



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS PROFESIONALES ARAGÓN

INTEGRACIÓN DE UNA SOLUCIÓN DE RESPALDO DE
INFORMACIÓN DIGITAL CON UN ARREGLO VIRTUAL
COMPARTIDO, AVC

T E S I S
QUE PARA OBTENER EL TÍTULO DE:
INGENIERO MECANICO ELECTRICO
EN EL AREA DE INGENIERIA ELECTRICA
E L E C T R O N I C A
P R E S E N T A:
ISRAEL ERICK PACHECO CABALLERO



FES Aragón

ASESOR:
ING. ELEAZAR MARGARITO PINEDA DIAZ

MÉXICO

2009



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

Objetivo.

Introducción.

CAPÍTULO 1. Almacenamiento, tendencias y ciclo de la vida de la información

- 1.1. Ciclo de vida de la información
- 1.2. La paradoja en la gestión de la información
- 1.3. Aspectos centrales en la administración de la información

- 1.4. La jerarquía en el almacenamiento
- 1.5. Tipos de respaldo de información
- 1.6. Tecnologías futuras

CAPÍTULO 2. Redes de almacenamiento

- 2.1. Interfaces de conectividad para dispositivos de almacenamiento
 - 2.1.1. Interfaz de pequeños sistemas de cómputo (SCSI)
 - 2.1.2. Interfaz de canal de fibra a (FC-AL)
- 2.2. Tipos de redes
- 2.3. Red de área de almacenamiento (Storage Area Network, SAN)

- 2.4. Tipos de conexión de dispositivos de almacenamiento
- 2.5. Componentes de una red de área de almacenamiento

CAPÍTULO 3. Medios magnéticos de almacenamiento

- 3.1. Disco magnético
- 3.2. Principios básicos de funcionamiento

3.3. Densidad de área

3.4. Cabezas de lectura-escritura

3.5. Servo-técnicas y características físicas

3.6. Codificación

3.6.1. Respuesta parcial parecido máximo

3.7. Arreglos de discos

3.8. Niveles de arreglos (RAID)

3.9. Cinta magnética

CAPÍTULO 4. Arreglo Virtual Compartido, AVC

4.1. V2X AVC

4.2. Componentes físicos

4.3. Subsistemas

4.4. Mapeo dinámico

4.5. Paridad dual

4.6. Reparación de un disco dañado

4.7. Pistas y volúmenes funcionales

4.7.1 Tabla de dispositivos funcionales

4.7.2. Directorio de pistas funcionales

4.8. Transferencia de datos

CAPITULO 5. Diseño e implementación de una solución de respaldo de información

5.1. Planteamiento del ambiente de respaldo

5.2. Integración del adaptador de ruta del servidor (HBA)

5.2.1. Sistema operativo Windows® 2003

5.2.2. Sistema operativo Solaris®

8

5.2.3 Sistema operativo HP-UX 11.x

5.3. Integración del switch de la red de almacenamiento

5.4. Integración del administrador del arreglo virtual compartido, AVC

5.4.1. Pre-requisit

OS

5.4.2. Instalación

5.5. Software de respaldo

Glosario

Conclusiones

Bibliografía

Objetivo.

Toda información esta condenada a desaparecer a menos que quede algún registro confiable, y no solamente basta que quede registrada, sino que pueda ser recuperada en tiempo y forma adecuada, para quien en determinado momento la requiera.

El valor que puede alcanzar esta información puede ser inapreciable, por ejemplo, los movimientos en las cuentas de los ahorradores de un banco, los inventarios de una tienda de autoservicio, el pago de los contribuyentes al gobierno, la información e historia académica de los alumnos de una universidad, el comportamiento accionario de una empresa. En cualquiera de estos y muchos otros ejemplos el valor que puede llegar a tener la información determina la misma existencia de la empresa, institución, o la estabilidad económica y social de una nación.

Es por ello que constantemente se diseñan medios y métodos más seguros y eficientes para manejar la cantidad de información que se genera todos los días en la vida cotidiana. El problema es que de entre tanta información generada es difícil determinar cuál puede tener en un momento determinado un valor excepcional, y cual no tendrá relevancia alguna. Esto es un factor principal debido al costo asociado al almacenamiento de la información, y aunque el costo de los medios de almacenamiento se ha reducido impresionantemente en comparación a los inicios de la industria, la cantidad de información que la humanidad genera actualmente supera por mucho cualquier pronóstico, ya que todas las emisoras de radio, televisión necesitan sus imágenes y sonidos para preparar sus programas, los gobiernos requieren de controles y estadísticas sobre su población y territorio, las escuelas y centros de investigación manipulan información que en determinados casos pueden alcanzar centenares de terabytes. Sin embargo, no toda esta información es útil, y muchas veces, la gran mayoría no vuelve a utilizarse nuevamente, el problema es que no se puede saber si en un momento dado cierta información se va o no a ocupar.

Aunque a lo largo de la historia la humanidad ha desarrollado diversos medios de llevar registro de los hechos y acontecimientos que le interesan, la capacidad y velocidad de ellos permaneció por mucho tiempo siendo muy modesta hasta hace poco tiempo, cuando la tecnología contemporánea aprovecha la capacidad de algunos metales de polarizarse ante un campo eléctrico variable, creando una nueva forma de registro que podía extenderse a velocidades muy altas y capacidades sumamente grandes. Si bien fue durante el siglo XIX que científicos como Maxwell o Faraday sentaron las bases del electromagnetismo, los medios magnéticos de almacenamiento de información empiezan a desarrollarse a partir de la segunda mitad de siglo XX, paralelamente a los sistemas de cómputo, y son por mucho los más empleados actualmente.

De tal forma, han salido al mercado soluciones en que se intenta disminuir el costo económico sin afectar en la medida de lo posible la capacidad y el desempeño de los medios de almacenamiento. Arreglos de discos de alta capacidad-baja velocidad, mezcla de componentes de fibra óptica con puramente electrónicos, y una constante e indeclinable densidad de grabación en un mismo espacio generan una competencia interminable entre los fabricantes de tecnología, en la que el estudiante universitario debe estar al tanto para integrarse exitosamente al ambiente laboral, constantemente más competitivo.

El presente trabajo tiene el objetivo de proporcionar las características más importantes y la implementación en un ambiente a gran escala de una poderosa solución de almacenamiento de información digital, el Arreglo Virtual Compartido, AVC por sus siglas (Shared Virtual Array, SVA), el cual se comporta de manera competitiva frente a las necesidades y restricciones que el manejo de información representa en estos días.

Siendo el subsecuente objetivo de esta tesis, mostrar la implementación de una solución de gran escala de respaldo de información. Sin llegar a ser tan grande como lo puede ser una solución para las instituciones bancarias, es entonces apropiada para instituciones públicas y privadas que manejen bases de datos y archivos que en conjunto rondan los 15 TB, con un crecimiento semanal de hasta 600 GB.

Introducción.

Para aseverar que la información sufre un ciclo de vida, en el capítulo 1 partimos de que el valor que para un individuo o entidad dado puede tener en un momento dado varía bajo ciertos escenarios muy comunes actualmente, por ejemplo, la obligación legal de conservar varios años registros de obligaciones tributarias empresariales, actualización de inventarios en autoservicios, operaciones bancarias. En estos y muchos otros casos la recuperación o pérdida de información puede representar un grave problema.

En el capítulo 2 se expresan las necesidades de transportar esta información digital desde su origen hasta su destino, por lo que se exponen las redes de almacenamiento digital que se han adecuando a diferentes interfaces y componentes para satisfacer las necesidades que demanda esta tecnología de información.

Es necesario recalcar que el presente trabajo no es un tratado sobre el citado ciclo de vida de la información, solo se expone debido a que es necesario para entender algunas de las características mas destacadas de AVC y entender el contexto en el que fue diseñado para trabajar. Dado que básicamente se trata de un arreglo de discos aunque con bastantes características muy particulares, en el capítulo 3 se hace una concisa explicación de lo que es el disco magnético y los arreglos de discos, para de ahí partir a describir que es y como funciona el AVC.

En el capítulo 4 se estudia al AVC no solo como una maquina en sí, sino como un componente crítico para manejar el cambio del valor de la información a través del tiempo, es decir, que se inserta en el contexto actualmente en boga en el medio de que la información sufre un ciclo de vida, termino usado en alusión a los procesos naturales en que el objeto de estudio pasa por diferentes etapas y estados hasta recobrar sus cualidades originales.

Finalmente en el capítulo 5, se expone la implementación del AVC en un entorno a gran escala. Cabe destacar que si bien una solución de respaldo empresarial es una acepción muy amplia en la que pueden caber diseños para respaldo de unos cuantos gigabytes a decenas de terabytes, el diseño de AVC aquí propuesto comprende 4.6 terabytes de capacidad nativa física, pero debido a las capacidades de compresión dispuestas por el hardware, tendremos una capacidad por compresión de 18.4 terabytes como máximo, cabe mencionar que la compresión final depende de los datos a comprimir, ya que imágenes o bases de datos son muy difíciles de comprimir, en tanto que archivos de texto plano u de otra índole pueden comprimirse en un rango superior al 60%.

La solución propuesta, dada su capacidad y características, se propone conectar a un sistema operativo Solaris 8, ya que a apreciación personal provee mejor estabilidad que Windows y se encuentra mayormente difundido que el mainframe zOS de IBM.

CAPÍTULO 1. Almacenamiento, tendencias y ciclo de la vida de la información

1.1. Ciclo de vida de la información

Es poco común hablar de un ciclo de vida de la información como puede hablarse de otros ciclos de vida como el del agua o el carbono, sin embargo, no es demasiado difícil intuir que cierta información en un momento dado puede tener un valor que disminuye con el tiempo hasta el momento en que vuelve a utilizarse, y que bajo ciertas circunstancias puede cobrar un valor inapreciable.

El entender qué pasa con la información, la cual puede ser de muy diversa naturaleza, a través del tiempo, está convirtiéndose en un aspecto crítico de su manejo efectivo. Es necesario determinar si el valor y uso de la información disminuye o aumenta con el tiempo, aunado a que anteriormente nunca se había almacenado tanta información como se hace en estos días, sin que en algunos casos sea muy claro en qué momento deba ser eliminada.

La probabilidad de volver a utilizar nuevamente la información es la base para el almacenamiento jerárquico de la información, principio sobre el cual se pueden diseñar sistemas que den un trato diferente a la información con respecto al tipo y tiempo. Sea de muy diversa su naturaleza, la cantidad de veces que la información generada es utilizada decrece significativamente con el tiempo, es decir, disminuye la frecuencia de acceso a ella.

Cuando el concepto de Cerca de Línea (Nearline¹) empezaba a ser aceptado hace quince años, la creencia común era que archivar la información era el paso directo antes de su eliminación, bajo esa línea de pensamiento, uno a dos años de almacenamiento eran considerados como un tiempo razonable para mantenerla accesible. Debido a nuevos reglamentos gubernamentales y necesidades especiales de televisoras o investigación

¹ Nearline (near on-line): Es un término usado en la ciencia de la computación para describir un medio de almacenamiento de información intermedio entre una tecnología que soporta rápido y frecuente acceso a la información generalmente representado por discos duros, y un medio de almacenamiento de respaldo de información a largo plazo que no necesariamente requiere un rápido acceso, este representado generalmente por cinta.

científica, la información creada tiene que preservarse indefinidamente, aumentando sin cesar al paso de los años.

A nivel de sistemas de almacenamiento esto significa que la información que con más frecuencia se utiliza debe ser almacenada en medios de acceso más rápidos, y por lo tanto más caros, y que conforme la frecuencia de su utilización va disminuyendo es recomendable migrarla a medios de almacenamiento más económicos, pero que en contraparte tienen tiempos de acceso más grandes. Esto se observa en la figura 1.1.

Debido a los aspectos anteriormente mencionados, la mayor parte de la información cumple un ciclo de vida en el cual su importancia crece o decrece con el tiempo y requiere medios de almacenamiento acordes a la disponibilidad que esta información debe tener y al costo que se deriva de conservarla. El concepto administración del ciclo de vida de la información de (Information Lifecycle Management, ILM), está convirtiéndose en un factor decisivo al momento de diseñar soluciones de almacenamiento de información.

Como es posible apreciar en la siguiente figura, el método más empleado cuando se requiere velocidad es la memoria de estado sólido, la que consiste en circuitos integrados en los que se almacenan secuencias de 0 y 1 lógicos, pero que tienen que estar constantemente energizadas y tienen un costo muy elevado, seguidas por el disco magnético de alto desempeño, el cual consiste en discos magnéticos que giran a una velocidad muy alta y memorias volátiles integradas.

Finalmente podemos encontrar discos de arreglo masivo barato (Masive Array of Inexpensive Disks, MAID) y cinta magnética, con tiempos de acceso de algunos segundos, a un costo mucho menor.

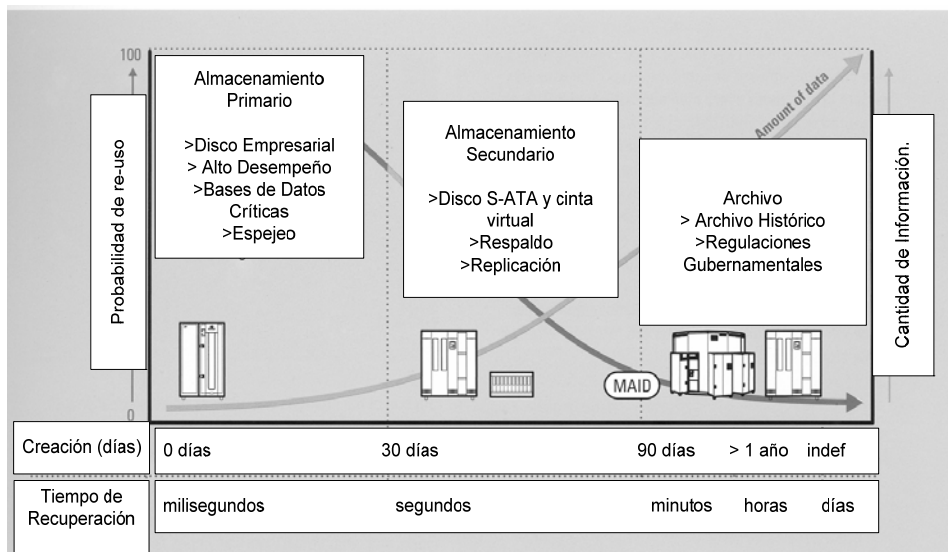


Figura 1.1.

Ciclo de vida de la información

1.2. La paradoja en la gestión de la información

Almacenamiento, ancho de banda y poder de procesamiento son abundantes y no hay barreras tecnológicas inmediatas a la vista.

El primer disco magnético comercial lanzado en 1956, el RAMAC 350 de IBM®, tenía 5MB de capacidad y medía 24 pulgadas de diámetro, con un precio de \$7,800 USD por MB, de tal forma que en aquel entonces este disco tenía un costo de \$39,000 USD. En la actualidad, el costo por MB no supera los \$0.04 USD.

La paradoja del manejo de la información define tres tecnologías fundamentales en la era de la información para procesar, almacenar y transmitir información. La tecnología de procesamiento reduce el tiempo requerido para procesar información, la tecnología de almacenamiento permite respaldar y recuperar la información, y la tecnología de ancho de banda permite que esta información sea transmitida a prácticamente cualquier parte donde el humano se encuentre.

Para representar gráficamente el comportamiento de esta paradoja, podemos tomar la figura 1.2 siguiente:

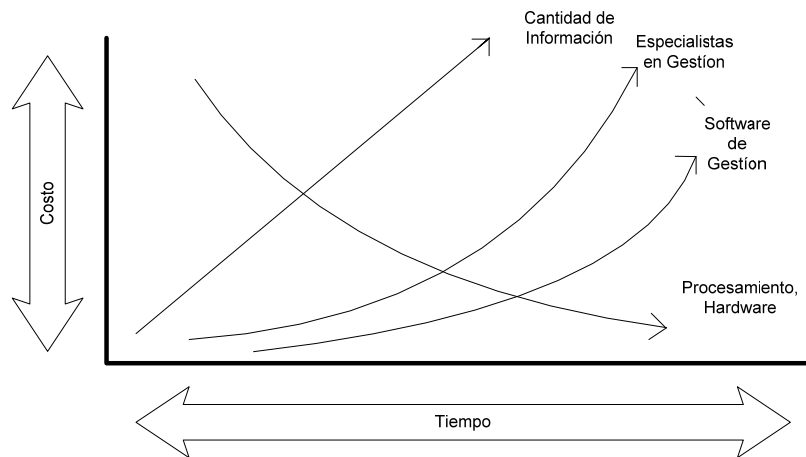


Figura 1.2.

La paradoja en la gestión de la información

De donde podemos determinar que; en tanto el costo de estas tecnologías a lo largo de los años tiende a cero, es decir, cada año es más barato almacenar un byte de información o procesarlo, el costo de manejarlo ha estado aumentando paulatinamente, ya que no ha crecido el desempeño de los medios de almacenamiento conforme ha crecido su capacidad, y la administración es cada vez más compleja debido a que hay que encontrar eficientemente la información específica entre un mar de datos que va acumulándose, en menores ventanas de tiempo, y con cada mejora en las tres tecnologías fundamentales parece que, paradójicamente, es más difícil manejar eficientemente la información.

- El precio del procesamiento de información por MB tiende a cero
- El precio de procesamiento por MB decrece de 25% a 35% cada año.
- La densidad de transistores en un circuito integrado se duplica cada 18 a 24 meses.
- El desempeño del procesamiento está mejorando del 30% al 35% cada año.
- El precio de almacenamiento de información por MB tiende a cero:

- La densidad de grabación en medios magnéticos se incrementa de 60% a 100% cada año.
- El costo de almacenamiento por MB decrece de 30% a 40% cada año.
- El costo de ancho de banda tiende a cero:
- A nivel mundial la demanda se duplica cada año.
- Los precios decrecen aproximadamente 35% cada año.
- La fibra falta de sustituir 80% del total de cableado instalado

Pero en contraposición surgen los siguientes problemas

- El costo de la administración del almacenamiento puede alcanzar hasta 5 veces el costo del hardware en ambiente de almacenamiento directamente conectado (Direct Attached Storage, DAS)
- El desempeño del disco tiene un incremento menor al 10% anual
- Las redes de transmisión son mucho más rápidas que los equipos a que van conectadas
- Cada vez es más difícil encontrar personal capacitado para controlar los nuevos ambientes de respaldo
- La cantidad de información digital crece mundialmente de 50% a 70% cada año. En un par de años, se habrá creado más información que toda la que se creó desde la primera civilización humana hasta ahora.

1.3. Aspectos centrales en la administración de la información

Hoy en día, un ambiente común de respaldo, ya sea en plataformas UNIX® o Windows®, tiene de 1 a 5 TB almacenados en disco, pero cada administrador tiene la capacidad de administrar eficientemente hasta un 1 TB. La demanda por almacenamiento está superando rápidamente a las herramientas para administrarlo, y se considera que en el 2007 de continuar así, se tendrá una capacidad de administrar 15 TB de un total de 60. Los negocios tendrán que verse obligados a administrar la información que en verdad tiene importancia.

En tanto que los datos continúan escalando en valor, seleccionar el sistema de almacenamiento que mejor empate las características de acceso de la aplicación que se pretende aplicar redituará en una mejor inversión. Cada aplicación de acuerdo a sus

características requerirá un medio de almacenamiento diferente. Por ejemplo, las bases de datos, compuestas por índices, apuntadores y tablas se ajustan mejor a un disco de estado sólido para un óptimo desempeño, en tanto que archivos largos de mediana disponibilidad pueden almacenarse en cinta magnética.

Como es de esperarse, la información cobra verdadera importancia al momento de recuperarse, por lo que una correcta forma de recuperación debe ser considerada. Se pueden identificar cuatro diferentes niveles de clasificación:

Información crítica: la falta de acceso a ella, aunque sea por un breve periodo, pone en riesgo a la empresa, por lo que comúnmente se utiliza la replicación en espejo y servidores en racimo o grupo, donde casi instantáneamente se puede estar de vuelta por un desastre, pero se ocupa el doble de recursos. Un ejemplo son las transacciones bancarias a través de los plásticos emitidos (Tarjetas de crédito/débito) o las realizadas en ventanilla.

Información vital: es parte del proceso normal de negocios y son tolerables pocos minutos para recuperarla. Los arreglos de discos y librerías robóticas de cinta caben en este segmento. Un ejemplo es la información de los pasajeros de una aerolínea.

Información sensitiva: se pueden tolerar algunas horas para recuperarla. Discos de bajo costo y cinta son las principales opciones. Un ejemplo es el censo de población.

Información no-crítica: puede reconstruirse en un periodo de varias horas o días. Típicamente consta de archivos personales, y puede usarse almacenamiento removible. Un ejemplo es el correo electrónico de un usuario, con más de un mes de antigüedad.

En la Tabla 1.1 se muestra un cuadro con la tecnología preferida de almacenamiento, en función del tipo de información.

Aplicación	Lectura/Escritura	Acceso Aleatorio/Secuencial	Tecnología Preferida
Base de Datos	L/E	A/S	Disco RAID (1,5,6)
Archivo Plano	L/E	A/S	Disco RAID (1,5,6)
Log	L/E	S	Disco RAID (1,5,6)

Archivo Sistema	L	A	Disco RAID (1,5,6)
Servidor Archivos	L	A/S	Disco NAS
Aplicación	Lectura/Escritura	Acceso Aleatorio/Secuencial	Tecnología Preferida
Respaldo Recuperación	L/E	S	Cinta, Disco RAID
Archivo	L	A/S	Cinta

Tabla 1.1.

Dispositivos de almacenamiento de acuerdo a la aplicación

1.4. La jerarquía en el almacenamiento

La industria del almacenamiento ha usado tradicionalmente una pirámide para mostrar la jerarquía de productos en función de su precio, desempeño y capacidades. El rango dentro de la jerarquía para los tres parámetros mencionados anteriormente es amplio, y existen algunos espacios vacíos como el del acceso aleatorio cercano a la línea, que será ocupado con librerías automáticas y discos de arreglo masivo inactivos. Los discos de estado sólido proveen el más amplio desempeño de Entradas/ Salidas (Input / Output, I/O) comparado con cualquier otro dispositivo y consiste en tecnología de memoria dinámica de acceso aleatorio (Dynamic Random Access Memory, DRAM) que elimina enormemente la latencia² en la transferencia de datos.

² Latencia: Es un tiempo de retraso entre el momento entre que algo es iniciado y el momento en que toma efecto o se vuelve detectable.

Los discos magnéticos contienen alrededor de 95% de toda la información crítica, son de muy alta confiabilidad y disponibilidad, y han claramente definido los niveles de desempeño aceptables para el resto de los medios de almacenamiento. En el nivel cercano de línea (Nearline) tenemos a las librerías automatizadas de cinta, arreglos de discos de arquitectura de tecnología avanzada - serial (Serial-Advanced Technology Architecture, S-ATA).

En la parte más baja tenemos los medios lejanos de línea (far-line) manualmente utilizados, es decir, medios de almacenamiento removibles, tales como disquetes, video discos e incluso papel en que se encuentran cartas y oficios. El concepto de jerarquizar fue empleado hasta hace poco tiempo, cuando la necesidad de soluciones integrales desplazó a sistemas aislados.

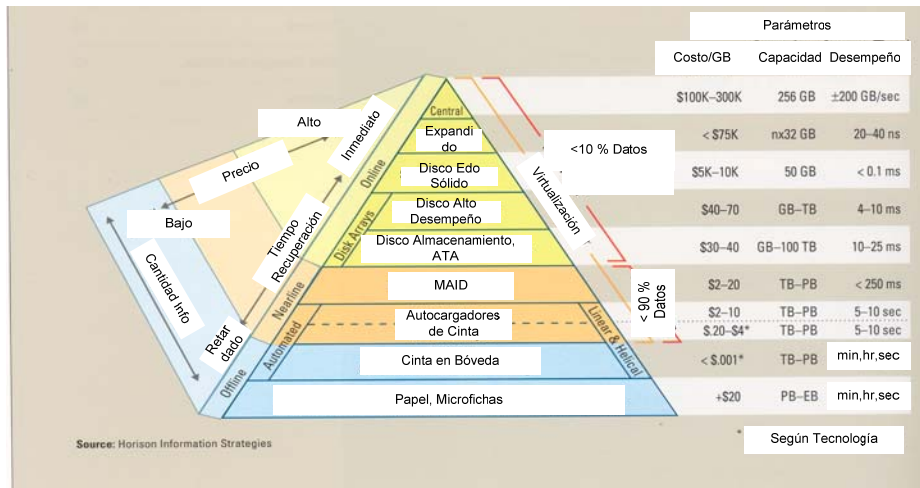


Figura 1.3.

La jerarquía del almacenamiento

Como se puede observar en la figura 1.3, 90% de la información se puede ubicar en medios baratos de baja disponibilidad, y sólo 10% requiere de medios de alta disponibilidad de costo mucho mayor. El costo de tener información de alta disponibilidad fácilmente puede deducirse considerando la seguridad, velocidad de procesamiento y redundancia en equipos de almacenamiento en que se debe invertir para tener al sistema trabajando.

En la actualidad el estándar de ultra alta disponibilidad requiere de 99.999% de confiabilidad, es decir, que el proveedor del equipo está comprometido a que en caso de que su equipo incurra el alguna falla, ésta no deberá superar el orden de 5 minutos al año, de lo contrario puede hacerse acreedor a penalizaciones económicas muy altas.

1.5. Tipos de respaldo de información

Como es de esperarse, la cantidad de espacio para almacenar información, por explosivo que sea su crecimiento, es limitada, y por lo tanto costosa. Ante ello surge la necesidad de discriminar entre la información que merece o no ocupar tal espacio, y cuál es la mejor manera de manejarla.

Del total de información que un usuario de red mantiene en su equipo personal, probablemente ni siquiera la tercera parte sea sensible a la empresa en alto grado, y probablemente la mitad no lo sea en grado alguno. Por otra parte se tiene el caso de las bases de datos, a las que accede mediante una sesión al servidor, donde se encuentra toda la información de sus clientes, inventarios y proveedores. Con la pérdida de esta información, inclusive la existencia misma de la organización puede estar en juego.

Ante ello se plantean 2 formas en que la información es generalmente respaldada:

Total: toda la información es respaldada. En caso de falla del servidor se puede cargar la información en otro y en poco tiempo estar nuevamente disponible. Cabe mencionar que no es un racimo o grupo (cluster), en que la falla de un servidor pasa los servicios a otro, ya que en caso de corrupción de información esta se presentará en ambos servidores, por lo que se debe recuperar la información del último respaldo no corrupto. Este tipo de respaldo es para los servidores y no para los usuarios, que suelen limitar sus respaldos a algunas carpetas o archivos.

Incremental: el software de control ordena únicamente respaldar los archivos que han sido recientemente creados o modificados, de tal forma que puede recuperarse la información mediante el respaldo total y los incrementales de menor tamaño y por lo tanto más rápido.

Es práctica común en la mayoría de las organizaciones en que la actividad disminuye los fines de semana, realizar un respaldo total los días de asueto donde pocos usuario y respaldos incrementales diarios en días laborales, que son mucho más rápidos. Después de una semana se hace un nuevo respaldo total y el espacio que ocupaban los incrementales es reutilizado para los incrementales de la semana en curso, de forma que se puede utilizar más o menos el mismo espacio de almacenamiento para tener los respaldos al día sin necesidad de estar comprando medios de almacenamiento constantemente.

1.6. Tecnologías futuras

Todos los fabricantes tienen proyectadas tecnologías que habrán de lanzar a futuro para competir en el mercado. En el área de almacenamiento se exponen brevemente un par que parecen muy prometedoras en desarrollo por IBM®.

Memoria magnética de acceso aleatorio (Magnetic Random Access Memory, MRAM): tecnología de celdas magnéticas microscópicas en desarrollo por IBM®, tienen muy bajo consumo de energía y es no volátil. Mucho más rápida que las memorias Flash, y se proyecta que será aproximadamente seis veces más rápida que las memorias DRAM que se encuentran actualmente en el mercado.

Ciempies (Millipede): es otra tecnología alternativa desarrollada por IBM®, en laboratorio ha alcanzado mil billones de bits por pulgada cuadrada, alrededor de 20 veces más que los discos magnéticos actuales. La información es representada al hacer perforaciones microscópicas de 10 nanómetros en un material plástico. Está proyectado para ser utilizado para aplicaciones personales en principio, ya que la tasa de transferencia de datos y la confiabilidad no es muy robusta, en función de esto, faltan aún algunos años para que sea utilizada en centros de cómputo.

CAPÍTULO 2. Redes de almacenamiento

2.1. Interfaces de conectividad para dispositivos de almacenamiento

Existen 2 características esenciales que buscamos cumpla cualquier dispositivo de almacenamiento: velocidad y capacidad de transportar grandes volúmenes de información ininterrumpidamente. Hay 2 interfaces que cumplen satisfactoriamente estos requerimientos y son utilizados casi exclusivamente por los fabricantes de equipos para almacenamiento masivo de información, que han adaptado sus equipos para transmitir los comandos de control en forma nativa en estas interfaces.

Interfaz de Pequeños Sistemas de Cómputo (Small Computer Systems Interface, SCSI)

Interfaz de Canal de Fibra (Fibre Channel, FC)

2.1.1. Interfaz de pequeños sistemas de cómputo (SCSI)

SCSI es un conjunto de estándares destinados construir un bus de E/S de alta velocidad independiente del dispositivo que se conecta, de tal forma que se pueden conectar al mismo cable discos duros, discos compactos y unidades de cinta, entre otros. El número de dispositivos va de 8 a 16 dependiendo del tipo de SCSI. Cada uno de los dispositivos conectados en un bus SCSI está identificado por un número de blanco o destino (target) y número de unidad lógica (Logical Unit Number, LUN). En caso de que se tenga varios dispositivos en el mismo bus a cada uno se le asigna un identificador (ID) SCSI, acto que se lleva a cabo mediante un selector externo (jumpers), o automáticamente si son subsistemas de uno más grande, por ejemplo, cartucheras de de lectura-escritura dentro de una librería de cintas.

Un dispositivo sofisticado, podrá crear unidades lógicas dentro de un mismo destino, incrementando el número de dispositivos que el sistema operativo del servidor al que se conecta pueda controlar, y que varía según el sistema operativo.

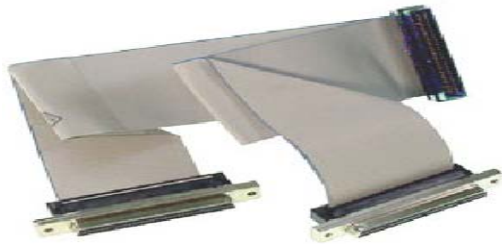


Figura 2.8.

Cable universal SCSI

El cable que se presenta en la figura corresponde a un cable universal SCSI de 68 pines, que el más usado.

La información viaja de y hacia el servidor mediante una adaptadora al bus del servidor (Host Bus Adapter, HBA) instalada en la tarjeta madre y es denominada iniciador (initiator), por ser quien indica qué dispositivo y cómo ha de trabajar. Como estándar tiene el SCSI ID 7.

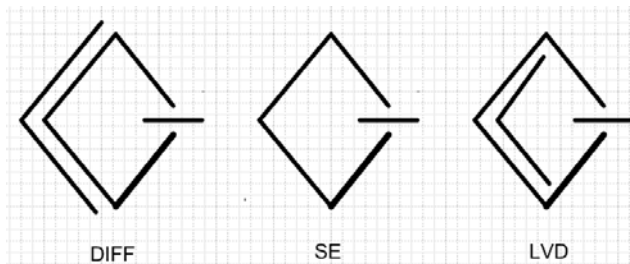


Figura 2.4.

Símbolos para tipos de SCSI

Es necesario distinguir los 3 tipos de señales que existen en SCSI, si bien actualmente SCSI diferenciador de bajo voltaje (Low Voltaje Differential, LVD) está muy cerca de ser el único tipo de señal SCSI existente en el mercado:

Terminación Única (Single Ended, SE): utiliza una señal de tensión de 5 Vcc contra tierra y la señal es la diferencia entre ambas, la distancia máxima del cable es de 6m, puede controlar hasta 15 dispositivos por bus.

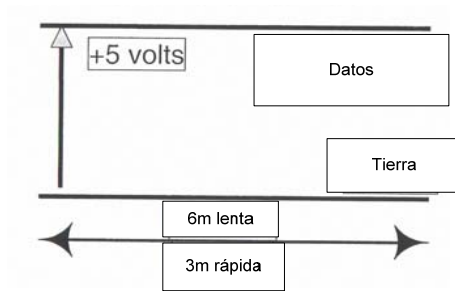


Figura 2.5.

Señal de Terminación Única (SCSI Single Ended)

Diferencial (Differential, DIFF): También conocido como diferencial de alto voltaje (HVD, High Voltage Differential). Emplea dos líneas de tensión y la señal es la diferencia entre ellas. La distancia máxima que puede cubrir es de 25 m, puede controlar hasta 15 dispositivos por bus.

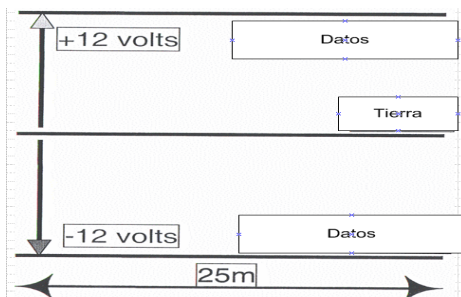


Figura 2.6.

Señal Diferencial (SCSI Differential)

Diferencial de Bajo Voltaje (Low Voltage Differential, LVD): introducido con el estándar SCSI-3, también llamado ultra 2 amplio (SCSI ultra 2 wide), es el medio comercial SCSI más veloz y el más utilizado en el campo, con una tasa de transferencia máxima de 80 Mb/s. La señal es la diferencia de voltaje entre dos líneas de 3.3 volts, puede controlar hasta 15 dispositivos por bus.

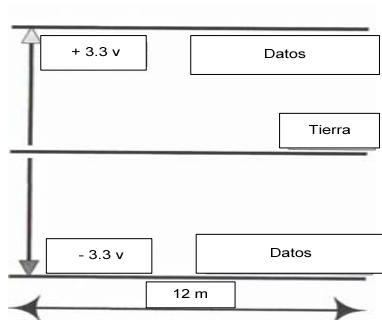


Figura 2.7.

Señal diferencial de bajo voltaje (Low Voltage Differential, SCS LVD)

Como podrá notarse en la siguiente tabla, SCSI es un interfaz muy rápida y en condiciones ideales puede alcanzar picos de hasta 288 GB/h. Es una muy buena opción para empresas pequeñas y medianas, además de económica, ya que una HBA SCSI tiene un costo de alrededor de \$300 USD pero presenta un par de serias desventajas:

Casi no realizan procesamiento sobre la información que trabajan, lo que resulta en un consumo considerable de los recursos del servidor que puede alentarlos, por lo que se utiliza para respaldos nocturnos.

Ancho de banda: si se tiene un solo dispositivo de almacenamiento el rendimiento es excelente, pero a medida de que se agregan nuevos dispositivos se van repartiendo el ancho de banda, de modo que si conectamos 4 dispositivos en LVD tendremos una velocidad máxima de 20 MB/s.

La conexión más utilizada es directamente una unidad de cinta o un pequeño arreglo de discos a uno de los servidores de mayor producción, de tal forma que cada servidor tendrá un espacio físico donde respaldar o recuperar su información. En caso de que algún servidor, unidad de cinta o volumen de disco falle, los demás trabajarán normalmente de manera independiente.

A lo largo del tiempo han ido evolucionando diferentes generaciones de SCSI, cada uno con grandes ventajas respecto al anterior.

La interfaz SCSI tiene en el mercado una muy alta compatibilidad y la gama de productos es muy amplia. Abarca no solamente dispositivos de almacenamiento de información, sino

lectoras de CD, impresoras y digitalizadores de imágenes, por lo que es difícil que salga del mercado en el corto plazo, de hecho constantemente diversos fabricantes tienen iniciativas de cómo mejorarlo utilizando medios de transporte de señal intermedios entre el cobre y la fibra óptica, que será estudiada más a detalle en su correspondiente apartado.

Todo lo anterior es debido a que SCSI no incluye instrucciones para la transmisión de información, sino que incluye comando de control para una amplia variedad de dispositivos tales como unidades de grabación de cinta magnética y robots automatizados de cinta magnética, también conocidos como librerías de cinta debido a que asemejan a estantes conteniendo los cartuchos de cinta magnética.

Generación SCSI	Tamaño del Bus (bits)	Velocidad de Reloj (MHz)	Salida (MB/s)	ANSI Documento
SCSI-1	8	5	5	SCSI: X3.131-1986
Fast SCSI	8	10	10	SCSI-2: X3T9.2
Fast Wide SCSI	16	10	20	SCSI-3 (SPI): X3.277
Ultra SCSI	8	20	20	SCSI-3 (SPI, Fast-20): X3.277-1966
Wide Ultra SCSI	16	20	40	SCSI-3 (SPI, Fast-20): X3.277-1966
Wide Ultra-2 SCSI	16	40	80	SCSI-3 (SPI-2): X3T10/1142D

Tabla 2.1.

Generaciones SCSI

La tabla anterior nos muestra la evolución que ha tenido la interfaz SCSI a lo largo del tiempo, duplicando el tamaño del bus y mejorando 16 veces la velocidad de salida. Cabe señalar que es difícil que continúe creciendo a este paso utilizando medios de cobre, ya que este material tiene características de atenuación y ancho de banda que no lo permiten.

Actualmente, se utiliza un cable SCSI universal que soporta señales LVD y HVD, la forma en que cada uno de los 68 pines está utilizado se representa en la tabla 2.2. Cabe mencionar que algunos periféricos por diseño del fabricante hacen modificaciones a este despliegue en los pines, generando una variedad de adaptadores específicos a cada producto que deberán de ser proporcionados por el mismo fabricante a fin de conservar la compatibilidad.

Señal	Pin	Pin	Señal
+DB(12)	1	35	-DB(12)
+DB(13)	2	36	-DB(13)
+DB(14)	3	37	-DB(14)
+DB(15)	4	38	-DB(15)
+DB(P1)	5	39	-DB(P1)
GROUND	6	40	GROUND
+DB(0)	7	41	-DB(0)
+DB(1)	8	42	-DB(1)
+DB(2)	9	43	-DB(2)
+DB(3)	10	44	-DB(3)
+DB(4)	11	45	-DB(4)
+DB(5)	12	46	-DB(5)
+DB(6)	13	47	-DB(6)
+DB(7)	14	48	-DB(7)
+DB(P)	15	49	-DB(P)
DIFSENS	16	50	GROUND
TERMPWR	17	51	TERMPWR
TERMPWR	18	52	TERMPWR
RESERVED	19	53	RESERVED
+ATN	20	54	-ATN
GROUND	21	55	GROUND
+BSY	22	56	-BSY
+ACK	23	57	-ACK
+RST	24	58	-RST
+MSG	25	59	-MSG
+SEL	26	60	-SEL
+C/D	27	61	-C/D
+REQ	28	62	-REQ
+I/O	29	63	-I/O
GROUND	30	64	GROUND
+DB(8)	31	65	-DB(8)
+DB(9)	32	66	-DB(9)
+DB(10)	33	67	-DB(10)
+DB(11)	34	68	-DB(11)

Tabla 2.2.

Valores cable universal SCSI 68 pines

2.1.2. Interfaz de canal de fibra (FC-AL)

FC-AL es una interfaz serial que utiliza típicamente como medio la fibra óptica, aunque también puede aplicarse sobre cobre con menor velocidad, con un ancho de banda enorme. En la actualidad podemos encontrar dispositivos de almacenamiento compatibles para enlaces de hasta 4Gbps¹, y se encuentra en desarrollo tecnología para 9Gps. Debemos recordar que cuando se liberó FC se pensaba que no existirían dispositivos tan rápidos, sin embargo, ahora las memorias de estado sólido lo son.

Trama

FC-AL transmite tramas de 2148 bytes, que se dividen de la siguiente manera:

4 bytes conforman el inicio de trama (Start of Frame, SOF), una instrucción que indica el principio del bloque.

24 bytes componen el encabezado (Header), que se encarga de contener los bytes que identifican a ese bloque, como el identificador de origen (SID Source ID), identificación de destino (DID Destination ID).

2112 corresponden a la carga (Payload), que es propiamente la información a transmitirse.

4 bytes corresponden a la verificación de redundancia cíclica (Cyclic Redundancy Check, CRC). Algoritmo de corrección de errores.

4 bytes para fin de bloque (End of Frame, EOF), que indica el fin del bloque.

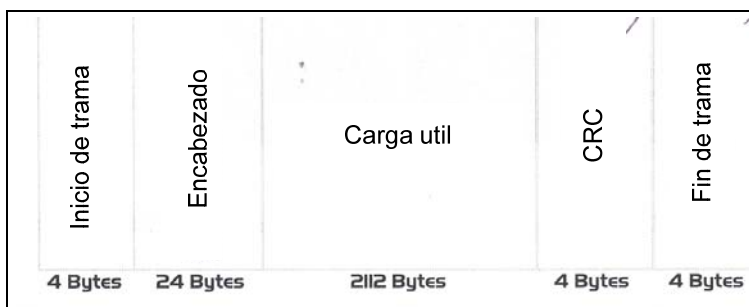


Figura 2.8.

¹ Gbps: gigabytes por segundo.

Trama FC-AL

Capas

Al igual que el modelo OSI, FC-AL se subdivide en capas que cumplen con una función específica. A continuación se detallarán cada una de ellas.

Capa 0 FC-AL: el nivel más bajo define la capa física, incluidas la fibra óptica, conectores y parámetros ópticos y eléctricos para una variedad de tasas de transferencias.

La capa física está diseñada para el uso de diversas tecnologías para encontrar el más amplio rango de requerimientos de sistema.

La capa 0 cuenta con un mecanismo de seguridad llamado control de fibra abierta (Open Fibre Control, OFC), debido a que la potencia de la luz excede los estándares de seguridad definidos para láser por las leyes norteamericanas. Si existe una condición de fibra desconectada y uno de los receptores no detecta enlace, el emisor de ese mismo dispositivo radia con menor potencia hasta que nuevamente se detecta conexión y ambos transmisores retransmiten con la potencia original.

Capa 1 FC-AL: aquí se realiza la codificación y decodificación serial, formación de caracteres especiales y control de errores. La información transmitida es codificada y a continuación pasa a un serializador y es transmitida por la fibra o el cable de cobre. La tasa de error de bit (Bit Error Rate, BER) debe ser menor a 10^{-12} .

Capa 2 FC-AL: las reglas de creación de tramas y los diferentes mecanismos para controlar las tres clases de servicios están definidos en esta capa.

Para auxiliar en el transporte de información a través del enlace, las siguientes tramas son definidas:

Conjunto ordenado: son palabras de 4 bytes que sirven para tener sincronía e indicar la condición del puerto receptor.

Trama: contienen la información a transmitir, direcciones de origen y destino e información de control.

Secuencia: está formada por una o más tramas relacionadas transmitidas de un puerto a otro. Cada trama dentro de la secuencia es numerada y sirve para detectar errores en la transmisión en las capas superiores.

Protocolo: protocolos para la transmisión, arriba de ellos pueden trabajar otros protocolos de mayor nivel. Pueden listarse protocolo de ingreso al fabric, protocolos de ingreso y salida de puerto_N².

Capa 3 FC-AL: Esta capa está dedicada a ofrecer servicios, los cuales pueden o no estar disponibles dependiendo la clase de servicio:

Cortar (Striping): Multiplica el ancho de banda al utilizar múltiples puertos_N en paralelo para transmitir (dependiendo la clase de servicio).

Grupos: La capacidad de más de un puerto para recibir una trama, esto disminuye la posibilidad de encontrar un puerto_N ocupado.

Multidifusión (Multicast): Transmisión a múltiples puertos en el fabric.

Capa 4 FC-AL: la capa 4, la de mayor nivel en la estructura y define la interfaces que pueden trabajar sobre FC. Especifica las reglas para que otros protocolos de alto nivel puedan ser “mapeados³” a las tramas de FC-AL y puedan utilizar a FC-AL como medio de transporte, algunos de ellos son:

- Interfaz de pequeños sistemas de cómputo (Small Computer System Interface, SCSI).
- Interfaz inteligente de periféricos (Intelligent Peripheral Interface, IPI).
- Protocolo de internet (Internet Protocol, IP)

² Puerto_N: punto de interconexión de un nodo fabric.

³ Se refiere al termino de poder agrupar o establecer una realización lógica con los diversos dispositivos que puedan ser reconocidos y operados como medio de transporte por las tramas FC-AL.

Capa FC	Nombre	Capa OSI	Nombre	Capa SCSI-3
4	Aplicación	7	Aplicación	Protocolos
		6	Presentación	Capa Alta
3	Servicios Comunes	5	Sesión	Protocolos
2	Creación de Tramas	4	Transporte	Capa Baja
		3	Red	Interconexión Física
1	Codificación	2	Enlace	
0	Variantes Físicas	1	Física	

Tabla 2.2.

Comparación entre las capas de los modelos FC-AL, OSI y SCSI-3

Clases de servicio en FC-AL

La forma en que se transmite la información depende de la clase de servicio en que se establece la conexión. Como podrá entenderse más adelante, las redes de almacenamiento se comunican en clase 1.

Clase 1: esta es una clase orientada a conexión, donde se transmiten tramas de confirmación de recepción o de no-recepción del destino al origen. Los puertos en esta conexión dedican el total del ancho de banda para la misma, sus características principales son:

1. Si el destino acepta la trama, se establece la conexión
2. Las siguientes tramas se transmiten en orden
3. Se reciben tramas de reconocimiento (acknowledgement, ACK) cuando una trama es recibida o cuando una trama esperada no llega
4. No se reciben tramas ocupado (busy, BSY))

Clase 1 es un servicio que provee conexiones dedicadas, proveyendo un equivalente a una conexión física dedicada. Una vez establecida una conexión clase 1, el enlace es preservado por el fabric. Esto garantiza el máximo ancho de banda entre dos puertos_N, por lo que es el más empleado en redes de almacenamiento.

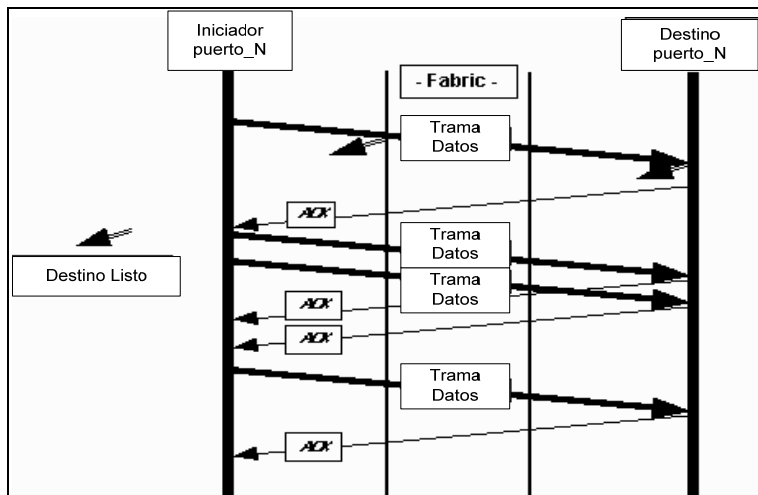


Figura 2.10.

Conexión clase 1

Fuente: <http://hsi.web.cern.ch/HSI/fcs/spec/overview.htm#b2>

Clase 2: clase no orientada a conexión entre 2 puertos con confirmación ACK de entrega o no-entrega, un puerto puede transmitir a diferentes puertos, entre sus principales características están:

1. Se asume que en receptor está listo
2. No se realiza conexión antes de la transmisión
3. Las tramas pueden no llegar en orden al destino

Clase 2 implica conmutación de tramas y permite la multicanalización de tramas de distintos orígenes en el mismo canal, sin embargo no se garantiza que sean entregadas por el fabric en orden y deberán de reensamblarse en el destino.

Ya sea clase 1 o 2 el receptor envía una trama de recibo o ACK confirmando la recepción de la trama. Si no se ha recibido la trama dentro de un tiempo específico se envía una señal de ocupado y el transmisor vuelve a intentarlo.

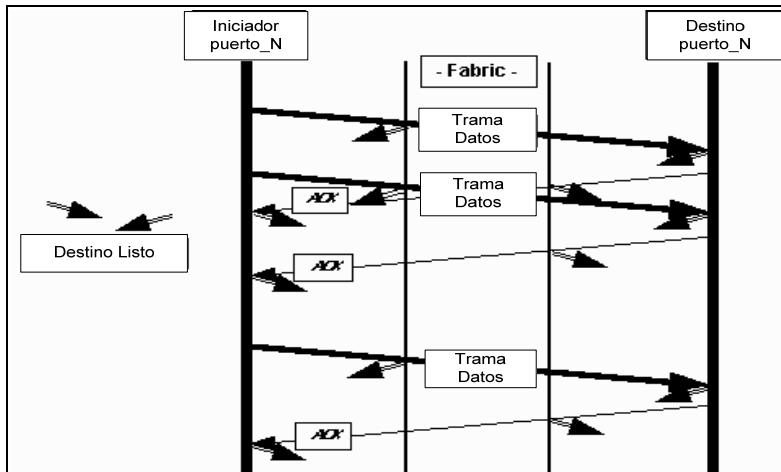


Figura 2.11

Conexión clase 2

Fuente: <http://hsi.web.cern.ch/HSI/fcs/spec/overview.htm#b2>

Clase 3: es la conexión más parecida a una red de área local, no existe notificación de entrega o no entrega. Sus principales características son:

1. La pérdida de tramas es manejada por protocolos superiores
2. Se asume que el receptor está listo
3. No se garantiza que ocurra una situación de ocupado

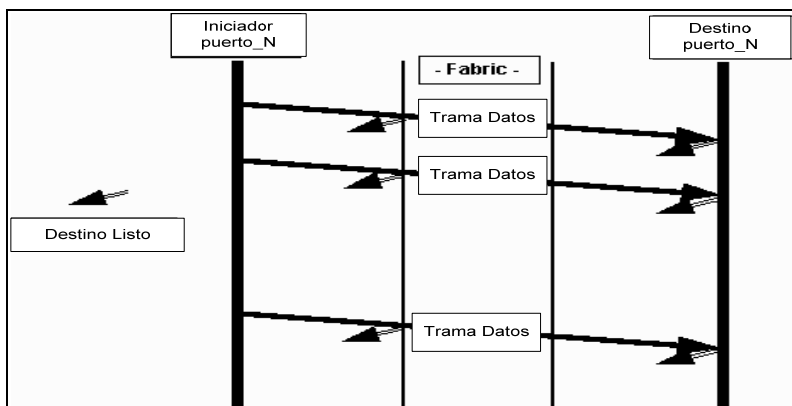


Figura 2.12

Conexión clase 3

Fuente: <http://hsi.web.cern.ch/HSI/fcs/spec/overview.htm#b2>

Suele ser común confundir canal de fibra con fibra óptica, sin embargo, el primero es un conjunto de estándares delimitados por la ANSI y el segundo es un medio físico por el que tramas FC-AL pueden transmitirse, y por mucho el más utilizado.

2.2. Tipos de redes

Se entiende por red una distribución de elementos de procesamiento de datos y periféricos, que pueden intercambiar información y recursos. Entre sus principales características se encuentran:

- Ambiente de conexiones no anticipadas
- Sobrecarga impredecible debido a colisiones de paquetes de datos y un medio compartido de transmisión
- Uso intensivo de software
- Cada nodo debe ser capaz de detectar condiciones erróneas en la red e intentar protegerse
- Las redes pueden interconectarse por medios físicos como cable y fibra óptica, además de radiofrecuencia.

Las redes típicamente han sido clasificadas por su cobertura como:

Red de área local (Local Area Network, LAN): es una red de comunicaciones utilizada por una sola organización en un espacio generalmente restringido a un inmueble. Pueden estar interconectadas por cable, fibra óptica e inalámbrica, y comunicarse mediante protocolos o “lenguajes”, en el que el más difundido es TCP/IP⁴.

Red de área metropolitana (Metropolitan Area Network, MAN): interconectan a varias LAN de distintas organizaciones y cubren el área geográfica de una ciudad o un pequeño

⁴ Protocolo de control de transmisión – Protocolo de Internet (Transmission Control Protocol - Internet Protocol, TCP-IP).

grupo de ellas. Debido a la cantidad de información y distancia entre equipos los medios de interconexión preferidos son la fibra óptica y las microondas.

Red de área amplia (Wide Area Network, WAN): interconectan regiones geográficas, países o continentes. Debido a su gran alcance es posible hablar del uso de satélites para su conformación, fibra óptica y microondas.

Conforme la industria del almacenamiento se consolida, un tipo de red se integra a las definiciones típicas citadas anteriormente. Este tipo de redes son parte central de nuestro trabajo y por ello se estudiarán a detalle, las redes de área de almacenamiento.

2.3. Red de área de almacenamiento (Storage Area Network, SAN)

Es una red dedicada a respaldar y recuperar información de gran tamaño en el menor tiempo posible. Es capaz de conmutar paquetes, sin embargo, una vez establecida la ruta más rápida entre dos equipos no la altera a menos que se presente la caída de un enlace o modificación de la red, esto de forma automática, y sirve para evitar colisiones que degraden la velocidad. El núcleo de la SAN es el módulo tejido (fabric)⁵, que es un switch o un grupo interconectado de ellos a donde se conectan los dispositivos de almacenamiento, también llamados “nodos”.

Cada switch tiene un identificador de dominio (Domain ID, DID), que lo identifica dentro del fabric, se discutirá mas ampliamente al módulo fabric en el apartado correspondiente.

Si bien pueden interconectar hasta 15 millones de dispositivos, debido a que las redes de almacenamiento son de un costo comparativo mucho mayor al de las redes de área local, difícilmente llegan a ser muy grandes, ya que mientras una tarjeta de red ethernet cuesta alrededor de \$15-300 USD dependiendo de el fabricante, una HBA para almacenamiento (que hace las veces de tarjeta de red) oscila alrededor de los \$1,000 USD. Esto se explica debido a la gran cantidad de procesamiento interno que realizan para no disminuir los

⁵ Tejido (fabric): El término hace alusión a una serie de hilos que se interconectan. La traducción al español rara vez es empleada.

recursos del servidor y así poder alcanzar velocidades de transferencia de hasta 4Gbps en los dispositivos comerciales más modernos, y respaldar cantidades información que pueden alcanzar varios TB diarios.

Las redes de almacenamiento comunican servidores que generan una alta cantidad de información, los usuarios comunes no están conectados directamente a ella, sin embargo mediante la red de área local pueden enviar su información por respaldar a un servidor y este puede mandar ese tráfico junto con el suyo.

Otra particularidad fundamental de las redes de almacenamiento es que no se desea que todos los dispositivos se comuniquen entre sí. De hecho los servidores se aíslan uno del otro y no hay comunicación entre ellos, solamente de un servidor a uno o varios dispositivos de almacenamiento, los dispositivos de almacenamiento pueden estar compartidos entre varios servidores, pero no hay tráfico entre los servidores a través de la SAN.

Es importante mencionar que un switch para redes de almacenamiento, es muy diferente a uno de comunicaciones, ya que en el switch de almacenamiento sólo buscamos crear enlaces lógicos estáticos entre nodos, es decir, buscamos que sólo ciertos dispositivos se comuniquen entre sí en cualquier momento. Físicamente el módulo fabric buscará la mejor ruta y ésta prevalecerá a menos que se interconectan nuevos nodos o un servidor altere su carga de trabajo. Los enlaces lógicos estáticos para agrupar dispositivos son llamados "zonas" y existen 2 formas diferentes para crearlas:

Nombre a lo amplio del mundo (World Wide Name, WWN)⁶: Cada puerto y cada nodo de FC tiene una cadena que lo identifica y es único dentro de la red de almacenamiento, por ejemplo, un nodo de tiene 2 WWN; una para el puerto A y una de nodo.

En el caso de nodos de puerto FC dual tenemos 3 WWN, una para el puerto A, otra para el puerto B y otra de nodo. En este caso sólo el puerto A esta activo y el puerto B sirve como redundancia, en caso de algunos software de respaldo puede darse de alto la WWN del nodo y éste elegir en qué puerto trabajar.

⁶ Aunque la traducción literal al español es "Nombre a lo Amplio del Mundo" también difícilmente se emplea en el campo de trabajo su traducción al español.

Cada WWN está compuesta de 16 cifras hexadecimales, por lo que tenemos 16^{16} combinaciones posibles. Cada dispositivo tiene un conjunto de WWN's único, aunque éstas pueden ser configurables para facilitar el reemplazo de un dispositivo dañado.

Al momento de crear la zona se especifica que WWN's tienen conectividad, e independientemente del puerto del switch al que estén conectados se establecerá el enlace, de modo que es el preferido cuando se contempla un escenario donde constantemente se modifica la distribución de los equipos, además de que facilita la planeación de la implementación, y se toma como estándar que la WWN del puerto A es la del nodo más uno, y la del puerto B la del nodo más dos, por ejemplo:

- WWNN (WWN de Nodo): 50:01:04:F0:0E:AD:C3:D9
 - WWNA (WWN puerto A): 50:01:04:F0:0E:AD:C3:DA
 - WWNB (WWN puerto B): 50:01:04:F0:0E:AD:C3:DB

Identificador de Puerto (Port ID): la otra forma de crear zonas es mediante puertos del fabric, es decir, se configura que un puerto dado de un switch tenga conectividad con otro puerto de ese mismo u otro switch. Es generalmente utilizado en implementaciones pequeñas o que no se planeen modificaciones en un largo tiempo, ya que al agregar un nuevo switch puede cambiar su identificador de dominio (Domain ID, DID) debido a que el switch que tiene la WWNN (WWN de nodo) más baja toma el DID 1, que es el principal, y altera las zonas antes creadas. Su nomenclatura es la siguiente;1,7: Switch DID 1, puerto 7

2.4. Tipos de conexión de dispositivos de almacenamiento

En un principio, los dispositivos de almacenamiento se conectaban directamente a los servidores siendo utilizados solamente por el equipo servidor, siendo conocido este tipo de conexión como almacenamiento directamente conectado (Direct Attached Storage, DAS). Poco tiempo después fue necesario, como sucedió con el resto de los equipos de cómputo, compartir y distribuir los recursos entre varios equipos interconectados en una red, para lo cual surgieron dispositivos de almacenamiento conectado a red (Network Attached Storage, NAS) que se conectaban directamente a las redes de área local como si se tratara de

⁷ 16^{16} elevado a la potencia 16, es decir, 18446744073709551616 WWN's únicas posibles.

impresoras o faxes, y pueden ser utilizados alternativamente por los equipos que se encuentran en la red.

Sin embargo, pronto surgió la necesidad de tener un mejor control sobre los dispositivos de almacenamiento y que estos tuvieran un mejor desempeño, debido a la gran cantidad de tiempo y recursos que necesitan los respaldos de información y para evitar mejorar su administración, surge la red de área de almacenamiento (Storage Area Network, SAN), estudiada más ampliamente en el apartado anterior, y que interconecta sus dispositivos o nodos mediante un módulo fabric, es decir, que los nodos están interconectados en modo fabric.

A continuación se estudiarán más ampliamente estos tipos de conexión.

Almacenamiento directamente conectado (Direct Attached Storage, DAS): los dispositivos interconectados en DAS transmiten la información directamente del servidor donde se genera hasta el dispositivo de almacenamiento dedicado. En este caso un único servidor tiene la tarea de enviar y controlar el tráfico, y en caso de estar conectado a una red de área local es el encargado de llevar la información que le llega de sus tarjetas de red a la tarjeta compatible con el dispositivo de almacenamiento. Puede o no tener un software para respaldos, y en caso de no tenerlo generará los respaldos mediante las herramientas del sistema operativo, como el Backup Manager de Windows® o la utilería “tar⁸” en los sistemas UNIX®, otros equipos podrán respaldar información transfiriendo sus archivos al servidor mediante los medios que su red soporte. En caso de tener software de respaldo se realizará lo anterior de modo automático instalando un software cliente en cada equipo. En caso de baja del servidor maestro ningún cliente podrá respaldarse.

⁸ Tar: se refiere en Informática a un formato de archivos ampliamente usado en entornos UNIX, identificados con la extensión **tar**. Además hace referencia al programa para la manipulación de archivos que es estándar en estos entornos. El formato fue diseñado para almacenar archivos de una forma conveniente en cintas magnéticas y de allí proviene su nombre, que proviene de arcvhivador en cinta (Tape **AR**chiver). Debido a este origen el formato está preparado para ser procesado linealmente, no contando con manera de extraer un miembro sin recorrer todo el archivo hasta encontrarlo.

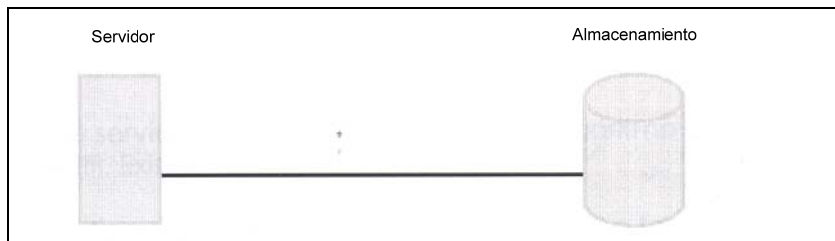


Figura 2.1.

Almacenamiento directamente conectado

Dentro de esta categoría podemos integrar al anillo arbitrado (Arbitrated Loop), el cual consiste en tener a varios dispositivos conectados entre sí y a su vez al servidor. Cada dispositivo debe de contar con un identificador de anillo (Loop ID) único y mediante el servidor recibirá o transmitirá información. La forma en que transmiten es mediante una trama de control, la cual va pasando de dispositivo a dispositivo, desde el emisor de una al receptor de otra, y de tal forma con las subsecuentes. Esto se hace mediante un concentrador (hub), de los cuales existen para fibra óptica. En el caso de SCSI por poder manejar varios nodos en un solo bus esto no es necesario.

Pueden conectarse en anillo arbitrado un máximo de 126 dispositivos, pero debido a los turnos que toman los dispositivos para comunicarse es muy difícil en la práctica que se conecten más de 5 dispositivos de esta forma, ya que empieza a degradarse la velocidad de los equipos. Debido a que van conectados al servidor y sólo tienen conectividad entre ellos, se dice que forma un "anillo privado", es decir, que un dispositivo fuera de este anillo no podrá acceder a él a menos que lo haga lógicamente a través del servidor.

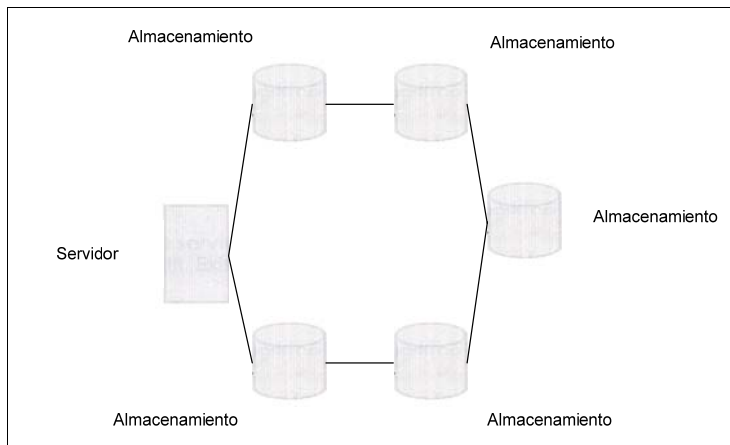


Figura 2.2.

Anillo privado

Almacenamiento conectado a red (Network Attached Storage, NAS): los dispositivos interconectados en NAS se encuentran integrados a la red de área local y pueden ser utilizados por cualquier equipo en la red, la baja del servidor maestro no implica detener la operación ya que existen diferentes rutas para llegar al dispositivo requerido según la red de área local lo permita, sin embargo ésta se realizará de forma desordenada, y tendrá prioridad el equipo que primero empiece a respaldar. Puede existir software de respaldo en uno o más servidores dejando a uno como maestro y a otros como redundancia en caso de falla, pero nada impide que un equipo se apodere del dispositivo de respaldo mediante herramientas del sistema operativo. Su costo no es mayor que el del dispositivo de almacenamiento que se desee utilizar, pero su mayor desventaja es que degrada la velocidad de la red significativamente y comúnmente no concluye adecuadamente los respaldos cuando el tamaño de los archivos ronda un GB.

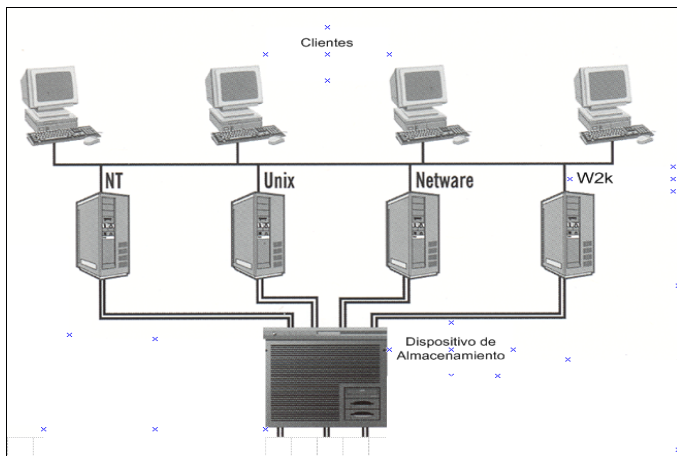


Figura 2.3.

Almacenamiento Conectado a Red

Fabric⁹: Es cuando una serie de dispositivos se interconectan mediante uno o más switches de almacenamiento. Si se tiene más de un switch puede conectarse uno con otro mediante una fibra óptica tendida entre dos puertos libres, denominado a este enlace inter-switch (Inter-Switch Link, ISL), a mayor número de ISL's mayor ancho de banda disponible para el tráfico entre switches. Existen algunas reglas para este tipo de enlaces que serán explicados en el apartado para switch. El switch con la WWNN (WWN de Nodo) más baja tendrá el identificador de dominio (Domain ID, DID) 1 y será el principal, sus funciones son:

Servidor de nombres de dominio (Domain Name Server, DNS): todo nodo puede conectarse una vez que se ha dado de alta ante el switch principal, y éste registrará en una tabla su WWN, si es único o parte de un loop, y a qué zona pertenece.

Administración: es el switch al que deberá conectarse para administrar al fabric, es decir, crear, eliminar y modificar zonas, fijar velocidad de puertos, habilitar servicios como Extended Fabric¹⁰, cargar microcódigo (Firmware)¹¹.

⁹ Literalmente tela o tejido en español, haciendo alusión a los hilos de una tela que se entrecruza.

¹⁰ Extended Fabric: conectar un switch hasta 10 Km de distancia; se deberá contar con el hardware adecuado.

Los servidores se conectan mediante una HBA de fibra óptica instalada en la tarjeta madre, de ahí una fibra óptica va hacia un puerto libre de un switch. Como anteriormente se ha mencionado, el número máximo de dispositivos que pueden ser conectados en un fabric es 15 millones, pero debido a su alto costo es difícilmente se aproxima.

Es importante mencionar que el módulo fabric funciona como una “caja negra”, ya que la forma en que se interconectan los switches es independiente de los servidores y no son detectados por el sistema operativo.

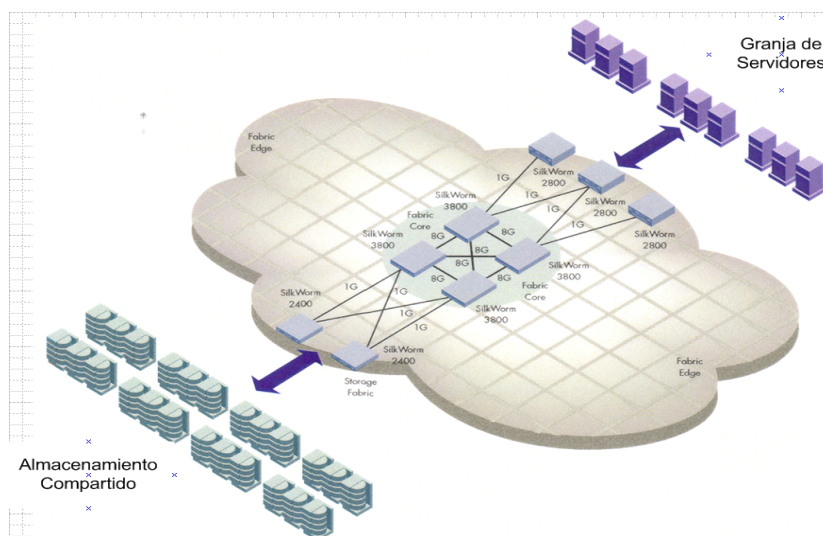


Figura 2.4.

Módulo fabric como núcleo de red de área de almacenamiento

Fuente: Global Learning Solutions, Documents on CD, 26ª.ed., Louisville U.S.A., Ed. Storagetek Literature Distribution, 2005, p 133.

Diversos dispositivos pueden ser conectados en anillo a uno de los puertos de un switch, en tal caso se formará un anillo arbitrado público, lo que hace posible que otro dispositivo dentro del fabric acceda a uno de esos dispositivos aunque no se encuentre dentro de ese anillo. Internamente el anillo se comportará como uno arbitrado privado.

¹¹ Microcódigo (Firmware): es sistema operativo interno del switch para administra sus recursos.

2.5. Componentes de una red de área de almacenamiento

Fibra óptica

La fibra óptica es un conducto transparente, generalmente elaborado de silicio o polímeros, recubierto por otro material transparente con un índice de refracción mayor, ocasionando que los rayos de luz regresen hacia el núcleo.

La fibra óptica se compone de la siguiente forma:

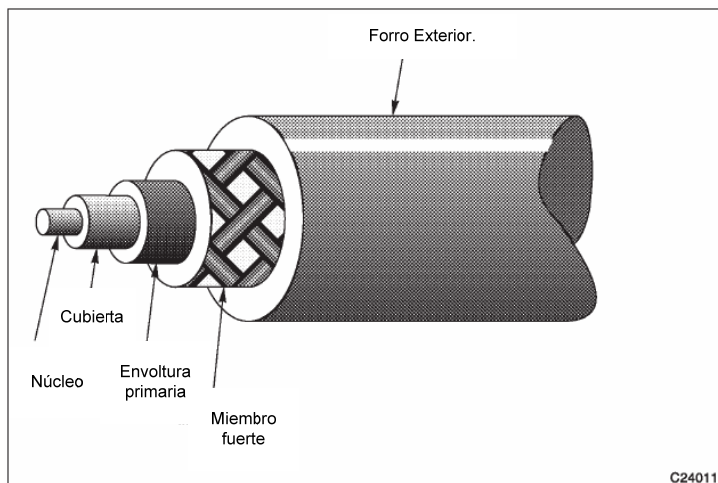


Figura 2.13.

Vista interna de una fibra óptica

Cubierta: opera como un espejo cilíndrico rodeando el núcleo. Debido al alto índice de refracción del núcleo y el bajo de la cubierta, la luz que viaja a través del núcleo es prevenida de escapar al rebotar en la cubierta.

Envoltura primaria: Sirve como protección para el núcleo de la fibra y la cubierta, ayuda a prevenir que la fibra sea demasiado frágil para su manejo e interacción con el ambiente.

Miembro fuerte: Provee fuerza a el cable y puede estar constituido por fibras de hierro o kevlar®.

Cubierta exterior: capa plástica protectora de color generalmente naranja encargada de servir de cubierta a las demás capas.

La información transferida a través de una fibra óptica inicia cuando el transmisor acepta información codificada digitalmente y la convierte en impulsos de luz. El receptor detecta el impulso, lo convierte en un impulso eléctrico y lo amplifica para recuperar la señal original.

Dependiendo de la longitud de onda la luz puede sufrir serias pérdidas, y las longitudes de onda donde ocurre menos son 820nm 1300nm y 1500nm. La transmisión se realiza en la fibra de dos diferentes formas, ya sea varios rayos de luz a la vez (multimodo) o uno solo (monomodo). A pesar de lo que pudiera inferirse en un primer término, la mayor velocidad de información se transmite por la fibra monomodo, esto debido a que a través de la distancia la frecuencia de la luz sufre una atenuación y el pulso se “ensancha”. A cierta distancia los pulsos de los diferentes rayos tienden a traslaparse y se tiene que aumentar la distancia entre ellos.

	Monomodo	Multimodo
Fuente de luz	Laser LED o LASER	Laser LED o LASER
Acoplamiento de la fuente	Difícil debido a las dimensiones del núcleo	Más fácil que el caso anterior
Potencia comunicada de la fibra	Muy poca	Mayor potencia transmitida
Ancho de banda	Limitado por la dispersión cromática, pero muy elevado	Limitado por la dispersión intermodal y mucho menos que la anterior
Dimensiones	Radio entre 5 y 10 micras	Radio de varias decenas de micras
Tipos según su estructura	De salto de índice	De salto de índice y de índice gradual

Las redes de área de almacenamiento se forman por interconexiones de fibra óptica en su totalidad actualmente, aunque debido al costo de los equipos de fibra nativos se han buscado nuevas opciones a través de cable de cobre.

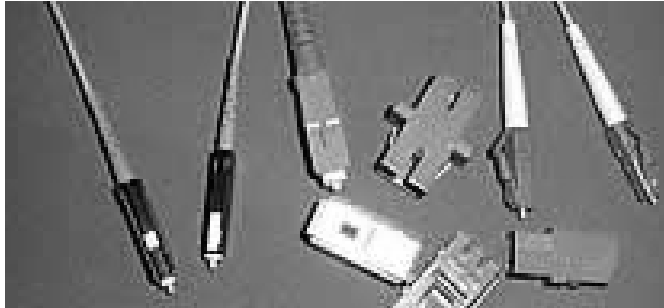


Figura 2.16.

Vista externa de una fibra óptica

Fuente: <http://www.fibra-optica.org/productos-fibra-optica/fibra-optica-estandar/productos-estandar.asp>

Dentro de la SAN pueden existir dispositivos SCSI, mismos que podrán ser interconectados a través de un convertidor denominado ruteador SCSI (SCSI router), denominado en la industria de esta manera por llevar la información convertida a SCSI a cada blanco (target) a que la información va destinado.

Adaptadora al bus del servidor (Host Bus Adapter, HBA)

Es una tarjeta electrónica encargada de convertir los impulsos eléctricos provenientes del procesador a impulsos que puedan transmitirse por el cable SCSI o la fibra óptica. La selección de la HBA depende de varios factores, los más importantes son:

Velocidad: la HBA deberá tener una velocidad coherente con la ventana de respaldo del dispositivo de almacenamiento, es decir, que se deberá usar una HBA con la velocidad suficiente para alimentar a los dispositivos de almacenamiento de forma que no se desperdicien los recursos de alta velocidad cuando la HBA envíe la información.

Compatibilidad: se debe verificar que la HBA tenga soporte, controladores actualizados para el servidor donde desea instalarse. Asimismo debe verificarse que el servidor donde desea instalarse tenga los recursos de memoria y procesador suficientes de modo que la HBA no sature el procesador y se pase o degrade el comportamiento del servidor.

Existen diferentes fabricantes de HBA, entre los que destacan Emulex®, Qlogic® y JNI® para la implementación de SAN. El costo de la HBA es considerablemente alto, de tal forma que al realizar el diseño de la SAN se deberá ser cauteloso en determinar si un servidor tiene la suficiente cantidad de información a respaldar o si conviene respaldarlos por Ethernet o SCSI LVD.



Figura 2.14.

HBA FC-AL

Fuente: <http://www.emulex.com/products/fc/9002l/ds.html>

Interruptor (Switch)

Es el componente que hace posible la interconexión de todos los dispositivos de la red de almacenamiento. Existen diferentes modelos y tamaños, desde los de nivel de entrada, de 8 puertos, hasta los directores, de 64 puertos o más. En el se interconectan lógicamente los dispositivos mediante la definición de “zonas”. La creación de zonas puede hacerse mediante dos formas:

Identificación de puerto (ID Port): En este método de hacer “zonas” se define en la configuración del switch qué puerto tiene comunicación con qué otro independientemente del dispositivo que a el se tiene conectado. Es utilizada generalmente para soluciones pequeñas, en donde unos cuantos dispositivos son conectados y estos son del mismo tipo, ya que facilita el reemplazo de un componente dañado con sólo desconectarlo y conectar el nuevo, sin embargo, puede llevar a conectar un dispositivo diferente al que originalmente se define, lo que puede producir efectos inesperados.

Nombre a lo largo del mundo (WWN Worldwide): consiste en un número de 16 cifras configurable en la mayoría de dispositivos, que debe ser único en toda la SAN. Tiene la ventaja de que el dispositivo puede conectarse en cualquier puerto y el switch automáticamente enviará el tráfico de ese dispositivo hacia aquellos que estén definidos en su zona; una ventaja adicional es que las tarjetas de fibra HBA generalmente se pueden enlazar de manera dedicada a un dispositivo mediante la WWN, de tal forma que al reiniciar el servidor se evita que existan cambios en la forma en que el sistema operativo detecta los dispositivos. La desventaja es que hay que cambiar la configuración del nuevo dispositivo en caso de que el original resulte dañado

Un switch al iniciar de un encendido sigue una secuencia lógica de pasos para iniciar sus puertos, esto es debido a que existen dos maneras diferentes de conexión de dispositivos, como loop y como fabric.

Los switches pueden conectarse entre sí formando un “fabric”, las líneas de interconexión entre ellos son denominadas enlace entre switches (Inter Switch Link, ISL) y sirven para incrementar el número de puertos disponibles en el fabric y el ancho de banda entre swiches, de tal forma que con las ISL suficientes se puede lograr que un dispositivo tenga una reducción en su desempeño apenas perceptible al pasar por un fabric de varios switches.

Puede conectarse a un puerto del switch un anillo de dispositivos, de esta forma cualquier dispositivo en el fabric podrá comunicarse con otro que esté en el anillo, formando uno público, que puede soportar hasta 255 dispositivos, pero todos ellos compartirán el ancho de banda del puerto del switch al que están conectados.

Los emisores de los puertos son removibles, en caso de que el emisor de luz se dañe, puede reemplazarse sin necesidad de reiniciar el switch y al hacerlo no cambia su WWN dicho puerto.



Figura 2.15.

Interruptor (Switch)

Fuente: <http://www.brocade.com>

Ruteador SCSI (SCSI router)

Es una alternativa cuando el costo de adquirir un nuevo switch (conmutador) es prohibitivo y se requiere de una velocidad de respaldo dentro del rango SCSI.

Los dispositivos son vistos por el fabric como si se tratara de un anillo público, con la diferencia de que este dispositivo separa y envía la información correspondiente a cada dispositivo, evitando en gran parte la disminución de la velocidad al interconectar un dispositivo tras del otro. Si bien la razón de que se le llame ruteador obedece a fines de mercadotecnia más que a técnicos, se le denomina de esta forma ya que dirige la información al dispositivo en cuestión directamente. Cuenta con una memoria volátil para incrementar la velocidad de transferencia de datos, ya que si un dato en específico es requerido de inmediato es enviado sin pasar por el dispositivo de almacenamiento.

Cuenta con un puerto FC que se conecta al switch conmutador, y 4 puertos SCSI (pueden variar según fabricante), para conectar los dispositivos SCSI, que pueden ser LVD o HVD.

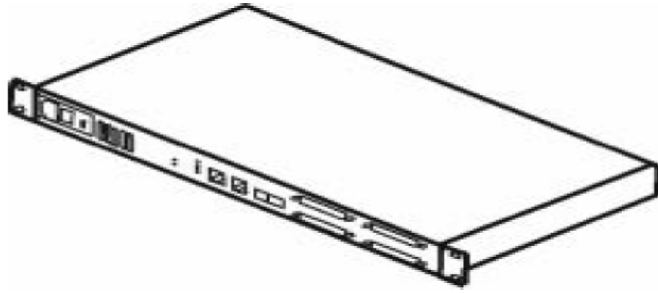


Figura 2.17

Ruteador SCSI.

Fuente: <http://www.support.storagetek.com/GlobalNavigation>

Servidores.

Tomamos como servidor al equipo de cómputo encargado de prestar una función a la red y no sólo al usuario que con ella trabaje. Aunque es muy difundido el uso de sistema operativo Windows®, en aplicaciones cada vez más grandes este sistema, que si bien es el más amigable, presenta serias desventajas en estabilidad, administración de recursos y velocidad de procesamiento. Una mejora en estos aspectos los presentan los servidores UNIX® en sus diferentes variedades, tales como Solaris® de Sun Microsystems®, HP-UX de Hewlett-Packard®. El sistema que mejor cumple con los requisitos antes mencionados es el mainframe¹² Z OS ® de IBM®, sin embargo sólo puede trabajar con hardware del propio fabricante y su costo es muy elevado. Un equipo de alta capacidad como la serie T-REX ® puede superar los 5 millones de dólares, y es empleado casi exclusivamente en instituciones bancarias.

Equipos de cómputo que generan la información a respaldar

Si bien una computadora personal puede hacer las veces de servidor, este debe de generar la suficiente cantidad de información para que sea justificable instalar una HBA en ellos. Generalmente están compuestos de varias celdas, que son tarjetas que albergan un grupo

¹² Mainframe: sistema operativo en el que todo el procesamiento se realiza en una computadora central, y los equipos conectados a él sólo realizan funciones de entrada-salida.

de procesadores, y que cada uno está destinado a proporcionar servicios de aplicaciones, bases de datos, directorio u otros (de ahí el nombre de “servidor”). A nivel de sistemas operativos la competencia es pareja, a diferencia de las computadoras personales que cuentan mayormente con sistema operativo Windows® ®®, y podemos encontrar:

HP-UX de Hewlett-Packard ®

Solaris® ® de Sun Microsystems ®

AIX de IBM® ®

Windows® ® de Microsoft ®

Z OS de IBM® ®

Software de respaldo

Aunque los sistemas operativos más difundidos cuentan con herramientas para realizar respaldos, no tienen la capacidad de realizarlos de forma automática, calendarizada, y prioritaria de acuerdo a los equipos.

Es por ello que existen diversas compañías dedicadas a programar software para cumplir con todas estas funciones, a continuación mencionaremos los más difundidos:

- BackupExec® y Netbackup® de Veritas Symantec®
- Networker® de Legato-EMC®
- Data Protector® de Hewlett-Packard®
- Galaxy® de Commvault®
- ArcServe® de CA®
- Tivoli Storage Manager® de IBM®

Aunque con sus diferencias y matices, en cada uno se definen los equipos clientes que se pueden respaldar a través de la red, los servidores que se respaldan a través de la SAN, los archivos a respaldar y los dispositivos automáticos para llevarlo a cabo. Asimismo elaboran reportes de los respaldos realizados y bitácoras de errores en caso de falla. Para ello uno de los servidores sirve como maestro y otros pueden tomar su lugar en caso de falla. Los clientes que pueden respaldarse muchas veces requieren la instalación de un software de

cliente, aunque en algunas otras basta configurar en el servidor maestro la ruta de los archivos a respaldar y contraseñas para acceder a los equipos.

Existen diferentes opciones de compra que comúnmente varían en función del número de equipos a respaldar.

CAPÍTULO 3. Medios magnéticos de almacenamiento

3.1. Disco magnético

Comúnmente conocido como disco duro, está compuesto de placas rígidas de aleación de aluminio o una mezcla de vidrio y cerámica recubiertos con una capa magnética. Hay disponibles unidades de disco duro con dos tamaños principales de diámetro, 5.25 y 3.5 pulgadas, aunque también existen de 2.5 y 17.5 pulgadas. Las unidades de disco duro se sellan herméticamente para mantener al disco libre de polvo.

Normalmente se apilan dos o más discos sobre un eje o pivote común, que hace que el conjunto gire a una velocidad de miles de revoluciones por minuto. Existe una separación entre cada disco, con el fin de permitir el montaje de un cabezal de lectura-escritura el extremo del brazo de acción.

Hay un cabezal de lectura-escritura en cada cara del disco, resultando en datos grabados en cada cara. El brazo de acción de la unidad sincroniza todos los cabezales de lectura-escritura para mantenerlos perfectamente alineados cuando se desplazan por la superficie del disco, con una separación de sólo una fracción de milímetro respecto al disco. Una pequeña partícula de polvo podría hacer que un cabezal se rompiera, dañando por consecuencia la superficie del disco.

Un disco duro está organizado en pistas y sectores. Cada pista está dividida en una serie de sectores, y cada pista y sector tiene una dirección física que el sistema operativo utiliza para localizar un determinado registro de datos. Normalmente, los discos duros tienen desde unos pocos cientos hasta unos pocos miles de pistas, donde hay un número constante de pistas por sector, utilizando los sectores externos una superficie mayor que los sectores internos. La disposición de pistas y sectores en un disco se llama formato.

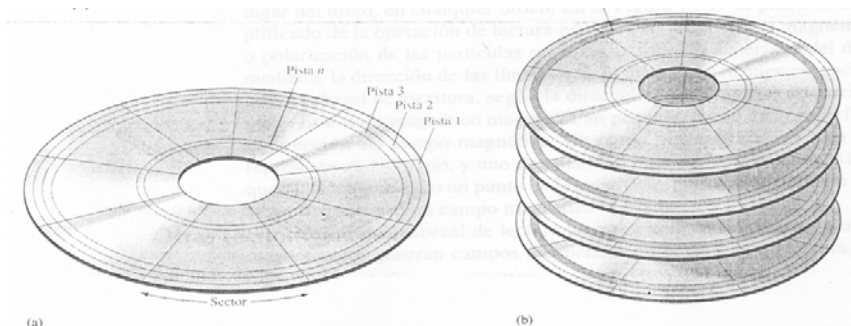


Figura 3.1.

Cilindros, pistas y sectores

Fuente: Floyd. Thomas, Fundamentos de Sistemas Digitales, 7^a.ed., Madrid España, Edit. Pearson Education, 2000, p 789.

3.2. Principios básicos de funcionamiento

Toda la información almacenada en un disco magnético esta grabada en pistas, las cuales son círculos concéntricos ubicados en la superficie de cada plato, muy parecidos a los círculos anuales de los árboles. Las pistas son numeradas, empezando de cero desde el exterior e incrementándose hacia el centro. Un disco moderno tiene decenas de miles de pistas en cada plato.

Se accede a la información al mover las cabezas de lectura-escritura del interior al exterior de los platos, movilizadas por el brazo de acción de las cabezas. Esta organización de la información aloja los datos de manera aleatoria sobre la superficie de los platos, llevando el registro de las ubicaciones físicas en donde la información es escrita en una tabla que reside en el disco y que se genera al formatear el disco, por ejemplo, en el sistema operativo Windows® la tabla de asignación de archivos (File Allocation Table, FAT, una vez que se desea recuperar la información se recurre a dicha tabla para conocer la ubicación de la información y las cabezas se posicionan a dicho punto. Las cabezas se mueven al mismo tiempo accionadas por el brazo de acción, mientras que los platos giran a una velocidad constante hasta llegar al punto en que se localiza la información a recuperar o respaldar.

Cada pista puede contener muchos miles de bytes de datos, por ello sería un desperdicio hacer a la pista la unidad más pequeña de almacenamiento, ya que archivos muy pequeños ocuparían un menor espacio que una pista, por lo tanto cada pista está partida en partes más pequeñas llamadas sectores. Cada sector almacena 512 bytes, además de algunos cuantos bytes extras utilizados para control interno del dispositivo y para detección o corrección de errores.

Una unidad de disco duro está hecha de múltiples platos, cada uno de los cuales utiliza dos cabezas para grabar y recuperar información, una en cada superficie del plato. Las cabezas se mueven en bloque de adentro hacia fuera, de tal forma que en un momento dado todas las cabezas se encuentran en el mismo número de pista en cada uno de los platos que componen a la unidad.

Debido a este arreglo, a menudo la localización de las cabezas no es referida a un número de pista sino a un número de cilindro. Un cilindro es, por lo tanto, el conjunto de todas las pistas en que las cabezas están localizadas, sin embargo, es igualmente correcto decir que una unidad de disco tiene cierto número de pistas o cilindros. Como puede fácilmente deducirse, al trazar imaginariamente la unión de las pistas del mismo tamaño a lo largo del eje, se pueden obtener cilindros imaginarios.

Un imán tiene dos polos, norte y sur, y energía magnética, llamada flujo, que va del polo norte al sur. La información es codificada en una serie de muy pequeños campos magnéticos que según se presentan norte-sur o sur-norte.

El disco duro es un dispositivo de acceso aleatorio, ya que los datos se encuentran dispersos en diferentes pistas del disco. La dirección o polarización de las partículas magnéticas sobre la superficie del disco se controla mediante la dirección de las líneas de flujo magnético producidas por el cabezal de escritura, según la dirección de un impulso de corriente en el devanado. Este flujo magnetiza un pequeño punto de la superficie del disco en la dirección del campo magnético.

Cuando un cabezal de lectura pasa por una superficie magnetizada, los puntos magnetizados generan campos magnéticos en el cabezal de lectura, lo que provoca impulsos de tensión en el devanado. La polaridad de estos impulsos depende de la dirección del punto magnetizado e indica si el bit almacenado es un 1 o un 0, a menudo los cabezales

de lectura y escritura se combinan en un solo ensamble pero se encuentran magnéticamente aislados.

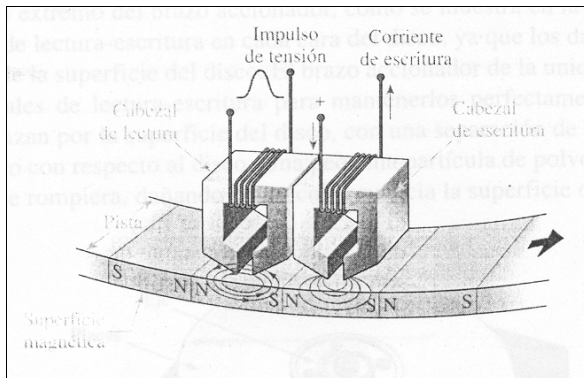


Figura 3.2.

Grabación en disco mediante campos magnéticos

Fuente: Fuente: Floyd. Thomas, Fundamentos de Sistemas Digitales, 7^a.ed., Madrid España, Edit. Pearson Education, 2000, p 789.

Llevar la cuenta de la ubicación de los archivos es el trabajo del controlador de dispositivos que gestiona el disco duro, algunas veces denominado BIOS de la unidad de disco duro. El controlador de dispositivo y el sistema operativo acceden a dos tablas que se utilizan para controlar la ubicación y los nombres de los archivos. La primera tabla se denomina FAT¹, que muestra qué partes del disco se han asignado a archivos específicos. La segunda tabla es el directorio raíz, que contiene el nombre de los archivos, tipo, fecha y hora de creación, cluster inicial y demás información relevante.

Los platos magnéticos se colocan sobre un eje de rotación mediante el cual giran uniformemente a una velocidad constante, en el extremo del brazo de acción se encuentra la cabeza de lectura-escritura. Dado que hay varios platos generalmente existe igual número de brazos, aunque dependiendo del fabricante se han construido brazos con doble cabeza, platos de doble superficie, cabezas independientes para lectura y para escritura, etc.

¹ File Allocation Table: tabla de asignación de archivos.

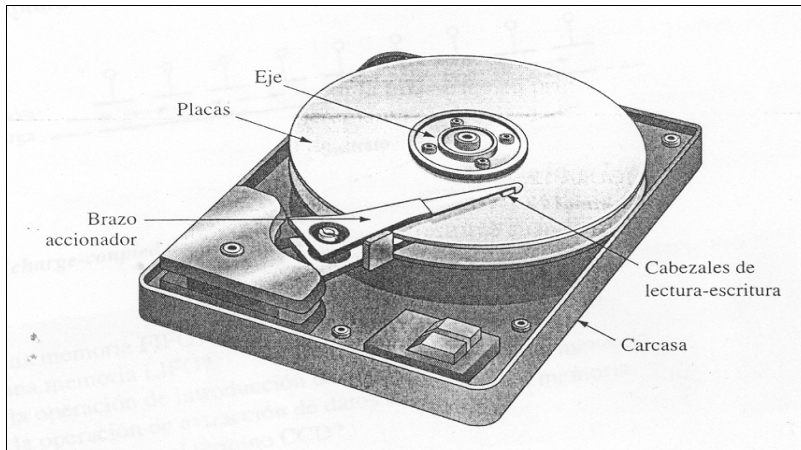


Figura 3.3.

Vista interna de un disco duro

Fuente: Floyd. Thomas, Fundamentos de Sistemas Digitales, 7ª.ed., Madrid España, Edit. Pearson Education, 2000, p 788.

El rendimiento de una unidad de disco duro viene determinado por varios parámetros básicos. Una operación de búsqueda consiste en el movimiento del cabezal de lectura-escritura hasta la pista deseada.

El tiempo de búsqueda es el tiempo medio de realización de dicha operación. Normalmente este tiempo va de 7 y 14 ms.

El periodo de latencia es el tiempo que tarda el sector deseado en colocarse debajo del cabezal, una vez que este se ha posicionado en la pista deseada. El periodo de latencia medio supone que el disco debe recorrer media revolución. Las velocidades de rotación son típicamente 3600 rpm, 4500 rpm, 5400 rpm, 7200 rpm, 10033 rpm, en estas últimas el periodo medio de latencia es inferior a 3 ms.

La suma del tiempo medio de búsqueda y el periodo medio de latencia es el tiempo de acceso de la unidad de disco.

3.3. Densidad de área

La densidad de área, también a veces llamada densidad de bit, se refiere a la cantidad de datos que pueden ser almacenados en una cierta área del disco magnético. Los platos magnéticos tienen dos superficies, pero sólo tomamos en cuenta lo que un área de ellas pueda contener, usualmente está expresada en bits por pulgada cuadrada, o BPSI². Siendo una medida bidimensional, se calcula como el producto de dos medidas unidimensionales.

Densidad de pista: es la medida de que tanto las pistas concéntricas están empaquetadas, cuántas pistas pueden ser colocadas en una pulgada de radio del plato. Por ejemplo, si tenemos un plato de 3.74 in. , eso nos da 1.87 in. Descontando el espacio interior que sostiene el plato y la parte más externa que tampoco puede ser usada, nos restaría 1.2 pulgadas de radio utilizable. Si en ese espacio contiene 22,000 pistas, la densidad de pista será aproximadamente de 18,333 pistas por pulgada.

Densidad lineal o de grabación: es la medida de que tan estrechamente están los bits empacados a lo largo de la pista. Dado que por ser concéntricas las pistas no tienen la misma longitud y de hecho no tienen la misma densidad, los fabricantes usualmente citan la densidad lineal más alta.

Como se ha expuesto anteriormente, la cantidad de información que se puede almacenar en un área específica se ha incrementado constantemente. En sus albores, el disco magnético para una PC era un plato que podía contener 4 MB por pulgada cuadrada, pero en estos días mediante cabezas magneto-resistivas de alta tecnología, se logró en 1993 en laboratorio superar los 130 GB por pulgada cuadrada.

Año con año la densidad de área se ha incrementado, venciendo las dificultades que conlleva hacer más pequeños los campos magnéticos polarizados en el disco para evitar que estos se traslapen, sin embargo, la densidad de área no puede incrementarse infinitamente.

² Bits Per Square Inch: bits por pulgada cuadrada.

Existe una nueva metodología llamada grabación perpendicular o vertical, en la cual la grabación se realiza en dos dimensiones, aunque aun debe ser mejorada. En un futuro puede incluso grabar información en 3 dimensiones, y en tal dado caso probablemente ya no sería necesario que el disco rotara.

Aún y con estas ventajas, físicamente no se puede exceder una barrera una vez que las partículas a polarizarse sean unos cuantos átomos. El máximo que se ha alcanzado en laboratorios hasta la fecha es de 50 Gb por pulgada cuadrada, y aunque las técnicas mencionadas anteriormente pudieran llevarla hasta un Pbit³, en ese momento prácticamente se estarían polarizando átomos en sí, y después ya no sería posible continuar. Si bien teóricamente es posible llevar hasta este punto la densidad de grabación, técnicamente es un gran desafío, ya que para detectar campos tan débiles los circuitos asociados deberán ser bastante más sensibles de los actuales, y evitar las interferencias o el ruido se antoja prácticamente imposible.

Si digitalizáramos toda la información que la humanidad ha generado desde sus comienzos hasta el siglo XIX seguramente obtendríamos un par de Pbits, pero durante el siglo XX y lo que va del XXI seguramente ya habremos superado esa cifra varias veces.

³ Petabit: 1,000,000,000,000,000 bits.

Es por ello que, aunque parece lejano, este límite empieza a ocupar a los científicos.

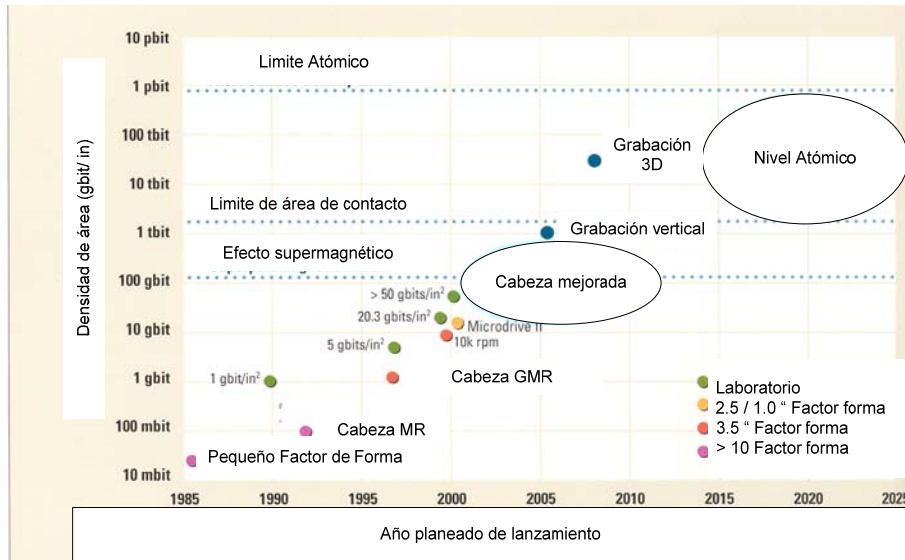


Figura 3.4.

Límite en la densidad de área

Fuente: Moore. Fred, New Game New Rules, 1ª.ed., Boulder U.S.A., Ed. Horison Information Strategies, 2003, p 46.

Dado que área es una medida de dos dimensiones, es el producto de otras dos dimensiones unidimensionales.

3.4. Cabezas de lectura-escritura

Los discos suelen tener de uno a cuatro platos, y teniendo una cabeza por superficie para lectura y otra para escritura nos da 16 cabezas máximo por disco, aunque puede haber algunas variaciones con algunos fabricantes.

Las cabezas no hacen contacto con los platos, y en esto se diferencian de las cabezas de disco flexible, VCR y unidades de cinta. La razón para esto es la alta velocidad a la que gira

el disco, y la constante necesidad de buscar de una pista a otra a lo largo del disco, lo que provocaría un desgaste sumamente rápido de platos y cabezas.

Como todos los componentes del disco magnético, las cabezas de lectura-escritura han sufrido múltiples modificaciones. La cabeza de escritura no ha alterado en mucho sus principios básicos de funcionamiento, si bien bastantes en la calidad de ensamble y materiales. Consiste en una bobina rodeando parcialmente un núcleo hueco ferroso en el que existe una delgada ranura, de de algunas micras de tamaño, por donde escapan las líneas de campo magnético que polarizan las partículas magnéticas en la superficie de los platos. Si bien las cabezas de lectura funcionaron en un principio en sentido inverso, el descubrimiento del efecto magneto-resistivo ha alterado enormemente su estructura, y hace que la lectura de una secuencia de bits sea más rápida que la escritura. Una cabeza magneto-resistiva emplea un material especial semi-conductivo que cambia su resistencia en la presencia de un campo magnético, en tanto la cabeza pasa sobre el plato, su resistencia se altera de manera correspondiente a la alteración de de los campos que se han polarizado en el plato. Un sensor es utilizado para detectar estos cambios en la resistencia, lo que permite que puedan leerse los bits.

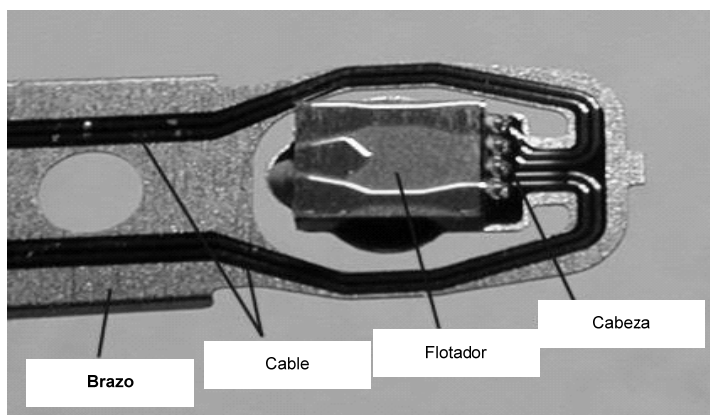


Figura 3.5.

Vista de una cabeza GMR

Fuente: <http://www.pcguide.com>

Las cabezas de los discos modernos flotan sobre la superficie de los discos y realizan todo su trabajo sin tocar el disco al que están magnetizando. El espacio entre las cabezas y el

plato es llamado altura flotante, o también hueco de cabeza. Las cabezas están montadas sobre resortes que sujetan a los deslizadores. Los deslizadores son los delgados soportes de las cabezas que las conectan con los brazos de las cabezas, y tienen la función de que cuando el disco está en reposo están en contacto con el plato impidiendo que la cabeza quede “bailando” dentro del disco. Cuando el disco se activa una delgada capa de aire se forma por debajo de los deslizadores, manteniendo una mínima pero importante separación del plato giratorio.

Debido a la mínima distancia entre el plato y los deslizadores, los discos son ensamblados en un cuarto sellado, donde el aire es filtrado para remover hasta las más finas partículas de polvo, y al contrario de la creencia, el interior de los discos no está en vacío, y lo más curioso, tampoco está sellado.

Un filtro especial de respiración es instalado, y sólo discos donde la altitud es mucha vienen sellados, debido a que no hay la suficiente presión para mantener a los deslizadores separados del plato.

Es realmente asombroso qué tan cerca están a la superficie de los discos las cabezas de lectura, que es alrededor de 0.5 micropulgadas, en tanto que un cabello humano tiene un grosor de unas 2,000 micropulgadas. Como se ha explicado, cuando la densidad de área se incrementa, los campos magnéticos son más pequeños y débiles. Para compensar esto, ya sea que las cabezas sean más sensitivas o se disminuya la altura flotante, y cada vez que dicha altura disminuye, cada aspecto mecánico debe ser ajustado para asegurarse que los platos no tienen irregularidades, el alineamiento de las cabezas y los platos es perfecto, y no se encuentra partícula alguna en el ambiente.

Una nuevo avance a las cabezas magneto-resistivas, son las cabezas gigantes magneto-resistivas o GMR. El término gigante no se refiere a su tamaño físico, de hecho físicamente son más pequeñas, sino por el efecto gigante magneto-resistivo, descubierto a finales de los años ochenta por Peter Gruenberg y Albert Fert, semejante al efecto magneto-resistivo, pero muchas veces más drástico al trabajar con grandes campos magnéticos y delgadas capas de varios materiales magnéticos, detectando bruscos cambios de resistencia en estos materiales, y si bien los materiales con que en un principio se experimentó no eran susceptibles para uso industrial, se sentaron las bases de esta tecnología.

Ingenieros y científicos en el centro Almaden de investigación de IBM® desarrollaron el producto comercialmente luego de probar cientos de materiales y métodos diferentes. Un importante avance fue el descubrimiento de que el efecto GMR opera bien en circuitos fotolitografiados⁴, lo que permitió instalar la cabeza de lectura directamente encima de la de escritura. Para 1997 IBM® introdujo esta tecnología al mercado, continuando con el constante incremento en la sensibilidad de estos dispositivos.

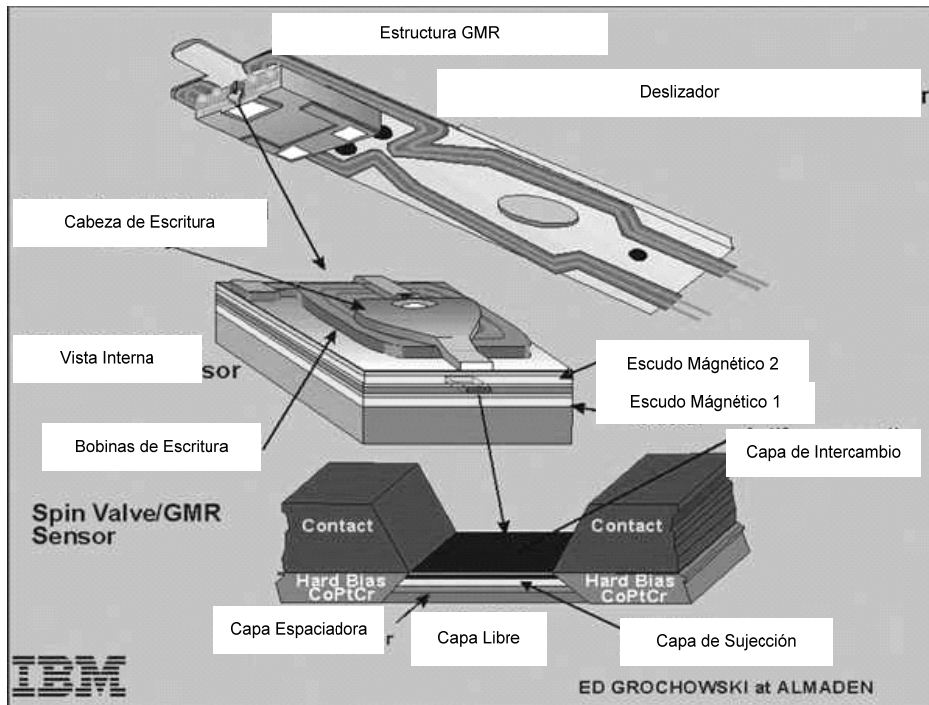


Figura 3.6.

Estructura de una cabeza GMR.

Fuente: <http://www.redbooks.ibm.com/>

Las cabezas GMR se componen de cuatro capas de material delgado apiladas en una sola estructura:

⁴ Mediante luz se dibuja una máscara mediante la cual las conexiones de los circuitos quedan grabadas.

Capa libre: esta capa funciona como sensor, está hecha de una aleación níquel-hierro y se desliza sobre la superficie que contiene los bits de información. Como su nombre lo indica, es libre de rotar en respuesta a las secuencias magnéticas en el disco.

Capa espaciadora: esta capa no es magnética, y es colocada entre las capas libre y de sujeción para separarlas magnéticamente.

Capa de sujeción: capa de cobalto está sujeta a una orientación magnética fija determinada por la capa de intercambio.

Capa de intercambio: constituida por un material antiferromagnético, típicamente hierro y manganeso.

Cuando la cabeza pasa sobre un campo magnético de cierta polaridad, la capa libre alinea sus electrones con los de la capa de sujeción, lo que crea una disminución en la resistencia de toda la cabeza. Cuando la cabeza pasa sobre un campo magnético de polaridad opuesta, los electrones en la capa libre giran y se desalinean de los de la capa de sujeción, lo que aumenta la resistencia en toda la estructura, de esta forma la resistencia cambia debido a la característica de girar la alineación de los electrones en la capa libre, por lo que IBM® ha llamado a esta estructura válvula de giro.

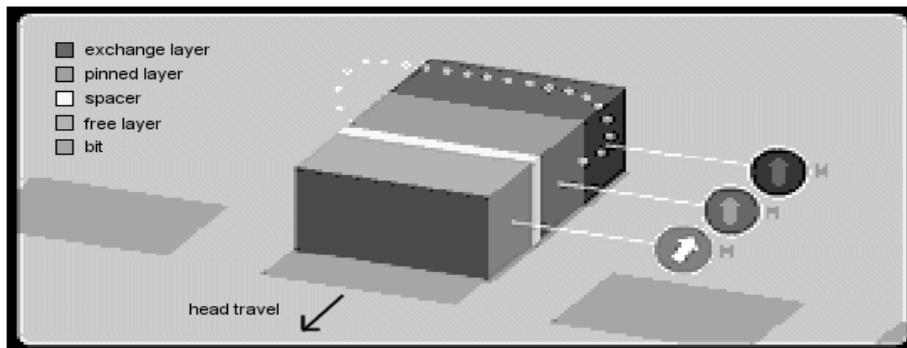


Figura 3.7.

Funcionamiento de una cabeza GMR

Fuente: <http://www.redbooks.ibm.com/>

Las cabezas GMR sufren una alteración en su resistencia de 5% a 8%, a comparación de 2% MR antecesoras. Se encuentra en sus etapas de desarrollo la tecnología CMR Colossal

Magneto-Resistiva, sin embargo, por estar en desarrollo la información sobre como serán es muy escasa., sin embargo, dado que se refieren igualmente al efecto magneto-resistivo, se anticipa que sólo se refiera a una mejora cualitativa en los materiales y método de ensamblaje, pero de modo de operación similar.

3.5. Servo-técnicas y características físicas

Los discos modernos utilizan un dispositivo llamado bobina de voz (voice coil)⁵ para mover el brazo de las cabezas hacia dentro y fuera de la superficie de los platos, y retroalimenta un servo sistema que dinámicamente posiciona las cabezas sobre el cilindro deseado. La bobina utiliza atracción y repulsión electromagnética. Una bobina es enrollada en una protuberancia metálica en el filo del juego de cabezas. Esto se monta dentro de un ensamble que contiene un imán permanente, y cuando se alimenta la bobina se genera una atracción o repulsión respecto al imán permanente. Al controlar la corriente se puede establecer la posición con precisión. El nombre de “bobina de voz” viene por la apariencia que guarda con la tecnología de los altavoces de audio.

El brazo es un servo sistema de ciclo cerrado, es decir, que realiza el movimiento, mide el resultado, calcula que tan lejos está del objetivo y con ello retroalimenta para realizar el siguiente movimiento. Las cabezas de lectura-escritura son las que indican la posición mediante códigos escritos en las pistas.

Si bien existen las técnicas de incrustación, donde los servo códigos están dentro de un sector en cada disco. También puede dedicarse un plato específicamente para ello, sin embargo, la técnica más empleada en estos días es la de servo código integrado, la cual consiste en dispersar los servo códigos a lo largo de las pistas de todos los platos. Los servo códigos y la información son leídas por las mismas cabezas. También, la necesidad por constante calibración termal se reduce al incluir los servo códigos junto con la información, ya que están a la misma distancia del centro y se expandirán y contraerán juntos con los cambios de temperatura. Los servo códigos están escritos en el disco desde que es fabricado. El equipo utilizado para ello es complejo y costoso y los códigos estarán allí por

⁵ Así denominada por tener semejanza al mecanismo que hace trabajar a los altavoces.

toda la vida del disco, la tarjeta controladora del disco evitará que sean reescritos por cualquier otra información.

Cuando los platos no están girando, los deslizadores descansan sobre los discos, así que al arrancar o detenerse rozan físicamente con ellos durante un breve periodo de tiempo. En tanto que platos y cabezas están diseñados con el conocimiento de que este contacto en algún momento deberá darse, es lógico pensar que esto deberá ser donde no haya información grabada. Por esta razón se reserva una pista que será la encargada de los despegues y aterrizajes de la cabeza, y apropiadamente es llamada zona de aterrizaje, y "estacionar" al acto de llevar las cabezas a esta zona. Los discos actuales con bobina de voz se estacionan automáticamente, y dependiendo del modelo, algunos de ellos incluso utilizan la energía cinética de los platos giratorios cuando la alimentación eléctrica es interrumpida. IBM® ha desarrollado una alternativa al estacionamiento de cabezas. En vez de permitir que las cabezas caigan al interrumpirse la corriente, las cabezas son levantadas completamente cuando el disco aún está girando mediante una rampa especial, y sólo después de eso los platos se detienen, el proceso ocurre a la inversa en el encendido.

Es común preguntarse por qué un solo brazo de acción controla las cabezas, si se pudieran incluir más para mejorar el desempeño. De hecho se intentó a inicio de 1990 por Conner Peripherals, sin embargo nadie más lo ha intentado. Existen varias razones por las que no se ha llevado a cabo, y aunque la complejidad en el diseño es mucha, el principal obstáculo no es técnico, sino de mercadeo. La duplicación de componentes lo harían muy caro, y la mayoría de los consumidores no pagarían por el incremento en el desempeño, ya que sólo se pueden abatir costos en producción en masa, y no hay un mercado lo suficientemente grande que pague el costo de componentes duplicados. Además, en la práctica, los arreglos de discos proporcionan muchas de las ventajas de desempeño al dividir la carga de trabajo entre varios discos.

Los platos magnéticos se componen de un sustrato y de una delgada capa que es la que en realidad almacena la información. El sustrato, que compone la mayor parte del plato, no tiene nada que ver con el proceso de lectura-escritura, y sólo sirve de soporte para el medio magnético.

Para ser apropiado, el material de sustrato debe de ser rígido, fácil de trabajar, ligero, estable, inerte magnéticamente, barato y disponible. El material más utilizado ha sido un compuesto del aluminio, el cual cumple este criterio.

Debido a la forma en que el plato gira con las cabezas apenas flotando sobre ellos, los platos deben ser muy suaves y planos. Una superficie que no es totalmente plana puede afectar seriamente las cabezas. Debido a que constantemente la distancia entre las cabezas y el plato se reduce, se desarrollan nuevos materiales para garantizar una superficie más llana. Al parecer componentes vítreos serán los sucesores del aluminio, que tienen varias ventajas comparativas:

Mejor calidad: la primera y más importante razón para cambiar al vidrio es que los platos son mucho más suaves y lisos que el aluminio, haciendo posible disminuir la altura de las cabezas

Mayor rigidez: el vidrio es más rígido que el aluminio por el mismo peso de material. Rigidez mejorada, una de las razones por las que el diámetro de los platos se está reduciendo, es importante para reducir el ruido y la vibración de discos que trabajan a alta velocidad.

Platos más delgados: la rigidez mejorada del vidrio también hace posible que los platos se vuelvan más delgados que con aluminio, permitiendo alojar más platos en el mismo espacio. También implica un menor peso, lo que ahorra también potencia del motor de giro y tiempo de inicio cuando se encuentra en descanso.

Estabilidad termal: cuando se somete al calor, el vidrio se expande mucho menos que el aluminio. Con algunos discos conteniendo 35,000 pistas por pulgada o más, incluso una pequeña expansión provoca que las pistas se alteren, y el servomotor tiene que realizar ajustes que disminuyen el desempeño.

Una desventaja clara del vidrio contra el aluminio es la fragilidad, particularmente cuando los discos son muy delgados, por ello se está experimentando con compuestos cerámicos. Uno de ellos es el MemCor, que es un vidrio con injertos cerámicos de alta pureza, que debido a las características inherentes de las cerámicas de dureza y resistencia al cambio de temperatura resulta en una composición idónea.

3.6. Codificación

Si bien, conceptualmente hablando, es simple ligar 0 y 1 a las polaridades N-S S-N, la realidad es mucho más compleja y no es posible establecer una correspondencia uno a uno, y hay 3 razones principales para ello:

1. **Campos vs. reverses:** las cabezas de lectura/escritura no están diseñadas para medir la polaridad real de los campos magnéticos, sino para reverses de flujo, los cuales ocurren cuando la cabeza se mueve de un área que tiene una polaridad norte-sur a otra que tiene una sur-norte y viceversa. La razón de que se prefieran reverses de flujo a campos magnéticos absolutos es simplemente que los reverses son más fáciles de detectar. Cuando la cabeza pasa sobre un revés en el flujo magnético se genera un pequeño pico de voltaje que puede ser detectado por un circuito. Conforme la densidad del disco se incrementa, es decir, que almacena una mayor cantidad de información en una superficie dada, se debe disminuir la intensidad de cada campo magnético para evitar que se interfieran entre ellos. Al ser menores los campos magnéticos su detección gradualmente se vuelve más difícil, lo que hace imposible la utilización de la polaridad real para codificar.
2. **Sincronización:** otra consideración es que debe haber algún método para marcar dónde un byte termina y el siguiente empieza. Incluso si pudiéramos usar las polaridades absolutas para marcar 1 y 0, ¿qué usaríamos para separar un byte de otro, si los campos se fusionan?
3. **Separación de Campos:** los campos magnéticos son aditivos, es decir, que si alineáramos una hilera de campos con la misma polaridad lo que obtendríamos es un solo campo sumamente fuerte y no varios campos pequeños mediante los cuales se pueda codificar información, sin mencionar que si los campos son mayores son más fáciles de interferirse o interaccionar con variables ajenas al disco.

Para poder establecer un método de codificación confiable debemos evitar tanto como sea posible alinear los campos con la misma polaridad para evitar que se fusionen, y para lograr que los bytes se diferencien debe de añadirse una señal de reloj muy precisa que debe estar sincronizada con la velocidad de giro de los platos.

A lo largo de los años la mayoría de dispositivos electrónicos ha sufrido una transformación al pasar de señales analógicas a señales digitales, esto para manejar

una nitidez superior en la señal y evitar la injerencia de ruidos externos que distorsionen la señal.

Curiosamente en la codificación de disco el efecto es el contrario, ya que debido a los reverses magnéticos es fácil construir una señal donde los picos sean fáciles de distinguir, es por ello que el método en boga de codificación que se expone en el siguiente tema, extrae de sus cabezas una señal analógica que luego convertirá en digital, este método es llamado "respuesta parcial parecido máximo".

3.6.1. Respuesta parcial parecido máximo

El método de codificación respuesta parcial parecido máximo o PRML⁶ es diferente a sus métodos antecesores que utilizaban *detección de picos*.

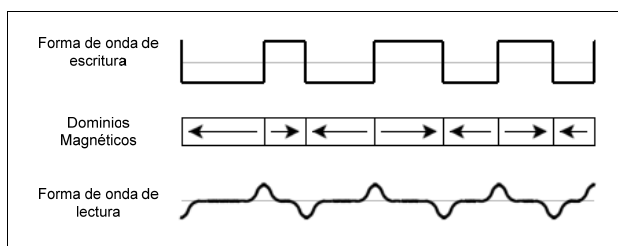


Figura 3.8.

Detección de picos

Fuente: <http://www.pcguides.com>

Los métodos de detección de picos funcionan bien en tanto los picos de voltaje son lo suficientemente grandes para ser separados del ruido de fondo de la señal. En tanto la densidad de información se incrementa, los reverses de flujo se hacen más continuos y la señal se vuelve mucho más difícil de analizar, ya que los picos se aproximan y empiezan a interferirse entre ellos, lo que puede ocasionar que no puedan distinguirse algunos bits y se

⁶ Partial Response, Maximum Likelihood.

corrompa la información. Debido a que esto debe evitarse a toda costa, se tiene que limitar la densidad de área de información, y esta densidad de área sólo puede incrementarse haciendo los campos magnéticos más débiles, pero en algún punto se vuelven tan débiles que los circuitos difícilmente puede interpretar dónde están los reveses de flujo.

Para combatir este problema un nuevo método fue desarrollado tomando un acercamiento diferente al problema, y cambia totalmente la forma en que la señal es leída y grabada de la superficie del disco. En vez de tratar de distinguir picos individuales para identificar reveses de flujo, una controladora utilizando PRML emplea un sofisticado muestreo de señal digital y algoritmos de procesamiento para manipular la señal análoga proveniente de las cabezas, llamada componente parcial, y determina la secuencia de bits que mejor se ajuste. Es decir, que las cabezas transmiten una señal analógica proveniente de los reveses de flujo magnético para después ser digitalizada.

Como toda técnica, el PRML se encuentra en constante evolución y modificación y es la tecnología en boga para la codificación de la información.

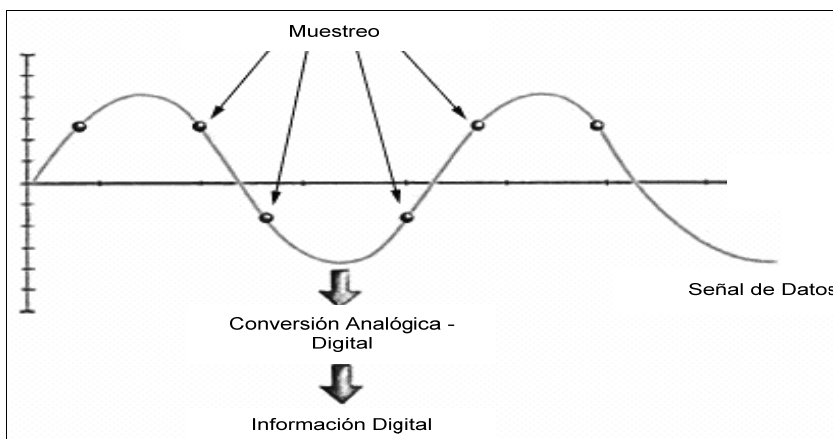


Figura 3.9.

Codificación PRLM

Fuente: <http://www.pcguides.com>

3.7. Arreglos de discos

Si bien el disco magnético tiene actualmente una capacidad superior a los 300 GB y una velocidad superior a las 15,000 rpm es insuficiente para ambientes empresariales o gubernamentales incluso de mediana escala. Por ello y por la necesidad de tener protección en caso de que algún dispositivo falle se desarrollan los arreglos de discos.

Básicamente estamos hablando de una serie de discos que van conectados a una tarjeta electrónica que se encarga de controlarlos y distribuir la carga de trabajo entre ellos. Dichas tarjetas a su vez deben de ir conectadas a una o más tarjetas procesadoras centrales que se encargan de todo el sistema en conjunto y su presentación hasta el servidor. Además deberá contener memoria caché, memoria no-volátil y tarjetas de interface para poder conectarse a los distintos medios de transmisión y tecnologías de servidores. Cuando un disco falla puede o no haber métodos para evitar que la información contenida en la unidad que falló se pierda.

Un arreglo de discos debe, además, tener mecanismos que prevengan una falla general del sistema al tener sistemas redundantes, es decir, que en caso de que se presente una falla en algún componente haya otro que pueda tomar su lugar de manera transparente evitando al máximo interrumpir la operación del equipo o degradando lo menos posible su funcionalidad ante los servidores.

Para lograr esto las fuentes de poder suelen ser duplicadas, así como los ventiladores, tarjetas controladores y prácticamente cualquier componente que ponga en riesgo la integridad del equipo. Los discos en sí, los cuales son de esperarse que puedan presentar primero algún problema debido al intenso uso que se hace de ellos, están generalmente protegidos por algún algoritmo de paridad que hace posible recuperar la información a partir de la paridad y los datos existentes. Dentro de los aspectos a considerar para un arreglo de discos está la necesidad presente y a futuro de espacio para respaldar información, la velocidad a que esta debe ser manejada y el grado de protección que se requiera para la información, ya que cuanto más valiosa sea esta información mejores sistemas de protección deberá tener.

Existen diferentes fabricantes de discos y de arreglos de discos. En el mercado mundial y nacional fácilmente puede verse compitiendo a Hitachi®, IBM®, EMC²®, Hewlett-Packard®,

Storagetek®, cada uno de los cuales ofrece características diferentes que en determinadas circunstancias puede ofrecerles ventajas sobre sus competidores.

Los arreglos de discos están orientados sobre dos aspectos principales: tener seguridad de que en caso de que un dispositivo falle la información contenida en él no se pierda; y mejorar el desempeño para respaldar y recuperar con mayor rapidez.

Esto último se logra al repartir un archivo muy grande en un arreglo con muchos discos, de forma que en cada porción del archivo hay una cabeza de lectura-escritura trabajando en él, con ello disminuyendo drásticamente el tiempo que tomaría a una cabeza posicionarse constantemente dentro de un solo disco, y si bien son unas pocas fracciones de segundo por operación, el resultado general de tener muchos cabezas trabajando sobre el mismo archivo puede ser muy notorio para el operador o proceso que lo esté requiriendo, mucho más notorio en aquellos discos que cuentan con pequeños buffer para incrementar la velocidad en cada una de las cabezas.

3.8. Niveles de arreglos (RAID)

Los niveles RAID son niveles de protección de la información que se encuentra en un arreglo de discos.

Nivel RAID 0: la información a escribir se reparte en los varios discos que componen el volumen, lo que significa que la información se escribe en paralelo, lo que mejora el desempeño en comparación a la escritura disco por disco. Cualquier disco que se pierda dentro del volumen significa la pérdida de la información.

Nivel RAID 1 (espejo): alto desempeño, ya que escribe paralelamente la información en dos discos diferentes. Es el nivel RAID con mayor disponibilidad, ya que en caso de pérdida de un disco su “espejo” procederá a tomar su lugar, también es muy seguro, ya que para que la información se perdiera tendrían que fallar simultáneamente un disco y su réplica, lo que es muy poco probable, sin embargo, sólo la mitad del arreglo de discos puede utilizarse para almacenar información, y la otra para hacer las réplicas, por lo que resulta muy costoso. Los discos en arreglo RAID 1 pueden soportar la pérdida de varios volúmenes, en tanto no fallen

un disco y su correspondiente réplica, si tal caso se presentara, se pierde toda la información.

Nivel RAID 2: toda la información es enviada a los discos en grupos de bits (típicamente medios bits o nibbles). Los bits son escritos en todos los discos del volumen en paralelo. El subsistema del arreglo genera códigos Hamming de corrección de errores, conocidos como ECC, y escribe dichos códigos a discos dedicados para ello. La principal ventaja de este arreglo es el acceso sincronizado, es decir, que actúa como si se tratara de un solo disco, pero se requiere dejar varios discos exclusivos para la generación de los códigos.

Nivel RAID 3: la información es escrita en todos los discos del volumen y la paridad es almacenada en un disco extra dedicado para este fin. Si un disco falla en el volumen, es posible recuperar la información leyéndola de los otros discos más la de paridad, lo que en tal caso sólo supone una degradación en el desempeño. Es el preferido para transferencia de archivos muy grandes tales como multimedia o diseño digital. La pérdida de más de un disco en el volumen conlleva la pérdida de toda la información.

Nivel RAID 4: la información se escribe en todos los discos del volumen en base a registros determinados, la información de paridad se almacena en un disco extra dedicado. Los discos no están sincronizados, lo que permite lecturas y escrituras simultáneas en distintos discos del volumen, esto mejora el desempeño para aplicaciones que utilizan pequeños bloques de información accedidos aleatoriamente, tales como bases de datos.

RAID Nivel 5: es el mejor para ambientes multi-usuario tales como respaldo de bases de datos o archivos de sistema el nivel de I/O es pequeño, contra una alta proporción de lectura. En este nivel los datos y la paridad están distribuidos en todos los discos del grupo. Para operaciones de actualización, se debe leer la información original, la paridad original, hacer nuevamente el cómputo de la paridad y rescribir la información y la paridad.

Si estando en RAID 5 falla un solo volumen dentro del grupo, todos los volúmenes asociados sufren una degradación en su desempeño, pero la protección de la información por paridad permite que la información siga disponible hasta que el volumen de repuesto haya regenerado al perdido, recobrando al fin nuevamente el desempeño original. La pérdida simultánea de dos volúmenes dentro de RAID 5 conlleva a la pérdida total de la información, aunque esto sea poco probable en un ambiente empresarial a mediana escala, por lo tanto este sistema es el más difundido en los arreglos de discos comerciales.

Nivel RAID 6: es una extensión del RAID 5, que permite la recuperación de información aún en caso de que dos discos fallen simultáneamente en el volumen, repartiendo dos cálculos de paridad entre todos los discos, pero el desempeño es ligeramente menor al RAID 5.

Nivel RAID 6+: elimina la degradación en el desempeño generado por los esquemas de redundancia de los niveles 5 y 6, al realizarse una actualización de información se calcula una nueva paridad y se escribe en un lugar diferente dentro del volumen, de forma que el lugar que ocupaba la paridad anterior permanece inválida hasta que nueva información llega a sobrescribirla.

3.9. Cinta magnética

La cinta magnética es una cinta plástica enrollada en uno o dos carretes recubierto de partículas ferromagnéticas susceptibles de polarización. El acceso es secuencial, a diferencia del disco que es aleatorio, por lo que la información se va almacenando del inicio de la cinta hasta el fin poniendo una etiqueta cada que se completa un archivo y modificando la etiqueta al inicio de la cinta donde se indica si un archivo dado está presente en esa cinta.

Si bien en un principio se empleaban carretes encargados de posicionar la cinta hasta el archivo dado, actualmente la cinta se encuentra dentro de un cartucho, el cual debe ser cargado en una unidad lectora de cinta para poder escribir o recuperar la información. Una vez que el proceso se ha completado el cartucho debe ser retirado y cargado otro que contenga el archivo deseado. Para llevar un control generalmente se destina un grupo de cartuchos para un día o semana, de tal forma que sea fácil ubicar en que cartucho se encuentra el archivo en cuestión dependiendo de la fecha en que se realizó su respaldo. El manejo de esta bitácora puede ser llevado por el operador, aunque el software de respaldo actualmente en el mercado cuenta con bitácoras y calendarios para ubicar los cartuchos donde se realizó o va a realizarse el respaldo deseado.

La cabeza de lectura-escritura si bien funcionan bajo principios similares a las de los discos, no tienen necesidad de ser tan pequeñas, ya que en el mismo espacio que ocupa una unidad de disco pueden embobinarse 500 ó 600 metros de cinta, además de que prácticamente están en contacto con el ambiente, ya que no se controla la presión interna

como en el caso de los discos, por lo que físicamente son más robustas, y casi la totalidad de los fabricantes sólo incluyen una cabeza de lectura-escritura por unidad de cinta.

Las unidades de cinta consisten básicamente en un servo mecanismo que recorre la cinta a la velocidad deseada y que es controlada por la cabeza de lectura-escritura, que puede ser fija o móvil. Al insertar el cartucho de cinta en la unidad ésta lo lleva hacia la cabeza y busca la etiqueta principal, que contiene información sobre los archivos que se han escrito. Una vez que ha encontrado la etiqueta principal se dice que la cinta “esta cargada”, y a partir de ese momento se puede empezar a trabajar con la cinta. El tiempo de carga y acceso a los archivos es superior al disco, debido a que conforme mayor porción de la cinta está siendo utilizada, mecánicamente tiene que llegarse hasta el último archivo escrito y de ahí continuar la escritura. Sin embargo, el costo de la cinta es mucho menor que el disco, y además los cartuchos pueden estar almacenados en alguna bodega hasta el momento de su utilización, cosa que no puede suceder con los discos que deben de estar constantemente energizados. De tal forma que la cinta es utilizada para archivos que piensan conservarse por largos periodos y que no están siendo utilizados constantemente.

Como se explicó en el Capítulo 1, la cinta está en el nivel más bajo del ciclo de vida de la información, pero en un momento dado su valor puede ser muy alto, por lo que debe cuidarse que el lugar donde se ubicarán los cartuchos esté bien ventilado, con temperatura entre 15°C y 25°C, libre de polvo y campos electromagnéticos. Bajo estas condiciones se puede garantizar la vida útil de los cartuchos de cinta por varias décadas.

Han existido diversos métodos de grabación, como escribir la información sobre una banda de la cinta magnética, sin embargo actualmente está imponiéndose la helicoidal. En este método el arreglo de cabezas sube y baja mecánicamente en tanto la cinta corre, de esta manera se forman franjas inclinadas en las cuales la información es escrita. Cada cabeza forma sus franjas, denominando a cada una de ellas canal. El formato ultrium 3, que es el que emplearemos en nuestra solución, maneja 8 canales. Adicionalmente se utilizan dos canales superiores y dos inferiores para dar control de la cabeza, estas franjas son escritas de fábrica y no pueden modificarse, en caso de desmagnetizarse la cinta quedará inservible.

Al igual que en disco se emplean cabezas GMR, existen dos pares de cabezas de lectura-escritura por canal, un par por cada arreglo en cada canal. Cuando la cinta gira en cierta dirección un par de cabezas se encargan de la lectura-escritura y sus correspondientes en

el otro arreglo se encargan de la verificación. En caso de discordar se regresa la cinta y se emprende nuevamente el proceso.

Las cabezas están incrustadas en un juego de cuchillas que sirven para eliminar residuos que pudieran estar adheridos a la cinta, además las están recubiertas por una delgada capa cerámica para evitar el desgaste.

La cabeza y los motores que embobinan la cinta se encuentran estrechamente coordinados para posicionarse en el archivo necesario. Al inicio de la cinta se escribe una etiqueta lógica con la información de todos los archivos presentes, de tal forma que se recorre la cinta alta velocidad hasta encontrar la marca del archivo correspondiente y empezar a trabajar con él, de tal forma es un medio secuencial, ya que se van agregando archivos uno tras del otro hasta llenar la cinta.

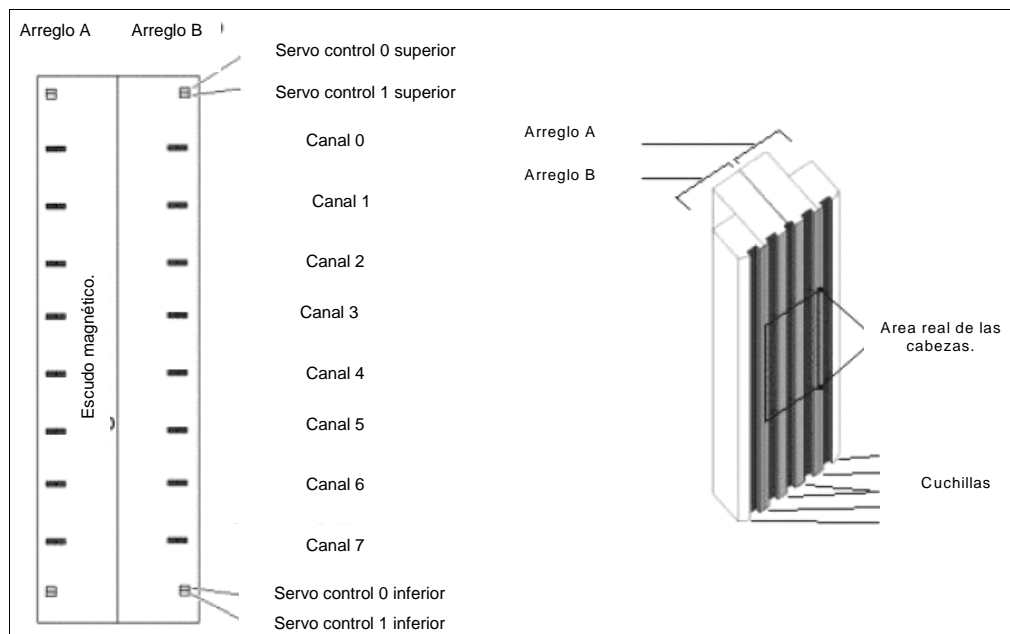


Figura 3.10.

Vistas interna y externa de las cabezas de lectura y escritura

Fuente: Global Learning Solutions, Documents on CD, 26ª.ed., Louisville U.S.A., Ed. Storagetek Literature Distribution, 2005, p 836.

CAPÍTULO 4. Arreglo Virtual Compartido, AVC

4.1. V2X AVC

El arreglo virtual compartido (Shared Virtual Array, SAV) es un arreglo de discos 6+, diseñado para cumplir con las funciones del manejo del ciclo de vida de la información detallado en el primer capítulo del presente trabajo.

Como se mencionó en el apartado 1.5. Capítulo 1, existen respaldos totales e incrementales de forma que la información que después de un tiempo no se necesita se elimina para liberar espacio, y se conserva la información importante por el plazo que sea necesario. Un ejemplo serían los precios de los productos, en los que de un par de años atrás no son importantes y que en todo caso pueden residir en los equipos clientes para personal administrativo que pudiera requerirlo. Para algunos registros fiscales y legales la ley requiere que se conserven hasta 10 años, sin embargo, a menos que se presente una auditoría no serán frecuentemente empleados, por lo que se pueden almacenar en cinta magnética, de menor costo y que no requiere estar constantemente energizada como los discos.

Esto es precisamente lo que realiza el AVC, la información de más uso reside en la memoria RAM integrada, lo que facilita tener un registro abierto y trabajar rápidamente con él hasta que después de cierto tiempo de no utilizarse se comprime y se almacena en disco. De acuerdo a las políticas definidas en el software de respaldo después de un tiempo determinado la información se elimina y se reemplaza por más actualizada o se migra hacia cinta magnética.¹

Dentro del AVC pueden generarse varias particiones y formatearse de manera adecuada a diferentes sistemas operativos que pueden coexistir en el AVC, reservando cada uno las unidades de disco que requieran.

¹ Incluso el manejo de las cintas magnéticas puede automatizarse formando una solución denominada administrador de almacenamiento virtual (Virtual Storage Manager, VSM), pero para ello se deben de incluir librerías robóticas y otros requisitos que escapan del alcance de esta tesis.



Figura 4.1.

V2X AVC de Storagetek®

Fuente: http://www.storagetek.com/products/product_page9.html

En la parte inferior del AVC podemos encontrar las tarjetas que comunican al equipo con el switch o los servidores. Estas tarjetas son las encargadas de comunicar internamente a todos los drives, ya que en su interior se encuentra una intrincada serie de anillos de fibra que enlazan de 5 en 5 los dispositivos, sin embargo, el AVC presenta una interfaz fabric en estas tarjetas al exterior, facilitando la interconexión con la red de área de almacenamiento.



Figura 4.2.

Tarjetas de conexión a servidores

4.2. Componentes físicos

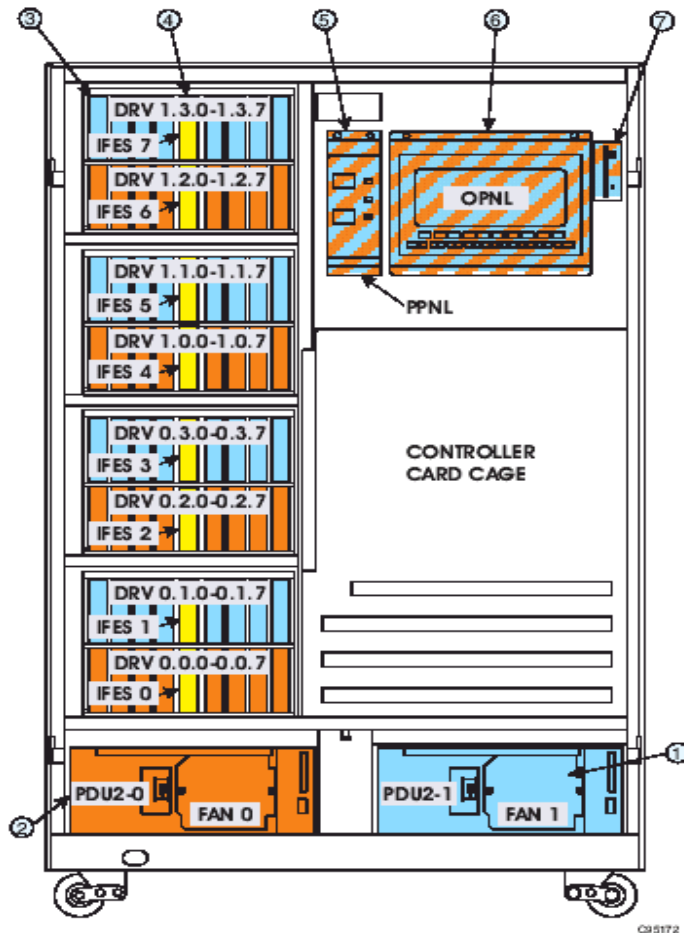
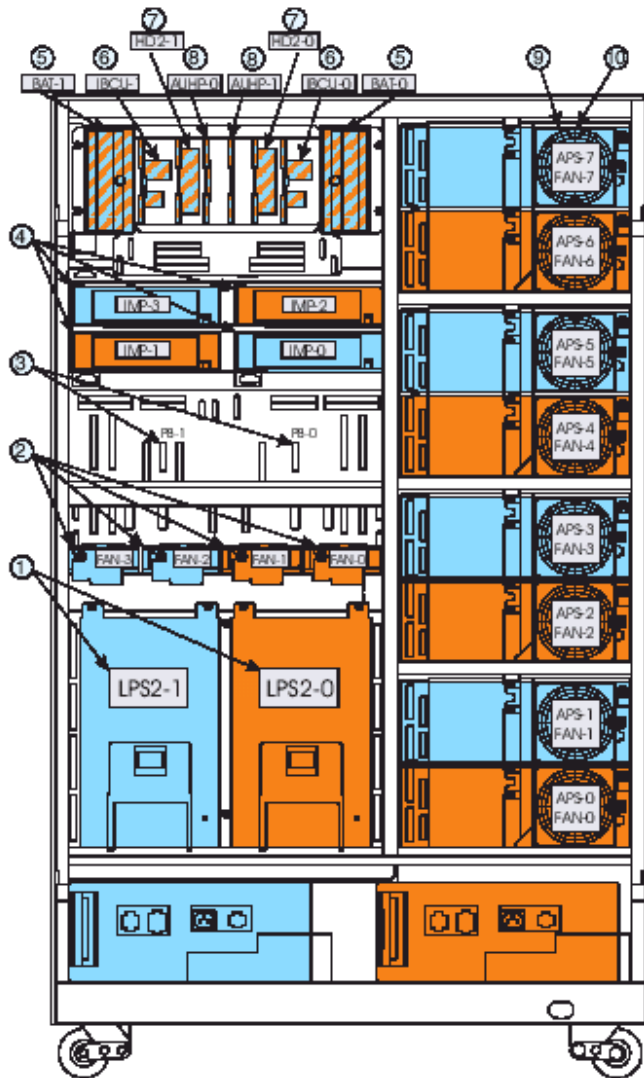


Figura 4.3.

Vista frontal

- 1)** Unidad de distribución de energía 1, (Power Distribution Unit, PDU 1), cuenta con ventiladores independientes, de aquí se distribuye la CA a los demás subsistemas.
- 2)** Unidad de distribución de energía 0, (Power Distribution Unit, PDU 0), cuenta con ventiladores independientes, de aquí se distribuye la CA a los demás subsistemas.
- 3)** Unidades de disco de fibra óptica de 73 GB y 300GB, hay 16 en cada arreglo horizontal con un máximo de cuatro arreglos.

- 4)** Tarjetas IFES, monitorean la alimentación eléctrica y ventiladores, sirven de conexión a las fibras creando anillos.
- 5)** Panel de control de alimentación, enciende o apaga el equipo y muestra el estado de la alimentación eléctrica
- 6)** Panel de operación, permite monitorear y controlar el equipo, modificar la configuración y reemplazar componentes dañados.
- 7)** Unidad de disco flexible, permite utilizar un disco de 3 ½ pulgadas para actualizar microcódigo, subir y bajar información de errores del equipo.



C25 185

Figura 4.4.

Vista posterior

- 1) Fuentes de poder para tarjetas lógicas y tarjeta madre C3
- 2) Ventiladores de las fuentes de poder para las tarjetas lógicas
- 3) Baterías BIOS para las tarjetas ISP, a diferencia de las baterías NVS no son reemplazadas rutinariamente.

- 4) Ventiladores de la jaula de tarjetas lógicas
- 5) Baterías CNV, proveen corriente directa para las tarjetas CNV por al menos 72 horas durante fallas de energía para proteger la memoria no-volátil.
- 6) Unidades de carga de las baterías: mantienen cargadas las baterías.
- 7) Unidades de disco de las ISP: unidades de disco para almacenar microcódigo funcional y de diagnóstico, registro de eventos e interfaz de soporte.
- 8) Tarjetas de conexión en caliente (hot plug)², alimentan y protegen contra descargas a los discos de las ISP, al panel de operación y a la unidad de disco flexible.
- 9) Fuentes de poder de los arreglos, proveen de energía a cada arreglo de discos, se compone de una fuente principal y una redundante en caso que la principal falle, un par para cada arreglo.
- 10) Ventiladores de las fuentes de poder de los arreglos, enfrían las fuentes de poder de los arreglos.



Figura 4.5.

Unidad de disco de 73 Gbyte

² Conexión en caliente (Hot plug): indica que se puede reemplazar sin apagar el equipo.

Fuente: Global Learning Solutions, Documents on CD, 26ª.ed., Louisville U.S.A., Ed. Storagetek Literature Distribution, 2005, p 836.

4.3. Subsistemas

Subsistema Hardware: unidad de control y almacenamiento integrada con 16 rutas de datos (data paths), y hasta 32 conexiones ESCON o 16 FC que pueden desempeñar hasta 32 operaciones I/O concurrentes.

Administrador del AVC (AVCA): software que expande la funcionalidad del AVC al agregar herramientas en el servidor.

Foto instantánea (SnapShot): herramienta que crea una copia virtual de cualquier volumen o data set en segundos sin espacio adicional y sin impactar el desempeño en que incurren todos los otros productos.

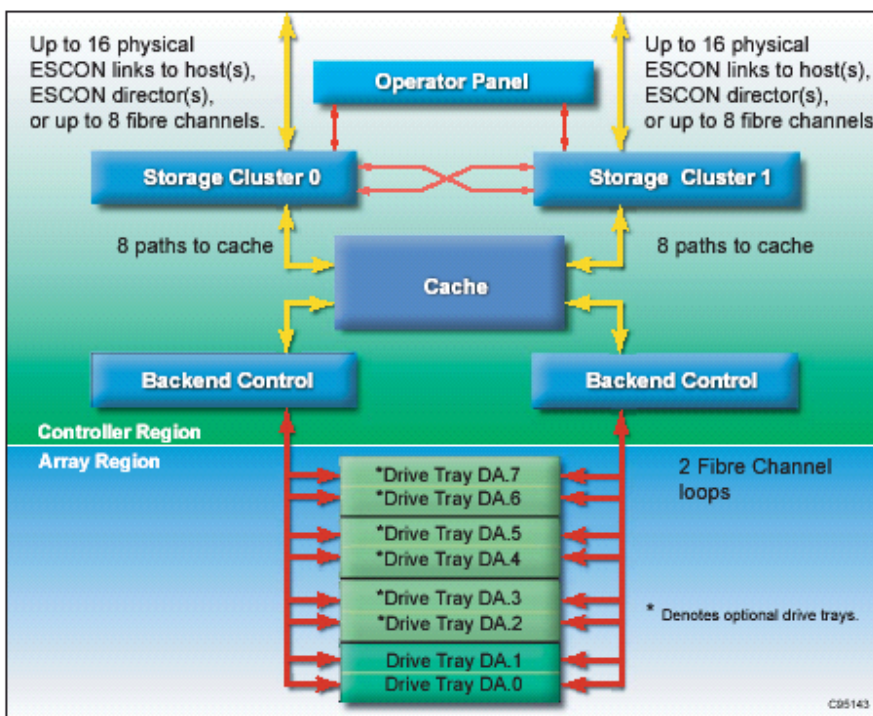


Figura 4.6.

Diagrama general de Entrada /Salida (I/O) del AVC V2X

Fuente: <http://www.support.storageetek.com>

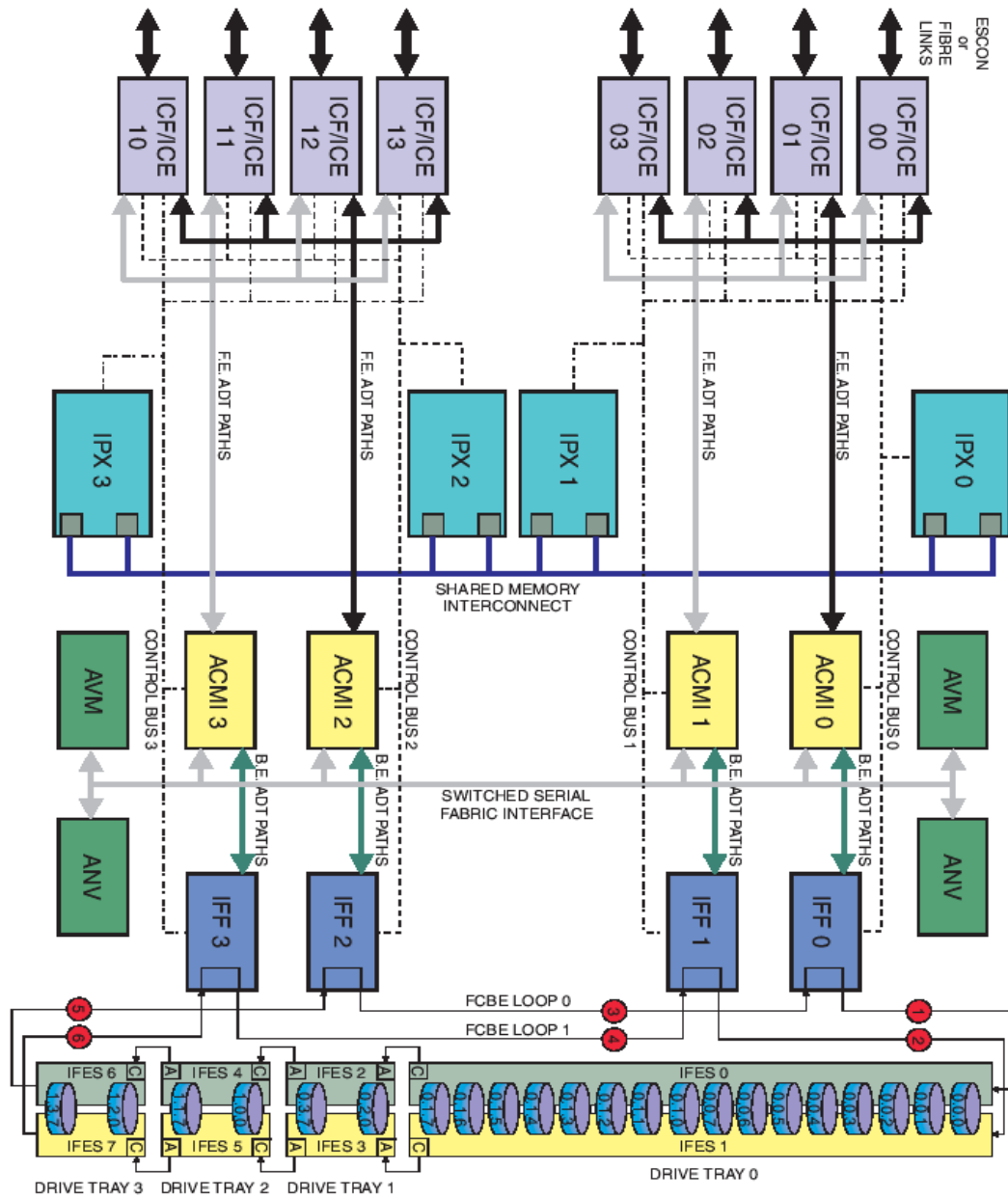


Figura 4.7.

Diagrama general V2X AVC

Fuente: Global Learning Solutions, Documents on CD, 26ª.ed., Louisville U.S.A., Ed. StorageTek Literature Distribution, 2005, p 133.

4.4. Mapeo dinámico

Cuando se configura un sistema AVC, el usuario define la configuración del sistema en unidades 3380 ó 3390, y para el servidor es aparente que estos son los dispositivos que está viendo y utilizando, pero no existen físicamente. El AVC dinámicamente hace un mapa o apunta la información a una serie de discos en el arreglo.

Una maniobra “destage”, es una transferencia de datos desde la memoria caché al arreglo de discos. Un requerimiento de escritura es enviado al dispositivos físicos y el ALB es llenado con datos del caché. El disco busca el cilindro especificado, cuando está la cabeza en el sector correcto la información es transferida del ALB al disco. En tanto la información es transferida del cache al disco, la información de redundancia es generada y acumulada en los buffers de paridad. Cuando una pista completa del arreglo ha sido escrita, la información de redundancia es también escrita al arreglo.

4.5. Paridad dual

En un arreglo de 15 dispositivos, que es también referido como 13+2+1, un disco es reservado como disco de repuesto y es globalmente disponible al sistema. Dentro del volumen, la información del usuario es grabada en pistas idénticamente direccionadas en todos excepto dos pistas, mismas que servirán para los dos niveles de redundancia. Un nivel de redundancia, denominado "P", es grabado en uno de los drives. Un segundo nivel de redundancia se calcula y pasa al otro disco como paridad "Q", sin embargo, no hay discos en específico para almacenar la paridad, la información es rotada entre todas las pistas de todos los discos en el arreglo.

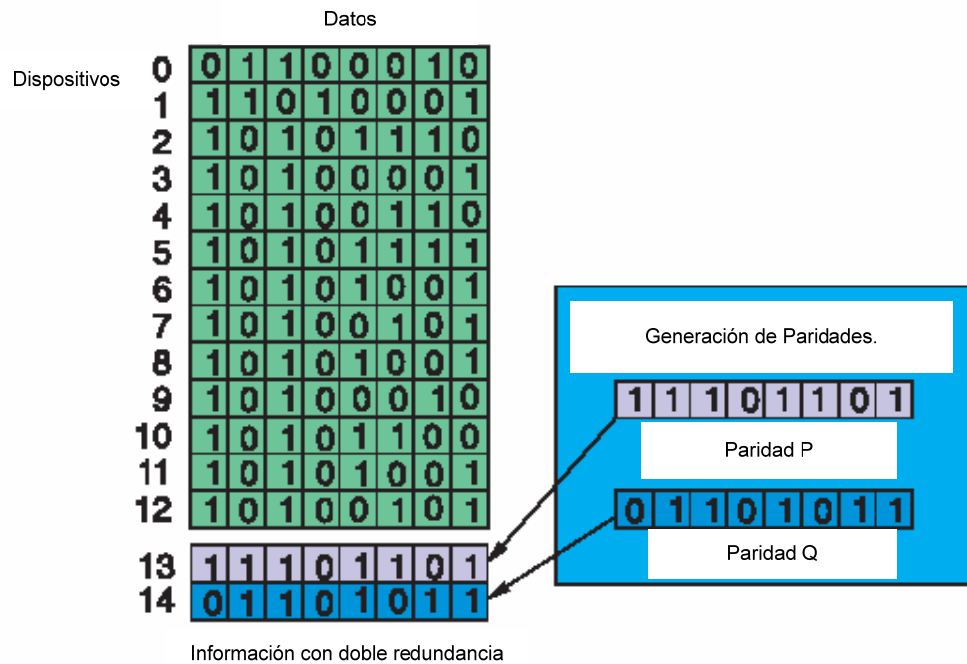


Figura 4.8.

Redundancia dual

4.6. Reparación de un disco dañado

Cuando un disco en el arreglo falla, la información del resto de los discos es utilizada para reconstruir la información del disco dañado en el disco de repuesto. Esta tarea corre de forma transparente para el servidor y no necesita intervención del usuario para completarse. Cuando el nuevo disco se cambie por el afectado pasará a ser el disco de repuesto. En caso de que un segundo disco llegara a fallar también puede proceder a reconstruirse, sin embargo, se podrá experimentar una reducción en el desempeño del 20%.

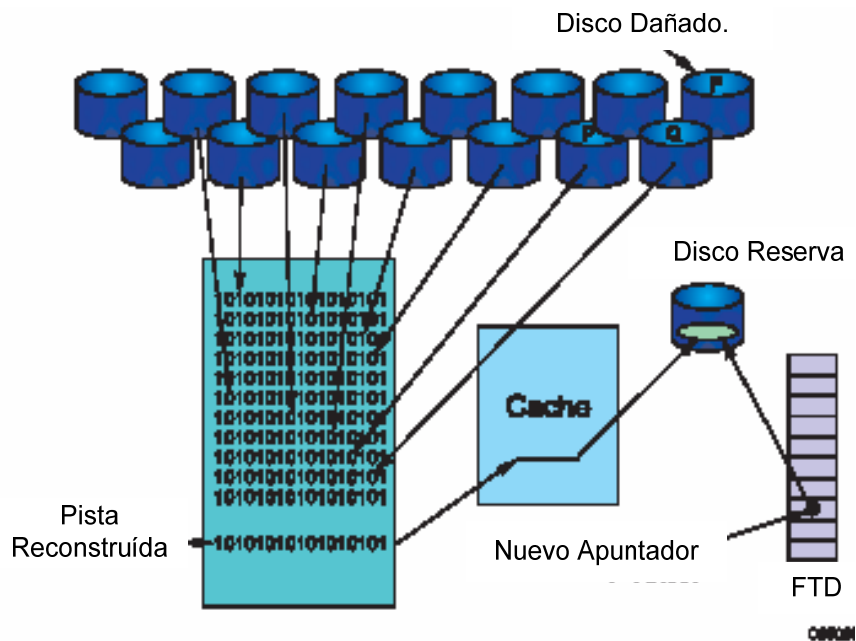


Figura 4.9.

Reconstrucción de discos

4.7. Pistas y volúmenes funcionales

El registro no es de un tamaño predeterminado, y es el AVC el encargado de distribuirlos a lo largo de los discos físicos, estos registros son denominados registros funcionales. Los registros funcionales se encuentran distribuidos a lo largo de varios discos, y la localización

de la física de los registros funcionales es alojada en el directorio de registros funcionales (Functional Track Directory, FTD), dicho directorio reside en la memoria base de la unidad de control. Asimismo, junto con la información almacenada se genera información de redundancia para la protección de la misma, esta información también se encuentra esparcida en el arreglo. Los dispositivos físicos están divididos en registros físicos, los cuales se componen de sectores físicos. El número de sectores requeridos para almacenar un registro funcional es determinado por que tan bien puede ser comprimido, el AVC siempre transfiere registros funcionales completos entre el caché y los dispositivos, además, el AVC nunca actualiza la información en el sitio donde es almacenada. Cuando un registro funcional es actualizado, se escribe a una nueva locación y la original es marcada como espacio libre.

Una característica más es que cuando un sector se detecta dañado, en este caso la ubicación física no vuelve a utilizarse y la información allí contenida se rescribe a otra ubicación física, todo esto sin correr ninguna tarea desde el servidor.

4.7.1 Tabla de dispositivos funcionales

El mapeo dinámico utiliza tablas de apuntadores que señalan a una dirección en los dispositivos físicos. Un conjunto de estos apuntadores es conocido como tabla de dispositivos funcionales, está alojada en la memoria compartida de la unidad de control, y contiene la configuración funcional del sistema, esto es, tipo de dispositivo, dirección, tamaño, tal como lo define el usuario. Una copia de esta tabla es respaldada en la unidad de control y en los discos.

4.7.2. Directorio de pistas funcionales

Una segunda tabla de apuntadores, conocida como directorio de pistas funcionales, tiene una entrada por cada pista funcional definida en el sistema. Se encuentra alojada en la

memoria base³ de la unidad de control. El sistema constantemente crea nuevas copias de dicha tabla y la sobrescribe a locaciones predefinidas en el arreglo. Durante el proceso de copia de la FTD, los cambios que en ella se realicen se registran en la memoria de almacenamiento no volátil (Non Volatile Storage, NVS) y son eventualmente escritos a disco.

4.8. Transferencia de datos

Toda la escritura al AVC son DASD escritura rápida a la memoria caché y NVS, la NVS actúa como respaldo de la memoria caché, cuando la operación está completada, el AVC informa al servidor para liberar el canal, y la información es encolada para ser manejada de la memoria caché hacia los discos.

Si llegara a presentarse un problema durante la escritura de la memoria caché a los discos, la información es recuperada al aplicar los cambios almacenados en la NVS. La compresión de la información se realiza en los circuitos que están en la ruta a la memoria caché, es decir, que al llegar a esta ya cuentan con compresión. Además, las transferencias sobre FC utilizan la asistencia en el formato de escritura (Record Format Assist, RFA), un estándar que provee las siguientes características:

- Compresión en bloques de información grandes
- Incrementa el desempeño
- Tamaño de bloque adaptable, que pueden ser de 512 B, 2-16 KB
- Mayor cantidad de unidades lógicas (Logical Unit Number, LUN)⁴

Los registros funcionales están distribuidos entre todos los dispositivos físicos disponibles, de esta forma la información es manejada por más brazos de acción y hay más rutas disponibles para el respaldo o recuperación de la información, de esta forma se maneja la información de manera paralela y se mejora el desempeño que puede tenerse al manejar un

³ Memoria base es memoria no disponible para las aplicaciones del servidor, mantiene información requerida para las operaciones del sistema.

⁴ Número de Unidad Lógica (Logical Unit Number, LUN): es un número lógico creado para manejar un dispositivo que se encuentra conectado a otro para llegar al servidor.

solo disco físico de gran capacidad, ya que la suma de los tiempos de latencia sería muy alta.

Este término en almacenamiento se refiere a que los volúmenes vistos desde el servidor no corresponden a un disco o grupo de discos físicos en el arreglo. En el AVC la información es almacenada en sectores (tracks) funcionales que contienen la información comprimida y que pueden estar repartidos en más de un disco físico, sin embargo, desde el punto de vista del servidor sólo existe un volumen en donde se almacena la información, y el manejo de la misma dentro del AVC es independiente para el servidor. La información del track es almacenada en el directorio de tracks funcionales FTD, el cual reside en la memoria de la unidad de control, al mismo tiempo en que la información se comprime y se aloja, se genera la información de redundancia misma que se reparte en varios discos físicos.

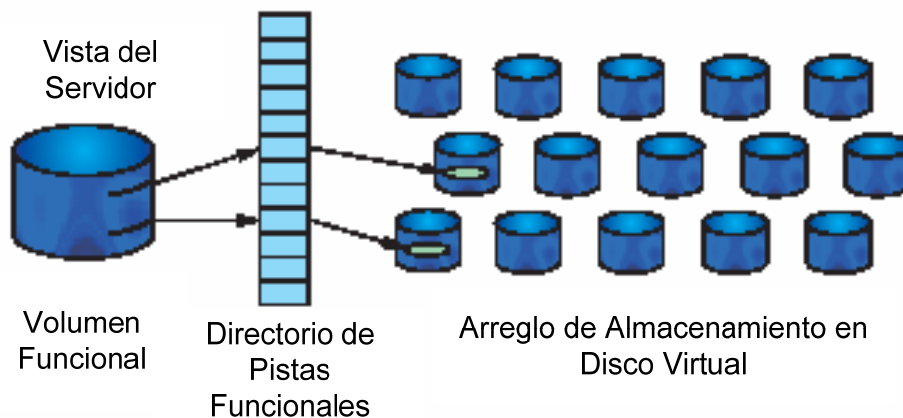


Figura 4.10.

Pistas y volúmenes funcionales

Desempeño: al estar repartida la información en varios discos físicos existen múltiples brazos de acción trabajando simultáneamente, de forma que el tiempo de latencia para leer o escribir una cierta cantidad de información se reduce al del disco de mayor latencia, es decir, que la operación puede realizarse en milisegundos a pesar que el tamaño del archivo sea muy grande.

Eficiencia en el manejo de espacio: convencionalmente un volumen tiene un espacio fijo predeterminado, lo que genera que cuando la información almacenada no cubre 100% del

espacio, como frecuentemente ocurre, se genera un desperdicio de espacio en cada volumen, que sumando todos los volúmenes es una cantidad considerable.

Los principales parámetros para monitorear un correcto comportamiento de nuestro equipo son:

- **Espacio colectado libre (Collected Free Space, CFS): debido** a que cuando el V2X AVC actualiza o renueva información lo hace sobre cilindros libres, la información anteriormente guardada queda reflejada como CFS, es decir, se sobrescribe al momento en que nueva información solicita espacio para almacenarse
- **Capacidad de carga neta (Net Capacity Load, NCL):** representa el porcentaje del total de capacidad del V2X utilizada para almacenar los tracks funcionales, es decir, donde reside la información del cliente. El NCL fluctúa durante el procesamiento de información debido a la escritura y borrado de información.

CAPITULO 5. Diseño e implementación de una solución de respaldo de información

5.1. Planteamiento del ambiente de respaldo

Se expone el escenario formulado en el alcance objetivo de la tesis, que puede tener cabida en una serie de departamentos muy marcados dentro de instituciones públicas y privadas que manejen bases de datos y archivos que en conjunto rondan los 15TB, con un crecimiento semanal de hasta 600 GB como las que citamos a continuación.

- ProceSAR
- Petroleos Mexicanos Exploración y Producción
- Petroleos Mexicano Dirección Corporativa de Finanzas
- Instituto Nacional de Estadística y Geografía
- KODAK Company
- Instituto Mexicano del Seguro Social
- TV Azteca Corporativo
- Grupo Imagen
- Televisa Corporativo

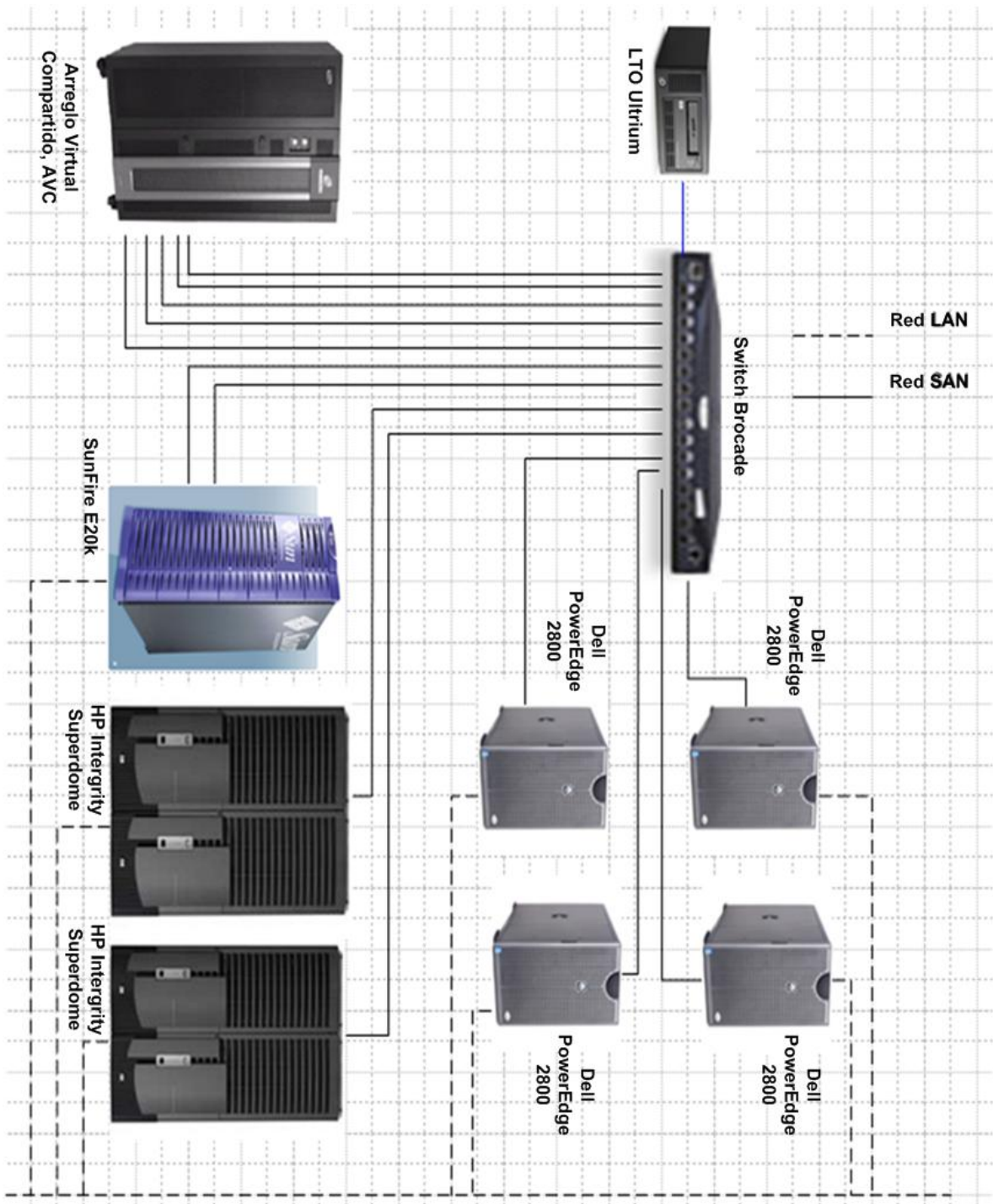


Figura 5.1.

Diagrama del centro de cómputo propuesto

El proceso de la conformación de nuestra red SAN se entiende de la siguiente manera:

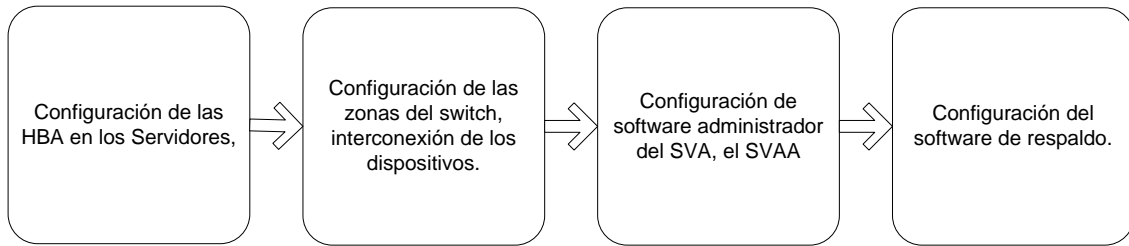


Figura 5.2.

Diagrama de flujo de la implementación de la red SAN

A continuación se detallan las características de los equipos y los procedimientos para su conexión en red SAN:

- **Sun FIRE E20K**

- Un servidor Sun Fire E20K con las siguientes características:
- Sistema operativo Solaris® 8.
- 4 dominios: (1)Producción, (2)Producción Cluster, (3)Administración y Finanzas e (4)Inventarios. Cada dominio utilizará una tarjeta Uniboard con 4 procesadores UltraSpark IV y 32 GB memoria RAM cada tarjeta Uniboard.
- 2 HBA FC-AL Emulex® LP9002L @ 2Gbps. La WWNN de la primer tarjeta es 10:00:00:00:c9:38:8d:00 y la de la segunda es 10:00:00:00:c9:38:8d:05.
- Motor de Base de Datos Oracle ® 9i en todos los dominios.



Figura 5.2.

Servidor Sun Fire E20K

Fuente: http://www.sun.com/servers/highend/sunfire_e20k/index.jsp

- **HP Integrity Superdome**

- Sistema operativo HP-UX 11i
- El primer equipo tiene un dominio: Transacciones, que agrupa a 4 celdas, cada una con 4 procesadores Intel Itanium 2 ® @ 1.6 GHz y un total de memoria de 64 GB RAM.
- El segundo equipo tiene dos dominios: Clientes y Proveedores, Nómina. Tiene una celda por dominio con 4 procesadores Intel Itanium 2 ® @ 1.6 GHz y 32 GB RAM.
- Una tarjeta Emulex® LP9002L @ 2Gbps con WWNN para el primer servidor 10:00:00:00:c9:38:8d:10 y 10:00:00:00:c9:38:8d:15 para el segundo.



Figura 5.3.

Servidor HP Integrity Superdome

Fuente: <http://www.hp.com/products1/servers/scalableservers/superdome/index.html>

- **Dell PowerEdge 2800**

- Sistema operativo Windows® 2003 ®.
- Procesador dual Intel Xeon ® @ 3.6 GHz en cada servidor, 12 GB RAM.
- 4 servidores:
 - Win1 SQL Server WWNN: 10:00:00:00:c9:38:8d:20
 - Win2 SQL Server WWNN: 10:00:00:00:c9:38:8d:25
 - Win3 Exchange WWNN: 10:00:00:00:c9:38:8d:30
 - Win4 Exchange WWNN: 10:00:00:00:c9:38:8d:35



Figura 5.4.

Servidor Dell PowerEdge 2800

Fuente: <http://www.dell.com>

5.2. Integración del adaptador de ruta del servidor (HBA)

5.2.1. Sistema operativo Windows® 2003

Primero ejecute el programa de instalación que viene en el CD que acompaña a la tarjeta.

1. Ejecute el archivo APInstaller.exe. Se desplegará el siguiente mensaje:



Figura 5.5.

Instalación de HBA Emulex LP9002 en Windows® 2003

2. Click **OK**
3. Se despliega la ventana de bienvenida
4. Clic **Siguiente**. Se despliega la ventana de controladores disponibles
5. El controlador más nuevo se encuentra pos omisión. Clic en **Siguiente**
6. Verifique el progreso de la instalación. Una vez que se complete la instalación un mensaje aparecerá en pantalla.

5.2.2. Sistema operativo Solaris® 8

1. Ingresar al sistema o cambiar de usuario a root “**-su root**”
2. Crear un directorio donde se alojará el archivo que descargó desde internet, o bien del CD que acompaña a la tarjeta. El controlador es un archivo comprimido de

extensión .tar , llamado lpfc-6.01f-sparc.tar el cual ubicaremos en el directorio “/emulex”

3. Cambiar al directorio donde está en archive comprimido:
 - a. # cd emulex
4. Extraer el archivo mediante la utilería “tar”
 - a. # tar xvf lpfc-6.01f-sparc.tar
5. Una vez que se ha descomprimido se debe instalar el paquete:
 - a. # pkgadd -d .
6. Se desplegarán las siguientes opciones.
7. Select package(s) you wish to process (or 'all' to process all packages). (default: all)
[?,??,q]:
8. Presione Enter.
 - a. Rebuild manual pages database for section 7d [y,n,?]:
 - b. “y”
 - c. Use IP networking over Fibre Channel [y,n,?]:
 - d. “n”
9. Do you want to continue with the installation of <lpfc> [y,n,?]:
 - a. “y”
7. Aparecerá el siguiente mensaje:
10. SCSI: If you are using lpfc to access disks, be sure to check the configuration file of your SCSI target driver presumably sd.conf to ensure that the driver will probe for all of the targets/luns in your environment.
11. Finalmente regresa al inicio, seleccione “q” para salir.
12. Select package(s) you wish to process (or 'all' to process all packages). (default: all)
[?,??,q]:
13. Se desplegará el siguiente mensaje:
 - a.

```
***IMPORTANT                                NOTICE                                ***  
  
This machine must now be rebooted in order to ensure sane operation.
```
14. Antes de reiniciar el equipo, considere que es posible que se deba modificar los archivos de configuración.
15. Si tiene conectados dispositivos que soportan LUN's, edite las entradas correspondientes en los archivos st.conf y sd.conf dentro de /dev.
16. Reinicie el sistema mediante los siguientes comandos.

- a. # sync
- b. # init 0
- c. # boot -- -r

5.2.3 Sistema operativo HP-UX 11.x

1. Instale la HBA Emulex en una ranura disponible en el sistema.
2. Ingrese al sistema como “**root**”.
3. Cargue el archivo “lpfc.depot” dentro del directorio “ /emulex”
4. Teclee:
 - a. cd Emulex
 - b. swinstall -s `pwd`/lpfc.depot -x autoreboot=true LPFC
5. El sistema reiniciará automáticamente y el controlador lpfc será incluido en el kernel de HP-UX.
6. Antes de que continúe, verifique si desea cambiar alguno de los parámetros por omisión, editando el archivo **adapCfg.h**

Nota: un nuevo archivo adapCfg.h es copiado en la ruta opt/lpfc/conf/ durante la instalación. Si está editado tendrá que salvarlo en un directorio diferente y editar en el nuevo archivo las entradas que hayan sido alteradas.

5.3. Integración del switch de la red de almacenamiento

1. Conectarse mediante un cable de red cruzado al switch.
 - a. La dirección IP estática de fábrica del switch es 10.77.77.77.
2. Una vez conectado el switch preguntará por usuario y password, los que son:
 - a. user: root
 - b. password: password
3. Para crear las zonas: ,notese que las cuatro tarjetas de los servidores Windows® ven sólo 2 tarjetas en el V2X AVC.
 - a. sw1:admin>zoneCreate “Sun1”,” 10:00:00:00:c9:38:8d:00;
20:06:00:60:69:c0:9f:00”
 - b. sw1:admin >zoneCreate “Sun2”,” 10:00:00:00:c9:38:8d:05;
20:06:00:60:69:c0:9f:05”

- c. sw1:admin >zoneCreate "HP1"," 10:00:00:00:c9:38:8d:10; 20:06:00:60:69:c0:9f:f9"
 - d. sw1:admin >zoneCreate "HP2"," 10:00:00:00:c9:38:8d:15; 20:06:00:60:69:c0:9f:fd"
 - e. sw1:admin >zoneCreate "Win1"," 10:00:00:00:c9:38:8d:20 20:06:00:60:69:c0:9f:a";
 - f. sw1:admin >zoneCreate "Win2"," 10:00:00:00:c9:38:8d:25; 20:06:00:60:69:c0:9f:a4"
 - g. sw1:admin >zoneCreate "Win3"," 10:00:00:00:c9:38:8d:30; 20:06:00:60:69:c0:9f:a9"
 - h. sw1:admin >zoneCreate "Win4"," 10:00:00:00:c9:38:8d:35; 20:06:00:60:69:c0:9f:a9"
4. Creamos una configuración donde incluimos las zonas anteriormente creadas.
 - a. sw1:admin >cfgCreate "solucion","Sun1;Sun2;HP1;HP2;Win1; Win2; Win3; Win4"
 5. Salvamos la configuración en la memoria no-volátil. Si existen otras configuraciones también serán salvadas.
 - a. sw1:admin >cfgSave
 6. Habilitamos la configuración que acabamos de salvar. Todas las operaciones del switch se interrumpirán y el switch reiniciará. Al levantar la configuración habilitada tomará efecto
 - a. sw1:admin >cfgEnable "solucion"
 7. De esta forma queda concluida la configuración del switch.

5.4. Integración del administrador del arreglo virtual compartido, AVC

5.4.1. Pre-requisitos

El administrador del AVCA es el programa encargado de hacer la configuración del AVC, es decir, definir particiones y formateo de las mismas, que luego serán empleadas por el software de respaldo de la SAN. El servidor deberá tener instalada al menos una tarjeta de

fibra óptica, en nuestro caso utilizaremos una LP9002L de Emulex sobre un servidor E20K de Sun Microsystems Solaris® versión 8. El AVC deberá ya estar encendido y con al menos una tarjeta ICF instalada.

El AVC deberá tener al menos el nivel de microcódigo: B01.02.00.00, y deberá reunir al menos las siguientes características:

1. Al menos una tarjeta ICF debe estar configurada para Fibre Channel, es decir, aparecer el valor "2" dentro del menú de configuración de interfase.
2. El AVC tenga un nombre asignado, esto se verifica dentro del menú de estado del sistema.
3. Debe haber al menos un arreglo de discos para producción, dentro de la configuración del sistema, entrar a estado de partición y verificar.
Nota: en el servidor, los dispositivos 3390-3 son identificados como SCSI A, y los dispositivos 3390-9 son identificados como SCSI B, si han sido definidos como SCSI en el panel de operación.
4. Ha sido asignado al menos un dispositivo ECAM privilegiado. Entrar al menú de configuración, después al de configurar subsistema y por último al de dispositivos funcionales, debe aparecer dispositivos 3390-3 o 3390-9 con "Y".
5. Asimismo, se deberá verificar en el servidor al que van conectados.
6. La HBA debe de estar configurada para acceder dispositivos funcionales.
7. Ingresar al sistema como "root", ingresar el comando # format, se deberá desplegar algo semejante a:

AVAILABLE DISK SELECTIONS:

0. c0t0d0 <ST39140A cyl 17660 alt 2 hd 16 sec 63>

/pci@1f,0/pci01/ide@3/dad@0,0

1. c1t0d0 <STK-9200-0401 cyl 10013 alt 2 hd 15 sec 48>

/pci@1f,0/pci@1/fibre-channel@1,1/sd@0,0

2. c1t0d1 <STK-9200-0401 cyl 10013 alt 2 hd 15 sec 48>

/pci@1f,0/pci@1/fibre-channel@1,1/sd@1,0

3. c1t0d3 <STK-9200-0401 cyl 3335 alt 2 hd 15 sec 48>

/pci@1f,0/pci@1/fibre-channel@1,1/sd@2,0

Specify disk (enter its number):

8. Revisar que los dispositivos aparezcan en el archivo /kernel/drive/sd.conf

Nota: es pertinente ejecutar un comando SCSI para verificar que los dispositivos responden y construyen sus archivos de sistema, lo que debe hacerse en el servidor de la siguiente manera:

- a. # init 0

ok reset-all

ok probe-scsi-all

ok boot -r

9. Dentro de la utilería de formato etiquetar el nuevo disco detectado.

- a. Selecting device_address

[disk formatted]

- b. Disk not labeled. Label it now? "y"

FORMAT MENU:

disk

type

partition

current

format

repair

label

analyze

defect

backup

verify

save

inquiry

volname

c. format> label

d. Ready to label disk, continue? Y

10. Cuando el etiquetado se complete, salir de la utilería de formateo

a. format> quit

5.4.2. Instalación

Una vez completados los pre-requisitos, continuar de la siguiente forma:

1. Antes de proceder se deberá verificar el espacio libre: se deberá contar con al menos 15MB de espacio libre, el directorio recomendado es:

a. /opt/Storagetek®/AVCA3.1.0

2. Ingrese al sistema como “**root**”

3. Inserte el CD de instalación

4. Cambie de directorio

a. # cd cdrom_dir

5. Ejecute el archivo de instalación

a. #./install

6. Inicia la instalación, despliega mensajes relacionados con el tipo de sistema y monitor.

a. Responda si desea ver el archivo README

Please hit enter to view README file.

Use ‘f’, ‘b’ and ‘q’ to forward, rewind and quit the README file respectively...?

If you did not get to read all parts of the README file, then manually view the cdrom_dir/README_AVCA_Solaris® file.

Please hit enter to start the actual installation process ...?

7. Una vez instalado el AVCA defina el tamaño de las particiones y el tipo de formato.

Una vez completado esto el AVC está listo para utilizarse, se mostrarán una serie de sumarios con una serie de opciones por omisión.

SnapShot	Installed	High Speed Data Mover	N/A
PPRC Status	Primary	Power PPRC	Installed

	Enabled	Installed	Maximum	Expansion
Number of I/O Interfaces	18	18	32	N/A
Non-volatile Storage (MB)	-	768	768	0
Effective Cache Size (MB)	-	98304	98304	0

Avg Cache Stages per Second 0.0 **Cache Fast Write Bypass** 0
Avg Cache Hits per Stage 61.0

Subsystem Activity Information

	Total I/O (x1000)	I/O per Second	% of I/O/sec	Average				I/O Service Time (MS)		
				MB/Sec	% of MB/Sec	Access Density	Cache Hit %	Total	Disc	Conn
Read	0.1	0.2	0.1	0.0	0.0	0.0	98.4	3859.7	1272.7	2587.0
Write	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Overall	229.9	374.4	-	0.0	-	0.3	98.4	2.1	0.7	1.4

Subsystem RW Ratio: 0.0:1 100.0 % Read 0.0 % Write

Subsystem Summary by Device Type

Device	Sum	Functional Device			Access	Virtual Efficiency Factor	
Type	Qty	GB InstId	GB Alloc	% Util	Density	(compression ratio)	
SCSIB	70	502.1	N/A	3.0	0.0	3.9:1	74.5%
DATA	8	68.1	N/A	0.0	0.0	0.0:1	0.0%
SCSIA	293	711.0	N/A	1.3	0.5	2.1:1	53.2%
33903	4	11.4	0.0*	0.0	0.0	0.0:1	0.0%
3380J	13	8.2	0.0*	0.0	0.0	0.0:1	0.0%

Figura 5.7. Vista Web de la configuración del AVC

I/O Interface Information

Cluster	I/O Interface			SCSI		Interface				
	ID	Enabled	Name	Domain	Type	Speed	I/O/Sec	% Active	Avg MB/Sec	
0	A	Y	ice000	-	-	Fibre	100.0	0.0	0.0	0.0
0	C	Y	WARTHOG	-	-	Fibre	100.0	0.0	0.0	0.0
0	E	Y	Atlas	-	-	Fibre	100.0	102.9	14.7	0.0
0	G	Y	-	-	-	Fibre	100.0	0.0	0.0	0.0
0	I	Y	pprclink	-	-	Serial	20.0	0.0	0.0	0.0
0	K	Y	-	-	-	Serial	20.0	0.0	0.0	0.0
0	M	Y	-	-	-	Fibre	100.0	48.7	7.0	0.0
0	O	Y	-	-	-	Fibre	100.0	78.4	11.3	0.0
1	A	Y	-	-	-	Serial	20.0	0.0	0.0	0.0
1	B	Y	-	-	-	Serial	20.0	0.0	0.0	0.0
1	C	Y	-	-	-	Serial	20.0	0.0	0.0	0.0
1	D	Y	-	-	-	Serial	20.0	0.0	0.0	0.0
1	E	Y	-	-	-	Fibre	100.0	0.0	0.0	0.0
1	G	Y	Rodent	-	-	Fibre	100.0	0.0	0.0	0.0
1	I	Y	-	-	-	Serial	20.0	0.0	0.0	0.0
1	K	Y	-	-	-	Serial	20.0	0.0	0.0	0.0
1	M	Y	-	-	-	Fibre	100.0	100.6	12.9	0.0
1	O	Y	-	-	-	Fibre	100.0	30.2	3.9	0.0

If the Interface Percent Active is greater than 50.0 for 240 contiguous minutes (4 hours) call your local SE support representative

Figura 5.8. Vista Web (II) de la configuración del AVC

5.5. Software de respaldo

El software de respaldo elegido para el desarrollo de nuestra solución es Networker® de Legato-EMC® maestro versión 7.1.2., que será instalado en el dominio de “producción” dentro del servidor Sun FIRE E20K®. Dicho dominio se encuentra en racimo o grupo (clúster)¹ con “Producción Back” para asegurar la continuidad del servicio.

Se debe correr el programa ejecutable para UNIX (shell) de instalación desde el disco compacto (CD) que contiene el software de respaldo. Se tomarán todas las opciones por omisión, y se darán de alta los siguientes tipos de licencias:

Software de Respaldo Legato	
Software	Networker v7.1.2
Licencias	1 Licencias Networker 16 licencias de Dedicate SAN 2 licencias de drives 1 Licencias de Cluster 2 Licencias de Agente SQL Server 2 Licencias de Agente Exchange
Número de Clientes en SAN	7

Tabla 5.1.

Licencias de Networker®

¹ Racimo o grupo (Cluster): par de servidores en que en caso de caída de uno el otro asume sus funciones.

Grupos de Cintas (Pools)²

Para designar grupos de cartuchos de los cuales disponen los diversos grupos de respaldos; se determinan los pools. Estos van creciendo en número de cartuchos de acuerdo a la demanda de espacio que se generan con los respaldos a los cuales son asignados.

Para tal tarea se conformarán los siguientes pools que por lo antes comentado, no cuentan con un número fijo de cartuchos, y conforme vayan siendo requeridos serán solicitados por el sistema.

Así como un Pool llamado estándar o “**Default**” el cual es creado por el sistema y es donde se respalda el Servidor Networker y otro llamado “**Índices**” en el cual se guarda el directorio del contenido de los cartuchos y forma parte vital de la operación del software:

- Índices.
- Default.
- Exchange.
- SQL.
- Informix.
- Oracle.

Una vez que hemos designado que tipos de servicios, pools y dispositivos tenemos, podemos arrancar la ventana del administrador de Networker®, que es de apariencia muy amigable, donde podemos explorar los equipos que tienen conexión al servidor y seleccionar los archivos a respaldar de cada uno.

La primera vez que nos conectemos a cada equipo cliente pedirá un usuario y contraseña con permisos de lectura y escritura para los archivos que deseemos trabajar, y en ocasiones sucesivas no será necesario a menos que por seguridad se fuerce dicha característica en el menú de opciones.

² Literalmente significa piscina o billar en español.

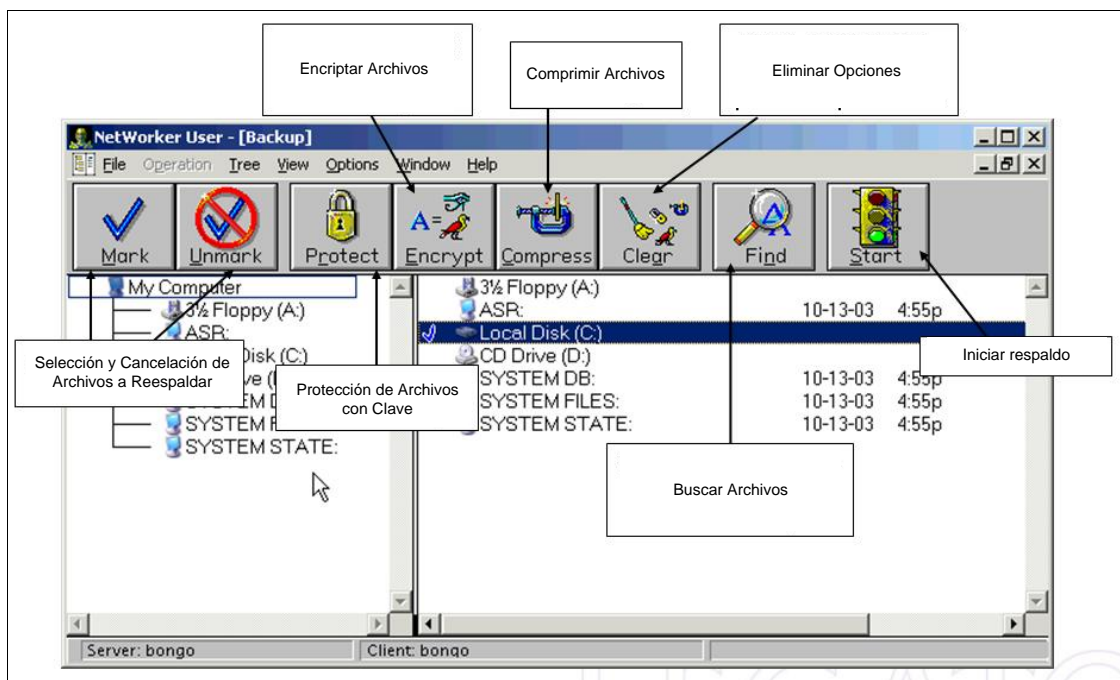


Figura 5.9.

Administrador de Networker®

Como es de esperarse, es deseable programar respaldos que periódicamente habrán de repetirse, para ello una vez que hemos creado y nombrado los respaldos podemos seleccionar en un calendario cada cuando habrá de repetirse y si el respaldo será total o incremental, de tal forma podemos garantizar que el espacio en los dispositivos de almacenamiento habrá de utilizarse de forma efectiva.

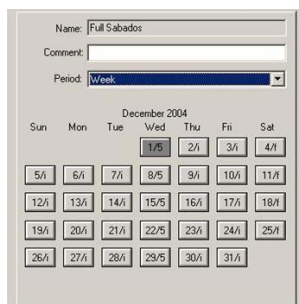


Figura 5.10.

Calendario de respaldos

Glosario

AIT: Advanced Intelligent Tape (Cinta inteligente avanzada) tecnología de 8 mm desarrollada por Sony que incluye una memoria EEPROM en el cartucho e indica la localización de los archivos, mejorando el tiempo de acceso a los datos. Utiliza cinta de metal evaporado.

Arreglo: grupo de dispositivos de almacenamiento controlados de tal manera que proveen mayores tasas de transferencia de datos, disponibilidad, o ambas. El arreglo puede ser tratado como uno a varios volúmenes por el sistema operativo dependiendo como se configure el subsistema.

ASIC: Application Specific Integrated Circuits (Circuitos Integrados de Aplicación Específica) circuito integrado para un solo propósito

Disponibilidad: La probabilidad que un sistema este disponible en un instante dado. En el contexto de un servidor, esta función expresa el porcentaje de tiempo que un sistema esta disponible para responder a peticiones del cliente

Bloque: Una cantidad de datos transferidos como una unidad en una sola I/O y es la menor cantidad de datos que pueden ser leídos o escritos a la vez.

Bus: Una colección de líneas de señal no interrumpidas que interconectan módulos de la computadora. Es una ruta de señal compartida por varios dispositivos.

Byte: Unidad de información digital que comprende ocho bits.

Canal: Término general para una ruta en la cual en la cual viajan las señales electrónicas.

CKD: Count, Key, Data (Cuenta, Entra, Datos) el nombre para el método para grabar físicamente información en un disco mainframe de IBM.

DAS: Direct Access Storage (Almacenamiento de Acceso Directo) nombre para los dispositivos de almacenamiento directamente conectado al servidor vía un canal I/O como SCSI o fibra óptica.

DAT: Digital Audio Tape (Cinta de Audio Digital) tecnología de grabación digital que usa un método helicoidal diagonal al filo de la cinta. La media es similar a los de audio.

Data Mining: extracción de información valiosa mediante algoritmos automatizados para descubrir relaciones escondidas y correlaciones.

Data Warehouse: una clasificación sistemática y comprensiva para coleccionar, filtrar y organizar información, así como hacerla disponible cuando sea necesario.

DLT: Digital Linear Tape (Cinta Digital Linear) método de grabación en cinta originalmente desarrollado por Quantum, actualmente modificado por diversas compañías.

DMA: Direct Memory Access (Acceso Directo a Memoria) método de manejar la transferencia de datos desde la memoria a un dispositivo periférico sin pasar por la unidad central de procesamiento.

DTDS: Disaster Tolerant Disk Systems (Sistemas de Discos Tolerantes a Desastres) nueva clasificación del buró consultivo para sistemas de almacenamiento RAID, que reemplaza las designaciones numéricas originales, incluye aquellas con el nivel extra “+”

DRAM: Dynamic Random Access Memory (Memoria Dinámica de Acceso Aleatorio) memoria volátil en circuito integrado que requiere de ciclos electrónicos de refrescamiento para preservar la información almacenada.

ESCON: Enterprise Systems Connectivity (Conectividad de Sistemas Empresariales)

FICON: Fibre Channel Connectivity. (Conectividad de Fibra Óptica) implementación de fibra óptica basada en los estándares de IBM, ofrece mejor desempeño que ESCON.

Firmware: categoría de chips electrónicos que mantienen su información sin alimentación electrónica.

HBA: Host Bus Adapter (Adaptador del Bus del Servidor) interfase que convierte el bus interno del sistema, tal como PCI o S-bus, a otro tipo de interfase como SCSI o FC

HSM: Hierarchical Storage Management (Administración Jerárquica del Almacenamiento) liberado por IBM en 1975 para computadoras mainframe y consiste en automáticamente almacenar la información en el dispositivo de menor costo siempre y cuando tenga el desempeño requerido por la aplicación.

LTO: Linear Tape-Open (Cinta Abierta Linear) tecnología abierta de cinta de media pulgada desarrollada por IBM, HP y Seagate. Consiste de dos formatos diferentes, ultrium y accelis.

LUN: Logical Unit Number (Número de Unidad Lógica) identificador de volumen de almacenamiento utilizado en unís y NT.

LVD: Low Voltage Differential (Diferencial de Bajo Voltaje) tecnología de señalización SCSI diseñada para aumentar la tasa de transferencia de datos.

MAID: Massive Array of Inactive Disks (Arreglo Masivo de Discos Inactivos) arreglo de discos de bajo costo que tiene en un determinado momento solo una parte de sus discos activos.

Magneto-resistivo: tecnología de grabación basada en la variación de la resistencia a la electricidad cuando un objeto entra en contacto con un campo magnético.

Conclusiones

Una de las principales conclusiones vertidas por esta tesis es que la virtualización de almacenamiento aumentará en las empresas que tengan centros donde apliquen la tecnología de la información. Entre las principales razones para virtualizar el almacenamiento figuran la reducción de costes, el ahorro de recursos humanos y un crecimiento del volumen de datos.

Esto se basa en la necesidad de diversificación de almacenamiento e intención de planificar el acopio basado en archivos como claves para implementar esta modalidad de virtualización. Partiendo de tecnologías base como los medios almacenamiento magnéticos y las diversas plataformas de sistemas operativos disponibles en el mercado. Falta fundamentar el ciclo de vida al cual esta destinada la información almacenada para ubicar la virtualización de almacenamiento, como una propuesta viable, como lo fue en este caso propuesto y trabajado en la tesis.

Otra conclusión es que el trabajo de tesis cumple uno de los objetivos establecidos en el capítulo 1, destacando la reducción de costes y disminución de recursos de tiempo asociados a la gestión de almacenamiento de datos; y como se instruye en el capítulo 2, se logra tener un recurso centralizado de información que otorga una administración viable a los datos y de implementación basada en recursos disponibles, siendo fundamental el uso del arreglo virtual compartido para cerrar el ciclo de vida la información propuesto.

La tecnología de AVC mostrada en el capítulo 4, permite el uso de menos unidades de disco duro que en un arreglo de discos tradicionales, haciendo el que la capacidad total de almacenamiento sea eficiente. En un arreglo de disco tradicional o medio magnético cualquiera como los estipulados en el capítulo 3, con frecuencia grandes cantidades de almacenamiento quedan sin uso; se dejan sin uso debido a cálculos exageradamente optimistas sobre el espacio necesario, concluyendo en este aspecto que una cantidad importante de almacenamiento está completamente inutilizada.

Como se visualiza en el capítulo 5, con el arreglo virtual compartido, se trata de un grupo común de almacenamiento, se puede simplemente reducir. Se puede tomar la capacidad que está sin uso en una aplicación y aprovecharla en otra aplicación que la necesite de acuerdo al ciclo de vida de la información, ya que utiliza su capacidad disponible casi al 100 por ciento después de ser sometido a un arreglo de protección de información.

Existe un adelanto en las empresas que usan arreglos de discos virtualizados, al descubrir la facilidad que representa administrar este tipo de arreglos de discos. Para que sea eficiente, un arreglo de disco virtualizado necesita software capaz de ayudar a administrarlo, y el manejo y implementación de este como se estipulo dentro del mismo capítulo 5.

Los sistemas de almacenamiento virtualizado le permiten ver y administrar todas sus unidades como si fueran un dispositivo grande con una utilidad de software que permite adecuar el uso del AVC en la mejor manera que se disponga.

Y si estas necesidades cambian en cualquier momento, el arreglo virtual compartido puede cambiar también.

En la integración de una solución de respaldo de información digital con un arreglo virtual compartido, AVC; se encontró que las mejoras globales más importantes son rendimiento/utilización, seguridad y automatización. También reveló que los principales beneficios que se han producido como resultado de las iniciativas de virtualización de almacenamiento son una provisión sencilla de hardware y el despliegue del software, un desarrollo más flexible y probado de entornos, así como la optimización en el flujo y ciclo de vida de la información.

El uso de arreglo virtual compartido en la infraestructura de tecnologías de la información, direcciona uno de los problemas actuales mas crecientes en los centros de datos de hoy. Este resuelve la dependencia entre las numerosas capas de tecnologías, creando transparencia, disponibilidad, seguridad y flexibilidad en el respaldo de información digital.

Bibliografía

Floyd. Thomas L, Fundamentos de sistemas digitales, 7a.ed., Madrid, España, Ed. Pearson Education, 2000, 1056 pp.

Kaufmann. Morgan, Data Mining. Concepts and Techniques, 1a.ed., Sn. Francisco, USA, Ed. Morgan Kaufmann Publishers, 2001, 550 pp.

Moore. Fred, New Game New Rules, 1ª.ed., Boulder U.S.A., Ed. Horizon Information Strategies, 2003, 84 pp.

Global Learning Solutions, Documents on CD, 26ª.ed., Louisville U.S.A., Ed. Storagetek Literature Distribution, 2005, 2386 pp.

Sun Learning Network, Special Edition CEC 2008 SUN BLUEPRINTS. Documents on CD Special Edition. Sun Microsystems, Inc. Noviembre, 2008.

Direcciones en internet:

<http://www.storagetek.com>

<http://support.storagetek.com/nr/system/access/ManualLogin.asp>

<http://www.snia.org/home>

<http://hsi.web.cern.ch/HSI/fcs/spec/overview.htm#b2>

<http://www.pcguide.com/ref/hdd/geom/tracks.htm>

<http://hsi.web.cern.ch/HSI/fcs/spec/overview.htm>

<http://www.seagate.com>

<http://www.emulex.com>

<http://www.hp.com>

<http://www.ibm.com>

<http://www.emc.com>

<http://www.qlogic.com>

<http://www.sun.com>

<http://www.brocade.com>

http://www1.us.dell.com/content/products/productdetails.aspx/pedqe_2800?c=us&cs=04&l=en&s=bsd

<http://www.redbooks.ibm.com/>

<http://www.monografias.com/trabajos13/fibropt/fibropt.shtml>

<http://www.fujitsu.com/downloads/EU/es/soporte/discosduros/UnpaseoporFibreChannel.pdf>

<http://www.enterate.unam.mx/Articulos/abril/fibraoptica.htm>

<http://www.fibra-optica.org/>

http://www.sun.com/servers/highend/sunfire_e20k/index.jsp

http://www.dothill.com/tutorial/html_tutorial/topic01.htm

<http://www-03.ibm.com/servers/storage/san/>

<http://www.microsoft.com/windowsserver2003/techinfo/overview/san.msp>

<http://www.networkworld.com/topics/sans.html>