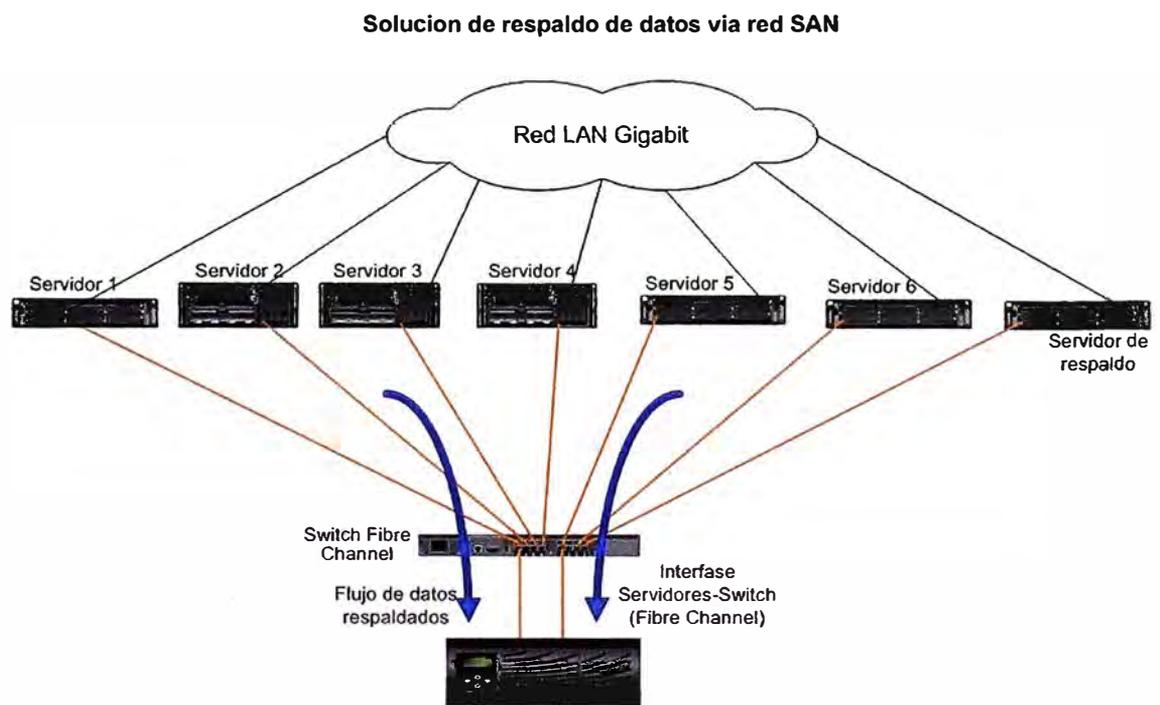


Figura 4.3.2.- Solución de respaldo de datos vía red SAN

En resumen la solución planteada presenta los siguientes beneficios:

- Elimina los cuellos de botella actuales como son, el servidor de respaldo, la red LAN, y la interfaz SCSI.
- Permite una óptima utilización de recursos al compartir la infraestructura SAN para el almacenamiento y el respaldo.
- Reduce las ventanas de backups, con lo cual elimina el problema actual y permite que la información de la empresa X3 continúe creciendo en el futuro.

CONCLUSIONES

1. Las redes de almacenamiento son cada vez más importantes debido al incremento de la información almacenada en el mundo y la también creciente demanda por accederla en todo momento.
2. El objetivo principal de las redes de almacenamiento es brindar una ruta rápida y segura para que los sistemas de procesamiento puedan acceder a la información almacenada en sistemas externos a los sistemas de procesamiento.
3. La primera tecnología de acceso a sistemas de almacenamiento externo fue la denominada SCSI, sin embargo a medida que la demanda fue en aumento, ésta resultó insuficiente; este hecho dio origen a nuevas tecnologías orientadas a redes, entre la que destaca la tecnología Fibre Channel.
4. La tremenda difusión de las redes IP está empujando la evolución de las redes de almacenamiento hacia esa dirección. Tecnologías como iSCSI, iFCP y FCIP, permiten a los sistemas de procesamiento acceder a sistemas de almacenamiento utilizando redes IP convencionales como medio de transporte. Es muy probable que, ante la aparición de la tecnología 10 Gigabit Ethernet, esta tendencia cobre mucha mayor fuerza.
5. Las redes de almacenamiento permite a muchas organizaciones utilizar eficientemente sus recursos de almacenamiento, mitigar los riesgos frente a desastres, ahorrar grandes cantidades de dinero al optimizar los procesos y reducir los costos comúnmente generados por la administración de la información.

GLOSARIO

- **ANSI:** Acrónimo de American National Standards Institute
- **FC:** Acrónimo de Fibre Channel
- **FCIP:** Acrónimo de Fibre Channel over IP
- **FICON:** Acrónimo de Fiber Conectivity
- **HBA:** Acrónimo de Host Bus Adapter
- **HIPPI:** Acrónimo de High Performance Parallel Interface
- **iFCP:** Acrónimo de Internet Fibre Channel Protocol
- **IPI:** Acrónimo de Intelligent Peripheral Interface
- **iSCSI:** Acrónimo de Internet SCSI
- **NAS:** Acrónimo de Network Attached Storage
- **RAID:** Acrónimo de Redundant Array of Inexpensive Disks
- **RFC:** Acrónimo de Request for Comments
- **SAM:** Acrónimo de SCSI Architecture Model
- **SAN:** Acrónimo de Storage Area Network
- **SAS:** Acrónimo de Serial Attached SCSI
- **SCSI:** Acrónimo de Small Computer System Interface
- **SFP:** Acrónimo de Small Form-factor Pluggable
- **SNIA:** Acrónimo de Storage Networking Industry Association

BIBLIOGRAFIA

1. Emmanuel C. Lallana and Margaret N. Uy, "The Information Age"
UNDP-APDIP, 2003
2. Storage Networking Industry Association
<http://www.snia.org>
3. David Deming, "SCSI – The Protocol for All Storage Architectures"
Storage Networking Industry Association, 2005
4. T10 Technical Committee, "Draft of Information technology - SCSI Architecture Model - 3 (SAM-3)"
American National Standards Institute, 2004
5. Ulf Troppens, Rainer Erkens, Wolfgang Muller, "Storage Network Explained"
John Wiley & Sons, 2004
6. Tom Clark, "Designing Storage Area Networks: A Practical Reference for Implementing Fibre Channel and IP SANs, Second Edition"
Addison Wesley, 2003
7. Dr. M. K. Jibbe, "Fibre Channel Technologies Current & Future"
Engineo Information Technologies Inc, 2006
8. T11 Technical Committee, "Fibre Channel – Physical Interfaces (FC-PI-4) (Draft)"
American National Standards Institute, 2006
9. T11 Technical Committee, "Fibre Channel – Framing and Signaling (FC-FS) (Draft)"
American National Standards Institute, 2003