



Universidad Nacional Autónoma de México
ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES ARAGON

**«Las Redes de Almacenamiento Inteligente como
Respuesta a las Demandas de Almacenamiento de
Datos de la Nueva Economía de Red»**

TESIS PROFESIONAL
QUE PARA OBTENER EL TITULO DE:
INGENIERO EN COMPUTACION
P R E S E N T A :
Sadi Israel Marroquín Román



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Agradecimientos

Agradezco primeramente al creador por haberme permitido lograr todos los objetivos que me he propuesto, por la vida y por el honor que me concedió de formar parte de la familia a la que pertenezco.

Quiero agradecer también a las personas que han dado forma e influido profundamente en mi vida con su amor, con su ejemplo, con su apoyo incondicional, en una palabra con su existencia:

A mi mamá

Por su tenacidad, cariño y por su sentido de responsabilidad y trabajo

A mi papá

Por su alegría, cariño, amor a la lectura y consejos

A Memo

Por su coraje, cariño, determinación y por su compañía a lo largo de todo este tiempo

A Paco

Por su audacia, cariño, sensibilidad y por su compañía a lo largo de todo este tiempo.

A Nono

Por su ternura, cariño, sencillez y sentido del humor

A Daniel

Por ser la nueva estrella que ilumina mi corazón

Un agradecimiento muy especial a mi corrector de estilo, Guillermo Marroquín Pérez

Indice

Lista de figuras	V
Lista de tablas	IX
Introducción	1
Capítulo I, Redes de Almacenamiento Inteligente	3
1.1 Ventajas de las Redes de Almacenamiento Inteligente (ISN's)	4
1.1.1 Protección de la información	4
1.1.2 Uso más eficiente de los recursos	8
1.1.3 Escalabilidad y manejabilidad	9
1.1.4 Alta disponibilidad	10
1.1.5 Capacidad de recuperación contra desastres	11
Capítulo II, Evolución de las tecnologías de almacenamiento y respaldo	13
2.1 El canal de Entrada/Salida (I/O)	13
2.2 Controladores de I/O	16
2.3 Topologías de conexión	20
2.4 El disco duro	22
2.5 La cinta magnética	24
2.6 Soluciones de cinta automatizadas	31
2.7 El disco óptico	32
Capítulo III, Determinando la mejor solución de almacenamiento	33
3.1 Desempeño	34
3.1.1 Carga de trabajo	34
3.1.2 Configuración y características de los discos	35
3.1.3 Características de las aplicaciones	44
3.1.4 Seleccionando un sistema de almacenamiento	45
3.2 Capacidad	46
3.3 Costo	47
3.4 Disponibilidad	47
3.5 Soporte	48
3.6 Ambiente de operación	49
3.7 Soluciones de respaldo	50
3.7.1 Identificando el ambiente a respaldar	50
3.7.2 Planeando una solución de respaldo	54
3.8 Crecimiento futuro	57



Capítulo 4, Redes y almacenamiento	59
4.1 La red	59
4.2 Modelos de almacenamiento	61
4.2.1 Almacenamiento conectado directamente al servidor	62
4.2.2 Almacenamiento conectado a la red	64
4.2.3 Redes de almacenamiento inteligente	67
Capítulo 5, El canal de fibra	71
5.1 Las capas del canal de fibra	71
5.1.1 Capa física	72
5.1.2 Capa de transmisión	75
5.1.3 Capa de entrega de datos	76
5.1.4 Capa de interfase con protocolos de alto nivel	79
5.2 Nombres y direcciones	79
5.3 Puertos	81
5.4 Topologías	84
5.4.1 Punto a punto	84
5.4.2 Ciclo arbitreado	85
5.4.3 Fábrica	91
Capítulo 6, Diseño, Implementación y Administración de ISN's	99
6.1 Objetivos técnicos	100
6.2 Definiendo la infraestructura de una ISN	102
6.2.1 Características de tráfico de las aplicaciones	103
6.2.2 Plataformas y almacenamiento	103
6.2.3 Concentradores de conexiones	104
6.2.4 Concentradores de conexiones inteligentes	105
6.3 Topologías de fábrica	107
6.3.1 Cascada	108
6.3.2 Malla	109
6.3.3 Anillo	111
6.3.4 Arbol	112
6.4 Disponibilidad de la información en una ISN	115
6.5 Aspectos a considerar al implementar ISN's	119
6.6 Prácticas recomendadas en ISN's	122
6.6.1 Planeando una ISN	122
6.6.2 Configurando una ISN	123
6.6.3 Escalando una ISN	124
6.7 Administración de ISN's	125



Capítulo 7, Productos y Soluciones Disponibles	131
7.1 Productos de hardware	131
7.1.1 Transceptores	132
7.1.2 Tarjetas adaptadoras	136
7.1.3 Concentradores de conexiones simples	138
7.1.4 Concentradores inteligentes	143
7.1.5 Puentes	148
7.1.6 Arreglos de discos	149
7.1.7 Multiplexores	153
7.1.8 Compuertas	155
7.2 Productos de software	156
7.2.1 Sistema operativo para fábricas	156
7.2.2 Traductor de ciclos	157
7.2.3 Zonificación	158
7.2.4 SES	159
7.2.5 Administración basada en navegador	159
Capítulo 8, Casos de Estudio	161
8.1 Aplicaciones	161
8.1.1 Respaldos	162
8.1.2 Agrupaciones de servidores	165
8.1.3 Redes de almacenamiento remotas	166
8.1.4 Replicación de datos	168
8.2 Ejemplos de Implementación	168
8.2.1 Consolidación de almacenamiento	169
8.2.2 Respaldo remoto	172
8.2.3 Copia remota	174
Conclusiones	175
Glosario	177
Bibliografía	189

Lista de figuras

Capítulo I

Figura 1-1 Modelo tradicional de respaldo y recuperación	5
Figura 1-2 Modelo de respaldo y recuperación a través de la red	5
Figura 1-3 Modelo de respaldo y recuperación basado en una ISN	6
Figura 1-4 Modelo de respaldo y recuperación basado en una ISN sin usar recursos del servidor	7
Figura 1-5 Esquema de una ISN simple en un ambiente heterogéneo	9
Figura 1-6 Esquema de agrupación de servidores a través de una ISN	11
Figura 1-7 Ambiente de recuperación contra desastres incorporando ISN's	12

Capítulo II

Figura 2-1 Conexión del canal del sistema con el canal de I/O	14
Figura 2-2 El controlador de I/O	17
Figura 2-3 Topología de canal	20
Figura 2-4 Topología de ciclo	21
Figura 2-5 Topología de fábrica	22
Figura 2-6 El disco duro	23
Figura 2-7 Unidades de disco con doble conexión	24
Figura 2-8 Cinta QIC	25
Figura 2-9 Cinta TRAVAN	26
Figura 2-10 Cinta de 4 mm	27
Figura 2-11 Cinta de 8 mm	27
Figura 2-12 Cinta DLT	28
Figura 2-13 Unidad de cinta Magstar	29
Figura 2-14 Unidad de cinta 9840	30
Figura 2-15 Cinta LTO	30
Figura 2-16 Soluciones de cinta automatizadas	31
Figura 2-17 Soluciones de respaldo basadas en la tecnología láser	32



Capítulo III

Figura 3-1 RAID 0	36
Figura 3-2 RAID 1	37
Figura 3-3 RAID 3	38
Figura 3-4 RAID 4	38
Figura 3-5 RAID 5	39
Figura 3-6 RAID 10	40
Figura 3-7 RAID 6	41
Figura 3-8 Unidad y ancho de columna	43

Capítulo IV

Figura 4-1 Modelos de almacenamiento	61
Figura 4-2 Almacenamiento conectado directamente al servidor	62
Figura 4-3 Servidor con múltiples conexiones SCSI	63
Figura 4-4 Almacenamiento conectado a la red	64
Figura 4-5 Anatomía del sistema operativo de un dispositivo NAS	65
Figura 4-6 Redes de almacenamiento inteligente	68

Capítulo V

Figura 5-1 Cables y conectores de canal de fibra	73
Figura 5-2 Formato de la trama del canal de fibra	76
Figura 5-3 Jerarquía de tramas, secuencias e intercambios	77
Figura 5-4 Nodo y puerto de nodo	79
Figura 5-5 Nombres y direcciones de puertos	80
Figura 5-6 Relación entre los puertos N y F en una red de canal de fibra	81
Figura 5-7 Proceso de comunicación entre un puerto NL y N	83
Figura 5-8 Ciclos públicos y privados	84
Figura 5-9 Topología punto a punto	84
Figura 5-10 Ciclo arbitreado	86
Figura 5-11 Ciclo arbitreado usando concentradores simples	86
Figura 5-12 Asignación de AL_PA's en un ciclo	88
Figura 5-13 Pasos en la secuencia de inicialización de un ciclo	89
Figura 5-14 Fábrica	91
Figura 5-15 Esquema de direccionamiento en una fábrica	92
Figura 5-16 Soporte de ciclo privado en un concentrador de conexiones inteligente	96
Figura 5-17 Zonificación	98



Capítulo VI

Figura 6-1 Agrupaciones de servidores usando concentradores simples	105
Figura 6-2 Fábrica en cascada	108
Figura 6-3 Fábrica en malla	110
Figura 6-4 Fábrica en anillo	111
Figura 6-5 Fábrica en árbol de tres niveles	113
Figura 6-6 Nivel 1 de disponibilidad en una fábrica	116
Figura 6-7 Nivel 2 de disponibilidad en una fábrica	116
Figura 6-8 Nivel 3 de disponibilidad en una fábrica	117
Figura 6-9 Nivel 4 de disponibilidad en una fábrica	118
Figura 6-10 Jerarquía de administración en una ISN	126
Figura 6-11 Arquitectura de administración de dispositivos	127
Figura 6-12 Administración cliente/servidor mediante HTTP	129

Capítulo VII

Figura 7-1 Convertidor de interfase a gigabit (GBIC)	133
Figura 7-2 Modulo de conexión de gigabit	134
Figura 7-3 Transceptor de pequeño factor de forma	135
Figura 7-4 Adaptador de interfase al medio	136
Figura 7-5 Tarjetas adaptadoras de canal de fibra	137
Figura 7-6 Diagrama funcional de una tarjeta adaptadora de canal de fibra	138
Figura 7-7 Concentrador de conexiones	138
Figura 7-8 Arquitectura interna de un concentrador simple	139
Figura 7-9 Concentradores inteligentes de Brocade y Qlogic	143
Figura 7-10 Uso de la topología de árbol para mejorar el desempeño	144
Figura 7-11 Puentes	148
Figura 7-12 Uso de los puentes FC-SCSI en soluciones de respaldo	148
Figura 7-13 Sistemas de discos RAID	150
Figura 7-14 Sistemas de discos JBOD	151
Figura 7-15 Dividiendo la tarjeta madre de un JBOD en ciclos separados	152
Figura 7-16 Diagrama de operación de multiplexores	153
Figura 7-17 Configuración de fábrica extendida con multiplexores	154
Figura 7-18 Configuración de fábrica extendida usando compuertas	155
Figura 7-19 Funciones y servicios ofrecidos por el sistema operativo para fábricas	157
Figura 7-20 Vista de una fábrica a partir de un navegador	160
Figura 7-21 Vista de un concentrador de conexiones inteligente	160



Capítulo VIII

Figura 8-1 Respaldo tradicional a través de una red LAN de producción	162
Figura 8-2 Transición a una implementación de respaldo sin usar la LAN	163
Figura 8-3 Implementación de un esquema de respaldos sin usar la LAN y los servidores	164
Figura 8-4 Agrupación de servidores usando ciclo arbitreado	166
Figura 8-5 Red de almacenamiento inteligente metropolitana	167
Figura 8-6 Esquema de replicación remota	168
Figura 8-7 Esquema de una solución de consolidación de almacenamiento	170
Figura 8-8 Esquema de configuración de una solución de respaldo remoto	173
Figura 8-9 Esquema de configuración de copia remota	174

Lista de tablas

Capítulo II

Tabla 2-1 Comparativo de las diferentes tecnologías de canal de I/O	16
Tabla 2-2 Comparativo de las diferentes implementaciones de SCSI	18
Tabla 2-3 Topologías de conexión de almacenamiento	20
Tabla 2-4 Capacidades de los diferentes modelos de cinta TRAVAN	25
Tabla 2-5 Características de los diferentes modelos de cinta DDS	26
Tabla 2-6 Características de los diferentes modelos de cinta de 8 mm	28
Tabla 2-7 Características de los diferentes modelos de cinta DLT	29

Capítulo III

Tabla 3-1 Recomendaciones de unidad de columna por carga de trabajo y tipo de acceso	42
Tabla 3-2 Velocidades de procesamiento y radios de compresión de diferentes tipos de archivos	52
Tabla 3-3 Capacidades y velocidades de transferencia de diferentes unidades de cinta	54
Tabla 3-4 Velocidades de transferencia de datos para diferentes tecnologías de red	55

Capítulo V

Tabla 5-1 Capas del canal de fibra	72
------------------------------------	----

Capítulo VI

Tabla 6-1 Objetivos de diseño	102
Tabla 6-2 Comportamiento de las topologías de ISN de acuerdo al tipo de acceso a los datos	114
Tabla 6-3 Número máximo de puertos y dispositivos por topología	115
Tabla 6-4 Niveles de disponibilidad en fábricas	118
Tabla 6-5 Factores de disponibilidad	119
Tabla 6-6 Zonificación por sistema operativo	124
Tabla 6-7 Zonificación por arreglo de discos	124

Capítulo VII

Tabla 7-1 Componentes de hardware encontrados en una ISN	132
Tabla 7-2 Distancias de conexión usando fibra óptica	133
Tabla 7-3 Características y aplicaciones de los GBIC's	134
Tabla 7-4 Características de tres tarjetas adaptadoras de canal de fibra	136
Tabla 7-5 Diferentes estados de un puerto en un concentrador	139
Tabla 7-6 Características de los concentradores simples	142
Tabla 7-7 Características de concentradores inteligentes de 8 y 16 puertos	146
Tabla 7-8 Características de concentradores inteligentes de 64 puertos	147
Tabla 7-9 Comparativo de sistemas RAID	151
Tabla 7-10 Comparativo de sistemas JBOD	153
Tabla 7-11 Características de un multiplexor	154
Tabla 7-12 Características de una compuerta	156
Tabla 7-13 Características de la zonificación	158
Tabla 7-14 Características principales de SES	159

Capítulo VIII

Tabla 8-1 Detalle de la configuración utilizada en un esquema de consolidación de almacenamiento remoto	171
Tabla 8-2 Niveles de fábrica extendidos	172

Introducción

Desde los inicios de la era de la computación, la capacidad de procesamiento de servidores, estaciones de trabajo y computadoras ha acaparado la atención de ingenieros, investigadores y desarrolladores, mientras que el almacenamiento y otras funciones de entrada/salida (I/O, *Input/Output*) han quedado relegadas a segundo plano. Este modelo de cómputo en donde la unidad de procesamiento es el elemento principal resolvió por mucho tiempo las necesidades de manejo de datos de las corporaciones, sin embargo, la gran explosión que ha tenido el Internet, la consolidación del esquema cliente/servidor, la proliferación de sistemas de Planeación de Recursos Empresariales (ERP, *Enterprise Resource Planning*) y de Sistemas de Soporte para la toma de Decisiones (*DataWarehousing*) así como las aplicaciones multimedia, han hecho que este paradigma esté cambiando hacia un modelo centrado en el almacenamiento.

Las demandas de almacenamiento de datos impuestas por la nueva economía de red exigen entre otras cosas gran capacidad de almacenamiento pero sobre todo escalabilidad para hacer frente a requerimientos inesperados de disco, disponibilidad total de la información, acceso a los datos desde cualquier dispositivo o plataforma de cómputo y facilidad de administración.

En el esquema tradicional de conexión de arreglos de discos, los servidores y los arreglos están unidos a través de conexiones dedicadas, en donde los datos de un disco en particular son normalmente accedidos por un solo servidor a través de un canal conocido como Interfase para Sistemas de Cómputo Pequeños (SCSI, *Small Computer System Interface*). La cantidad de datos que puede manejar ese servidor está limitado por el número de discos soportado por el canal y el número de canales soportado por el servidor, esto tiene una enorme repercusión en términos de escalabilidad aún y cuando el servidor tenga la capacidad de soportar grandes cantidades de disco, ya que cuando esta capacidad llega a su límite la única manera de agregar más almacenamiento es adquiriendo un nuevo arreglo de discos y por supuesto un nuevo servidor al cual estos discos estarán conectados, esto duplica el costo cuando la verdadera necesidad es el almacenamiento.

En términos de disponibilidad de la información, si el servidor o cualquiera de sus conexiones a disco fallan, el acceso a los datos se pierde, para ambientes de misión crítica esta posibilidad de pérdida de datos es inaceptable. Esto ha llevado a muchas empresas a invertir considerables sumas de dinero para prácticamente duplicar en algunos casos su infraestructura de cómputo con el afán de protegerse.

En cuanto al acceso a la información desde diferentes plataformas de cómputo, las soluciones multipuerto disponibles son muy caras de implementar o propietarias en su diseño, lo cual las hace difíciles de administrar, esto obliga a que se adquiera almacenamiento para cada plataforma específica.

Las Redes de Almacenamiento Inteligente (ISN's, *Intelligent Storage Networks*) constituyen la base de un nuevo paradigma de cómputo centrado en el almacenamiento, al introducir la flexibilidad de las redes como medio de conexión entre servidores y sistemas de discos, eliminando las conexiones dedicadas. Además, al estar basadas en la arquitectura de canal de fibra que proporciona un óptimo ancho de banda, rutas de datos redundantes y capacidad de grandes distancias se satisfacen una gran cantidad de requerimientos incluyendo desempeño, escalabilidad, flexibilidad de conexión y alta disponibilidad.

Sin embargo, el resolver problemas de disco usando la tecnología de redes de almacenamiento inteligente no es una tarea trivial. El diseñar redes de almacenamiento requiere un análisis cuidadoso de las aplicaciones y de la habilidad de los componentes utilizados para su implementación. Algunas características y capacidades que existen como estándares del canal de fibra pueden no existir todavía como productos reales. Los estándares por sí mismos pueden definir algunas capacidades abstractas que tengan poco que ver con implementaciones del mundo real y, desafortunadamente, dependiendo de la filosofía de mercadotecnia de los proveedores de cómputo, algunos fabricantes de componentes de ISN pueden exagerar en la capacidades de sus productos para ganar ventaja en el mercado.

El propósito de este trabajo de tesis es el de dar a conocer la tecnología de las Redes de Almacenamiento Inteligente, las soluciones disponibles hoy en día pero también las limitaciones, dando un enfoque práctico de su implementación. La tesis está dividida en tres partes:

La primera parte inicia con una breve descripción de lo que son las redes de almacenamiento inteligente y sus principales características, se explican además las necesidades de cómputo que las empresas tienen en la actualidad y se resaltan algunas de las ventajas de las ISN's, se habla también acerca de la evolución que las tecnologías de almacenamiento y respaldo han tenido, así como sus tendencias. Aquí se introducen varios conceptos generales y se explica cómo determinar la mejor estrategia de almacenamiento y respaldo basados en las premisas de desempeño, capacidad, escalabilidad, interoperabilidad y confiabilidad.

La segunda parte trata en detalle la tecnología de las Redes de Almacenamiento Inteligente y su interrelación con los conceptos de redes y canal de fibra, se explican los diferentes modelos de almacenamiento disponibles a la fecha, sus principales características, así como sus ventajas y desventajas para que puedan ser tomadas en cuenta al momento de elegir una solución de almacenamiento; finalmente, se cubren los principales elementos que se deben considerar para diseñar, implementar y administrar una solución de ISN.

La tercera y última parte, discute los productos y soluciones disponibles en el mercado y se abordan también casos prácticos de implementación con algunos de estos productos.

I

Redes de Almacenamiento Inteligente

Durante la última década, la gran cantidad de cambios que han habido en el área de la computación así como la globalización que han tenido los negocios gracias al éxito de Internet, han originado que los requerimientos de almacenamiento de datos crezcan de manera exponencial. Estos requerimientos han tenido un profundo efecto en la mayoría de los centros de datos de hoy en día. Por lo que las empresas están teniendo que hacer frente a crecimientos desmesurados en la información mientras tratan de minimizar costos. Como resultado, muchas organizaciones están buscando formas efectivas de asegurar la confiabilidad y la disponibilidad de su información sin que esto implique que se tengan que hacer gastos excesivos y al mismo tiempo que se pueda asegurar el retorno de la inversión.

Para permanecer exitoso en un medio tan dinámico, las organizaciones necesitan sistemas de almacenamiento confiables que puedan manejar y proteger su información más crítica. Estos sistemas deben ser capaces de adaptarse rápidamente a crecimientos de datos inesperados a la vez que se garanticen los mayores niveles de disponibilidad posibles, lo que resulta muy difícil de lograr con el modelo tradicional de almacenamiento. Una alternativa para lograr esto la constituyen las redes de almacenamiento inteligente.

Las redes de almacenamiento inteligente (ISN, *Intelligent Storage Networks*) son ambientes de red independientes de la red o redes de producción, en donde se conectan dispositivos de almacenamiento y cinta junto con los servidores o equipos de cómputo, formando así una red adicional destinada exclusivamente al manejo y tratamiento de la información. El objetivo principal de esta nueva red es la de proporcionar una nueva infraestructura de cómputo escalable y confiable en donde el almacenamiento no dependa más de los servidores y que a la vez pueda cumplir con los requerimientos de alta disponibilidad y desempeño que demandan las aplicaciones de hoy en día, brindando además un ambiente centralizado para manejar de una manera más efectiva los recursos disponibles y la información.

1.1 Ventajas de las ISN's

El objetivo principal de las Redes de Almacenamiento Inteligente es el de proporcionar una mejora sustancial en el manejo de la información y de los recursos de cómputo. A continuación se mencionan los principales requerimientos que se tienen en los centros de datos de las empresas y la manera en la que las ISN's pueden dar respuesta a estas necesidades.

- Protección de la información
- Uso más eficiente de los recursos
- Escalabilidad y manejabilidad
- Alta disponibilidad
- Capacidad de recuperación contra desastres

1.1.1 Protección de la información

Conforme la información se convierte en el valor más importante de una empresa, asegurar su integridad se vuelve más crítico que nunca. Una de las mejores maneras para lograr esto consiste en tener una o varias copias adicionales de la información las cuales puede ser hechas a disco o a cinta - lo que comúnmente se conoce como respaldos, entre más información se tenga y dependiendo de la infraestructura y políticas de la empresa mayor será el tiempo que se tome un respaldo y más complicado será llevarlo a cabo, en la actualidad muchas organizaciones se enfrentan al reto de tener que respaldar cada vez mayores volúmenes de información mientras que el tiempo del que se dispone para hacer esto se reduce o no existe sobre todo en el caso de las corporaciones que tienen que dar servicio las 24 horas del día, los 365 días del año.

Los modelos tradicionales de respaldo y recuperación consisten de un sistema de discos y de cintas dedicados para cada servidor, con cada equipo de cómputo respaldando su propia información. Este diseño ofrece una deficiente utilización de los recursos ya que aunque el sistema de cintas de un equipo esté disponible no puede ser usado por otro servidor. Además, cada plataforma de cómputo tiende a usar sus propias herramientas lo que complica la administración así como el proceso de respaldo y recuperación, esto sin considerar que no se puede tener una certeza de la integridad de la información respaldada ya que esto es responsabilidad de cada uno de los administradores de los diferentes equipos y generalmente es una tarea que casi siempre es relegada a segundo plano.

La figura 1-1 ejemplifica el modelo tradicional de respaldo y recuperación en un ambiente de cómputo heterogéneo.

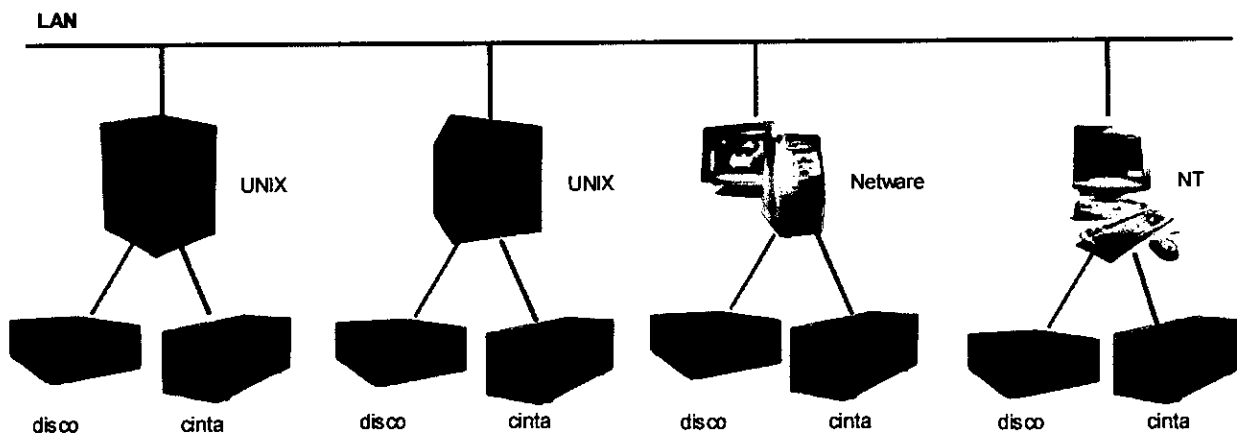


Figura 1.1 Modelo tradicional de respaldo y recuperación

Una alternativa a este esquema involucra la implementación de una solución de respaldo y recuperación basada en la red. Este modelo involucra el uso de un servidor dedicado para controlar los sistemas de cintas junto con un software que controla el proceso de respaldo y recuperación. El servidor de respaldos recibe los datos de los otros equipos a través de la red y los almacena en su propio disco o los manda directamente a cinta. Este esquema centralizado proporciona un mejor uso de los recursos a la vez que facilita la administración.

La principal desventaja de este esquema es que la red se vuelve un potencial cuello de botella convirtiéndose en el factor principal que impide cumplir con los requerimientos de tiempo para hacer el respaldo de la información. Esto sin considerar que se pueden presentar problemas de desempeño si se usa la red de producción. Es por eso que en la mayoría de las soluciones basadas en este esquema se opta por tener una red dedicada para el proceso de respaldo y recuperación, aunque esto tiene un costo asociado, el cual muchas veces no es significativo comparado con las pérdidas que se pueden tener al degradarse el desempeño de una red de producción. La figura 1-2 ilustra el modelo de respaldo y recuperación basado en la red.

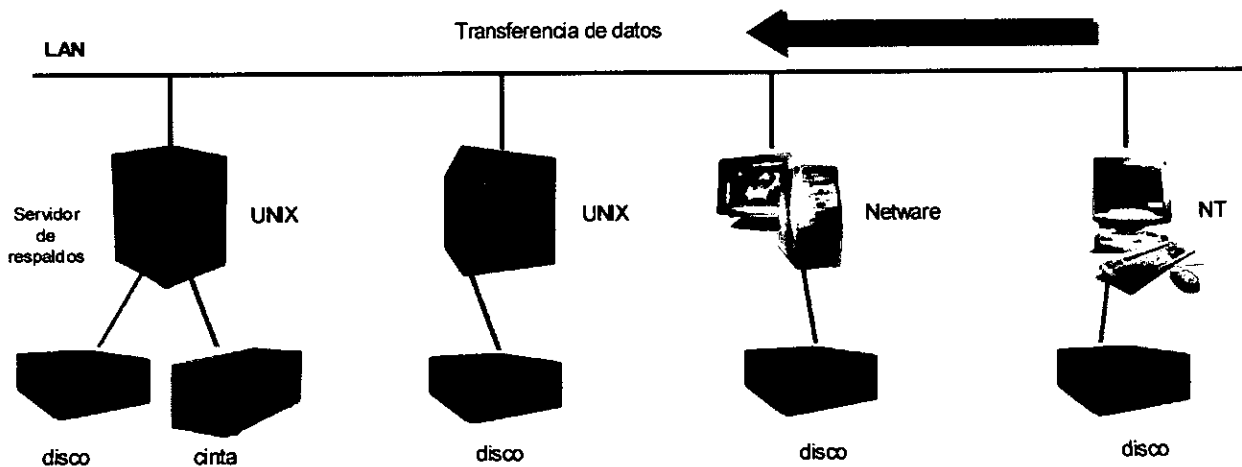


Figura 1-2. Modelo de respaldo y recuperación a través de la red

En contraste, las redes de almacenamiento inteligente pueden simplificar el proceso de respaldo y recuperación a la vez que se aseguran los tiempos disponibles para hacer esto, gracias a las altas velocidades que manejan los concentradores de conexiones inteligentes (*switches*) usados para implementar ISN's. Adicionalmente el canal de fibra que es el medio físico utilizado en las redes de almacenamiento está diseñado para transportar grandes bloques de datos de una manera más eficiente y confiable que las redes basadas en el Protocolo de Internet (IP, *Internet Protocol*).

El quitar la Red de Área Local (LAN, *Local Area Network*) del esquema de respaldo y recuperación proporciona muchas ventajas. Las unidades y bibliotecas de cintas conectadas a una red de almacenamiento pueden ser configuradas de tal forma que cada servidor envíe sus datos directamente a la cinta sin tener que pasar por la red de producción, haciendo uso del mismo *software* de respaldos utilizado en un ambiente de red basado en IP. Las ISN's permiten grandes transferencias de datos entre servidores, sistemas de almacenamiento y cintas dejando la LAN solo para propósitos de comunicación entre servidores, tal como se muestra en la figura 1-3.

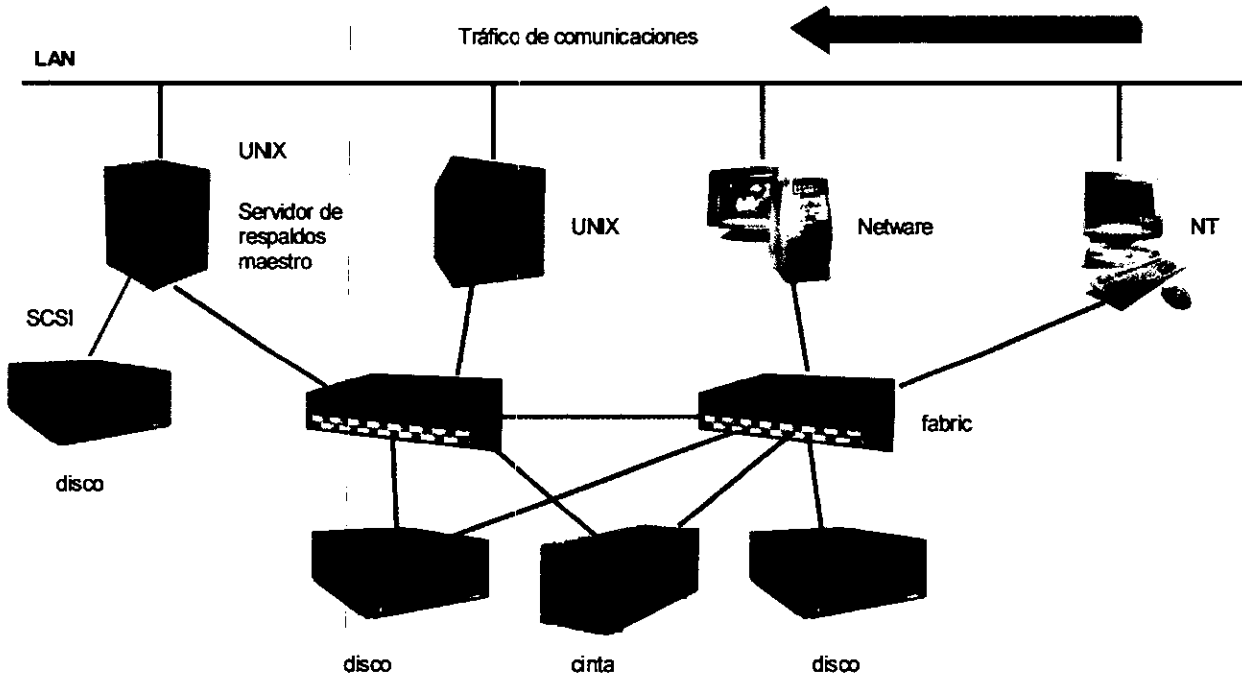


Figura 1-3. Modelo de respaldo y recuperación basado en una ISN

Como resultado se tiene una solución más rápida que no interfiere con la red de producción y en donde está garantizada una velocidad de transferencia de datos de 100 MB/s aún y cuando se vayan agregando más nodos a la configuración, la solución también es más escalable comparada con las otras dos ya que el agregar un nuevo sistema de discos o de cintas no implica que se tenga que agregar otro servidor. En este esquema se tiene un mejor aprovechamiento de los recursos ya que cualquier nodo puede hacer uso de cualquier dispositivo que esté disponible a la vez que se logra eficientar la administración al mantener un esquema centralizado.

Otra implementación que aún está en desarrollo consiste en liberar totalmente a los servidores de la tarea de hacer un respaldo, esta tecnología es conocida como *server-free backup*. En este modelo los datos son transferidos directamente entre los dispositivos de almacenamiento (discos y cintas) sin intervención de los servidores. Este proceso es realizado por una tecnología que es implementada en algún dispositivo de ISN o incluso en los mismos sistemas de almacenamiento. Este esquema permite que se reduzcan significativamente los ciclos de procesador requeridos normalmente en una operación de respaldo a la vez que el impacto a la operación es mínimo. De esta manera se mejora considerablemente el desempeño ya que se libera tanto a la red de producción como al servidor de la tarea de respaldos.

La figura 1-4 ilustra este esquema en donde todas las tareas requeridas para manejar un respaldo ahora son manejadas desde un dispositivo ISN.

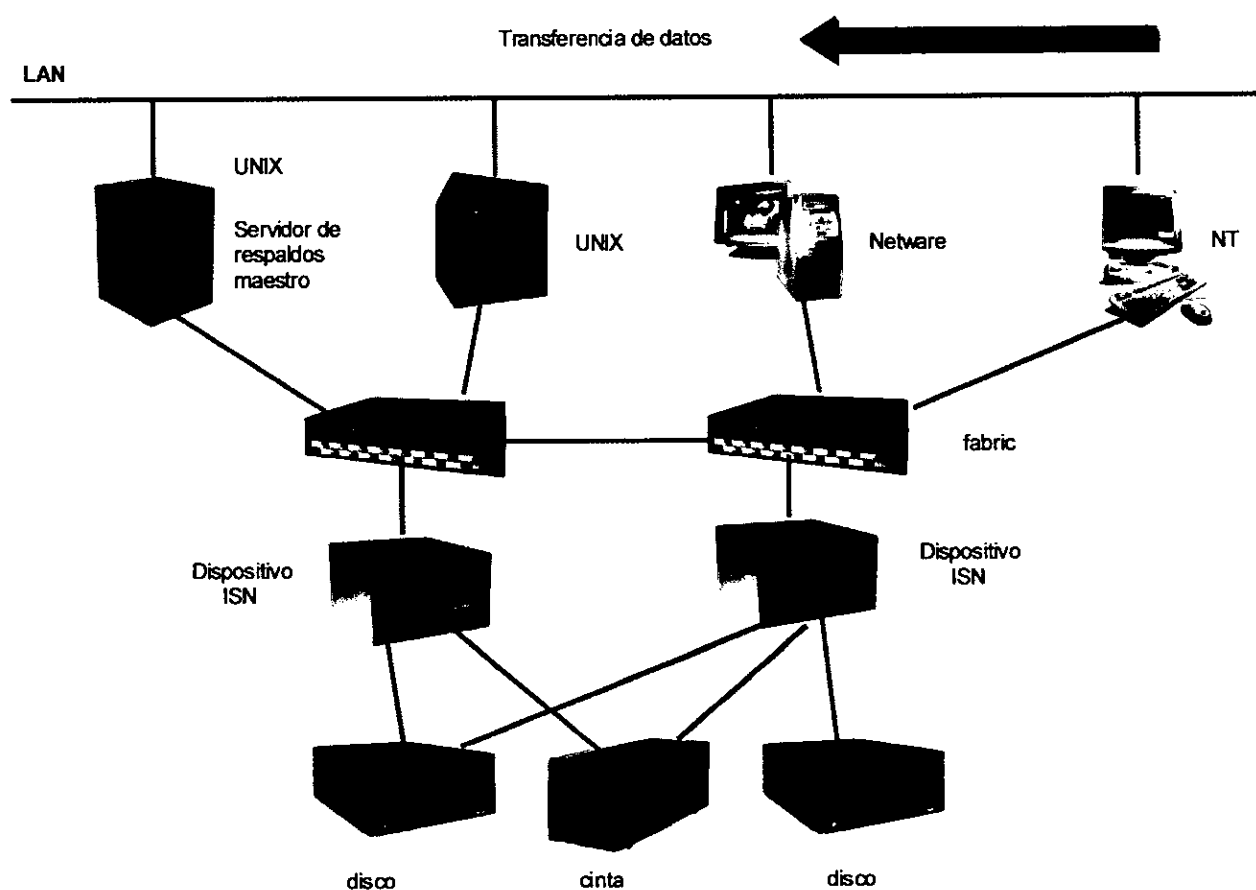


Figura 1-4. Modelo de respaldo y recuperación basado en una ISN sin usar recursos del servidor.

1.1.2 Uso más eficiente de los recursos

Tradicionalmente, las organizaciones han conectado sus sistemas de almacenamiento directamente a los servidores, esto debido principalmente a razones técnicas. Como ya se mencionó, esta implementación resulta en una pobre utilización de los recursos de almacenamiento ya que el espacio disponible en disco de un servidor no puede ser usado por otro que se esté quedando sin este recurso.

La capacidad de almacenamiento en el esquema tradicional depende del sistema de discos, pero también depende del servidor y de su capacidad de conexiones, por lo que es muy común tener que agregar otro servidor a la configuración cuando tal vez el único requerimiento tiene que ver con espacio en disco. Este esquema puede llegar a complicar la tarea de administración debido al potencial número de servidores y sistemas de almacenamiento que se pueden llegar a tener en caso de un crecimiento sustancial en los volúmenes de información.

Las redes de almacenamiento inteligente, por otro lado, proporcionan un ambiente de almacenamiento flexible en donde la forma en la que los recursos son agregados y administrados cambia considerablemente ya que no se tiene que comprar un servidor para conectar un sistema de almacenamiento cada vez que se necesita más espacio en disco.

Al permitir una conectividad de cualquier servidor a cualquier arreglo de discos las ISN's mejoran considerablemente la compartición de recursos. Este esquema de sistemas abiertos permite la selección del mejor equipamiento en términos de servidores y arreglos sin importar la marca del fabricante. Es importante mencionar, sin embargo, que en la actualidad una implementación de redes de almacenamiento requiere que se utilicen siempre concentradores de conexiones inteligentes de un mismo proveedor ya que hasta el momento no se han terminado de definir los estándares que permitan una comunicación adecuada entre concentradores de conexiones inteligentes de diferentes proveedores.

La figura 1-5 muestra la manera en la que varias plataformas de cómputo pueden compartir recursos de almacenamiento y cinta a través de una red de almacenamiento inteligente.

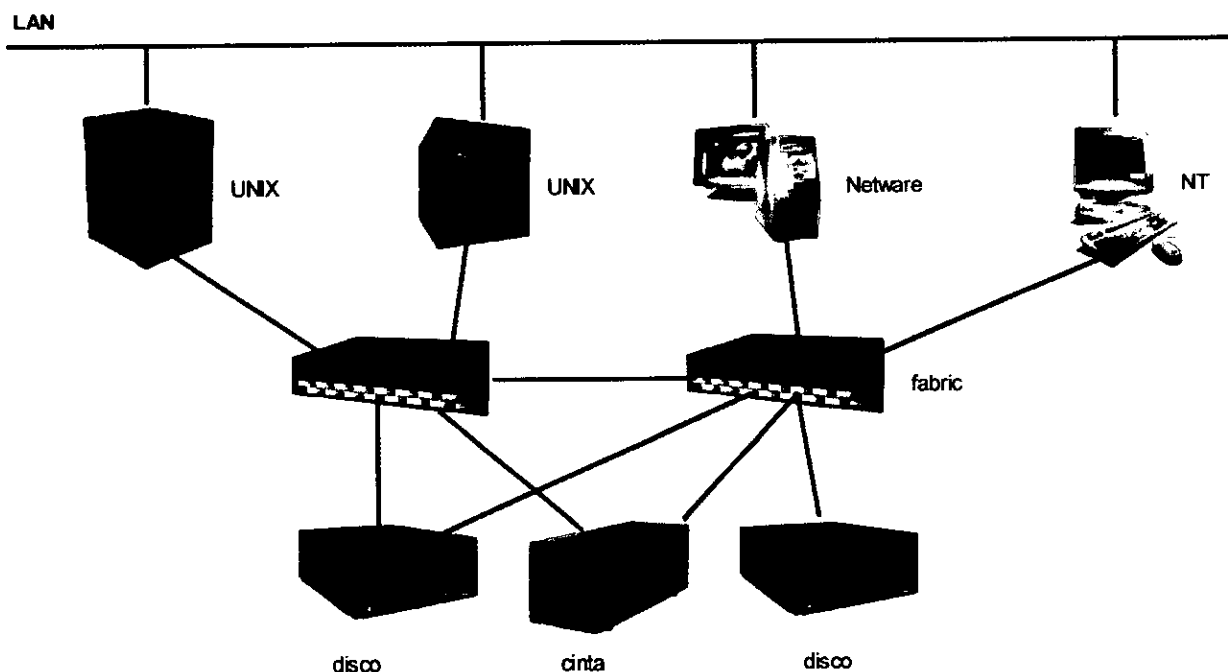


Figura 1-5. Esquema de una ISN simple en un ambiente heterogéneo

1.1.3 Escalabilidad y Manejabilidad

Las redes de almacenamiento inteligente también proporcionan una forma de crecer la infraestructura de cómputo en términos de servidores, sistemas de almacenamiento y cintas conforme el número de usuarios crezca, sin necesidad de que haya un impacto en el desempeño. Este modelo permite agregar nuevos recursos sin interrumpir las operaciones en el resto del sistema, de esta manera, una empresa puede comenzar con una configuración pequeña y agregar nuevos servidores con o sin almacenamiento asociado, lo único que se necesita hacer es agregar más concentradores de conexiones inteligentes conforme el número de nodos en la red se incremente. Para simplificar la administración, las ISN's automáticamente aprenden y se ajustan a las nuevas topologías que se formen como resultado de la adición de nuevos recursos.

Las ISN's también mejoran la administración de los recursos de almacenamiento a través de la centralización. Por ejemplo, las herramientas de *software* de las ISN's permiten a las corporaciones organizar el espacio en disco, replicar datos, respaldarlos y monitorear todos los recursos de manera continua. Estas capacidades son críticas ya que por un lado permiten garantizar óptimos niveles de operación y administración mientras permiten prevenir y detectar fallas, corregirlas e identificar nuevas necesidades.

1.1.4 Alta disponibilidad

Debido a contratos de servicio, regulaciones de la industria u otros requerimientos las organizaciones de hoy en día demandan los mayores niveles de disponibilidad posibles en sus sistemas. De hecho, incidentes en el manejo de información que anteriormente eran vistos como problemas menores ahora pueden tener un gran impacto en el negocio. Para dar respuesta a estos requerimientos de alta disponibilidad, las redes de almacenamiento inteligente están diseñadas para facilitar un ambiente altamente confiable que prevenga o pueda minimizar el impacto en el negocio que pueda tener alguna falla en la infraestructura de cómputo que soporta la operación.

Algunos de los beneficios clave de las ISN's en términos de alta disponibilidad incluyen la redundancia interconstruida de los componentes usados en su implementación, las capacidades de ruteo automático de tráfico y las características de transferencia de la operación de un dispositivo con falla a otro que esté disponible (*failover*).

Las diferentes opciones de conectividad existentes permiten que se puedan desarrollar redes de almacenamiento inteligentes sin un solo punto de falla mientras que un componente de *software* puede ser usado para detectar automáticamente problemas en la red y rutear el tráfico alrededor de una conexión fallida para asegurar así una ruta confiable para los datos.

Adicionalmente los componentes utilizados en la implementación de ISN's proporcionan capacidades de reemplazo en caliente (*hot swap*) que permiten la instalación, configuración y operación de dispositivos sin comprometer la operación. Combinando múltiples concentradores de conexiones inteligentes y el *software* apropiado es posible construir redes de almacenamiento inteligentes altamente escalables con los mayores niveles de disponibilidad.

Las ISN's también pueden mejorar sus características de alta disponibilidad mediante el soporte de agrupaciones de servidores (*clusters*). Las agrupaciones de servidores son usadas para asegurar que las aplicaciones puedan continuar corriendo en el evento de que un servidor falle. Los ambientes tradicionales de agrupaciones de servidores involucran dos servidores compartiendo un sistema de discos. Si uno de los servidores falla, el otro toma la carga del nodo que falló y de esta manera la aplicación puede continuar corriendo. El servidor que queda disponible accesa los datos a través del disco compartido. Esto representa un diseño poco flexible ya que generalmente está limitado a dos servidores compartiendo un sistema de discos, uno de los cuáles permanece sin hacer nada hasta que es requerido en caso de que el otro servidor falle. Adicionalmente, los servidores y los dispositivos de almacenamiento se encuentran generalmente cerca uno del otro, lo cual solo da una protección limitada en el caso de desastres.

La figura 1-6 muestra como se puede implementar una solución de agrupación de servidores con varios nodos utilizando una red de almacenamiento inteligente.

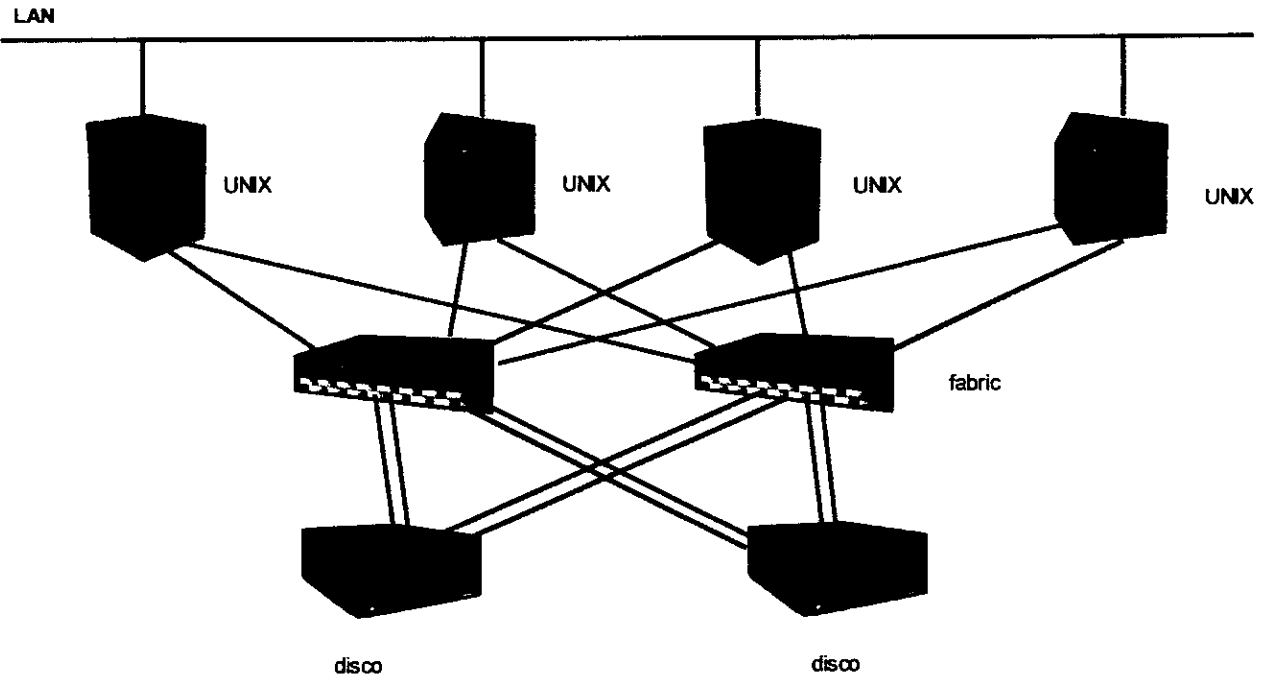


Figura 1-6. Esquema de agrupación de servidores a través de una ISN

Con un esquema de agrupación de servidores basado en ISN, muchos servidores pueden compartir el almacenamiento. De hecho, en algunas implementaciones cualquier servidor puede manejar la carga adicional cuando uno de los servidores se encuentra indisponible asegurando de esta manera que no existan recursos desperdiciados. Las redes de almacenamiento inteligentes también permiten el soporte de conexiones a grandes distancias -en el orden de decenas de kilómetros, lo que permite implementar efectivos programas para prevenir desastres.

1.1.5 Capacidad de recuperación contra desastres

Un factor crítico en cualquier esquema de cómputo que pueda asegurar los mayores niveles de operación es la habilidad de protección y recuperación de datos en caso de un desastre. Las ISN's proporcionan la base para soportar ambientes en donde los datos deben ser transferidos automáticamente a otro sitio remoto si se presenta una contingencia. Como resultado, un esquema de protección contra desastres basado en redes de almacenamiento inteligentes puede proporcionar un modelo más eficiente de respaldo y recuperación así como una conectividad extendida y una mayor tolerancia a fallas tal como se puede apreciar en la figura 1-7.

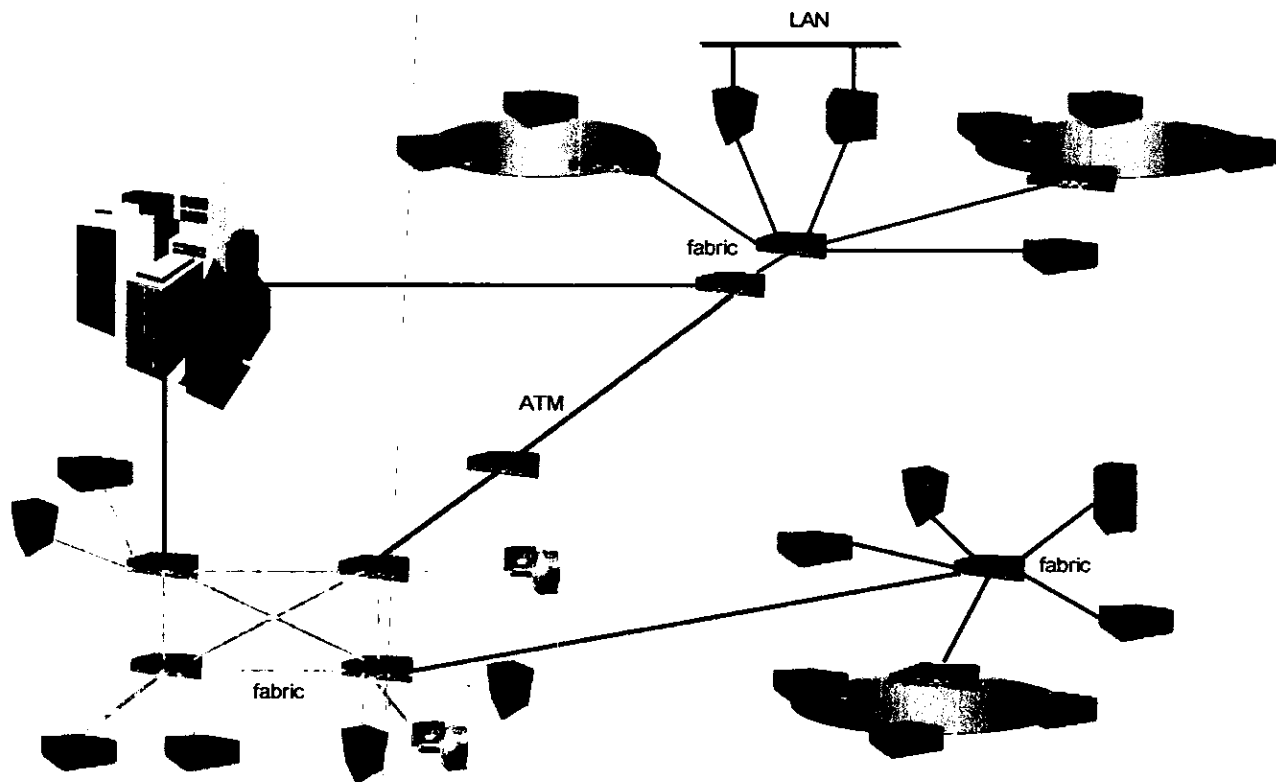


Figura 1-7. Ambiente de recuperación contra desastres incorporando ISN's

Una solución de recuperación contra desastres depende fundamentalmente de los requerimientos específicos del negocio en términos de que tanto tiempo pueden estar los servicios fuera de línea. Obviamente muchas corporaciones no pueden permitirse minutos de inactividad sin sufrir serias consecuencias. Algunos otros pueden tolerar incluso días de no operación. Por lo tanto, un plan de recuperación de desastres necesita ser desarrollado acorde a los requerimientos de disponibilidad, a la infraestructura de cómputo que se tenga así como a las políticas de negocio de las empresas.

Las ISN's proveen un componente clave en una solución de recuperación contra desastres protegiendo la información sobre grandes distancias o permitiendo configuraciones en espejo en sitios remotos. La tecnología de canal de fibra proporciona una conectividad garantizada en distancias de 10 o más kilómetros, lo que es clave para este tipo de implementaciones.

II

Evolución de las tecnologías de almacenamiento y respaldo

El almacenamiento de datos es un término que abarca un amplio rango de componentes de *hardware*, desde el canal del sistema, el canal de I/O y los controladores de entrada y salida (I/O) hasta los medios de almacenamiento como son la memoria de acceso aleatorio (RAM, *Random Access Memory*), la unidad de disco duro, la cinta magnética, el disco magneto óptico, el disco compacto (CD-ROM, *Compact Disk - Read Only Memory*) y el disco de video digital (DVD, *Digital Video Disk*). Cada componente tiene sus propias características, capacidades y limitantes, las cuales necesitan ser entendidas cuando se está diseñando una arquitectura de almacenamiento.

2.1 El canal de I/O

El canal de I/O es el encargado de mover la información desde el canal del sistema a los dispositivos periféricos y viceversa. Adicionalmente, se encarga de transportar datos entre las diferentes tarjetas conectadas a éste. Entre el canal de I/O y el canal del sistema se encuentra un dispositivo electrónico llamado puente (*bridge*) que se encarga de manejar el intercambio de datos entre estos dos canales. La figura 2-1 muestra la interconexión del canal del sistema y el canal de I/O.

Si el procesador de una computadora se pudiera comparar al motor de un automóvil, entonces el canal de I/O sería análogo a la transmisión de un carro, el cual transfiere el poder del motor a los neumáticos y hace que el coche se mueva. Un sistema que no pueda manejar eficientemente las transferencias de datos entre canales no podrá proporcionar un desempeño adecuado sobre todo cuando la carga de trabajo es mucha.

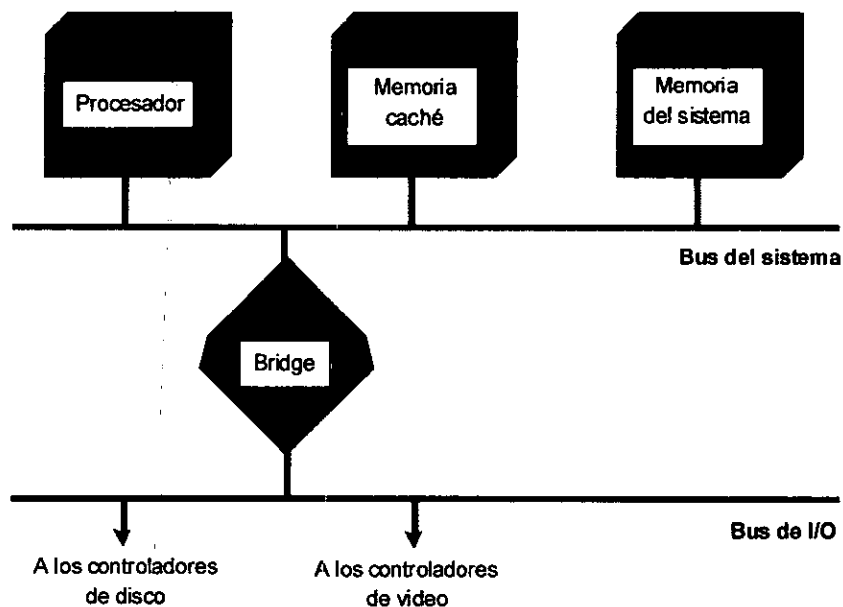


Figura 2-1. Conexión del canal del sistema con el canal de I/O

En la actualidad existen diferentes tecnologías de canal de I/O, algunas con mayor aceptación que otras, a continuación se mencionan las principales.

- PCI
- c-PCI
- VESA
- ISA
- Microcanal
- EISA
- S-Bus

PCI

El canal de Interconexión de Componentes Periféricos (PCI, *Peripheral Component Interconnect*) se ha convertido en el estándar de la industria debido a su bajo costo, desempeño y utilidad para una amplia gama de aplicaciones incluyendo almacenamiento. Con el inicio de la familia de procesadores 486, Intel comenzó a integrar en su arquitectura la tecnología PCI permitiendo al mismo tiempo que otras tecnologías existentes se integraran a este nuevo diseño. Posteriormente otras compañías comenzaron a participar en el desarrollo de PCI a tal grado que hoy en día compañías como IBM, Sun Microsystems y Compaq la han adoptado a pesar de que tenían sus propias implementaciones de canal de I/O.

c-PCI

El canal compacto PCI es un derivado del canal estándar PCI pero enfocado fundamentalmente para aplicaciones de tipo industrial y telecomunicaciones. Las computadoras en estos ambientes típicamente requieren de una alta resistencia y tolerancia a fallas debido a las condiciones en las que tienen que operar. Una de las características más interesantes de c-PCI es que fue diseñada para soportar la conexión de tarjetas sin tener que dar de baja o apagar el equipo.

VESA

El canal VESA que significa canal de la Asociación de Estándares Electrónicos para Video (*Video Electronic Standard Asociation*) fue el predecesor de PCI en el mercado, este proporcionaba mucho mejor desempeño que los anteriores canales de I/O basados en PC. Algunas de sus principales ventajas en desempeño vinieron de tener una interfase de I/O directamente al procesador lo que resultó en transferencias más rápidas de datos, sin embargo, habían restricciones sustanciales en el número de tarjetas soportadas lo que impidió que pudiera competir contra PCI.

ISA

El canal conocido como Arquitectura Estándar de la Industria (*ISA, Industry Standard Architecture*) nació con las PC's, fue construido para transferir datos entre dispositivos y procesadores relativamente lentos. Con la aparición de la IBM PC-AT, este canal fue mejorado a 16 bits de datos y a una velocidad máxima de 8 MHz, sin embargo, conforme se incrementaron las velocidades de los procesadores y el número de aplicaciones creció, este canal se volvió obsoleto.

Microcanal

En 1987, IBM anunció su arquitectura de microcanal que fue diseñada para operar a velocidades de reloj mucho mayores a las que habían en aquella época, desafortunadamente esta tecnología introdujo un nuevo conector físico que lo hizo incompatible con las tarjetas y computadoras basadas en ISA de aquel entonces lo que ocasionó que nunca llegara a tener la popularidad que IBM esperaba. Con el tiempo esta tecnología se llegó a usar en sistemas Unix sobre todo de IBM pero hoy en día es prácticamente una tecnología muerta.

EISA

EISA significa ISA mejorado (*enhanced ISA*) y fue desarrollado por Compaq pero implementado por varias compañías de PC's como una alternativa a la arquitectura de microcanal de IBM. EISA permitió que tanto las tarjetas ISA como EISA fueran instaladas en el mismo sistema compartiendo la misma arquitectura de *bus* lo que la colocó en el mercado como una tecnología ampliamente aceptada. Sin embargo, con la aparición de PCI que también ofrece compatibilidad con las tarjetas ISA y EISA pero con mucho mejor desempeño, EISA está tendiendo a desaparecer.

S-Bus

S-Bus es un canal de I/O usado en productos de Sun Microsystems. Cuando apareció, S-Bus proporcionaba mejor desempeño que los canales basados en PC's incluyendo las primeras implementaciones de PCI, sin embargo, no llegó a alcanzar la popularidad necesaria para volverse un estándar en la industria y fue alcanzado por PCI en términos de desempeño. En la actualidad Sun está cambiando a PCI porque le resulta más barato de implementar y por sus ventajas en desempeño.

La tabla 2-1 muestra las principales características de las diferentes tecnologías de canal de I/O.

	ISA	EISA	MCA	VESA	PCI 2.0	c-PCI	PCI 2.1
Velocidad de reloj	8.3	8.33	10	33	33	33	66
Número de ranuras soportadas	8	8	10	3	3	8	2
Ancho del canal (bits)	16	32	64	32	32	64	64
Velocidad de transferencia (MB/s)	16	64	160	160	132	132	264
Tarjetas multifunción	no	no	si	si	si	si	si
Autoconfiguración	no	no	si	no	si	si	si

Tabla 2-1. Comparativo de las diferentes tecnologías de canal de I/O

Conforme los procesadores continúen incrementando su desempeño, los canales de I/O que los alimentan de datos tendrán que mejorar como ha pasado en los últimos años. En la actualidad se está trabajando en el desarrollo de una nueva tecnología de canal de I/O llamada *System I/O* que es el resultado de la combinación de dos iniciativas de canal de I/O de IBM, Compaq, HP, Intel y Sun.

2.2 Controladores de I/O

Después del canal de I/O el siguiente paso en la ruta física de entrada y salida de datos es el controlador de I/O también conocido como tarjeta adaptadora (HBA, *Host Bus Adapter*). Este es implementado como una tarjeta que se coloca en alguna de las ranuras de expansión del canal de I/O del servidor o como un dispositivo electrónico integrado en la tarjeta madre. Para la operación de un controlador de I/O se necesita un componente de *software* que corre en dos lugares: en el *firmware* del controlador y en el sistema como un manejador de dispositivo tal como se muestra en la figura 2-2.

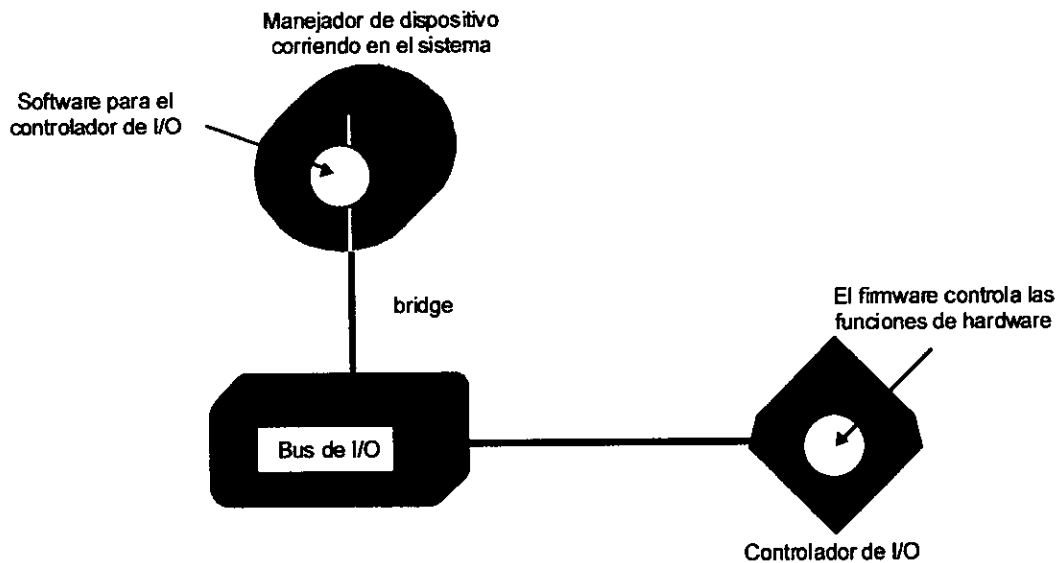


Figura 2-2. El controlador de I/O

El *firmware* es un *software* específico que está optimizado para la circuitería particular del controlador. Este se encarga de las funciones más básicas del controlador, tales como encendido, pruebas y detección de errores.

El manejador de dispositivo dirige el modo de operación y la transferencia de datos a través del controlador de I/O. Los manejadores de dispositivo generalmente soportan diversas funciones del sistema operativo así como de otros productos de *software*. Existen diferentes clases de controladores de I/O en el mercado enfocados al área de almacenamiento los cuáles se describen a continuación.

- ATA e IDE
- SCSI
- SSA
- Canal de Fibra

ATA e IDE

La Conexión de Tecnología Avanzada (ATA, *Advance Technology Attachment*) y la Electrónica de Unidad Integrada (IDE, *Integrated Drive Electronics*) son dos términos que se usan indistintamente, sin embargo, ATA es la tecnología de controlador e IDE es la tecnología de disco. ATA es una interfase de desempeño moderado y de bajo costo diseñada principalmente para unidades de disco de computadoras personales ya que no soporta operaciones de I/O simultáneas en múltiples discos por lo que resulta lento para sistemas operativos multitareas. Aunque su desempeño en un solo dispositivo es bastante bueno —16.6 MB/s— ATA no es considerado óptimo para usarse en aplicaciones que utilizan múltiples accesos a disco, sin embargo, se continúa trabajando para desarrollar el controlador de I/O ATA para soportar ambientes de I/O intensos.

SCSI

La Interfase para Sistemas de Cómputo Pequeños (SCSI, *Small Computer System Interface*) es una familia de protocolos que comparten una interfase paralela usada para conectar servidores y dispositivos periféricos. Desde la primera implementación de SCSI, la SCSI-1, han sido muchas las mejoras que se han hecho en cuanto a velocidad de transferencia, número de dispositivos soportados y ancho del canal mientras que se ha mantenido compatibilidad con previas implementaciones. La tabla 2-2 muestra las variantes de SCSI así como sus principales características en términos de velocidad y ancho del canal.

Tipo de SCSI	Velocidad del canal	Ancho del canal	Dispositivos soportados
	MB/s	bits	
SCSI-1	5	8	8
Fast SCSI	10	8	8
Fast Wide SCSI	20	16	16
Ultra SCSI	20	8	4 y 8
Wide Ultra SCSI	40	16	4, 8 y 16
Ultra SCSI 2	40	8	8
Ultra Wide SCSI 2	80	16	16

Tabla 2-2. Comparativo de las diferentes implementaciones de SCSI

SCSI proporciona ventajas significativas en desempeño para sistemas con muchas peticiones de I/O ya que permite que se puedan realizar varias operaciones al mismo tiempo. La señalización eléctrica entre el controlador de I/O y los dispositivos es mucho más rápida que la de los componentes mecánicos utilizados en unidades de disco y cinta. SCSI aprovecha esta diferencia de tiempo para manejar varias operaciones en múltiples dispositivos simultáneamente.

Debido a que en SCSI la información viaja en forma paralela, los cables utilizados en esta tecnología son complejos comparados con otros controladores, adicionalmente existen restricciones en la longitud de los cables (hasta 25 metros) lo que hace que SCSI no pueda ser implementado en esquemas de conexión que requieren de grandes distancias.

En la actualidad SCSI constituye la tecnología más usada para la conexión de arreglos de discos y servidores, sin embargo, las desventajas que tiene en términos del número de dispositivos soportados, distancia y complejidad en los cables están favoreciendo la adopción de tecnologías como el canal de fibra.

SSA

La arquitectura de almacenamiento serial (SSA, *Serial Storage Architecture*) fue desarrollada por IBM como una tecnología de red de almacenamiento enfocada fundamentalmente a unidades de disco. SSA proporciona una interfase de disco serial que usa una conexión bidireccional formando un ciclo, los datos y comandos enviados de un adaptador SSA pueden viajar en cualquier dirección alrededor del ciclo a sus dispositivos de destino. Si se presenta alguna interrupción en el ciclo, SSA puede reconfigurar automáticamente el sistema para mantener la conexión. Esta configuración de ciclo permite que las unidades de disco puedan ser agregadas o quitadas sin ninguna interrupción en la operación.

Debido a esta doble conexión se puede tener una velocidad de transferencia de datos de hasta 80 MB/s con un soporte de hasta 192 unidades por adaptador y una distancia máxima de conexión entre el servidor y las unidades de disco de 25 metros. De hecho, en algún momento SSA llegó a ser considerada como una tecnología capaz de competir con el canal de fibra, sin embargo, a pesar de todas estas ventajas el hecho de ser considerada una tecnología propietaria no ha permitido que SSA se desarrolle plenamente.

Canal de Fibra

En la actualidad, el canal de fibra constituye una de las tecnologías de mayor adopción para conexión de arreglos de discos, dentro de sus principales características se encuentran la velocidad de transferencia de datos que es de 100 MB/s, la posibilidad de poder ser usada en diferentes ambientes y aplicaciones al permitir el transporte de diversos protocolos incluyendo IP y SCSI, además de su aplicación en ambientes que demandan gran cantidad de imágenes y gráficos al ser capaz de transmitir paquetes de datos muy grandes.

Otra característica interesante del canal de fibra es el soporte de múltiples topologías de conexión que la hacen atractiva tanto para redes de datos como para redes de almacenamiento, a la vez que se soportan distancias de conexión de hasta 10 kilómetros. En cuanto al número de dispositivos que pueden ser conectados en un adaptador de fibra también existen ventajas ya que en estos momentos se tiene soporte para hasta 127 dispositivos mediante concentradores de conexiones simples o hasta 15.5 millones de dispositivos utilizando concentradores de conexiones inteligentes comparado con el máximo de 16 que soporta SCSI.

Si hay alguna desventaja que en estos momentos tenga esta tecnología, esta es el precio y la infraestructura que se requiere para que sea implementada, sin embargo, el canal de fibra constituye la base para las redes de almacenamiento inteligente por todas las ventajas que ofrece.

2.3 Topologías de conexión

La topología usada para conectar dispositivos de almacenamiento a sus respectivos servidores es una variable importante que puede determinar el desempeño y la escalabilidad de una solución de cómputo. En la actualidad son tres las principales topologías que se utilizan tanto para redes de datos como para sistemas de almacenamiento tal como se puede ver en la tabla 2-3.

	Topología de canal	Topología de ciclo	Topología de fábrica
Redes de datos	Arcnet	Token Ring	10 Base T
	Ethernet ancho	FDDI	100 Base T
	Ethernet delgado		1000 Base T
Redes de almacenamiento	SCSI Paralelo	SSA	ATM
	ATA	Firewire	ESCON
	Canal de fibra	Canal de fibra	Canal de fibra (fábrica)
	Punto a punto	Ciclo	

Tabla 2-3. Topologías de conexión de almacenamiento

Topología de Canal

La topología de canal (*bus*) utiliza tecnologías de transmisión serial y paralela. En lo que a almacenamiento se refiere la transmisión paralela goza de mayor popularidad por razones de desempeño pero también ha ocasionado que los cables sean más complejos y de menor longitud para asegurar la integridad de las señales. Esta topología está basada en el acceso compartido a un cable común. Cada sistema o dispositivo que quiera comunicarse por el canal necesita esperar su turno para tomar el control. La figura 2-3 ilustra el esquema de conexión de esta topología.

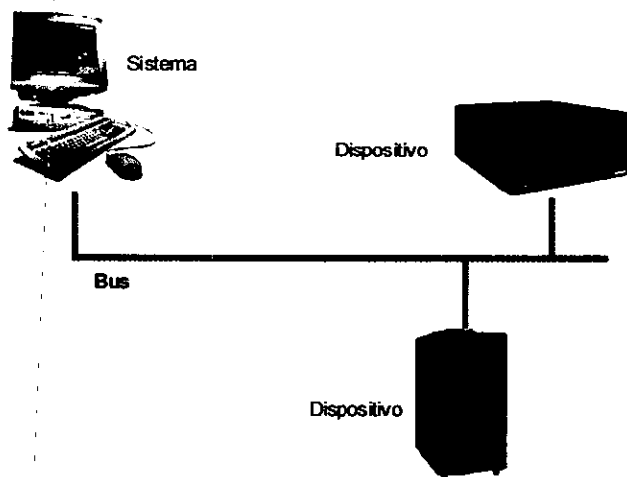


Figura 2-3 Topología de canal

Topología de Ciclo

La topología de ciclo es similar a la de canal en el sentido de que se requiere de algún mecanismo que determine quien controla el ciclo. Sin embargo, la tecnología de ciclo tiene la ventaja de que puede ser implementada como una ruta doble de tal manera que el tráfico pueda viajar en dos direcciones. Generalmente se utiliza un concentrador para lograr este tipo de conexión aunque también se puede lograr sin necesidad de este dispositivo.

Otra diferencia entre la tecnología de canal y de ciclo es que esta última utiliza transmisión serial en lugar de paralela, lo que permite que los cables y conectores sean más pequeños y que puedan ser de mayor longitud.

La figura 2-4 muestra la manera en que tres dispositivos pueden ser conectados a través de un concentrador en una topología de ciclo.

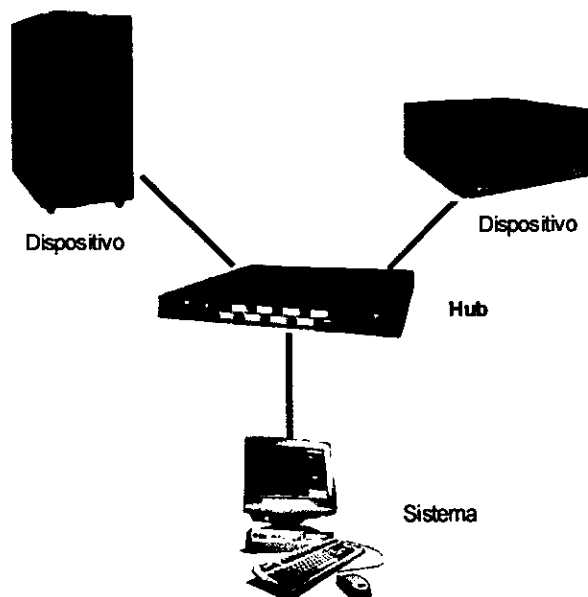


Figura 2-4 Topología de ciclo.

Topología de Fábrica

La topología de fábrica al igual que la de ciclo utiliza transmisión serial, sin embargo, los concentradores inteligentes usados en su implementación no utilizan una conexión compartida. El concentrador permite una conexión dedicada por cada dispositivo de almacenamiento a los diferentes servidores, por esta razón no es necesario ningún mecanismo para el control del canal ya que múltiples pares de entidades se pueden comunicar al mismo tiempo. A diferencia de la topología de ciclo si se requiere del concentrador para llevar a cabo la conexión.

La figura 2-5 muestra una conexión típica de dispositivos periféricos y servidores a través de un concentrador de conexiones inteligente.

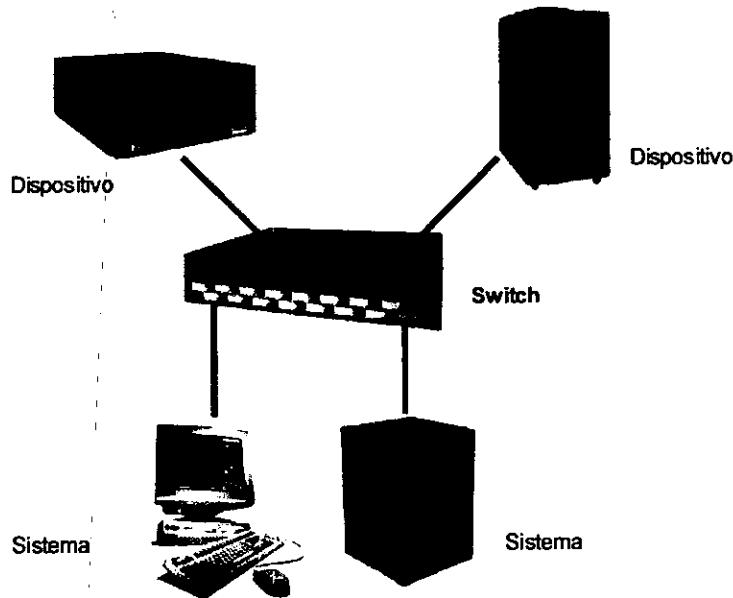


Figura 2-5 Topología de fábrica

2.4 El disco duro

La unidad de disco duro se ha convertido en el dispositivo más popular de almacenamiento de datos ya que permite que estos se encuentren siempre accesibles. Desde su primera aparición en 1955, el disco duro ha venido de ser un periférico caro, complejo y de gran tamaño encontrado solo en ambientes de grandes centros de datos (*mainframes*) a un dispositivo barato capaz de albergar una decena de *Gigabytes* en una pulgada cuadrada con un alto grado de confiabilidad.

Cuando se piensa en un disco duro, la primera característica que salta a la mente es la capacidad de éste, sin embargo, el desempeño y la confiabilidad juegan un papel sumamente importante como con cualquier otro dispositivo de cómputo.

La figura 2-6 muestra algunos de los modelos de discos duros que existen en el mercado actualmente.



Figura 2-6 El disco duro

Tecnología Magneto resistente

Esta es la última tecnología usada para almacenar datos en los discos duros la cual ha permitido soportar capacidades del orden de *Gigabytes* en una pulgada cuadrada, consiste básicamente en un material magneto resistente (una aleación ferromagnética) cuya resistencia eléctrica cambia a la presencia de un campo magnético. Cuando se aplica dicho material a la cabeza lectora de una unidad de disco es posible detectar y leer campos magnéticos débiles generados por muchos bits almacenados en una superficie muy pequeña.

Velocidad de rotación

Entre mayor sea la velocidad de rotación de una unidad de disco es menor el tiempo que la cabeza de éste debe esperar para comenzar a leer o escribir datos y mayor la velocidad en la que estos pueden ser transmitidos. Mientras que esta cantidad de tiempo puede ser corta, en el rango de entre 2 y 10 milisegundos, cuando cada operación de disco es afectada, el impacto en el desempeño puede ser más que notorio. En la actualidad la velocidad de rotación de la mayoría de los discos es de entre 10000 y 15000 revoluciones por minuto por lo que cualquier velocidad menor a esta debe ser desechada como opción.

Tiempo de búsqueda

Después de la velocidad de rotación este es el segundo aspecto más importante en el desempeño de un disco. El tiempo de búsqueda es el tiempo que le lleva a la cabeza de un disco para moverse de una pista a otra cuando se están leyendo o escribiendo datos.

Espacio de memoria

Todas las unidades de disco tienen espacios de memoria (*buffers*) para almacenar de manera temporal ciertos datos que se están leyendo o escribiendo. Estos espacios de memoria proporcionan una forma en la que el controlador de I/O puede trabajar con el disco para transferir datos aunque este se encuentre realizando otra función. Entre mayor sea el espacio de memoria mejor será el desempeño del disco.

Conexión Dual

Conforme las tecnologías de almacenamiento se han desarrollado, la necesidad de proporcionar rutas de acceso alternas a los discos se ha incrementado. La conexión dual consiste en tener más de una conexión física a un disco tal como se muestra en la figura 2-7, esta doble conexión hace que el desempeño mejore a la vez que se tiene una mayor protección para los datos ya que en el evento de que una ruta de I/O o controladora fallen, la otra puede continuar accediendo a la información. En la actualidad las unidades de disco con doble conexión están disponibles para las tecnologías de canal de fibra y la arquitectura de almacenamiento serial.

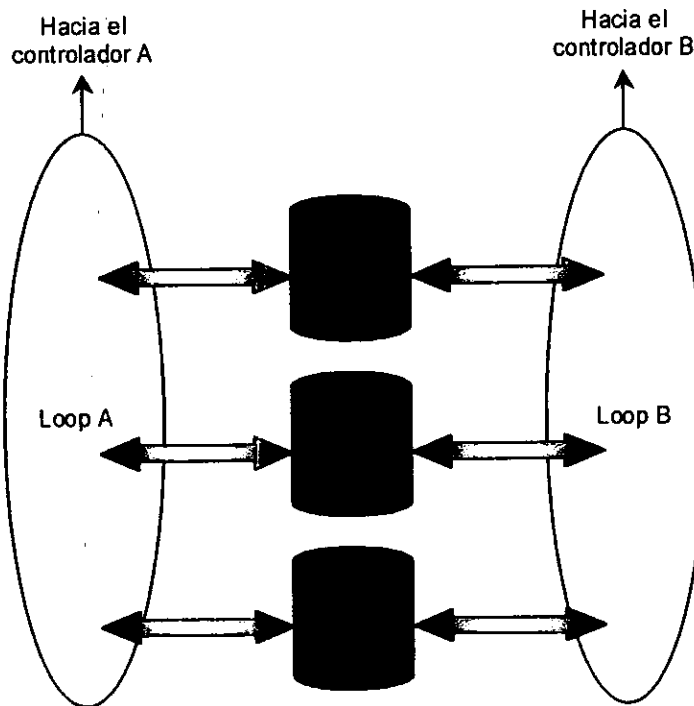


Figura 2-7. Unidades de disco con doble conexión

2.5 La cinta magnética

La cinta magnética fue reemplazada por el disco duro como medio de almacenamiento primario al inicio de los años 60's, desde entonces ha sido ocupada como medio de almacenamiento secundario y se ha consolidado como el medio preferido para respaldar información.

En la actualidad existen diferentes formatos de cinta los cuales en base a sus características de desempeño y capacidad se han colocado en diferentes nichos de mercado cubriendo desde las necesidades más elementales de protección de datos a nivel de escritorio hasta los ambientes de respaldo empresariales más complejos.

QIC

El cartucho de cuarto de pulgada (QIC, *Quarter Inch Cartridge*) utilizado principalmente en ambientes de computadoras de escritorio emplea una técnica de grabación lineal en la que los datos son escritos en pistas paralelas que corren a lo largo de la longitud de la cinta. El número de pistas y la longitud total determinan la capacidad nativa de la cinta. La mayoría de las unidades de cinta QIC utilizan un mecanismo de lectura-escritura que consiste de una sola cabeza de escritura colocada entre dos cabezas de lectura. Esta configuración permite a la unidad verificar las escrituras mientras estas se realizan.

QIC está tendiendo a desaparecer incluso del mercado de las PC's ya que no ha podido adaptarse de manera adecuada a los rápidos cambios que han habido en las capacidades de los discos duros. En la figura 2-8 se puede apreciar la imagen de una cinta tipo QIC.

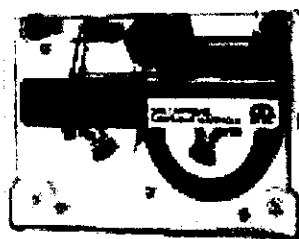


Figura 2-8. Cinta QIC

TRAVAN

El formato TRAVAN desarrollado por Imation (antes 3M) está teniendo bastante aceptación como una tecnología de reemplazo de QIC ya que ofrece mayores capacidades de almacenamiento manteniendo compatibilidad con el formato QIC a un precio bastante accesible. Compañías como IBM, Hewlett Packard, Acer, Toshiba, Data General, NCR, Unisys y Hitachi están incorporando esta tecnología como parte de sus ofrecimientos de cinta para el mercado de computadoras personales. La tabla 2-4 muestra las características de los diferentes modelos de cinta TRAVAN

Modelo	Capacidad	Velocidad
	GB	MB/seg
TRAVAN-1	0.8	0.25
TRAVAN-2	1.6	0.25
TRAVAN-3	3.2	0.5
TRAVAN-8	8	1.2
TRAVAN-NS1	8	1.2
TRAVAN-NS2	20	2

Tabla 2-4. Capacidades de los diferentes modelos de cinta TRAVAN

La figura 2-9 muestra la imagen de una cinta TRAVAN



Figura 2-9. Cinta TRAVAN

DDS, 4 mm

La tecnología de cinta de 4 mm fue originalmente desarrollada por Phillips como un formato de audio digital conocido como Cinta de Audio Digital (DAT, *Digital Audio Tape*). En 1998 Hewlett Packard y Sony se unieron para definir el estándar de grabación de esta tecnología en el ambiente de cómputo al que denominaron Almacenamiento de Datos Digitales (DDS, *Digital Data Storage*). Actualmente DDS se encuentra en su cuarta generación. La tabla 2-5 sintetiza la evolución que DDS ha tenido, así como las características de los diferentes modelos.

Modelo	Capacidad GB	Velocidad MB/seg
DDS	2	0.55
DDS-1	2 a 4	0.55 a 1.1
DDS-2	4 a 8	0.55 a 1.1
DDS-3	12 a 24	1.1 a 2.2
DDS-4	20 a 40	3

Tabla 2-5. Características de los diferentes modelos de cinta DDS

Las unidades DDS utilizan una tecnología de grabación denominada *Helical Scan* en la que los datos son grabados en pistas que corren diagonalmente sobre la cinta la cual se mueve a una velocidad relativamente lenta comparada con la tecnología de grabado lineal lo que permite una densidad de grabación más alta ya que se reduce la interferencia magnética entre pistas permitiendo que estas puedan ser posicionadas más eficientemente. Las unidades de cinta de 4mm tienen dos cabezas para lectura y dos para escritura. Para que se pueda leer o escribir sobre una cinta, esta se mueve 90° alrededor de la circunferencia del tambor y las cabezas se mueven en forma de espiral de abajo hacia arriba de la cinta. La figura 2-10 muestra la imagen de una cinta de 4 mm.



Figura 2-10. Cinta de 4 mm

8 mm

La cinta de 8 mm utiliza la misma técnica de grabado usada en la cinta de 4 mm, sin embargo, a diferencia de las unidades DAT, las unidades de 8 mm utilizan una configuración de más cabezas, una de lectura, una de escritura y una servo cabeza alrededor del tambor así como una cabeza adicional de borrado. La figura 2-11 muestra la imagen de una cinta de 8 mm.

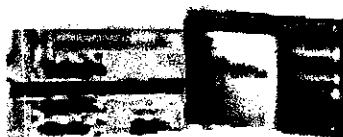


Figura 2-11. Cinta de 8 mm

Existen dos diferentes implementaciones de la cinta de 8 mm las cuales utilizan diferentes tecnologías de compresión de datos. Exabyte impulsa la tecnología *Mammoth* la cual utiliza cassettes estándar de 8 mm, mientras que Seagate y Sony han desarrollado un cartucho diferente con un formato propietario conocido como Cinta Inteligente Avanzada (AIT, *Advanced Intelligent Tape*). Aunque ambas tecnologías tienen la misma velocidad de transferencia de datos y las dos soportan el uso de la tecnología AME (*Advanced Metal Evaporated*) AIT utiliza una memoria permanente llamada Memoria en Cassette (MIC, *Memory in Cassette*) la cual proporciona un mejor desempeño que la implementación de Exabyte. En cuanto a la capacidad de almacenamiento Mammoth soporta 20 GB sin compresión mientras que AIT soporta 25 GB. Recientemente Sony anunció la segunda generación de su tecnología AIT con una capacidad de 50 GB y una velocidad de transferencia de 6 MB/s. La tabla 2-6 muestra las capacidades y velocidades de transferencia de los diferentes modelos de cinta *Mammoth*.

Modelo	Capacidad	Velocidad
	GB	MB/seg
Mammoth-LT	14	2
Mammoth-1	20	3
Mammoth-2	60	12
Mammoth-3	120	18

Tabla 2-6 Características de los diferentes modelos de cinta de 8 mm.

DLT

La cinta lineal digital (DLT, *Digital Linear Tape*) fue desarrollada a mediados de los 80's por Digital Equipment Corporation como una solución de cinta de media pulgada para la Micro VAX. La tecnología fue adquirida por Quantum Corporation en 1994 y ahora ellos licencian DLT a todos los fabricantes de unidades y bibliotecas de cintas.

La tecnología DLT utiliza la técnica de grabación lineal también conocida como serpentina. La cinta lleva los datos en canales paralelos que corren a través de una cabeza estacionaria, cada uno de los canales pasa sobre su propio elemento de escritura dentro de la cabeza. Al aumentar el número de canales se aumenta la velocidad de transferencia de datos. La figura 2-12 muestra la imagen de la unidad de cinta DLT.



Figura 2-12 Cinta DLT

La cinta lineal digital utiliza una sola ruta de cinta y opera a una tensión baja constante en conjunto con una cabeza estacionaria ayudando a minimizar el desgaste de cabezas y cintas. DLT ofrece varias características avanzadas como la implementación completa del conjunto de comandos de SCSI-2 y diagnósticos interconstruidos, el costo por MB de DLT es comparable al de la tecnología de 8 mm ofreciendo mayor capacidad y desempeño, es por eso que este formato de cinta goza de bastante popularidad en ambientes de respaldo empresariales. La tabla 2-7 muestra las características de los diferentes modelos de cinta DLT.

Modelo	Capacidad	Velocidad
	GB	MB/seg
DLT 2000XT	15	1.25
DLT 4000	20	1.5
DLT 7000	35	5
DLT 8000	40	6

Tabla 2-7 Características de los diferentes modelos de cinta DLT

Magstar

Magstar es el último modelo en la línea de cintas de media pulgada de IBM el cual tiene una capacidad de almacenamiento de 10 GB ofreciendo conectividad con ESCON y SCSI lo que permite que esta familia pueda ser usada tanto en ambientes de grandes centros de datos como en sistemas abiertos Unix.

Por ejemplo, el modelo 3590 de Magstar, está diseñado para el mercado de grandes centros de datos y para aplicaciones especiales mientras que el Magstar MP que también está enfocado al mismo nicho utiliza el mismo método de grabación lineal que 3590 pero emplea un formato y diseño completamente diferentes. Magstar MP utiliza una cinta de 8 mm en vez de cartuchos de media pulgada. Este fue diseñado como una unidad de cinta de acceso rápido enfocado especialmente para el mercado de bibliotecas de cinta. La figura 2-13 muestra una unidad de cinta Magstar.



Figura 2-13 Unidad de cinta Magstar

9840

La tecnología 9840 fue diseñada por StorageTek en conjunto con Imation con un formato compatible al de la cinta 3480 de IBM. Esta unidad de cinta de media pulgada tiene una capacidad nativa de 20 GB con una velocidad de transferencia de datos de 10 MB/s. A diferencia de los cartuchos 3480, en 9840 no hay necesidad de sacar la cinta de su cartucho para leer o escribir lo que permite que se tengan mejores tiempos de acceso. 9840 es también la respuesta de Storage Tek al Magstar MP de IBM ya que tiene un excelente tiempo de carga y descarga de cinta por lo que también está enfocado básicamente al mercado de bibliotecas de cinta.

El costo de 9840 es similar al de Magstar y al igual que los ofrecimientos de IBM la unidad de cinta 9840 ofrece conectividad con SCSI y ESCON. La figura 2-14 muestra una unidad de cinta 9840



Figura 2-14 Unidad de cinta 9840

LTO

LTO (*Linear Tape Open*) es una iniciativa de Seagate Technology, IBM y Hewlett Packard para definir un formato de cinta basado en especificaciones abiertas utilizando lo mejor de las tecnologías existentes, estas compañías están buscando crear dos formatos de cinta intercambiables entre múltiples proveedores. Sin embargo, será difícil que alguno de estos dos formatos sea compatible con alguna de las cintas existentes.

Este formato de cinta está basado también en la técnica de grabación lineal y para su implementación se utilizan dos tecnologías. Para ambientes en donde se requiere de gran capacidad y confiabilidad se utiliza Ultrium el cual tiene las mismas características de forma que la cinta DLT usando un diseño de una sola bobina manejando capacidades de hasta 100 GB y una velocidad de transferencia de datos de entre 10 y 20 MB/s.

Por otro lado para ambientes en donde se requiere de acceso rápido con un alto grado de confiabilidad la tecnología empleada es Accelis cuyo factor de forma está asociado con Masgtar MP. Accelis tiene un diseño de doble bobina lo que le permite tiempos de acceso menores a los 10 segundos, con una capacidad de 25 GB y una velocidad de transferencia de datos de entre 10 y 20 MB/s. La figura 2-15 muestra la unidad de cinta LTO.



Figura 2-15 Cinta LTO

2.6 Soluciones de Cinta Automatizadas

Debido al incremento en los volúmenes de información y a la necesidad de eficientar y reducir los costos asociados con el respaldo y recuperación de los datos, los sistemas de cinta automatizados han ganado bastante popularidad en nuestros días.

Aunque existen diferentes variantes de sistemas de respaldo en general estos caen en tres categorías: apiladores, autocargadores y bibliotecas.

Los apiladores (*stackers*) consisten de una sola unidad de cinta y un mecanismo para intercambiar y cargar las cintas. Los autocargadores (*autoloaders*) adicionalmente tienen la habilidad de cargar cintas específicas en demanda. Las bibliotecas de cintas automatizadas (ATL's, *Automated Tape Libraries*) utilizan un método de operación mucho más sofisticado que el de los autocargadores ya que permiten la manipulación de múltiples unidades de cinta incorporando robots para lograr esto, además cuentan con controladores y *software* que les permiten responder a múltiples peticiones de servicio solicitados por diferentes sistemas. Las ATLs también utilizan lectores de código de barras como mecanismos de identificación.

La figura 2-16 ilustra los diferentes tipos de soluciones de cinta automatizados que existen en la actualidad.

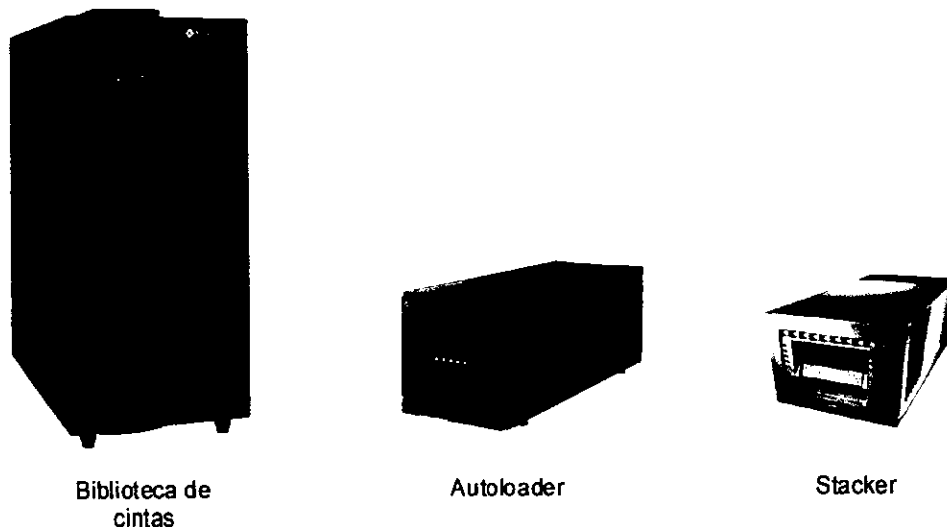


Figura 2-16. Soluciones de cinta automatizadas

2.7 El disco óptico

La tecnología de disco óptico abarca una amplia gama de productos que utilizan luz para leer y escribir datos dentro de los que se encuentra el Disco Compacto de Memoria de Solo Lectura (CD-ROM, *Compact Disk - Read Only Memory*), el Disco de Video Digital (DVD, *Digital Video Disk*) y los discos Magneto Ópticos (MO). Los discos magneto ópticos cambian la polaridad de un campo magnético en un medio de grabación utilizando magneto y láser lo que permite que este medio de almacenamiento sea borrable y escribible a diferencia del CD-ROM o DVD-ROM.

En la actualidad existen variantes de CD's y DVD's que permiten que también se les pueda grabar información, estas variantes son el CD-RW y el DVD-RAM los cuales utilizan un mecanismo de grabación llamado de cambio de fase el cual involucra el uso de un láser para transformar una sección del medio de grabación y un mecanismo de reflectividad para poder modificar datos binarios.

Los discos magneto ópticos tienen varias ventajas sobre el CD-RW y el DVD-RAM ya que grandes densidades de datos pueden ser almacenadas usando láser de onda corta, por otro lado los discos MO tienen mayor confiabilidad lo que los constituye como un excelente medio de almacenamiento de imágenes, de hecho en el mercado existen soluciones similares a las ATL's pero para discos magnéticos llamados *jukeboxes* los cuales permiten el manejo de grandes volúmenes de información, por lo que este tipo de soluciones está enfocado fundamentalmente a grandes ambientes de cómputo. La figura 2-17 muestra las diferentes soluciones de respaldo basadas en la tecnología láser.

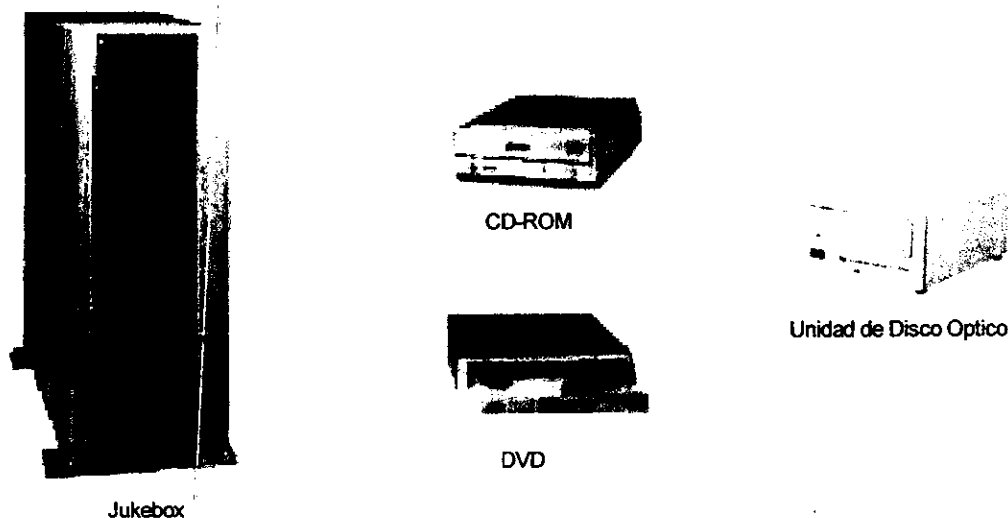


Figura 2-17 Soluciones de respaldo basadas en la tecnología láser

III

Determinando la mejor solución de almacenamiento

Antes de elegir una solución de almacenamiento es necesario identificar plenamente los requerimientos y la funcionalidad que se necesita, así como el impacto costo/beneficio que la solución tendrá no solo en el área informática sino en la empresa en general. Sin embargo, en el proceso de analizar requerimientos de cómputo a menudo se omiten aspectos importantes que pueden traer consigo serias consecuencias y esto no es una excepción cuando se habla de almacenamiento.

En el proceso de evaluar requerimientos de almacenamiento a menudo la capacidad y el desempeño son los criterios principales, desafortunadamente, estos solo cubren una porción de las necesidades de manejo de datos.

En general se pueden distinguir ocho áreas funcionales que pueden ayudar a determinar una solución de almacenamiento total que incluya no solo la configuración del arreglo de discos, sino cualquier otro componente necesario de *software*, infraestructura de comunicaciones, políticas de respaldo, planes de recuperación contra desastres, etc. Dichas áreas son:

- Desempeño
- Capacidad
- Costo
- Disponibilidad
- Soporte
- Ambiente de Operación
- Soluciones de Respaldo
- Crecimiento futuro

Parte del proceso de evaluación de requerimientos de almacenamiento consiste en asignar las prioridades correspondientes a cada una de estas áreas. Aunque, se puede afirmar que el desempeño y la capacidad son los primeros aspectos que se deben tener en cuenta cuando se está pensando en una solución de almacenamiento ya que ésta deberá cumplir cuando menos con un criterio mínimo de desempeño y con una capacidad mínima en disco, ajustes a la configuración inicial se pueden hacer después en base al crecimiento futuro, costo y nivel de disponibilidad, pero un estimado inicial es importante.

3.1 Desempeño

En general el ancho de banda (*bandwith*) y el número de operaciones de entrada/salida por segundo (IOPS) son los parámetros más utilizados para medir el desempeño de un sistema de almacenamiento, sin embargo, estos parámetros pueden variar dependiendo de la carga de trabajo de la aplicación y de la configuración y características del arreglo de discos.

3.1.1 Carga de trabajo

La carga de trabajo de una aplicación está definida por:

- Patrón de acceso
- Tamaño de las operaciones de entrada y salida
- Nivel de carga
- Rango de búsqueda o tamaño de archivo

Patrón de acceso

Es la unidad de trabajo básica a ser realizada por un sistema de almacenamiento, está definida por tres características: La primera es si el acceso es a un dispositivo crudo (*raw device*) o a un sistema de archivos (*file system*). El acceso a dispositivos crudos es a través de un controlador de dispositivo orientado a carácter (*byte*), el cual no hace uso de la paginación de memoria virtual, mientras que el acceso a sistemas de archivos es a través de un controlador de dispositivo orientado a bloques de datos, el cual si hace uso de la paginación de memoria virtual. El acceso a dispositivos crudos es mejor en términos de desempeño ya que la aplicación se hace cargo directamente del acceso sin pasar por el sistema operativo pero la información es más difícil de manipular a no ser que sea directamente desde la aplicación.

La segunda característica de un patrón de acceso es si este es secuencial o aleatorio, el acceso secuencial significa que cada operación de entrada y salida continúa donde se hizo la última operación, el acceso aleatorio significa que cada operación de I/O involucra la búsqueda de una nueva dirección en donde se realizará la siguiente operación. La tercera característica define el tipo de operación: lectura o escritura.

Por ejemplo, una operación de escritura secuencial sobre un dispositivo crudo puede ser representada como: RSW (*Raw Sequential Write*)

Tamaño de I/O

También se le conoce como tamaño de bloque, este representa el número de *bytes* que son transferidos en una sola operación de I/O, dependiendo del tipo de aplicación el tamaño de I/O varía, por ejemplo, para aplicaciones de procesamiento de transacciones en línea (OLTP, *On Line Transaction Processing*) se tienen I/O's pequeños mientras que para aplicaciones de soporte para toma de decisiones (*Data Warehousing*) los tamaños de I/O son grandes.

Nivel de carga

Es el número de peticiones de I/O concurrentes.

Rango de búsqueda

Define el rango de direcciones de volumen lógicas en donde se harán las operaciones de I/O.

Por lo tanto, una carga de trabajo puede estar representada por: 6, 2 KB RRR 2 GB que quiere decir 6 peticiones de I/O de 2 KB de una operación de lectura aleatoria sobre un dispositivo crudo en un rango de búsqueda de 2 GB.

3.1.2 Configuración y características de los discos

La configuración y características de desempeño de un arreglo de discos están dadas por:

- Nivel de RAID
- Tipo de RAID
- Unidad de columna
- Ancho de columna
- Ancho de banda
- Número de operaciones de I/O por segundo

Nivel de RAID

En 1987, como respuesta a las necesidades de mejorar el desempeño, disponibilidad y confiabilidad de los métodos de almacenamiento en arreglos de discos existentes en aquellos días científicos de la Universidad de California, Berkeley, desarrollaron un método de almacenamiento de datos al que denominaron Arreglo Redundante de Discos Independientes, (RAID, *Redundant Array of Independent Disks*). Esta tecnología goza de gran popularidad en nuestros días ya que hace más fácil el manejo de grandes cantidades de datos ya que varios discos físicos se pueden ver como una sola unidad lógica proporcionando además ventajas significativas en términos de desempeño, confiabilidad y disponibilidad.

Los científicos de Berkeley propusieron cinco niveles de RAID para proporcionar un balance en desempeño y protección de datos, ahora existen más como resultado de la combinación de niveles así como de algunas variaciones, a continuación se describen los niveles de RAID más populares:

- RAID 0
- RAID 1
- RAID 3
- RAID 4
- RAID 5
- RAID 0+1
- RAID 5+3
- RAID 6

RAID 0

Aunque RAID 0 no es, estrictamente hablando un nivel de RAID ya que no proporciona ningún método de protección de la información, tal como lo especifican las siglas, muchos lo consideran como uno de los niveles de RAID y es bastante popular sobre todo cuando se usa en conjunto con RAID 1. En RAID 0, los datos son esparcidos o distribuidos en múltiples discos pero no existe ninguna protección de la información, tal como se muestra en la figura 3-1. Las transferencias de datos primero son divididas y después son escritas en segmentos llamados columnas (*stripes*) a cada uno de los discos. La escritura y sobre todo el acceso a la información es bastante rápido ya que múltiples operaciones de I/O pueden ser procesadas en paralelo, sin embargo, no existe ningún método de verificación de errores o protección de la información durante las escrituras a discos por lo que no hay ninguna tolerancia a fallas, la pérdida de cualquier disco en el arreglo significa la pérdida de la información. Este tipo de RAID se recomienda para almacenar información que no sea de misión crítica.

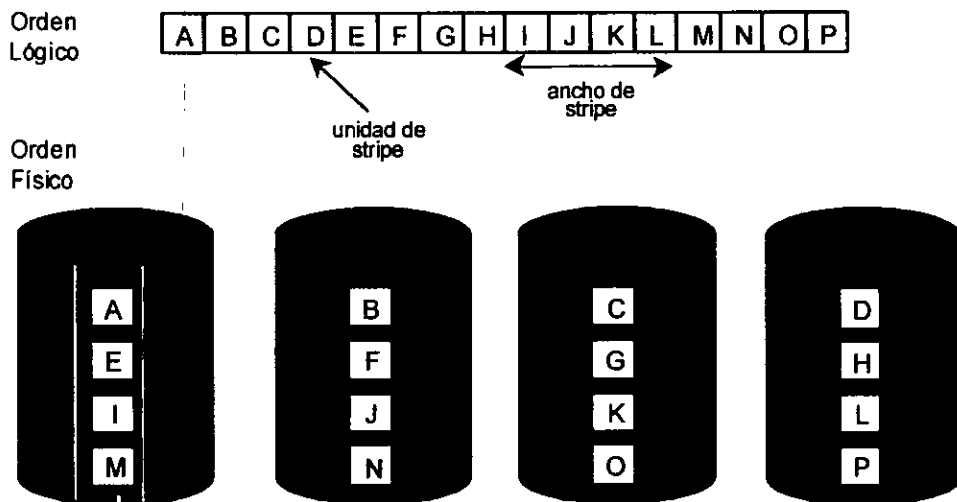


Figura 3-1. RAID 0

RAID 1

El RAID 1 consiste en copiar los datos escritos de un disco a otro denominado disco espejo tal como se ilustra en la figura 3-2, RAID 1 es altamente tolerante a fallas puesto que la actividad de I/O puede continuar sin interrupciones aún en el evento de que un disco falle ya que las operaciones de entrada/salida se transfieren al disco espejo cuando se presenta esta situación. Una vez que el disco que falló es reemplazado, los datos son copiados del disco espejo al nuevo disco y después se resincronizan para continuar operando de manera normal. RAID 1 no proporciona el mejor desempeño ya que se utilizan métodos de verificación y corrección de errores en la escritura a discos para garantizar la exactitud de la información, lo que decrementa la velocidad de escritura, por otro lado, esta es la solución de RAID más cara ya que se necesita el doble de espacio en disco del que se requiere para almacenar la información. Este tipo de RAID se recomienda para aquellos que buscan total redundancia en disco, sin importarles mucho el costo y el desempeño.

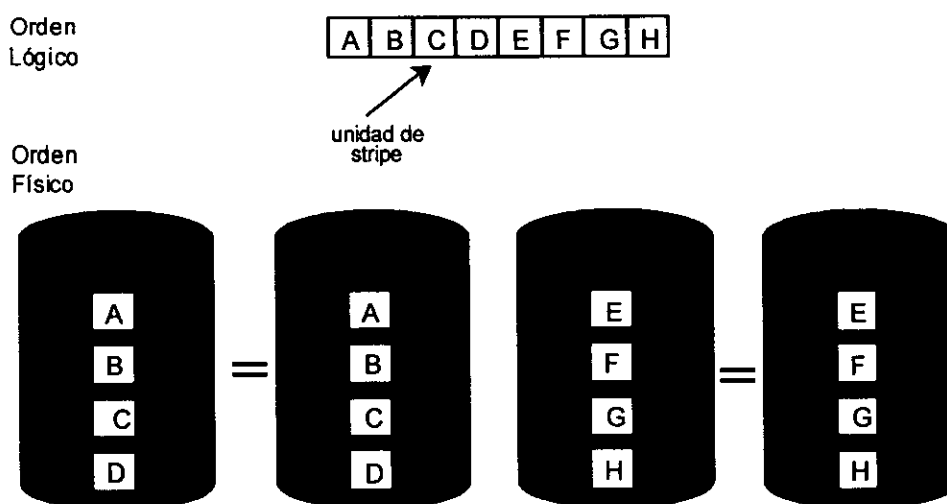


Figura 3-2. RAID 1

RAID 3

RAID 3 utiliza una unidad de disco para protección de datos en donde se almacena la información para verificación y corrección de errores. Los datos son distribuidos a través de todos los discos excepto por el disco de protección de datos, también conocido como disco de paridad. Esta información de paridad es almacenada para mantener la integridad de la información y también es guardada en columnas (*stripes*), cada una de las cuales contiene la información correspondiente a la columna en los discos de datos. La figura 3-3 ilustra la forma en la que RAID 3 es implementado.

En el evento de que un disco falle, la información contenida en ese disco es reconstruida a partir de la información almacenada en el disco de paridad. El arreglo puede continuar operando en este modo mientras que no falle un segundo disco. Esta implementación de RAID tiene buen desempeño en ambientes de escritura intensa y es mucho menos costosa que RAID 1 pero no ofrece la mejor protección a los datos.

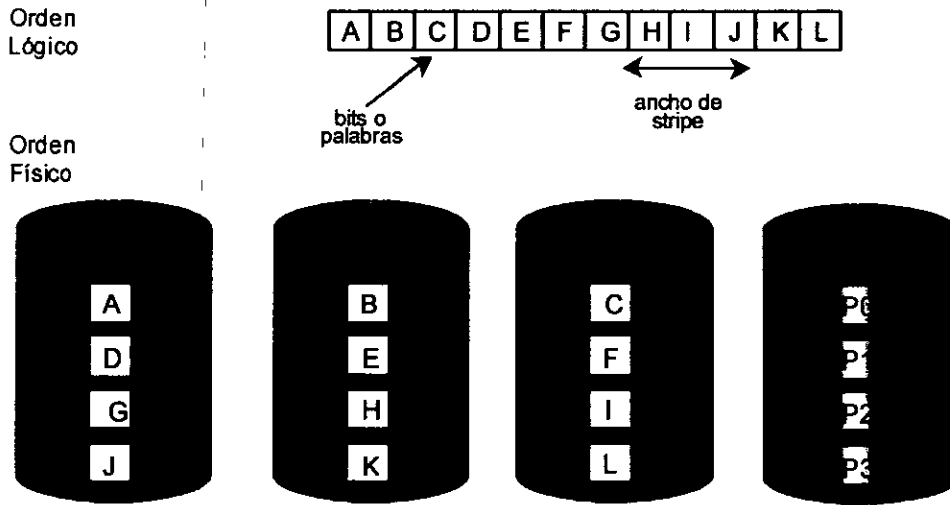


Figura 3-3. RAID 3

RAID 4

RAID 4, es similar en concepto a RAID 3, sin embargo, éste tiene una unidad de columna mayor, la cual hace que los discos operen de una manera más independiente que en RAID 3. Esta característica sacrifica algunas de las capacidades de desempeño que se tienen en RAID 3 en favor de un acceso más rápido en aplicaciones de lectura intensa tal como consultas a bases de datos. RAID 4 constituye el peor desempeño en escritura que cualquier otro nivel de RAID, esto se debe a la concentración de información de verificación de errores en un solo disco. La figura 3-4 muestra la manera en la que RAID 4 es implementado.

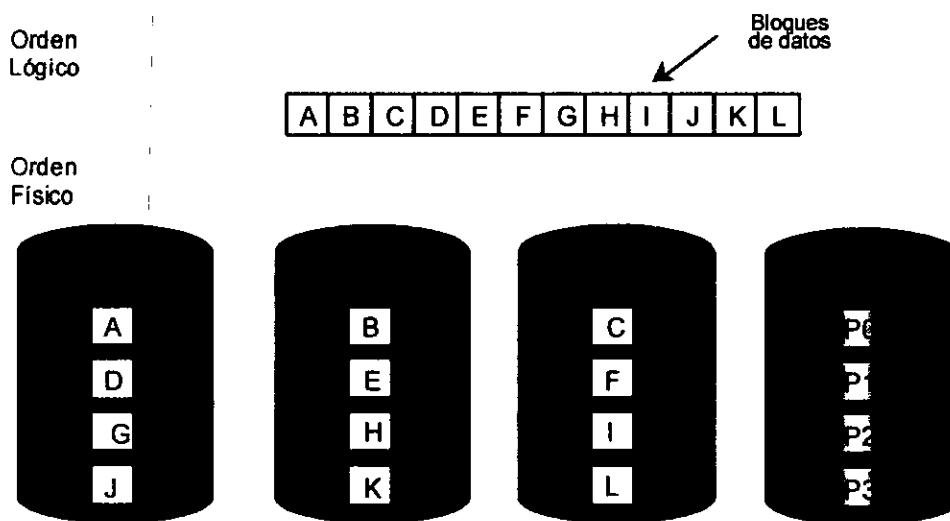


Figura 3-4. RAID 4

RAID 5

En RAID 5, la información de paridad es distribuida a través de todos los discos y al igual que en RAID 3 y RAID 4 se utiliza el mismo esquema de decodificación de la información de paridad para permitir que el arreglo no se vea afectado si una unidad de disco falla.

Ya que los datos y la información de paridad se almacenan a través de todos los discos, para recuperarse de la falla de cualquiera de estos es importante que la información de paridad de un disco en particular no se encuentre en el mismo disco en donde se tienen almacenados los datos.

Este tipo de RAID es bueno en ambientes con grandes peticiones de I/O pero no si se trata de muchas escrituras y sobre todo si estas son pequeñas ya que cualquier operación de escritura con un pequeño bloque de datos involucra que muchos discos sean accedidos y que se calcule nueva información de paridad. Este proceso de calcular y registrar nueva información de paridad en escrituras representa carga adicional que degrada el desempeño.

Hoy en día RAID 5 es una de las implementaciones de RAID más populares porque no es muy costosa, ofrece buen desempeño para un buen número de aplicaciones además de la protección a la información.

La figura 3-5 muestra la manera en la que se distribuyen los datos y la información de paridad en un RAID 5.

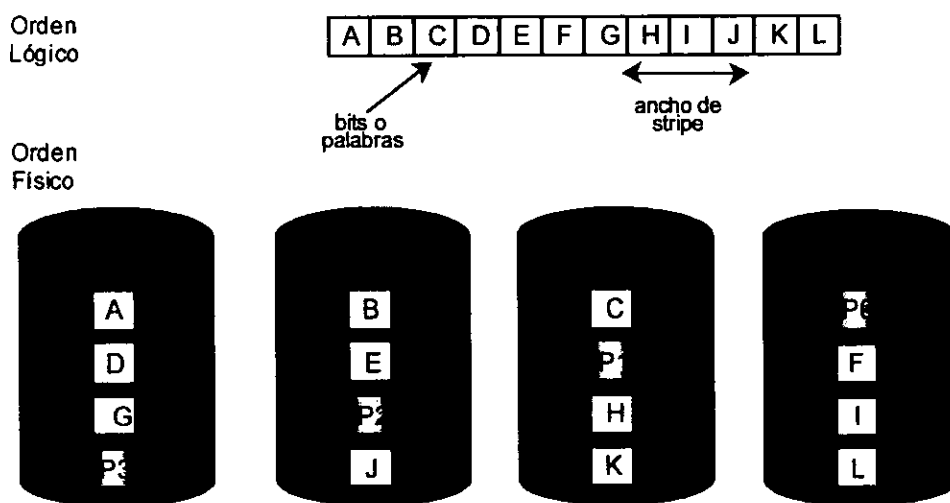


Figura 3-5. RAID 5

RAID 0+1

A la combinación de RAID 0 y 1 se le conoce como RAID 0+1 o RAID 10, esta combinación tiene ventajas interesantes ya que se combina la velocidad del RAID 0 con la redundancia que proporciona RAID 1, como resultado se tiene un sistema que es muy eficiente en operaciones de lectura/escritura y que es altamente tolerante a fallas. RAID 0+1 es tan costoso como RAID 1 pero tiene mucho mejor desempeño.

Actualmente la mayor parte de los productos de manejo de volúmenes implementan primero RAID 1 y luego RAID 0 ya que tiene ventajas en términos de confiabilidad y velocidad de recuperación que si se implementa de forma inversa. Esta implementación de RAID es altamente recomendable en ambientes de misión crítica. La figura 3-6 muestra la forma en la que RAID 10 es implementado.

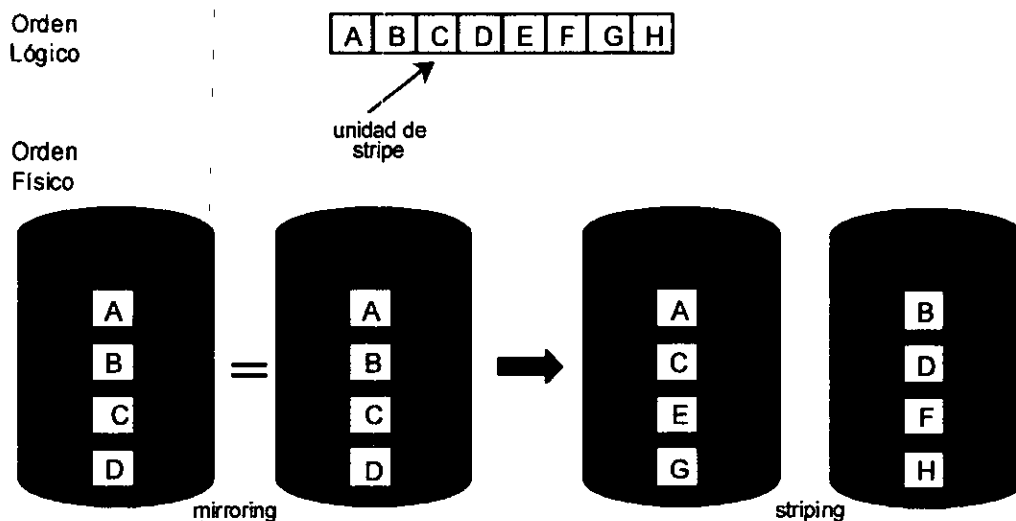


Figura 3-6. RAID 10

RAID 6

RAID 6 es una extensión de RAID 5, esta extensión consiste en un segundo esquema de paridad distribuida que mejora la tolerancia a fallas del arreglo. En RAID 6 además de los datos escritos a través de todos los discos y de la información de paridad también distribuida se calcula un segundo conjunto de datos de paridad que también se escribe a todos los discos, tal como se muestra en la figura 3-7. Esta configuración permite que más de un disco pueda fallar y que el arreglo continúe operando, las desventajas de esta implementación son la complejidad en el diseño y la cantidad de cálculos de paridad que se tienen que hacer, esto ocasiona que RAID 6 tenga muy pobre desempeño en operaciones de escritura.

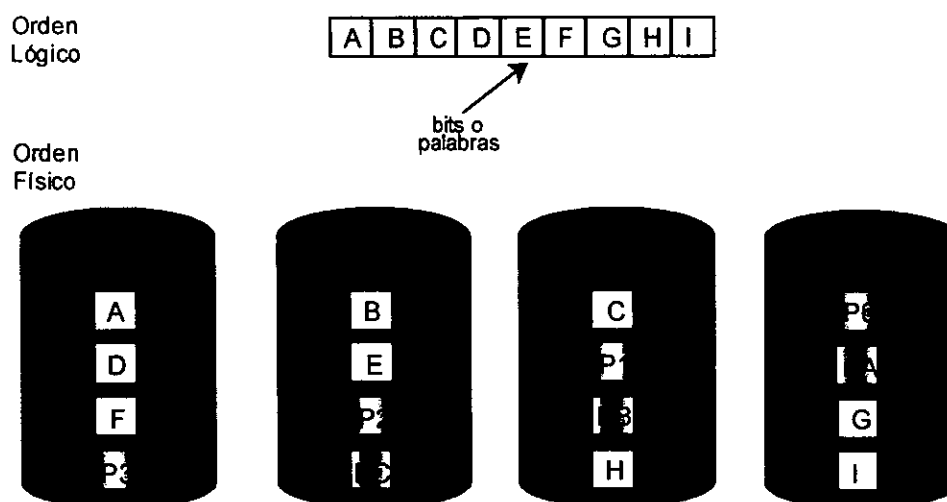


Figura 3-7. RAID 6

RAID por software

En este modelo, la implementación de RAID es realizada a través de un *software* que se instala en el servidor que tiene acceso a los discos, el *software* en el servidor controla toda la funcionalidad de RAID así como las operaciones de administración y sincronización de datos.

La principal ventaja de implementar RAID por *software* es la flexibilidad ya que se pueden tener varios niveles configurados y los grupos de RAID pueden ser expandidos a múltiples controladores de disco, cosa que no sucede con las implementaciones de RAID por *hardware*, también es posible cambiar las configuraciones de RAID de manera relativamente sencilla si las necesidades de almacenamiento de datos cambian.

En general el RAID por *software* usa más recursos del sistema que su contraparte por *hardware*, es por eso que en los niveles de RAID que requieren de muchos cálculos de paridad, se trata de evitar el uso de RAID por *software* ya que el cálculo de paridad es una tarea que consume muchos recursos y que puede degradar el desempeño del equipo. En general, el RAID por *software* es bastante utilizado para implementar RAID 0 y RAID 1 en los cuáles no se necesitan hacer cálculos de paridad.

RAID por hardware

En el RAID por *hardware* todo el procesamiento físico y lógico de las operaciones de I/O se libera del servidor y se lleva a cabo en el procesador que trae el controlador de RAID, que además se encarga de hacer los cálculos de paridad. En ambientes de escritura intensas con tamaños de I/O pequeños, en donde además se requiere de niveles de RAID en donde se hacen cálculos de paridad la solución de RAID por *hardware* es la ideal.

Existen dos maneras de implementar RAID por *hardware*, la primera es a través de un controlador externo que forma parte del arreglo de discos, el cual tiene un procesador y en algunas implementaciones coprocesadores o motores dedicados para hacer los cálculos de paridad y el manejo de las operaciones de I/O, estos controladores tienen además memoria caché para mejorar aún más el desempeño así como baterías para mantener esta memoria el tiempo suficiente para copiar la información a disco, en caso de pérdida de corriente. La otra implementación de RAID por *hardware* es a través de controladores basados en tarjetas adaptadoras (HBA's, *host bus adapters*), que en general tienen las mismas características que los controladores externos pero estos dependen de la plataforma así como de los controladores de dispositivo del sistema operativo en donde son configurados.

Unidad de columna

También conocido como tamaño de segmento o tamaño de columna es el bloque de datos más pequeño que puede ser leído o escrito a un disco físico antes de moverse al siguiente disco en una columna de RAID, el tamaño de la unidad de columna depende del tipo de aplicación, afortunadamente los dos tipos de implementaciones de RAID: *hardware* y *software*, permiten modificar el tamaño de la unidad de columna para ajustarse a los patrones de acceso y tamaños de I/O de las aplicaciones que usarán el arreglo, sin embargo, esto debe hacerse antes de que los datos se escriban a los discos, ya que una vez que los discos tienen información la única manera de cambiar la unidad de columna es respaldar la información, reconfigurar el arreglo de discos y volver a escribir la información.

La tabla 3-1 muestra algunas recomendaciones del tamaño de la unidad de columna dependiendo de la aplicación y el tipo de acceso.

carga de trabajo	tipo de acceso	tamaño de I/O	unidad de columna	unidad de columna
		óptimo	RAID 1	RAID 5
OLTP	RRR, RRW	2 KB	16 KB	64 KB
OLTP/ISP/ NFS v2	FRR, FRW	8 KB	64 KB	64 KB
NFS v3	FRR, FRW	32 KB	128 KB	128 KB
DSS/DW/HPC	RSR, RSW, FSR, FSW	16 KB	16 KB	16 KB
DSS/DW/HPC	RSR, RSW, FSR, FSW	64 KB	n/a	32 KB
DSS/DW/HPC	RSR, RSW, FSR, FSW	128 KB	16 KB	n/a
DSS/DW/HPC	RSR, RSW, FSR, FSW	256 KB	n/a	128 KB
DSS/DW/HPC	RSR, RSW, FSR, FSW	512 KB	256 KB	n/a
DB/FS LOGS	RSW, RSR	2 KB	16 KB	n/a

Tabla 3-1. Recomendaciones de unidad de columna por carga de trabajo y tipo de acceso

Ancho de columna

El ancho de columna (*stripe width*) es el número de discos físicos usados en una columna de RAID sin considerar la información de paridad, como regla entre más grande sea la columna mejor es el desempeño, sin embargo, el determinar el ancho de columna depende de los requerimientos de I/O y de consideraciones de *hardware*. La figura 3-8 muestra la diferencia entre la unidad de columna y el ancho de columna.

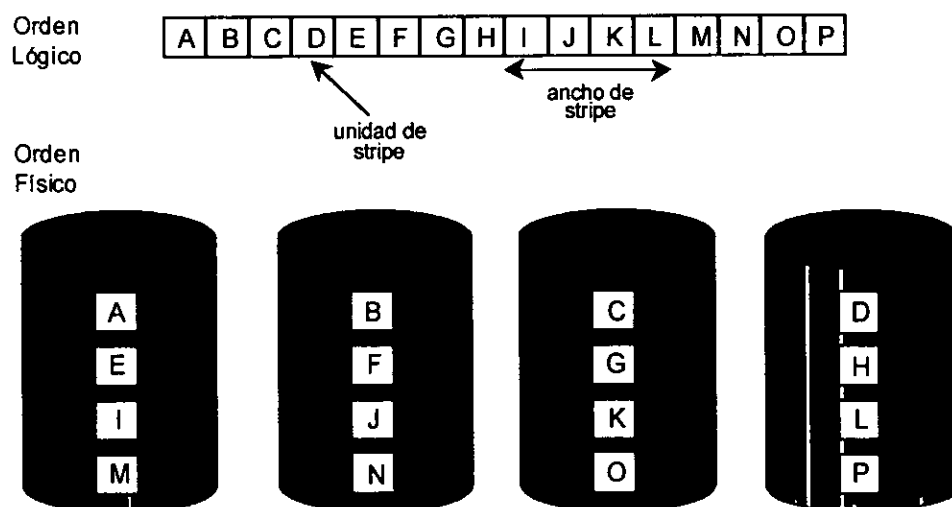


Figura 3-8 Unidad y ancho de columna

Ancho de banda

El ancho de banda es la cantidad de datos que puede ser transmitida por un canal de comunicaciones en una unidad dada de tiempo, generalmente está medida en MB/seg y está íntimamente ligada con el tipo de interfase de conexión utilizada tanto por el sistema de almacenamiento como por los discos mismos.

Número de operaciones de I/O por segundo

Es el número de señales de computadora de recepción (entrada) y transmisión (salida) que pueden ser manejados a lo largo de la ruta de conexión entre el canal del procesador y las unidades de disco.

3.1.3 Características de las aplicaciones

La configuración de un sistema de almacenamiento en términos del nivel y tipo de RAID así como de la unidad y ancho de columna dependen del tipo de trabajo que tendrá la aplicación en turno, por lo que es muy importante conocer las características de las aplicaciones según la carga de trabajo.

Procesamiento de Transacciones en Línea (OLTP, Online Transaction Processing)

Son aplicaciones orientadas a operaciones que se traducen en transacciones tales como agregar, añadir y eliminar registros. Las aplicaciones de tipo OLTP se caracterizan porque los tamaños de I/O son relativamente pequeños (entre 2 KB y 16 KB) con accesos predominantemente aleatorios, en donde se requieren tiempos de respuesta menores a 15 ms y el porcentaje de escrituras es de más del 30% del total de operaciones de I/O, algunos ejemplos de aplicaciones de tipo OLTP son las aplicaciones de punto de venta y servicio a clientes, aplicaciones de Planeación de Recursos Empresariales (ERP, *Enterprise Resource Planning*), pruebas de rendimiento y bases de datos.

Proveedores de Servicio Internet (ISP, Internet Service Providers)

Estas aplicaciones también se caracterizan por tener tamaños de I/O pequeños (entre 2 y 8 KB) en donde generalmente se accesan archivos completos, los tiempos de respuesta son del orden de menos de 15 ms y el componente de escritura es de poco más del 30%, algunos ejemplos de estas aplicaciones son Netscape Server, WWW, Excite, Infoseek y World Net.

Sistema de Archivos de Red v2 (NFS, Network File System v2)

NFS v2, implementa las funciones necesarias para compartir archivos, es un protocolo sin estado y se caracteriza porque los tamaños de I/O son de 8 Kb, el componente de escritura es de entre 10 y 30% y los tiempos de respuesta están por debajo de 50 ms. NFS versión 2, está cayendo en desuso y en general su uso debe ser evitado si está disponible la versión 3.

Sistema de Archivos de Red v3 (NFS, Network File System v3)

Una de las características principales de NFS 3 es que las operaciones de escritura son mucho más rápidas que en la versión anterior y se mantiene el mismo modo sin estado de NFS versión 2.

Es importante recordar que en NFS existen dos tipos de operaciones: Operaciones en datos y operaciones en atributos, cada entrada en un sistema de archivos tiene un número de características que describen el archivo o el acceso a él: directorios, ligas, permisos, tamaños, etc. Existen ciertas aplicaciones que hacen más uso de las operaciones en datos mientras que otras hacen mayor uso de las operaciones en atributos.

Las operaciones de NFS versión 3 que son intensas en los atributos se comportan de manera similar a NFS versión 2, algunos ejemplos de estas aplicaciones son los servidores de archivos, aplicaciones de desarrollo de *software*, versiones anteriores de servidores de correo y el protocolo de Sistema Común de Archivos de Internet (CIFS, Common Internet File System).

En NFS versión 3 con operaciones intensas en los datos, los tamaños de I/O son de aproximadamente 32 Kb y debido a sus mejoras en desempeño es muy común encontrar operaciones en archivos de más de 1 MB, el componente de escritura es de entre 20 y 40%. Algunos ejemplos de este son las aplicaciones de visualización científica y las aplicaciones geológicas sobre NFS.

Aplicaciones de Soporte de Decisiones (Data Warehousing)

Estas aplicaciones están más orientadas a información que a transacción, se caracterizan por tener accesos a disco secuenciales con tamaños de I/O grandes del orden de 64 a 2048 KB, el componente de lectura es de más del 90% del total de las operaciones. Ejemplos de aplicaciones de este tipo son las aplicaciones de investigación de mercados y bases de datos como Oracle, Informix y DB2.

Cómputo de Alto Desempeño (HPC, High Performance Computing)

Este tipo de aplicaciones se caracteriza por I/O's grandes del orden de 1 a 16 MB, en general entre más grande sea el tamaño de I/O mejor es el desempeño, algunos ejemplos de estas son las aplicaciones de Modelado Molecular, Modelado del Elemento Finito, Análisis Sísmico, Aplicaciones Meteorológicas y Geológicas.

3.1.4 Seleccionando un Sistema de Almacenamiento

En general las aplicaciones de tipo OLTP se benefician de los sistemas de almacenamiento que tienen un buen rendimiento en términos del número de operaciones de entrada y salida por segundo, por lo que si se está pensando en un sistema de almacenamiento para una aplicación de este tipo el parámetro de desempeño a considerar deben ser los IOPS. Por otro lado para mejorar el desempeño y proteger los datos la opción de RAID más apropiada es RAID 0+1, aunque, RAID 5 puede ser una alternativa para reducir costos, pero debería optarse por una implementación a través de *hardware* debido a las operaciones de paridad que se tienen que hacer en este tipo de RAID.

Las aplicaciones de soporte de toma de decisiones (*Data Warehousing*) sacan el mayor provecho de los sistemas de almacenamiento que ofrecen el mejor ancho de banda, RAID 0+1 es la recomendación para un ambiente de misión crítica en este caso y aunque una implementación de RAID por *hardware* es la ideal, una implementación por *software* es bastante buena sobre todo para RAID 0, 1 ó 0+1.

Es importante considerar cuando se está pensando en el nivel de RAID, que es posible combinarlos incluso en el mismo sistema de almacenamiento dependiendo de las necesidades que se tengan, por ejemplo, en una aplicación de base de datos las tablas puede estar en RAID 5 y los índices en RAID 1 además, si hay ciertos datos que no son críticos y el desempeño es una prioridad también se puede emplear RAID 0.

En cuanto a la unidad de columna la recomendación es elegir unidades pequeñas para aplicaciones con tamaños de I/O pequeños y unidades grandes para aplicaciones con tamaños de I/O grandes, por ejemplo, para una aplicación de tipo OLTP una unidad de columna puede ser de 16 KB mientras que para una aplicación de tipo HPC un tamaño de columna de 256 KB es el adecuado.

Otro aspecto importante a considerar cuando se habla de desempeño es la capacidad de procesamiento del servidor para satisfacer la carga de I/O que se necesite así como las necesidades de crecimiento futuras. Actualmente, existen sistemas de almacenamiento que permiten crecer en tres dimensiones: capacidad, desempeño y disponibilidad lo que constituye una garantía porque en el pasado nada más se podía crecer en capacidad lo que ocasionaba problemas de desempeño a medida que se agregaba más disco.

3.2 Capacidad

Cuando se habla de capacidad es importante resaltar que no es suficiente con tener un estimado de la cantidad de disco que se requiere en el momento de hacer la elección, es necesario también tener un estimado del crecimiento. Es por eso que es determinante que al elegir un sistema de almacenamiento este tenga la escalabilidad para hacer frente a requerimientos de disco inesperados. Actualmente existen proveedores que ofrecen un crecimiento modular en donde es posible agregar arreglos a un mismo gabinete sin requerir conexiones adicionales al servidor y sin que el desempeño se vea degradado, esto es un aspecto interesante a considerar, sin embargo, la limitante de crecimiento estará dada por la capacidad de conexiones del servidor.

Adicionalmente, dependiendo del nivel de RAID a implementar las necesidades de disco crecerán, por ejemplo, para implementar un RAID 1 ó 0+1 es necesario considerar el doble de capacidad de disco de la que se requiere para almacenar los datos, por otro lado en un RAID 5 se requiere aproximadamente el 30% del espacio total en disco para almacenar la información de paridad, además cuando los discos son formateados y se crean sistemas de archivos se ocupa aproximadamente el 10% de la capacidad de ellos.

3.3 Costo

El costo es por supuesto la tercera variable más importante a considerar después del desempeño y la capacidad. De hecho, existe una regla en el ámbito de cómputo que dice: de los tres elige dos. Dependiendo de las prioridades o del presupuesto que se tengan se le tendrá que dar preferencia a dos de estas variables, por ejemplo, el costo por *gigabyte* se va abajo cuando las densidades de los discos suben, aunque las operaciones de I/O también tienden a bajar, por ejemplo si se tienen 72 GB de datos distribuidos a través de diferentes geometrías de discos, se puede esperar aproximadamente cuatro veces más el número de IOPS y ancho de banda sobre ocho discos de 9.1 GB que sobre dos discos de 36 GB, sin embargo, muchos proveedores de sistemas de almacenamiento coinciden en que esta diferencia de desempeño entre discos de diferentes densidades está disminuyendo conforme las tecnologías de discos mejoran.

Cuando se habla de costo, no es suficiente con considerar el precio por *Gigabyte*, hay que tomar también en cuenta el costo de instalación, servicio, configuración, consultoría y capacitación, sin mencionar las licencias de *software* y cualquier otra adecuación que sea necesaria en la infraestructura de cómputo.

3.4 Disponibilidad

La continuidad en los negocios no depende del acceso a la red o del acceso al servidor, depende del acceso a la información, es por eso que cualquier solución de almacenamiento de misión crítica debe garantizar óptimos niveles de disponibilidad.

Aunque ya se ha discutido ampliamente acerca de los niveles de RAID, estos son absolutamente necesarios para poder asegurar la disponibilidad de un sistema de discos, en general, para ambientes de misión crítica RAID 0+1 debería ser siempre implementado, en caso de que el precio sea una limitante, RAID 5 puede ser usado en su lugar. Además, cuando se está configurando un nivel de RAID se debe considerar también la configuración de discos de reemplazo en caliente (*hot spare*), estos discos no contienen datos ni información de paridad, de hecho, la mayor parte del tiempo permanecen en un estado pasivo. Cuando se tiene un disco de reemplazo en caliente en una configuración de RAID y uno de los discos que contiene datos falla, la información de éste es copiada al disco de reemplazo en caliente el cual toma el lugar del disco que falló sin que se interrumpa el servicio, en cuanto el disco que falló es reemplazado, el disco de reemplazo en caliente toma nuevamente su estado pasivo, listo para entrar en acción en caso de que otro disco falle. Los discos de reemplazo en caliente pueden ser configurados para actuar solo en un arreglo de discos o incluso para funcionar a través de diferentes arreglos o controladoras de discos, esta capacidad varía dependiendo del proveedor.

Una buena solución de almacenamiento que garantice los mayores niveles de disponibilidad posibles debe contar con componentes redundantes, estos componentes pueden variar dependiendo del modelo y proveedor, pero es común encontrar soluciones que tengan fuentes de poder y enfriamiento redundantes, con conexiones redundantes a la corriente eléctrica y en el caso de soluciones con controladora de RAID, con doble controladora y memoria caché en espejo así como baterías para proteger la información almacenada en la memoria caché en caso de una interrupción eléctrica. Un aspecto adicional a esta característica de redundancia es la capacidad de conexión en caliente de las partes la cual agiliza significativamente la reparación de una falla.

Aunado a esto, algunos sistemas tienen la capacidad de múltiples conexiones al servidor que junto con un componente de *software* pueden ser configurados para mejorar el ancho de banda, para hacer balanceo de cargas y para mejorar la disponibilidad ya que en el evento de que una conexión falle se puede tener una de respaldo que garantice la continuidad del acceso a la información.

Cuando se tienen que hacer comparaciones del nivel de disponibilidad que ofrecen diferentes sistemas de discos, en ocasiones se toma como punto de comparación el tiempo entre fallas (MTBF, *Mean Time Between Failures*), sin embargo, no es el mejor punto de comparación ya que esta medida solo está basada en *hardware* y si se considera que del 100% de fallas que se dan en un sistema de cómputo aproximadamente solo el 20% son atribuibles al *hardware* se está olvidado el 80% restante, además, es muy difícil hacer comparaciones ya que el MTBF está íntimamente ligado con el número de componentes del sistema, por lo que cuando se están haciendo comparaciones es más recomendable considerar los niveles de disponibilidad en lugar del tiempo entre fallas. Otro punto importante es que la disponibilidad no solo está dada por el *hardware* sino además por el *software*, el nivel de soporte, la experiencia y conocimiento de los operadores, el uso de partes en sitio así como del diseño de buenos planes para prevenir y combatir situaciones de contingencia.

3.5 Soporte

Este punto está íntimamente ligado con la disponibilidad, sin embargo, es tan importante que se debe tratar por separado. Aunque ya se tenga identificado el sistema de almacenamiento con la configuración que se necesite en términos de alta disponibilidad y desempeño no debe pasarse por alto el aspecto de servicio. Existen diversos modelos de servicio que se pueden ajustar a demandas específicas, algunos de los aspectos a considerar son:

- Monitoreo, que se pueda llevar a cabo de forma remota ya sea por parte del cliente o incluso del mismo proveedor.
- Tiempos de respuesta y de solución requeridos, en donde incluso se puedan considerar sanciones al proveedor en caso de que estos no se cumplan.
- Definición de responsabilidades en cuanto a quien debe solucionar problemas, reportar fallas o hacer pruebas con el equipo.

- Esquema de soporte, 24 horas al día, 7 días a la semana, 365 días al año, si así se requiere.
- Posibilidad de tener refacciones en el sitio, para minimizar los tiempos de solución.
- Consultoría de parte del proveedor para establecer políticas de respaldo y planes de recuperación contra desastres.

3.6 Ambiente de operación

Generalmente este punto se pasa por alto y es sorprendente ver como se puede tener ya el equipo en el centro de cómputo solo para darse cuenta de que las condiciones de operación en cuanto a energía, enfriamiento o espacio físico no se cumplen, o que simplemente no es posible llevar el equipo hasta el centro de cómputo porque no cabe en el elevador del edificio.

En general las características de operación que se deben tener en cuenta son:

- Requerimientos de energía
 - 220 volts / 110 volts
 - Consumo total de energía medido en unidades de potencia el cual puede variar dependiendo del proveedor y del tipo de arreglo de discos
 - Fuentes de poder ininterrumpibles (UPS's, *Uninterrupted Power Supply*)
 - Fuentes de toma de energía redundantes que permitan obtener la energía desde lugares diferentes para garantizar la disponibilidad de energía a los equipos.
- Requerimientos de enfriamiento
 - Temperaturas de operación que generalmente van de 5 a 35°C
 - Humedad
- Espacio físico requerido
- Espacios de servicio necesarios
 - Además del espacio que ocupa el equipo es necesario considerar un espacio libre para cuestiones de servicio y mantenimiento el cual debe ser el necesario para que por lo menos una persona pueda realizar alguna actividad de servicio o mantenimiento en el equipo.
- Peso
- Infraestructura requerida
- Requerimientos de distancia
 - En caso de que el sistema de almacenamiento y el servidor estén en lugares distintos
- Conectores necesarios
- Aspectos seguridad y acceso

Aunque estos aspectos no parecen tener un impacto en la selección de una solución de almacenamiento en ocasiones pueden influir para terminar eligiendo una solución con menor desempeño o más cara por no tenerlos en cuenta.

3.7 Soluciones de Respaldo

La necesidad de hacer respaldos y recuperar la información de una manera rápida y confiable ha alcanzado un nuevo nivel de importancia debido al incremento dramático en la cantidad de datos que las compañías de hoy en día manejan así como el impacto que el manejo de la información puede tener en el éxito o fracaso de un negocio. Cualquiera estaría de acuerdo en afirmar que el respaldar la información que se maneja en una empresa no es una opción es más bien una obligación, pero no es suficiente con tener una biblioteca de cintas que permita hacer respaldos, es necesario hacer un estudio cuidadoso de las necesidades de respaldo que se tienen para definir la mejor solución posible.

3.7.1 Identificando el ambiente a respaldar

Para poder hacer un buen estudio que permita determinar la mejor solución de respaldos en términos del servidor o servidores de respaldo necesarios, de la capacidad y características de las bibliotecas de cintas, de la infraestructura de comunicaciones y funcionalidad del *software* es necesario entender el ambiente que será respaldado.

Volumen de datos

Uno de los primeros aspectos a considerar es la cantidad total de datos que será respaldada, en este sentido existen ciertos aspectos como el número de archivos a respaldar o la longitud de las rutas de directorios que pueden hacer crecer el tamaño de los índices de la base de datos de respaldos y que puede tener un impacto en el desempeño. En esta base de datos de respaldos se mantiene un registro por cada instancia de archivo, dicho registro mide entre 150 y 200 *bytes*, aunque puede variar dependiendo del *software* de respaldo que se use, esto significa que para respaldar un millón de registros de una base de datos se necesitará entre 143 y 191 MB de espacio en disco para guardar los índices de la información de respaldos.

Otro aspecto importante es el tamaño y periodicidad de los cambios. Estos determinan la frecuencia con la que deben hacerse los respaldos y la cantidad de datos escrita durante respaldos incrementales, aquí es importante determinar si se deben hacer respaldos completos de la información o solo se debe respaldar la información que está cambiando.

Mientras que es más rápido guardar solo las porciones que cambian, es también más fácil recuperar directorios o sistemas de archivos de respaldos completos que de respaldos incrementales, esto se debe al proceso de recuperación de datos: Para recuperar información de un respaldo incremental primero se necesita recuperar un respaldo completo y después los respaldos incrementales hasta la última versión, este proceso puede ser relativamente lento y poco deseable, por lo que la opción de realizar respaldos completos o incrementales dependerá de lo que es más importante: hacer respaldos rápidos que no afecten la operación o recuperar rápidamente la información en caso de una emergencia.

Como regla general se puede estimar aproximadamente el doble de tiempo para recuperar información comparado con el tiempo que se lleva respaldar los mismos datos. La principal razón de esta discrepancia de tiempo se debe a la naturaleza de las lecturas comparado con las escrituras. Otra razón es el tiempo que se lleva recuperar determinados archivos especialmente cuando se tienen grandes bases de datos con millones de registros.

Una buena práctica cuando se está pensando en la capacidad de la biblioteca de cintas es que esta sea configurada para soportar de tres a cinco veces el volumen total de datos, esto es, para asegurar la capacidad de poder hacer múltiples respaldos de la información: totales e incrementales con diferentes niveles de periodicidad.

Tipos de datos

El tipo de datos a respaldar está íntimamente ligado con el nivel de compresión que se puede esperar al realizar un respaldo, no es posible tener el mismo factor de compresión en archivos de texto que en archivos gráficos. El nivel de compresión depende en mucho de la redundancia en los datos, algunos tipos de datos, video, por ejemplo, tienen muy poca redundancia que eliminar, por lo tanto estos datos tendrán muy poca compresión y no se recomienda usar esta característica por que no ofrece muchos beneficios para este tipo de datos en particular. Debido a la gran variedad de datos así como a las limitantes propias de las unidades de respaldo no se pueden esperar los niveles de compresión especificados por el fabricante. Sin embargo, no se puede negar que la compresión tiene dos ventajas fundamentales: aumentar la velocidad en la que los datos son procesados y compactar los datos escritos a cinta de manera que esta pueda almacenar más información.

La figura 3-2 muestra las velocidades de procesamiento y el radio de compresión que se puede esperar para diferentes tipos de archivos.

modo	velocidad de procesamiento	radio de compresión
ninguno	1:1	1:1
texto	1.46:1	1.44:1
JPG	0.93:1	0.92:1
base de datos	1.60:1	1.57:1
servidor de archivos	1.60:1	1.63:1
servidor de web	1.57:1	1.82:1

Tabla 3-2. Velocidades de procesamiento y radios de compresión de diferentes tipos de archivos

Estructura de archivos

Otro aspecto relacionado con el desempeño al realizar respaldos es si el respaldo se hará sobre dispositivos crudos o sistemas de archivos. El respaldo sobre dispositivos crudos no involucra mucha intervención de la aplicación o del sistema operativo, además el acceso a disco es largo y secuencial lo que minimiza la carga del sistema, sin embargo, no se pueden recuperar archivos específicos de un dispositivo crudo ya que este es tratado como una sola unidad, por lo que para recuperar una sola porción de datos se tiene que recuperar el dispositivo crudo completo lo que toma más tiempo.

En contraste el respaldar sistemas de archivos agrega más carga al sistema ya que los datos son almacenados primero en la memoria virtual y después transferidos a cinta, sin embargo, en el proceso de recuperación es posible recuperar porciones específicas de datos sin tener que recuperar el sistema de archivos completo.

Actualmente existen diferentes implementaciones de sistemas de archivos como el acceso directo de I/O que permite un desempeño parecido al de los dispositivos crudos, esta variante de sistema de archivos no copia los datos a memoria virtual antes de realizar el respaldo como se hace comúnmente lo que reduce ciclos de reloj y uso de memoria que se traducen en mayor rapidez al hacer el respaldo.

Origen de los datos

El saber de donde vienen los datos puede ayudar a planear un buen esquema de respaldos. Por ejemplo, cuando el servidor donde están los datos es el encargado de hacer los respaldos, la complicación de la configuración de la red es eliminada y la planeación se puede enfocar exclusivamente a las capacidades de procesamiento del servidor y de la biblioteca de cintas. Cuando esta es la situación, los respaldos son mucho más eficientes cuando se evita la copia de datos a memoria virtual usando I/O directo o respaldando el dispositivo crudo en lugar del sistema de archivos. La memoria del sistema es otro factor importante y también se puede optimizar su uso al eliminar la copia de los datos a memoria virtual, si esto no es posible se debe configurar el sistema para que las peticiones de memoria sean más agresivas cuando se está realizando el respaldo.

En cambio, cuando es necesario respaldar información almacenada en clientes remotos, que es la situación más frecuente, es necesario enfocarse en los requerimientos de red para poder cumplir con las ventanas de respaldo.

No todas las redes se comportan de la misma manera, diferentes tecnologías de red tienen diferentes propiedades. Por ejemplo, la tecnología *Ethernet* tiende a ser barata y por lo mismo muy usada, pero su ámbito de aplicación tiende a ser limitado a redes de área local, además la naturaleza de ésta causa que el ancho de banda se degrade conforme más nodos son agregados a la red. Por otro lado, ATM (*Asynchronous Transfer Mode*) y FDDI (*Fiber Distributed Data Interface*) tienden a ser más usadas en ambientes de red de área amplia (WANs, *Wide Area Networks*) ya que la degradación es menor bajo cargas de trabajo pesadas, sin embargo, éstas utilizan conexiones de fibra óptica que las hacen más difíciles y caras de implementar.

En una red típica empresarial es muy común encontrar múltiples segmentos y varias tecnologías de red. El ancho de banda disponible puede variar considerablemente, por lo que es de suma importancia determinar el ancho de banda que se tiene para cada ruta entre servidor y cliente, otro aspecto a considerar es la segmentación de la red, que usada adecuadamente puede ayudar a disminuir la carga conforme se tengan más clientes.

Ventanas de respaldo

La ventana de respaldo es el tiempo del que se dispone para respaldar los datos. Esta es una de las variables más importantes a considerar en el diseño de una solución de respaldo ya que probablemente constituye el requerimiento principal a ser cumplido.

Un aspecto a considerar para cumplir con el tiempo que se tiene para respaldar los datos es la factibilidad de hacer respaldos completos e incrementales. Es una buena idea hacer respaldos completos semanal o mensualmente y los incrementales diariamente para minimizar el tiempo requerido para respaldar los datos.

Cuando el tiempo del que se dispone para hacer respaldos es mínimo es importante considerar si la no disponibilidad de los datos es aceptable, de esta manera se puede implementar un esquema de respaldos en donde los usuarios tengan acceso al sistema sin tener acceso a los datos mientras se hace el respaldo, si es necesario que los datos estén disponibles todo el tiempo entonces el factor a sacrificar es el desempeño si esto es permisible, en el caso de las bases de datos por ejemplo, existen soluciones de *software* que permiten que la base de datos entregue una fotografía de las tablas de la base datos con diferentes bloques de datos (*streams*) para minimizar el tiempo requerido de respaldo y permitir que la información continúe siendo accesada por lo usuarios pero con una degradación del tiempo de respuesta.

3.7.2 Planeando una solución de respaldo

El sistema de respaldos

En general, los dispositivos de respaldo son fáciles de predecir y el determinar el mejor sistema de respaldos está dado básicamente en función de las especificaciones del dispositivo en cuanto a capacidad y desempeño, de hecho muchos coinciden en afirmar que la dificultad de estos estriba en la instalación y configuración más que en la planeación.

Para ambientes de misión crítica la mejor opción en términos del sistema de respaldos la constituyen las bibliotecas de cintas las cuales ofrecen múltiples unidades de cinta y capacidades de respaldo de cientos de *gigabytes* con un mecanismo que permite manipular las cintas de forma automática. En dichos ambientes en donde se requiere tener el mayor nivel de disponibilidad posible incluso en el esquema de respaldos es más recomendable tener un par de bibliotecas de cintas de tamaño medio o pequeño que una sola grande.

La determinación del número y tipo de unidades de cinta así como la interfase a usar es vital para poder cumplir con los requerimientos de desempeño, no se debe olvidar que diferentes unidades de cinta tienen diferentes características, capacidades y formatos que deben ser considerados. La tabla 3-3 muestra las características de capacidad y velocidad de transferencia de datos de los formatos de cinta más populares en el mercado.

Dispositivo	Capacidad (MB)		Velocidad de Transferencia	
	Sin Comprimir	Comprimido	Sin Comprimir	Comprimido
	1:1	1.4:1	1:1	1.4:1
DDS-4	24	30	2	2.8
Exabyte 8900	20	28	3	4.2
DLT 7000	35	49	5	7
DLT 8000	40	73	6	9
STK 9840	20	28	9	12.6
STK Redwood	50	71.5	10.5	14.7

Tabla 3-3. Capacidades y velocidades de transferencia de diferentes unidades de cinta

La red

Cuando se tiene que planear una infraestructura de red para realizar respaldos, el objetivo es configurar suficiente ancho de banda entre el punto donde se encuentran los datos y el punto en donde se encuentra la biblioteca de cintas. La mejor solución de almacenamiento es aquella en la que se puede tener una red dedicada a los respaldos. La tabla 3-4 muestra las velocidades de transmisión para diferentes tecnologías de red.

Tecnología	Velocidad Teórica	Velocidad Nominal
	Mbits/seg	MB/seg
Modem	0.056	0.004
ISDN	0.128	0.010
Frame Relay 512	0.512	0.039
T-1	1.54	0.115
T-3	44.7	3.4
Ethernet	10.0	1.0
Fast Ethernet	100.0	8.0
Gigabit Ethernet	1000.0	50.0
FDDI	100.0	8.0
ATM 155	155.0	11.0
ATM 622	622.0	50.0

Tabla 3-4. Velocidades de transferencia de datos para diferentes tecnologías de red

Una vez que ya se tiene estimado el ancho de banda que se puede tener en la red así como en el sistema de cintas y una vez que se conoce la cantidad de datos a respaldar se puede determinar el tiempo que se tomaría transferir los datos desde el origen hasta el destino con la siguiente fórmula: $\text{Tiempo (segs)} = \text{Tamaño de los datos (MB)} / \text{Ancho de Banda (MB/seg)}$

Por otro lado, mientras que nuevos dispositivos de almacenamiento solo necesitan ser conectados y configurados, cambiar una infraestructura de red puede significar hacer muchos esfuerzos económicos que muchas veces necesitan ser planeados con años de anticipación, por lo que en muchas ocasiones, es necesario adaptarse a la infraestructura de red existente y aquí el objetivo es colocar los servidores de respaldo en los lugares en donde se pueda aprovechar el mayor ancho de banda posible y en donde esto no sea posible implementar una solución de respaldo híbrida: local y remota.

El sistema de discos

Después de la red, el sistema de discos constituye el segundo cuello de botella que se puede tener en una solución de respaldos. Todos los aspectos de configuración del sistema de discos en cuanto a la unidad de columna, el nivel y tipo de RAID así como las capacidades de ancho de banda y I/O repercuten también en la velocidad a la que se pueden respaldar y recuperar los datos.

Distribución de los dispositivos de respaldo

Como una regla el dispositivo de respaldos debería estar tan cerca como sea posible de la información. En caso de que las unidades de respaldo estén lejos de donde residen los datos, la red necesita ser capaz de soportar las velocidades de transferencia necesarias para moverlos, de lo contrario será un continuo dolor de cabeza.

Una forma efectiva de implementar una solución de respaldos sobre todo en grandes ambientes de cómputo es conectar los dispositivos de cinta en los servidores donde residen los mayores volúmenes de información.

Servidores de Respaldos

La configuración de un servidor de respaldos se puede simplificar siguiendo reglas simples como el considerar un procesador por cada 4 ó 5 unidades de cinta de la biblioteca de respaldos, otro aspecto a considerar es el número de ciclos de reloj requeridos para mover datos a determinada velocidad, en el caso de algunos servidores Unix por ejemplo, estudios han demostrado al momento de escribir este trabajo de tesis que se requieren entre 5 y 10 Mhz para mover 1 MB/seg de datos, esto quiere decir que si se requiere hacer un respaldo local a cinta con una velocidad de transferencia de 35 MB/seg se necesitarán 350 Mhz de poder de CPU. En el caso de la memoria se pueden seguir las reglas de configuración que dicte cada proveedor del equipo, una muy buena práctica consiste en desactivar la copia a memoria virtual.

El software

El *software* determina como será usado el *hardware* para hacer el respaldo y recuperación de la información. El *software* no tiene ningún impacto en el desempeño del respaldo pero si en la manera en la que este se haga. En la actualidad existen diferentes productos de respaldo que ofrecen diversas características y ventajas dependiendo del ambiente a respaldar por lo que se deberá tener cuidado de seleccionar el producto que realmente cubra con las necesidades que se tengan.

Ambiente de operación

Al igual que con cualquier otro dispositivo de cómputo las bibliotecas de cintas necesitan operar en un ambiente controlado de temperatura y humedad por lo que se debe contar con las condiciones adecuadas para evitar fallas en los dispositivos de respaldo o cintas. Otra consideración importante es la limpieza de las cintas y dispositivos de respaldo que generalmente se realiza a través de una cinta limpiadora. Equipos de respaldo que operen en condiciones de polvo, por ejemplo, cerca de impresoras o fuera de las especificaciones de operación recomendadas deben ser limpiados más frecuentemente.

Soluciones de Protección de datos

Una solución de respaldo debe dar respuesta a las preocupaciones reales y objetivos del cliente. Por ejemplo, una buena solución para recuperar archivos borrados accidentalmente no es la mejor solución para una recuperación contra desastres. En ocasiones una solución de respaldo tradicional en donde los respaldos se realizan a través de la red puede no satisfacer completamente las expectativas de protección de datos. Es cierto que tanto los dispositivos de cinta como el *software* han mejorado considerablemente en los últimos años en términos de desempeño, capacidad y confiabilidad pero también es cierto que los volúmenes de información y las necesidades de disponibilidad han crecido a la par.

En la actualidad existen soluciones de protección de datos que se pueden considerar como una alternativa o como una solución conjunta al esquema tradicional de respaldos, una de estas alternativas de solución la constituyen las utilerías de copia instantánea que permiten tomar fotografías de la información en un momento determinado del tiempo para poder usar después esta copia en actividades secundarias como es el caso de respaldos o pruebas sin interrumpir la operación normal del sistema. La ventaja de esta técnica es que una vez que se ha hecho la copia de los datos se mantiene un archivo que registra los cambios entre el disco primario y secundario y solo se copian las diferencias que existen entre ambos cuando se requiera hacer otra copia instantánea, esto reduce significativamente el tiempo de respaldo ya que la copia resultante se puede utilizar para hacer el respaldo local.

Los métodos de protección contra desastres son otra gran preocupación de los administradores de un centro de datos, la mejor solución en estos casos es tener un sitio alternativo que pueda salir a la operación en caso de que el sitio primario quedara fuera de servicio, para este sentido existen diferentes soluciones una de ellas quizá la más compleja la constituyen las soluciones de agrupaciones de servidores a distancia en las que se pueden tener dos sitios equipados de la misma manera tanto en *hardware* como en *software* y configurados de tal forma que exista una comunicación constante entre ellos para que en el caso de que uno de los servidores falle el otro pueda tomar total control de la operación, este tipo de solución es bastante costoso ya que implica redundancia en todos los componentes de *hardware* y *software* así como una buena infraestructura de comunicaciones sobre todo si los dos sitios se encuentran bastante separados entre sí.

Otra alternativa de solución consiste en utilizar replicación de datos en tiempo real, esta solución resulta ser menos costosa pero debe haber intervención manual para que el sitio remoto pueda tomar control en caso de un desastre.

3.8 Crecimiento futuro

Una decisión sobre una solución de almacenamiento hecha el día de hoy será algo con lo que se tendrá que vivir por lo menos durante tres años, por lo que un plan de crecimiento en capacidad y funcionalidades futuras es algo que no debe olvidarse como parte de cualquier estudio de almacenamiento.

Los requerimientos más importantes a considerar en términos de crecimiento son:

- Escalabilidad y desempeño.
- Adopción de nuevas tecnologías.
- Conectividad y soporte a ambientes heterogéneos.

Escalabilidad y desempeño

Anteriormente los planes de crecimiento se hacían por cinco años ahora este período se ha reducido a tres o incluso menos tiempo, aunque esto también depende del presupuesto, sin embargo, aún si el presupuesto es limitado es necesario considerar como parte de éste los recursos necesarios para hacer frente a necesidades futuras de almacenamiento. Hasta hace poco se solía hacer estimaciones del crecimiento futuro en base a la experiencia del pasado, esto ya no funciona. El crecimiento explosivo de los datos inherente a esta nueva economía de red ha cambiado dramáticamente el proceso de hacer estimaciones de crecimiento. De hecho algunas empresas ni siquiera tienen una experiencia pasada de crecimiento, por lo que como parámetro se puede utilizar la experiencia de compañías similares en cuanto al giro de negocio, infraestructura de cómputo, tipo de aplicaciones y número de usuarios para hacer un estimado. Por eso es importante que la solución de almacenamiento a elegir permita una buena escalabilidad incluso en el mismo gabinete, requiriendo el mínimo de conexiones posibles y en donde no se vean comprometidos el desempeño y la disponibilidad.

Adopción de nuevas tecnologías

Sin lugar a dudas la industria de cómputo es el área que evoluciona más rápidamente y en donde se presentan más cambios. Este es otro punto a considerar cuando se está eligiendo una solución, ya que la visión del proveedor y un claro plan de evolución de los productos que ofrece pueden ser la diferencia entre comprar un producto que se vuelva obsoleto en un período relativamente corto de tiempo o entre un equipo que sea capaz de adaptarse a los cambios de la tecnología.

Conectividad y soporte a ambientes heterogéneos

Prácticamente todas las corporaciones en la actualidad poseen un ambiente heterogéneo en donde convergen diferentes plataformas de cómputo, diferentes sistemas operativos y aplicaciones por lo que es muy importante que una solución de almacenamiento sea abierta en términos del soporte a estándares. En ocasiones también se vuelve necesaria la capacidad de tener un sistema de discos que se pueda conectar a diferentes servidores con soporte a diferentes sistemas operativos, sin embargo, las soluciones que permiten esta conexión multiplataforma son propietarias y costosas y realmente no comparten los mismos volúmenes entre los diferentes servidores, algunos de estos proveedores utilizan técnicas de copia de datos para permitir el acceso de los datos entre servidores, otros utilizan algún mecanismo de *software* como el que se utiliza en las soluciones de alta disponibilidad para poder compartir volúmenes entre servidores. Las tecnologías que permiten la conexión de un sistema de discos a varias plataformas de cómputo todavía están mejorando y las Redes de Almacenamiento Inteligente constituyen una de las principales promesas para lograr esto de la mejor manera.

IV

Redes y Almacenamiento

El gran incremento en los volúmenes de información que las empresas están experimentando, la necesidad de responder rápida y eficientemente a miles de usuarios conectados a través del Internet, la proliferación de grandes aplicaciones que consumen cada vez más recursos, la necesidad de compartir la información a través de diferentes aplicaciones y plataformas de cómputo y la obligación de mantener el servicio que se presta de forma ininterrumpida a la vez que se trata de minimizar costos han puesto en tela de juicio las capacidades del esquema tradicional de almacenamiento en donde los arreglos de discos se encuentran directamente conectados a los servidores y han hecho que surgan nuevas alternativas de solución como es el caso del modelo de almacenamiento conectado a la red y de las redes de almacenamiento inteligente.

Para poder entender estas nuevas alternativas de solución es importante conocer las características que las tecnologías de red ofrecen tanto para servidores como para sistemas de discos y la forma como las redes están siendo usadas como parte de estos nuevos esquemas de almacenamiento.

4.1 La Red

Independientemente del medio utilizado, de la topología, protocolo o de su cobertura geográfica las redes comparten una serie de características que también son la base para los nuevos esquemas de almacenamiento.

Transporte Serial.

Las redes utilizan un medio de transporte serial. Los ceros y unos digitales que constituyen los datos son transmitidos secuencialmente a través de las redes. La transmisión serial permite que los datos puedan viajar a través de grandes distancias y a una velocidad constante. La transmisión sobre fibra óptica, por ejemplo, puede extenderse a decenas de kilómetros. Por el contrario, la transmisión paralela debe usar múltiples líneas para enviar grupos de datos concurrentemente lo que provoca que el medio de transmisión sea más grande y complejo en su diseño.

Método de Acceso

Un dispositivo en la red necesita de un método de acceso para obtener el control del medio de transporte. *Ethernet* utiliza un método de sensado de portadora y de detección de colisiones, mientras que *Token Ring* proporciona una señal que puede ser solicitada por un solo usuario a la vez. Independientemente del tipo de transporte, debe existir un control de acceso para asegurar la integridad de los datos.

Direccionamiento

Cada dispositivo en una red debe tener una identidad única la cual es establecida a través de una dirección. Una tarjeta de red *Ethernet*, por ejemplo, tiene una dirección única conocida como control de acceso al medio (MAC, *Media Access Control*) la cual es asignada por el fabricante, esta dirección es usada para el intercambio de información junto con una dirección de red asignada por el administrador.

Empaquetado de Datos

Los datos enviados a través de la red son manejados en forma de paquetes. Un archivo gráfico, por ejemplo, debe estar dividido en múltiples paquetes para que pueda ser transmitido a través de la red. Cada paquete contiene una porción del archivo original así como información adicional que permite su transmisión, cuando todos los paquetes que conforman un dato determinado alcanzan su destino, el dato original es obtenido quitando la información adicional que se utilizó para su transmisión y uniendo todos los paquetes que lo forman.

Ruteo de paquetes

Las redes consisten de múltiples segmentos que son unidos a través de ruteadores o concentradores de conexiones. La segmentación física de una red es necesaria para evitar cuellos de botella, por lo que si un usuario de un segmento de red necesita comunicarse con otro en un segmento diferente el dato debe ser manejado por un ruteador para lograr este propósito. El manejo del ancho de banda en los segmentos de red y la colocación apropiada de ruteadores para la comunicación es fundamental en el diseño de redes.

Además de la comunicación de diferentes segmentos de red, el ruteo también permite que se puedan tener accesos redundantes, de tal manera que si una conexión falla se pueda seguir teniendo acceso a través de la otra conexión.

4.2 Modelos de Almacenamiento

Gracias a las ventajas que ofrecen las tecnologías de red han aparecido dos modelos de almacenamiento adicionales al esquema de almacenamiento tradicional, estos tres modelos están definidos básicamente por su forma de conexión, pero también por el nivel de disponibilidad, escalabilidad, desempeño y facilidad de administración que ofrecen.

- Almacenamiento Conectado Directamente al Servidor
- Almacenamiento Conectado a la Red
- Redes de Almacenamiento Inteligente

La elección de un modelo en particular dependerá de la funcionalidad y características que se estén buscando, aunque en algunos casos la mejor solución de almacenamiento será una combinación de los tres modelos. La figura 4-1 muestra la manera en que estos tres esquemas se ven en un ambiente empresarial.

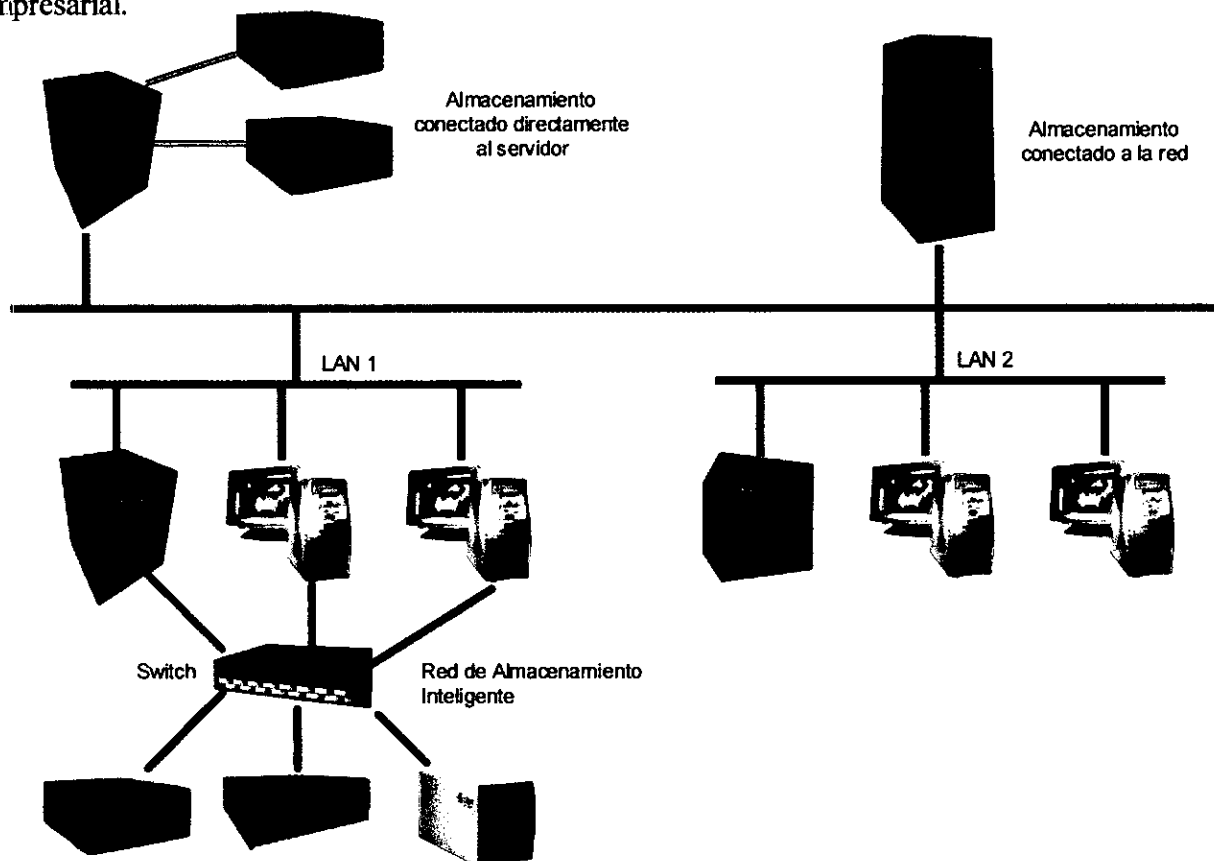


Figura 4-1. Modelos de Almacenamiento

4.2.1 Almacenamiento Conectado Directamente al Servidor

Como se muestra en la figura 4-2, este modelo también conocido como DAS (*Direct Attached Storage*), está representado por una topología de red simple en la que todos los componentes de almacenamiento son locales a cada nodo. Los servidores, estaciones de trabajo y PC's son agregados alrededor de la red para crear un ambiente en el que todas las aplicaciones corren en su propio servidor, por lo tanto cada sistema de almacenamiento puede ser configurado de acuerdo a las características propias de la aplicación obteniéndose ventajas significativas en términos de desempeño.

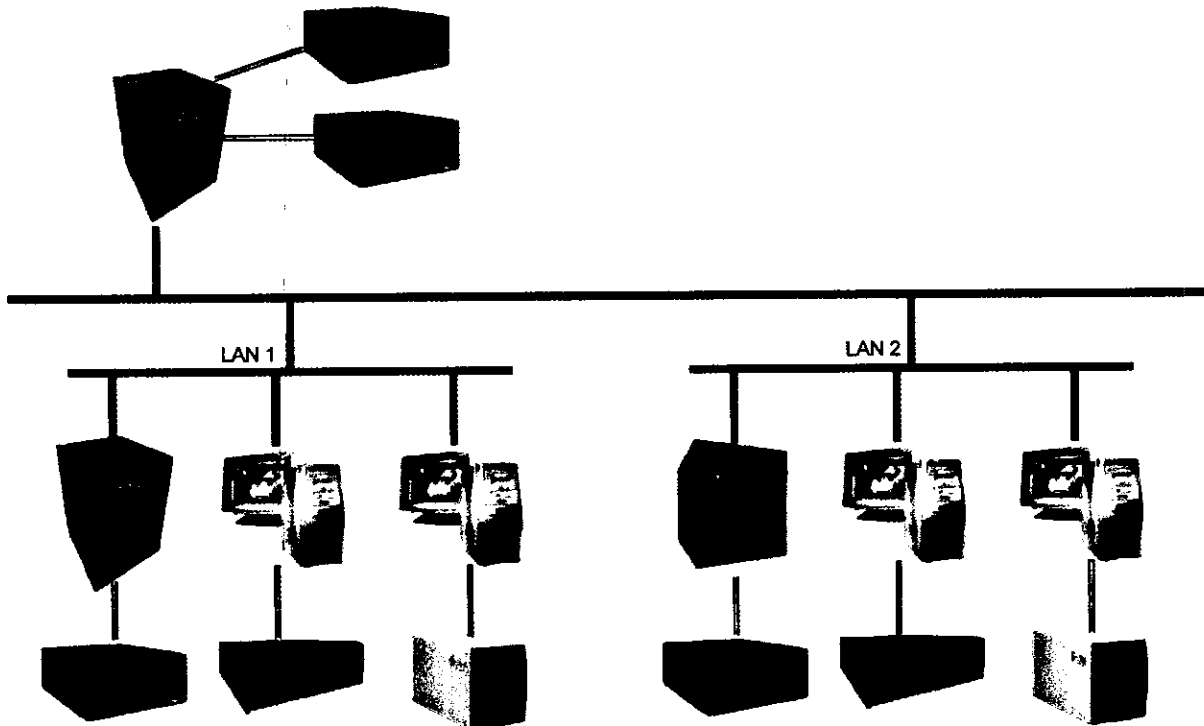


Figura 4-2 Almacenamiento conectado directamente al servidor

El método más común para conectar dispositivos de disco directamente al servidor es a través del canal SCSI aunque algunas instalaciones también utilizan el canal de fibra, sin embargo, SCSI es el medio predominante. La tecnología SCSI utiliza un medio de transmisión paralela por lo que existe una limitante en la distancia a la que puede estar el servidor y el arreglo de discos de entre 15 y 25 metros. Adicionalmente por la forma de transmisión, los cables tienden a ser complejos y es necesario un cableado preciso y terminadores en los puertos que no se utilicen para que se asegure la integridad en la transmisión de los datos. Es posible tener conectados múltiples dispositivos en el mismo canal a través de una conexión en cascada, sin embargo, también existe un límite en este número el cual varía dependiendo de la implementación de SCSI siendo el número máximo 16, por lo tanto si se requieren grandes capacidades de almacenamiento serán necesarias múltiples tarjetas SCSI tal como se muestra en la figura 4-3, el uso de múltiples canales SCSI también es recomendable por desempeño pero esto hace más compleja y costosa la configuración.

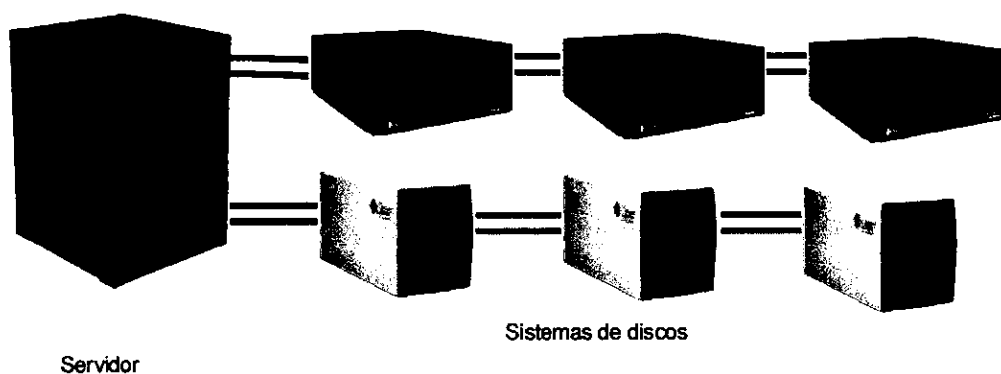


Figura 4-3. Servidor con múltiples conexiones SCSI

La responsabilidad de administrar los datos en este modelo reside en el administrador local del sistema (el usuario final en el caso de las PC's o estaciones de trabajo), los cuales deben asegurarse de la disponibilidad, accesibilidad, desempeño y protección de los datos almacenados. Esto tiene grandes desventajas ya que en la mayoría de los casos los usuarios no cuentan con el conocimiento y experiencia necesarios para hacerse cargo de estas tareas, incluso cuando se sabe como hacerlo, este tipo de tareas de administración a menudo son dejadas en segundo plano. Por otro lado, el costo de pertenencia de los equipos de cómputo se incrementa considerablemente al tener una persona dedicada por cada equipo.

En cuanto a la disponibilidad ésta dependerá de la factibilidad de tener conexiones redundantes, de las características propias del sistema de discos y del servidor así como de la configuración de RAID que se tenga. Es posible tener una configuración en espejo entre sistemas de almacenamiento, lo cual es una muy buena idea si se requiere tener un buen nivel de protección para los datos. En el caso de que se necesite el mayor nivel de protección posible una solución de agrupación de servidores (*cluster*) es lo más recomendable.

En lo que a escalabilidad se refiere, ésta dependerá de la capacidad de crecimiento del sistema de discos así como del número de conexiones de I/O que soporte el servidor, en la medida que se alcancen esos límites el agregar más disco significará tener que agregar un servidor más al ambiente de cómputo lo que implica una doble inversión.

En términos de respaldos, debido a que toda la información es local y no se hace mucho uso de la red, la infraestructura de comunicaciones tiende a ser simple en su diseño y fácil de administrar, sin embargo, para poder respaldar la información de cada nodo se necesitará de algún dispositivo de respaldo también local lo que dificulta la implementación de políticas de respaldo por lo que la información constantemente correrá el riesgo de perderse o de dársele mal uso, a menos que exista una persona o grupo de personas y un esquema centralizado de respaldos dedicado a salvaguardar la información contenida en todos los nodos de la red.

Por otro lado, el florecimiento de Internet y de aplicaciones como ERP así como del modelo cliente/servidor hacen que cada vez sea más necesario que la información sea compartida por diferentes plataformas, equipos y aplicaciones, lo cual es muy difícil con este esquema.

En la actualidad, esta forma de conexión es la más utilizada al grado de que más del 90% de los sistemas de discos son DAS, de acuerdo a los analistas este porcentaje tenderá a disminuir en los próximos años por las ventajas que parecen ofrecer los otros dos modelos, sin embargo, este seguirá vigente por mucho tiempo más y su uso seguirá siendo recomendado cuando no se necesite compartir información, cuando los niveles de disponibilidad y escalabilidad ofrecidos por este modelo sean suficientes y cuando se quiera tener un sistema de discos configurado específicamente para las necesidades de la aplicación y de los datos que serán almacenados en ella.

4.2.2 Almacenamiento Conectado a la Red

Este modelo también conocido como (NAS, *Network Attached Storage*) se caracteriza por tener el sistema de discos conectado directamente a la red de producción tal como se muestra en la figura 4-4. Este esquema permite que el almacenamiento esté mejor distribuido a través de la red, ya que es posible configurarlo de tal manera que esté en donde se necesite. Algunos analistas, consideran que este esquema distribuido es mucho mejor en términos del mejor aprovechamiento de los recursos que el tener los discos conectados directamente a los servidores o incluso que tener una red de almacenamiento, ya que con estos dos últimos esquemas se puede generar más tráfico en ciertos segmentos de la red que en otros.

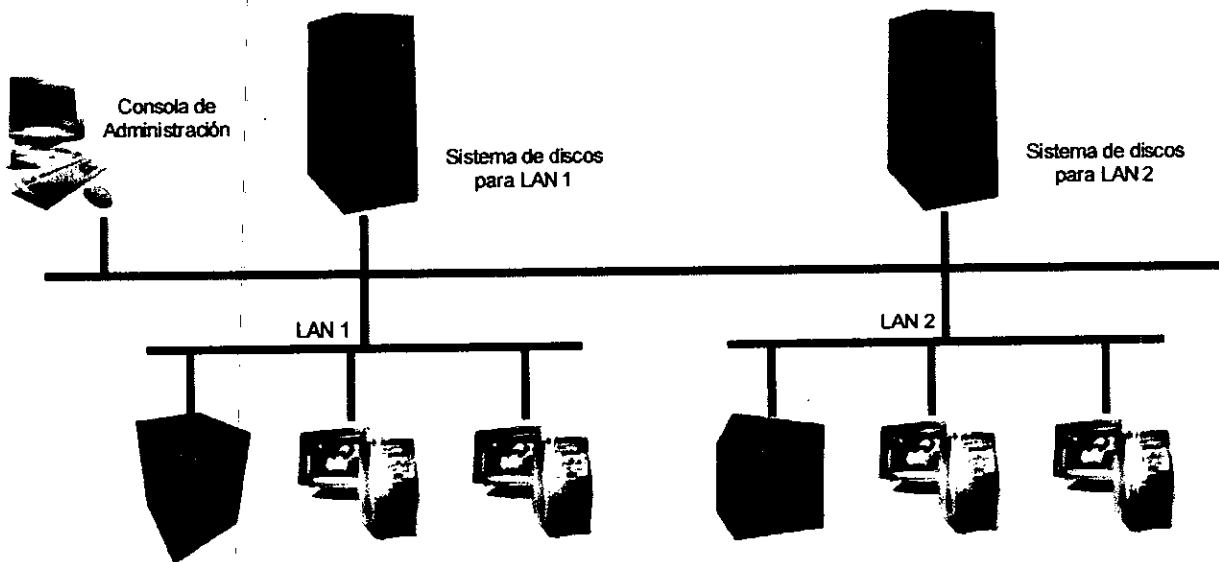


Figura 4-4. Almacenamiento conectado a la red

Este modelo resuelve muchos de los problemas asociados con la conexión directa al servidor aunque también tiene sus puntos débiles. Un dispositivo NAS libera al almacenamiento de su dependencia al servidor a la vez que elimina los problemas de cableado y distancia inherentes a la tecnología SCSI. El acceso a éste es prácticamente universal ya que teóricamente cualquier dispositivo que pueda ver el arreglo a través de la red será capaz de accederlo a través de algún protocolo de acceso de archivos remoto.

Un sistema NAS opera como un servidor en una relación cliente/servidor, tiene su propio procesador, sistema operativo o micronúcleo (*microkernel*) optimizado para manejar las operaciones de I/O. Para las aplicaciones corriendo en la red, un sistema NAS es como un servidor ordinario, para cualquier cliente parece como un disco duro más.

Los dispositivos NAS caen dentro de la categoría de clientes delgados ya que están diseñados exclusivamente para manejar una tarea: las operaciones de I/O, aunque en realidad, estos sistemas de discos hacen mucho más que eso ya que dentro de sus características se incluyen herramientas de administración y monitoreo basadas en el WEB y en el protocolo SNMP, además de la incorporación de *software* para el manejo de volúmenes y configuración de diferentes niveles de RAID así como controladores y manejadores para las diferentes interfaces de red y arreglos de discos y en algunos casos agentes de respaldo y recuperación. La figura 4-5 muestra los diferentes componentes de *hardware* y *software* de un sistema NAS.

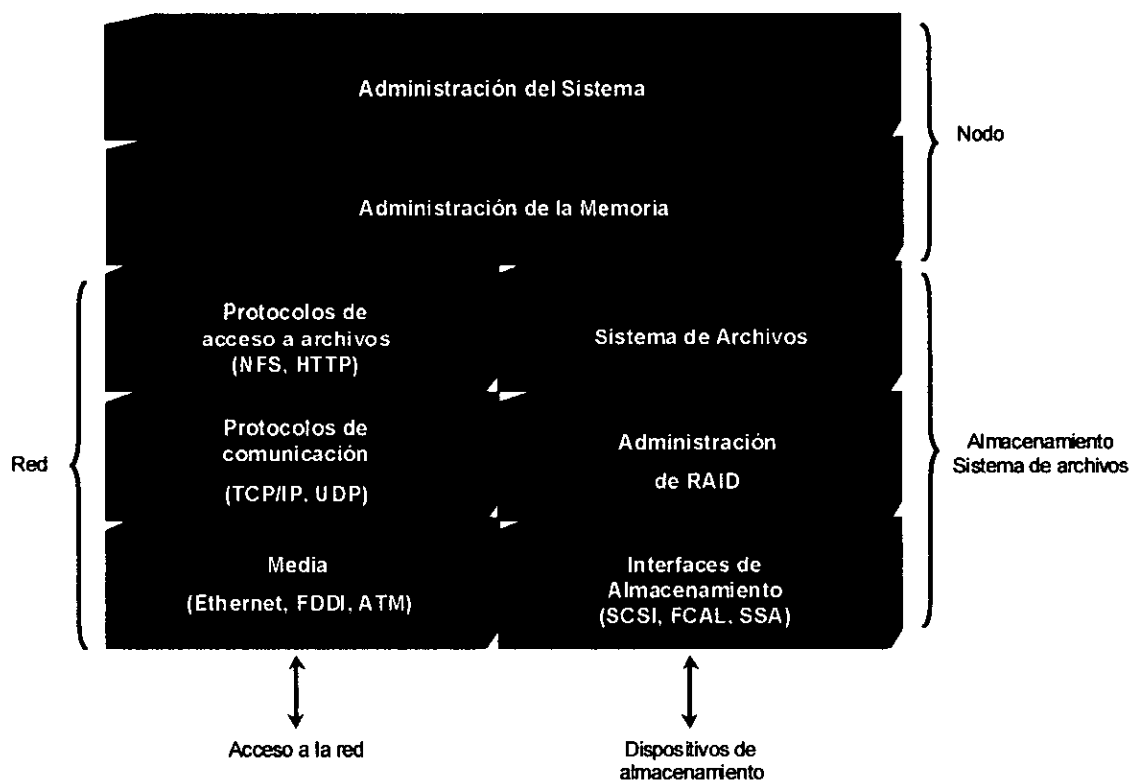


Figura 4-5. Anatomía del sistema operativo de un dispositivo NAS

Técnicamente, un sistema de archivos de un dispositivo NAS tiene dos componentes: El sistema de archivos local y un protocolo de acceso de archivos a través de la red usado para permitir el acceso remoto de tal manera que los usuarios puedan almacenar y manipular sus archivos como si el disco fuera local al equipo de cómputo.

Una de las características más atractivas de los sistemas NAS es la posibilidad de utilizarlos como un medio de almacenamiento común para diferentes plataformas. Aunque en la actualidad la mayoría de los dispositivos NAS soportan NFS que es el protocolo de acceso de archivos de red predominante en el ambiente UNIX y CIFS que es el protocolo utilizado por ambientes Windows, existen algunos sistemas NAS que están diseñados para operar en ambientes homogéneos. En el caso de los sistemas NAS que soportan tanto NFS como CIFS existen diferentes implementaciones, una de ellas utiliza una región específica para cada ambiente, por lo que realmente no se comparte la información sino que se configuran dos áreas separadas de almacenamiento dentro del dispositivo NAS, la segunda permite que se utilice tanto NFS como CIFS sobre una misma partición o área de disco, pero es necesario hacer manipulaciones en los archivos para poder conservar las características propias de manejo de archivos que tiene cada plataforma, la tercera consiste en utilizar un sistema de archivos propietario para el almacenamiento en disco y en hacer la correspondiente conversión cuando un cliente solicita un determinado archivo.

Probablemente uno de los puntos más débiles de este modelo de almacenamiento lo constituye el desempeño al ser un dispositivo de almacenamiento de propósito general, especialmente si se pretende utilizarlo para diferentes aplicaciones. Por ejemplo, una aplicación para toma de decisiones es muy diferente en el manejo de las operaciones de I/O que una aplicación de tipo OLTP lo cual hace muy difícil que este esquema pueda ser usado en ambos ambientes ya que el desempeño se puede ver comprometido. Una alternativa es implementar este modelo en un esquema centralizado/distribuido, centralizado en cuanto al tipo de aplicación y distribuido en cuanto a las necesidades de disco y uso de la red que se tengan. Esta es la principal razón por la que muchos definen a un dispositivo NAS como a un servidor de almacenamiento o un servidor de archivos conectado a la red ya que trabaja directamente sobre sistemas de archivos.

En cuanto a la administración esta es mucho más eficiente desde el proceso de instalación por las características incorporadas como es el caso del uso del protocolo DHCP el cual permite que el sistema de discos descubra su propia dirección IP, también es posible la administración a través de un navegador y debido a que este es un sistema de discos que puede ser compartido por muchos nodos en la red es posible tener un solo administrador dedicado exclusivamente a los dispositivos NAS en vez de varios como en el modelo anterior.

En lo que se refiere a la disponibilidad ésta dependerá de la red y del sistema de discos en sí, en ambientes en donde se requieren altos niveles de disponibilidad es una buena idea tener una red dedicada para los dispositivos NAS, esto también mejorará el desempeño considerablemente, aunque implica un gasto adicional. La mayoría de los sistemas NAS tienen características de redundancia como sus contrapartes que son conectadas directamente a los servidores, pero es una buena idea buscar el sistema de discos con mayor tolerancia a fallas posible. También es posible tener dos dispositivos NAS para que en caso de que uno falle se pueda seguir teniendo acceso a través del otro, esta característica requerirá de algún esquema de replicación o copia simultánea para poder operar.

La escalabilidad está dada en función de las características de crecimiento propias del dispositivo NAS aunque cuando se llega al límite no es necesario adquirir un servidor adicional, basta con agregar otro dispositivo NAS a la red, sin embargo, es importante considerar que a medida que se agregue más disco a la red el ancho de banda de esta puede verse comprometido especialmente si se trata de una red *Ethernet* o *Fast Ethernet*, una alternativa cuando se requieren manejar grandes volúmenes de información a través de este modelo es utilizar ATM o *Gigabit Ethernet*.

En cuanto a los respaldos, en la mayoría de los productos NAS existentes se cuenta con un agente que permite llevar a cabo esta función, sin embargo, en la mayoría de los casos es preferible tener una biblioteca de cintas directamente conectada al sistema de discos. Este, de hecho, es el esquema más recomendado ya que de esta manera se puede respaldar la información sin tener que hacer uso de la red. Esto se vuelve más relevante en la medida en la que se tengan más datos que respaldar.

Finalmente es importante resaltar que en la actualidad las soluciones NAS están siendo utilizadas con mayor frecuencia para resolver problemas asociados con tener espacio en disco para diferentes plataformas y en la mayoría de los casos no constituyen la mejor solución en términos de desempeño debido a la carga de trabajo que pueden ocasionar en la red, por eso en la mayoría de los casos se recomienda tener una red dedicada a estos dispositivos.

4.2.3 Redes de Almacenamiento Inteligente

Aunque tanto el modelo de NAS como el de redes de almacenamiento inteligente (ISN, *Intelligent Storage Network*) utilizan la red como medio de conexión y existen muchas similitudes entre ellos, su forma de operación es distinta. Una solución NAS trabaja sobre sistemas de archivos mientras que una solución ISN o SAN como también se le conoce trabaja sobre dispositivos de bloque, es importante mencionar que estas dos tecnologías no compiten entre sí sino que por el contrario son complementarias y es posible que en un futuro no muy lejano se puedan llegar a fusionar.

La diferencia fundamental en la forma de conexión entre una solución NAS y una ISN es que la primera puede utilizar la infraestructura de red existente para poder operar, esto quiere decir, que un dispositivo NAS puede ser conectado directamente a la red de producción a través de cualquier tecnología estándar de red que ya esté operando como puede ser *Ethernet*, ATM, FDDI, etc. Por otro lado, una solución ISN involucra el implementar una red dedicada exclusivamente para el almacenamiento de datos tal como se ve en la figura 4-6, como se puede ver en esta nueva red son conectados los sistemas de discos junto con los servidores, estaciones de trabajo o PC's que los utilizarán. Esto no quiere decir que una solución NAS no pueda tener su propia red dedicada de almacenamiento, la diferencia estriba en que una ISN necesita forzosamente de una infraestructura de red adicional y una NAS no.

Otra diferencia importante es que la conexión del sistema de discos en una solución ISN es directa a la red mientras que en una solución NAS realmente el que se conecta a la red es el servidor el cual a su vez tiene conectado el sistema de discos a través de un canal el cual puede ser SCSI o Fibra.

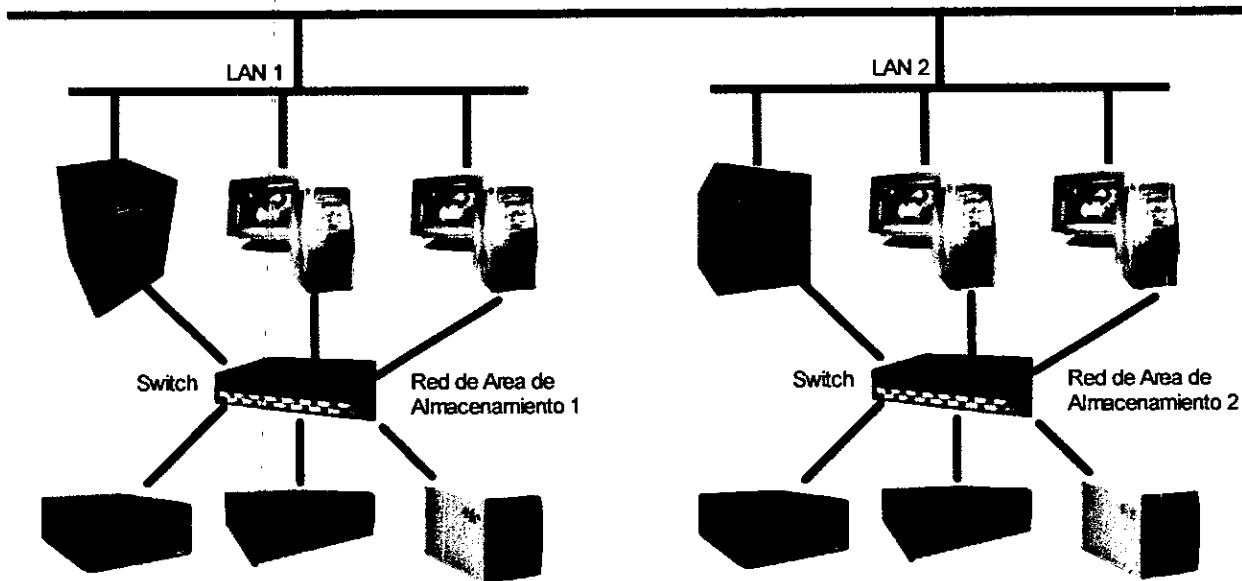


Figura 4-6. Redes de Almacenamiento Inteligente

En términos de desempeño este modelo ofrece muchas ventajas, por principio de cuentas está basado en la tecnología de fibra óptica la cual ofrece una velocidad de transmisión de datos de 100 MB/s comparado con Ultra SCSI que es de 40 MB/s. Además, por la naturaleza misma de este modelo se tiene una red dedicada para el almacenamiento desde su concepción lo que evita que se tenga que compartir la red de producción. Por otro lado, es posible tener diferentes sistemas de almacenamiento conectados a la red y configurados para cada necesidad específica que se tenga, por ejemplo, se pueden tener tres diferentes sistemas de almacenamiento configurados para tres aplicaciones diferentes: Toma de decisiones, OLTP y NFS sin tener que dedicar un solo arreglo de discos para las tres.

Adicionalmente el uso de concentradores de conexiones inteligentes permite que el desempeño sea aún mejor ya que no se comparte un canal para todos los dispositivos conectados a éste sino que se tiene un canal dedicado por cada dispositivo conectado.

En cuanto a la facilidad de administración y uso, los proveedores han prometido que esta será relativamente sencilla, pero la realidad es que por ser una tecnología de aparición reciente los administradores tendrán que sufrir un poco hasta que haya alcanzado su madurez. Por el momento muchos de los concentradores de conexiones inteligentes (*switches*) existentes en el mercado soportan el protocolo SNMP y pueden ser administrados a través del Web.

Para poder administrar una solución de este tipo es necesario que los administradores del sistema y de los arreglos de discos ahora también cuenten con conocimientos de red ya que la implementación de una ISN requiere de los diferentes componentes que se encuentran en una red común: concentradores de conexiones, compuertas, puentes, etc. Por lo que la administración en este modelo va más allá de la configuración de RAID o del manejo de volúmenes. Sin embargo, este esquema permite un esquema de administración centralizado el cual reduce considerablemente los costos.

En términos de disponibilidad es posible tener esquemas de alta redundancia en donde se tengan dobles conexiones y por las características propias de la tecnología de canal de fibra es posible tener esquemas de recuperación de desastres en donde se tengan servidores y sistemas de almacenamiento localizados en diferentes sitios, lo que resulta bastante provechoso ya que se pueden tener esquemas de redundancia que no solo involucren el disco, sino los servidores y las aplicaciones.

La capacidad de crecimiento es una de las características más atractivas de este modelo ya que ya no se depende más de la capacidad de conexión del servidor o de las limitantes impuestas por el canal de I/O. Si se necesita más disco basta con agregarlo al concentrador de conexiones o adquirir otro adicional pero no es necesario agregar otro servidor que se encargue de manejar el nuevo disco. Otra ventaja importante es que el concentrador de conexiones inteligente que es el elemento fundamental para implementar una ISN tiene un soporte teórico de 16 millones de nodos lo cual elimina la preocupación de la escalabilidad.

El soporte a múltiples plataformas y la posibilidad de usar diferentes sistemas de almacenamiento independientemente de la marca es una de las mayores promesas de este modelo. En este momento todavía esto no es una realidad total ya que todavía no se han definido completamente los estándares de Fibra Optica así como la manera en que se comunicarán los diferentes elementos de una red de almacenamiento, pero se están haciendo esfuerzos importantes para que esto se lleve a cabo. Algunas compañías están desarrollando sus propios estándares para que todos los componentes que estén involucrados en una red de almacenamiento puedan convivir sin mayores problemas, una de las iniciativas más importantes es *Jiro* la cual está siendo promovida por Sun Microsystems y respaldada por un gran número de proveedores de cómputo.

Como se puede ver este modelo centralizado en la red tiene muchos beneficios, sin embargo, en este momento resulta hasta cierto punto difícil de implementar y costoso pero esto es parte del precio que se tiene que pagar por ser una tecnología nueva y por la escasa información que existe al respecto. En los siguientes capítulos se abordarán los detalles detrás de las redes de almacenamiento inteligente de tal forma que primero pueda ser considerada como una solución viable y después se puedan conocer los elementos necesarios para diseñar e implementar una ISN de acuerdo a las necesidades específicas que se tengan.

V

El Canal de Fibra

Aunque es posible implementar Redes de Almacenamiento Inteligente (ISN's, *Intelligent Storage Networks*) utilizando otras tecnologías diferentes a la del canal de fibra, para muchos, las palabras canal de fibra e ISN son sinónimos ya que prácticamente todos los componentes que se necesitan para integrar una ISN están basados en la tecnología de canal de fibra.

Desde su aparición en 1988, el canal de fibra fue desarrollado para ser una tecnología de red escalable de alta velocidad y muy flexible, pero no fue originalmente diseñada como una tecnología para redes de almacenamiento. De hecho, por muchos años los desarrolladores del canal de fibra creyeron que esta tecnología alguna vez reemplazaría a las redes Ethernet y FDDI pero no consideraban la posibilidad de utilizarla para crear redes de almacenamiento.

5.1 Las capas del canal de fibra

Los estándares del canal de fibra definen una arquitectura de multicapas que determina la manera en la que los datos se mueven a través de la red como se muestra en la tabla 5-1. La capa superior, FC-4, establece la interfase de comunicación entre el canal de fibra y las aplicaciones de más alto nivel. El protocolo SCSI, por ejemplo, debe mapear los dispositivos de canal de fibra a dispositivos lógicos para que puedan ser accesados por el sistema operativo. FC-3 está todavía en desarrollo y probablemente incluirá las definiciones necesarias para encriptación y compresión de datos entre otras cosas. La capa FC-2 define como los bloques manejados por las aplicaciones son segmentados en secuencias de tramas (*frames*) para el envío a otras capas. Esta capa también incluye varias clases de servicio y mecanismos de control de flujo. Las dos capas de más abajo, FC-1 y FC-0 se enfocan en el transporte de datos a través de la red. FC-1 define los medios de codificación y decodificación de datos para su transmisión así como la estructura de comandos para acceder el medio de transporte. FC-0 establece los estándares para diferentes tipos de medios, distancias de cableado permisibles y señalización.

Capa	Función	Uso
FC-4	Mapeo a protocolos de mayor nivel	SCSI-3, IP
FC-3	Servicios varios	En construcción
FC-2	Entrega de datos	Tramas, control de flujo, clase de servicio
FC-1	Transmisión	codificación 8b / 10b, detección de errores
FC-0	Interface física	características de cables conectores y del medio de transmisión

Tabla 5-1. Capas del canal de fibra

5.1.1 Capa física

El canal de fibra puede correr sobre un medio óptico o de cobre. Debido a su inmunidad al ruido y a su capacidad de soportar grandes distancias por la potencia de la señal el medio óptico es preferido, aunque también el medio de cobre es utilizado sobre todo en las interfases de los sistemas de almacenamiento para reducir costos de implementación y para reducir el tamaño de las unidades de disco.

Los cables de canal de fibra están formados por dos componentes: uno para transmitir y el otro para recibir. Como resultado, estos cables están formados por dos hilos separados dentro del mismo medio. En el caso de los cables de cobre, estos tienen dos pares de hilos de par trenzado, en cada par, un conductor lleva la señal y el otro es un hilo a tierra. El conector utilizado es un DB-9 de 9 pines y la longitud máxima soportada usando cables de cobre es de 30 metros.

Los cables ópticos son referenciados por el modo o las frecuencias de onda de luz que pueden transmitir. El cableado multimodo tiene un núcleo de transmisión de 50 ó 62.5 micrones y lleva una luz láser de onda corta que permite una longitud máxima de 175 metros para el núcleo de 62.5 micrones y de 500 metros para el núcleo de 50 micrones. La fibra monomodo es construida con un núcleo de 9 micrones y lleva una luz láser de onda larga. Con un diámetro del núcleo mucho menor y una sola señal, este tipo de fibra soporta distancias de hasta 10 kilómetros. Los conectores utilizados en los cables ópticos se conocen como conectores SC. La figura 5-1 muestra los tipos de cables y conectores de canal de fibra más utilizados.

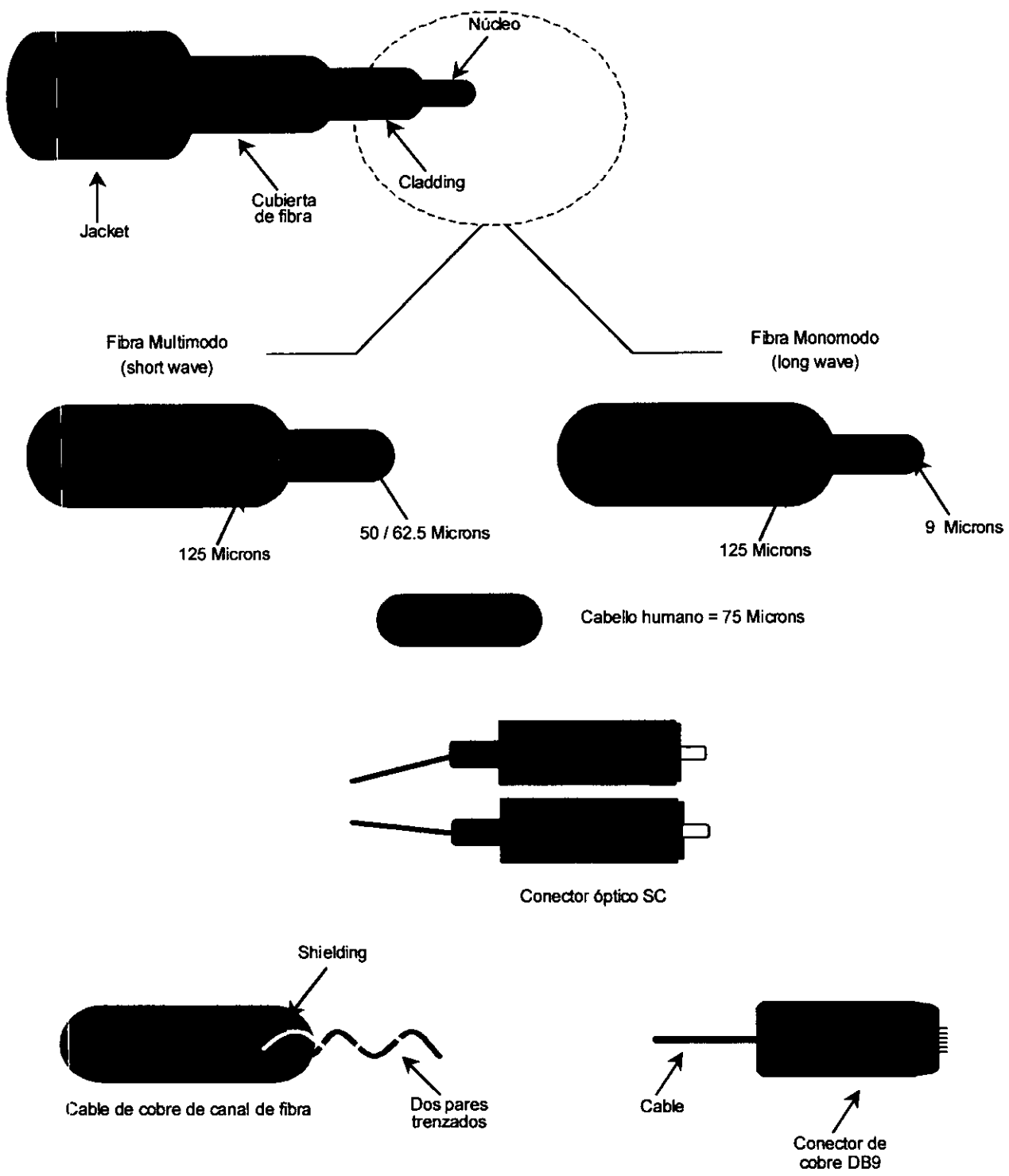


Figura 5-1. Cables y conectores de canal de fibra

Tanto los arreglos de discos como los servidores utilizan señales eléctricas para operar, cuando se utiliza un medio óptico de transmisión es necesario convertir los pulsos ópticos a eléctricos para que las señales puedan ser interpretadas por el servidor o por el sistema de discos. Esto se logra a través de transceptores (*transceivers*) los cuales también constituyen las interfases de conexión.

Uno de los primeros transceptores desarrollados fue el GLM (*Gigabaud Link Module*) el cual está disponible tanto para fibra óptica como para cobre, este tipo de transceptor ha ido cayendo en desuso pero todavía se puede encontrar en el mercado. Los GLMs son módulos semipermanentes que se fijan a la tarjeta adaptadora o tarjeta madre dentro de un sistema de discos, servidor o concentrador de conexiones. Una de las principales desventajas de estos módulos es el mantenimiento ya que el equipo debe ser apagado para poder reemplazar o reiniciar un módulo. Además de la interfase, los módulos GLM incorporan componentes de reloj, recuperación de datos y circuitería de serialización.

El transceptor de mayor uso en la actualidad es el Convertidor de Interface a Gigabit (GBIC , *Gigabit Interface Converter*), El GBIC fue desarrollado por Compaq, Sun, Amp y Vixel y se ha convertido en el estándar de la industria. Los GBIC's no tienen los componentes adicionales que tienen los GLM's pero son dispositivos modulares de reemplazo en caliente. Esta característica permite que los GBIC's puedan ser insertados y quitados sin tener que apagar el equipo además de que facilita el reemplazo del medio. Los GBIC's están disponibles tanto para señales de onda corta (500 m) como para señales de onda larga (10 km) con conectores SC para la versión de fibra y conectores DB-9 para la versión de cobre. La tendencia en el diseño de los GBICs es la de incluir componentes adicionales que permitan hacer diagnósticos y monitoreo.

Un tercer transceptor es el llamado «1x9» el cual se monta permanentemente al adaptador o tarjeta madre. Este módulo es más barato que el GLM o el GBIC pero no permite la mezcla de diferentes medios o interfases. En el caso de los concentradores de conexiones basados en este tipo de transceptor, la falla de uno de los módulos ocasionará la falla general del concentrador de conexiones.

Para aquellos casos en los que se requiera de una solución basada totalmente en un medio óptico pero en donde se tengan algunos dispositivos con interfases de cobre, existen adaptadores que permiten la conversión del medio de cobre al de fibra, estos adaptadores se conocen como Adaptadores de Interfase del Medio (MIAs, *Media Interface Adapters*) sin embargo, los MIA's agregan un componente adicional y por lo tanto otro punto de falla, pero son un riesgo que a veces vale la pena correr para eliminar los problemas de interferencia electromagnética y de distancia que tiene el medio de cobre.

5.1.2 Capa de Transmisión

Codificación de datos

La transmisión y recepción apropiada de una serie de *bytes* no sería posible si los *bits* de datos simplemente fueran metidos uno a la vez en algún medio de transporte. Sin un esquema de codificación sería imposible detectar cuando un *byte* de datos termina y cuando comienza el siguiente y sería imposible recuperar cualquier otra señalización adicional incluida en la cadena de datos. Este problema es aún más acentuado cuando se está hablando de velocidades de transmisión del orden de *gigabits* por segundo como es el caso del canal de fibra en donde se debe asegurar que cada *bit* que conforma un *byte* es el correcto para garantizar la precisión en los datos que se están transmitiendo. Para lograr esto la tecnología de canal de fibra utiliza un algoritmo de codificación desarrollado por IBM conocido como 8b/10b el cual consiste en convertir los *bytes* de 8 *bits* en *bytes* de 10 *bits*, esta técnica permite que en el proceso de codificación se agreguen además caracteres de comandos especiales.

Conjuntos ordenados

Para mover datos a través de la red, el canal de fibra utiliza una sintaxis de comandos conocida como conjuntos ordenados (*ordered sets*). Los conjuntos ordenados caen en tres categorías generales. La primera categoría es usada para indicar el inicio y el fin de las tramas de datos (*frames*) así como la clase de servicio usada. El canal de fibra define 11 tipos de inicio de trama (SOF, *Start of Frame*) y 8 de fin de trama (EOF, *End of Frame*). El tipo específico de SOF o EOF usado indica al equipo que lo recibe que acción necesita ser tomada. La segunda clase de conjuntos ordenados es conocida como señales primitivas, las cuales son usadas para indicar acciones o eventos en el transporte. Por ejemplo, la señal primitiva CLS es usada en configuraciones de ciclo arbitreado para indicar el fin de una transacción. Si una señal primitiva se pierde o corrompe esta debe ser regenerada a través de protocolos de más alto nivel. Un tercer tipo de conjuntos de datos es conocido como secuencias primitivas las cuales son usadas para indicar o iniciar cambios de estado en el transporte. A diferencia de las señales primitivas, las secuencias primitivas requieren de que por lo menos tres palabras consecutivas sean reconocidas antes de que se lleve a cabo cualquier acción. Un disco de canal de fibra por ejemplo, puede enviar una serie de secuencias primitivas de inicialización de ciclo (*loop*) cuando es configurado en una topología de ciclo pero ningún nodo responderá hasta que por lo menos tres de estas señales sean recibidas.

Aunque la mayoría de los conjuntos ordenados son comunes a las tres topologías de canal de fibra que existen, algunas señales y secuencias primitivas han sido definidas específicamente para la topología de ciclo arbitreado y la de fábrica. Una topología compartida como es el caso del ciclo arbitreado requiere de comandos especiales para obtener el control exclusivo del transporte mientras que una conexión dedicada como es el caso de la topología de fábrica asume que el transporte está siempre disponible y por lo tanto no se requiere negociar el acceso.

5.1.3 Capa de entrega de datos

Tramas

Para mover datos de un dispositivo de canal de fibra a otro los bloques de datos manejados por los protocolos de más alto nivel deben ser organizados en paquetes para su transmisión a través del medio de transporte. En la nomenclatura del canal de fibra, estos paquetes de datos se conocen como tramas (*frames*).

El primer componente de una trama es un conjunto ordenado que marca el inicio de éste (SOF). Este delimitador define como ya se había mencionado la clase de servicio usado y si la trama es la primera en una serie o simplemente es una más en una serie de tramas relacionadas. A continuación del SOF, se encuentra una cabecera de 24 *bytes* que contiene las direcciones de origen y destino, información de control, tipo de datos y posición de la trama dentro de una serie de tramas secuenciales. Después de la cabecera viene la unidad de datos que puede estar formada de 0 a 2112 *bytes*. Esta capacidad de longitud variable permite que el canal de fibra se puede adaptar a los requerimientos de diferentes tipos de aplicaciones.

La integridad de los datos dentro de una trama es responsabilidad de un verificador de redundancia cíclica o CRC que consta de 32 bits. El cálculo del CRC es realizado antes de que los datos pasen por el proceso de codificación 8b/10b y el CRC por si mismo es también codificado junto con el resto de los contenidos de la trama. Después del CRC aparece el conjunto de datos de fin de trama (EOF) que sirve para indicar que la trama está completa. El tipo de EOF usado depende de si la trama es la última o es una más de una serie de tramas. La figura 5-12 ilustra los diferentes componentes que integran una trama.



Figura 5-2. Formato de la trama del canal de fibra

Una vez que la trama es ensamblada con todos sus componentes, la entrega de las tramas es realizada por una jerarquía de protocolos constituida por secuencias e intercambios tal como se muestra en la figura 5-3. Una secuencia puede incluir una o más tramas relacionadas, tal como un simple archivo escrito a disco via múltiples tramas, mientras que un intercambio puede incluir una o más secuencias sin relación. Dos dispositivos en comunicación por ejemplo, pueden tener varios intercambios establecidos al mismo tiempo, con identificadores de intercambios y secuencias para separar el tráfico. Esta estructura de intercambios y secuencias de tramas maximiza la utilización de la comunicación entre dos dispositivos.

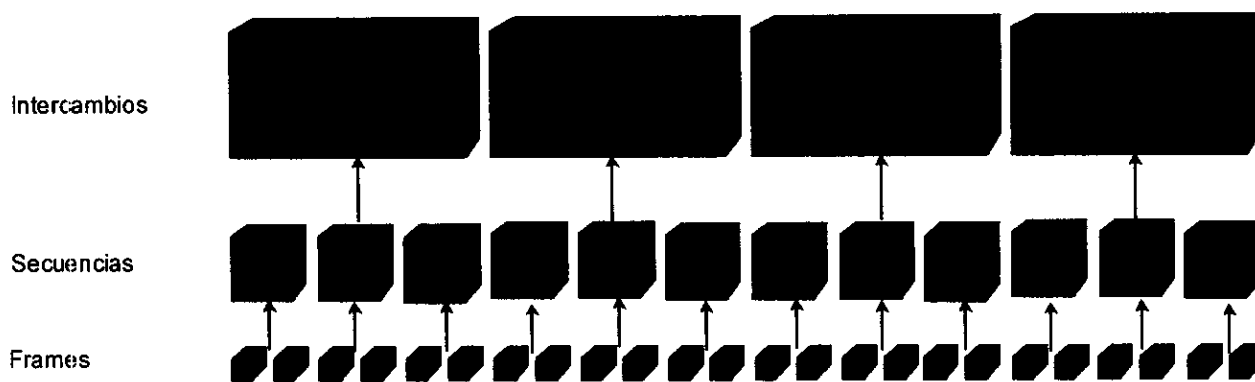


Figura 5-3. Jerarquía de tramas, secuencias e intercambios.

Clase de Servicio

Las aplicaciones pueden requerir diferentes anchos de banda, diferentes niveles de conectividad y de seguridad en la transmisión de los datos. Por ejemplo, las aplicaciones de respaldo necesitan de una conexión dedicada y sostenida así como de todo el ancho de banda disponible, mientras que las aplicaciones de OLTP no necesitan de un ancho de banda dedicado pero se benefician enormemente si existe una notificación por cada transferencia que se haga. Para cumplir con estos diferentes requerimientos, los estándares de canal de fibra proporcionan cinco clases de servicio aunque no todas están implementadas en productos reales.

La clase de servicio 1, define una conexión dedicada entre dos dispositivos con notificación de cada trama que se envía. En esta clase de servicio dos dispositivos que se están comunicando ocupan todo el ancho de banda disponible hasta que se termina la conexión. Esta conexión dedicada también asegura que todas las tramas serán recibidas en el mismo orden en el que fueron enviadas por lo que la velocidad para recuperar los datos es muy buena. En el caso de que la conexión entre un servidor y un sistema de discos sea a través de un concentrador de conexiones inteligente, este debe mantener un canal dedicado de 100 MBps entre los puertos y no se puede llevar a cabo ninguna otra tarea con ellos mientras dure la comunicación. Cuando se tienen múltiples conexiones las peticiones necesitan ser encoladas hasta que sea liberada alguna conexión esto implica que se necesite de mecanismos de memoria adicionales (*buffers*) en el conentrador de conexiones para guardar la información de las peticiones que están en espera. Debido a la complejidad adicional y al costo asociado con esta funcionalidad, la clase 1 generalmente no es soportada en concentradores de conexiones (*hubs y switches*)

Por otro lado, la clase de servicio 2 no requiere de una conexión dedicada entre los dispositivos que se están comunicando pero si proporciona notificación de cada trama que se envía. En las configuraciones con concentradores de conexiones inteligentes esta característica de no conexión dedicada permite que el concentrador pueda avanzar tramas sin tener que dedicar ancho de banda a la transacción. Sin embargo, múltiples rutas o mucho tráfico pueden ocasionar que las tramas sean recibidas en orden aleatorio y por lo tanto agregar carga adicional para recuperar los datos. Para evitar esto, una conexión de clase 2 puede ser configurada para que la entrega de tramas sea en orden, este problema no se presenta en ambientes de ciclo arbitreado ya que solo los dispositivos que se están comunicando ocupan el ciclo durante el tiempo que dure una transacción. La clase 2 es ideal para aplicaciones de misión crítica que demandan de un alto grado de integridad en los datos, también es ideal para aquellas aplicaciones que manejan muchas transacciones con datos pequeños.

La clase 3 también es de no conexión dedicada pero a diferencia de la clase 2 no existe notificación de la información que se envía. Esta clase opera de manera similar al protocolo UDP/IP en topologías LAN y depende de protocolos de mayor nivel para recuperarse de cualquier falla que ocurra en la transmisión de los datos. En esta clase se sacrifica la integridad de la información para reducir el tráfico y mejorar el desempeño. La clase de servicio 3 es ampliamente usada en ambientes de ciclo arbitreado, incluso para aplicaciones de misión crítica, esto es debido a la naturaleza temporal de las conexiones en una topología de ciclo. Las configuraciones con concentradores de conexiones inteligentes en ocasiones evitan usar esta clase por la ausencia de un mecanismo de notificación de tramas.

Para aplicaciones multiclase, el canal de fibra proporciona una característica especial llamada Intermix. Intermix permite que cualquier ancho de banda que no sea usado en una conexión de clase 1 pueda ser usado por conexiones de clase 2 ó 3. Existen otras dos clases de servicio pero no han sido implementadas en productos todavía, los concentradores de conexiones inteligentes de canal de fibra típicamente soportan solo servicios de clase 2 y 3, el soporte a otras características está determinado como siempre por el mercado, pero conforme pase el tiempo seguramente se verá mayor funcionalidad en los diferentes productos de canal de fibra.

Control de flujo

Para prevenir que un dispositivo sea incapaz de manejar el arribo de muchas tramas, el canal de fibra ofrece varios mecanismos de control de flujo basados en un sistema de créditos. Cada crédito representa la habilidad de un dispositivo de aceptar una trama adicional. Si el dispositivo que está recibiendo los datos no envía créditos entonces el dispositivo que los está enviando dejará de enviar más tramas. En la práctica, este esquema de créditos está basado en el número de espacios de memoria (*buffers*) que un nodo mantiene para almacenar los datos que está recibiendo.

Para eficientar el uso de la transmisión de datos se debe asegurar que múltiples tramas puedan ser enviadas consecutivamente. Esto se logra si el dispositivo que va a recibir la información emite suficientes créditos antes de que una transacción comience, otra manera es la de utilizar la capacidad de *full duplex* del canal de fibra para enviar créditos adicionales mientras que las tramas están todavía siendo recibidas.

5.1.4 Capa de interface con protocolos de alto nivel

Uno de los principales atractivos del canal de fibra es su habilidad de poder transportar una gran variedad de protocolos de alto nivel. Al hecho de poder adaptar estos protocolos al medio de fibra se le conoce como mapeo de protocolos. El mapeo de protocolos describe la localización y el orden de los mensajes que transmiten los protocolos de alto nivel dentro de las tramas en una transmisión entre dos dispositivos de fibra.

Dentro de los protocolos de alto nivel que ya han sido mapeados o que se planean mapear al canal de fibra se encuentran el protocolo de Interfase de Sistemas de Cómputo Pequeños (SCSI , *Small Computer System Interface*), el Protocolo de Internet (IP, *Internet Protocol*), La Arquitectura de Interface Virtual (VIA, *Virtual Interface Architecture*), La Interfase Paralela de Alto Desempeño (HIPPI, *High Performance Parallel Interface*) y ESCON.

5.2 Nombres y direcciones

Para garantizar la identidad única de los elementos que conforman un ambiente de canal de fibra se utiliza una jerarquía de nombres y direcciones. En la terminología del canal de fibra un dispositivo de comunicación es conocido como un nodo. Una tarjeta en un servidor, por ejemplo, constituye un nodo de canal de fibra. Generalmente, un nodo puede tener una o varias interfaces físicas conocidas como puertos del nodo o puertos N. La figura 5-4 muestra la diferencia entre un nodo y un puerto de nodo en la terminología de canal de fibra.

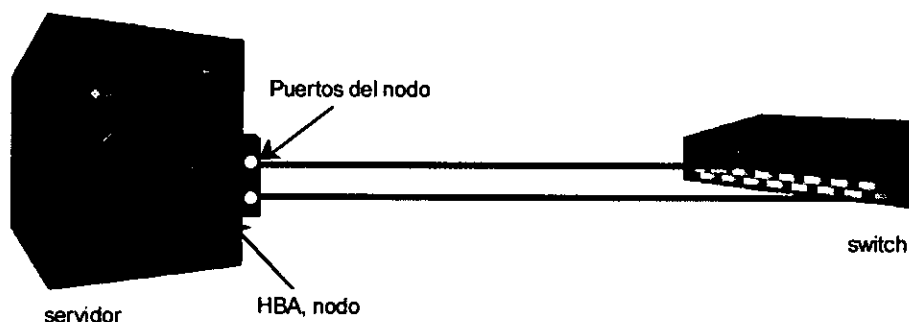


Figura 5-4. Nodo y puerto de nodo

ESTA TESIS NO SALE
DE LA BIBLIOTECA

Cada nodo tiene un nombre fijo (*Node_Name*) de 64 bits asignado por el fabricante para su identificación en la red a la que pertenece, este nombre se conoce como nombre mundial (*WWN, World Wide Name*) el cual es almacenado en memoria no volátil y generalmente se encuentra estampado en la superficie del producto en forma de número serial. A un puerto N en un nodo también se le asigna un nombre de puerto de 64 bits (*Port_Name*). Ambos nombres siguen un formato establecido por la IEEE. Esta convención de nombres permite que cada nodo y sus puertos asociados sean únicos y accesibles incluso en redes de almacenamiento complejas.

De manera similar al esquema de direcciones MAC de 48 bits en *Ethernet*, la convención de nombres de canal de fibra permite la administración global o local de cualquier dispositivo en una ISN. A diferencia de las direcciones MAC, sin embargo, los WWN no son usados para el transporte de tramas a través de la red. Además de su nombre, a cada puerto se le asigna dinámicamente una dirección de 24 bits o identificador de puerto el cual si es usado para el transporte de las tramas. Las direcciones de 24 bits de dos dispositivos comunicándose son agregadas a las cabeceras de las tramas para identificar la fuente y el destino de comunicación. La figura 5-5 muestra como se hace la asignación de nombres y direcciones a un nodo y a sus respectivos puertos.

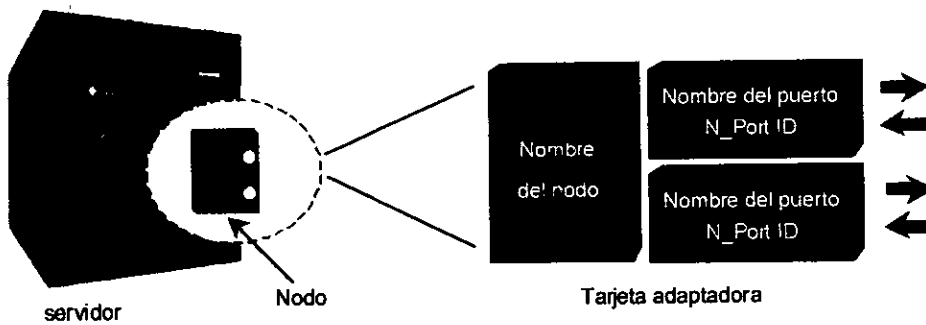


Figura 5-5. Nombres y direcciones de puertos

Este sistema de nombres y direcciones tiene varias ventajas. Los nombres de 64 bits proporcionan una forma única de identificación a través de la red. Sin embargo, si se usaran estos nombres para el ruteo de datos se agregaría carga significativa a la red, al usar direcciones de puertos de 24 bits se optimiza el uso de la red mientras que se permiten tener más de 16 millones de direcciones. Además de esta optimización de la red, este esquema de manejar direcciones de 24 bits elimina la administración manual de direcciones permitiendo que la misma topología las asigne. En ambientes de fábrica, el concentrador de conexiones (*switch*) es el responsable de asignar una dirección a cada dispositivo que se registra por primera vez. El concentrador mantiene una tabla de los WWN's de los dispositivos con sus correspondientes direcciones. En ambientes de ciclo arbitreado, los dispositivos por si mismos realizan una rutina de selección de direcciones que asegura que cada dispositivo en el ciclo tenga una dirección única. El direccionamiento dinámico elimina el elemento de falla humana y proporciona más flexibilidad al hacer cualquier cambio en una red de almacenamiento inteligente.

5.3 Puertos

Existen cinco tipos de puertos en las redes de canal de fibra

- Puertos N
- Puertos F
- Puertos L
- Puertos NL
- Puertos FL
- Puertos E
- Puertos G

Los puertos N, L y NL son usados por los nodos en una red de almacenamiento, mientras que los puertos F, FL, E y G son implementados en concentradores de conexiones inteligentes.

Puertos N y F

Los puertos N están representados básicamente por computadoras y sistemas de discos que accesan a la red de canal de fibra. Estos puertos tienen la responsabilidad de iniciar y terminar el proceso de envío de tramas, sin los puertos N no habría tráfico en la red.

Los puertos F están implementados en los concentradores de conexiones de canal de fibra para realizar tareas de administración y proporcionar servicios de conectividad a los puertos N. Por ejemplo, si un puerto N trata de conectarse a otro puerto N, pero el puerto de destino está ocupado o inaccesible, el concentrador de conexiones puede encolar la comunicación hasta que la ruta esté disponible. La figura 5-6 muestra la distribución de los puertos N y F en una red de canal de fibra.

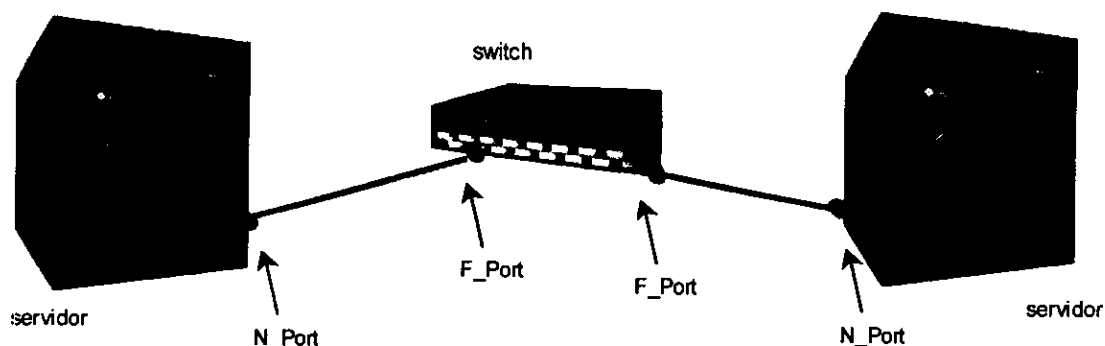


Figura 5-6. Relación entre los puertos N y F en una red de canal de fibra.

Los puertos N y F tienen una relación de uno a uno. Solo un puerto N está conectado a un puerto F externo en el concentrador de conexiones inteligente. Otros puertos N en el concentrador pueden comunicarse con un puerto N en particular iniciando comunicaciones con este a través de sus respectivos puertos F en el concentrador. Los puertos N se encuentran en comunicación constante con los puertos F, independientemente de que ellos estén enviando o recibiendo tramas. Durante períodos de inactividad, los puertos N envían tramas IDLE a sus respectivos puertos F en el concentrador. Estas tramas IDLE establecen una señal (*heartbeat*) entre el puerto N y F que permite que cualquier falla en la comunicación sea identificada inmediatamente.

La comunicación sobre una red de canal de fibra requiere de un proceso de registro de múltiples etapas en la que los puertos N primero establecen conexiones con sus respectivos puertos F en el concentrador y después con los nodos N con los que desean comunicarse.

Puertos L

A diferencia de las redes con concentradores inteligentes (*switches*) los nodos en las redes con topología de ciclo comparten una estructura de cableado común. Para esta topología existe otro tipo de puerto conocido como puerto L. Mientras que los puertos N fueron diseñados para comunicarse con puertos F en los fabricas, los puertos L fueron diseñados para iniciar comunicaciones directamente con otros puertos L en un ciclo.

Debido a que varios nodos comparten el mismo medio de transmisión es necesario algún método para manejar colisiones y para permitir el acceso a los nodos que necesitan comunicarse. La topología de ciclo de canal de fibra realiza esto a través de un mecanismo de arbitraje. Al igual que el mecanismo de arbitraje en SCSI el puerto con la mayor prioridad que participa en un ciclo de arbitraje gana el derecho de comenzar a transmitir, sin embargo, el arbitraje del ciclo se realiza a través del intercambio de tramas a diferencia del arbitraje en SCSI que se lleva a cabo a través de señales eléctricas.

Puertos FL y NL

Cuando los ciclos fueron agregados a la familia de topologías de canal de fibra, se hizo necesario permitir a los puertos N y F comunicarse con puertos L en redes con topología de ciclo, de ahí surgieron los puertos FL y NL.

Un puerto FL es implementado en un concentrador para que de esta manera pueda participar como un nodo especial en una red con topología de ciclo. Por definición la red de ciclo reserva una dirección para un solo puerto FL. En otras palabras, no pueden haber dos concentradores activos comunicándose al mismo tiempo en un ciclo.

Un puerto NL es un puerto que reside en un ciclo y que tiene las capacidades de un puerto N y L. Es decir, un puerto NL soporta el proceso de registro en un ambiente de fábrica así como el de arbitraje de un ciclo. Esto permite que diferentes dispositivos en fábricas y ciclos arbitrados puedan comunicarse entre sí.

La figura 5-7 muestra el proceso de comunicación requerido cuando un sistema en un ciclo inicia una comunicación con un arreglo de discos en una fábrica, como se puede observar este proceso involucra cuatro pasos:

1. Arbitraje del ciclo para ganarlo
2. Registro a la fábrica a través del puerto FL en el concentrador
3. Registro con el nodo de destino N
4. Registro de proceso si es necesario.

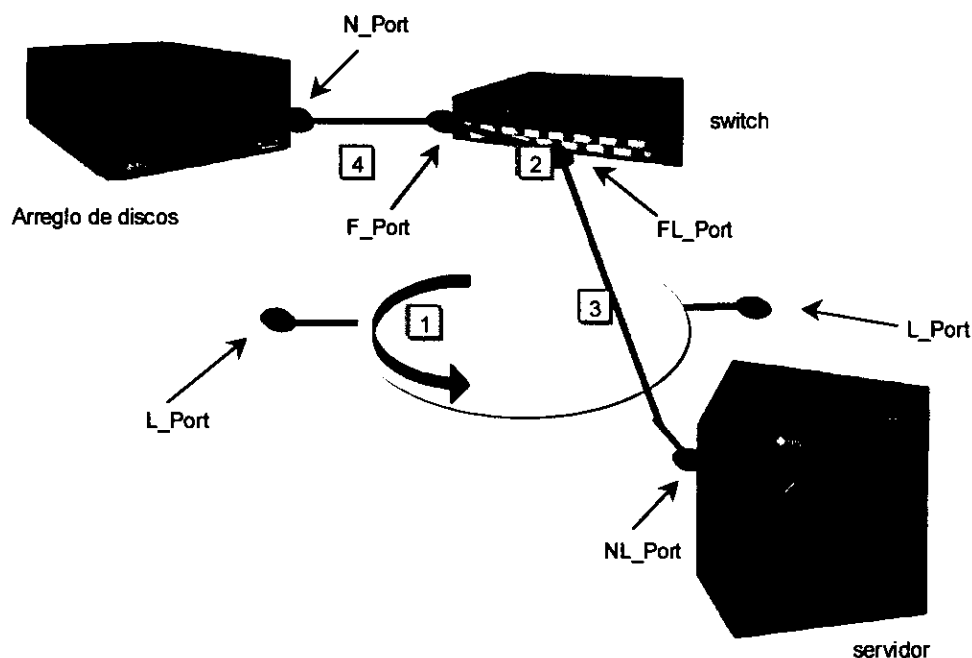


Figura 5-7. Proceso de comunicación entre un puerto NL y N

Ciclos privados y públicos

Existen dos tipos de redes con topología de ciclo las cuales están definidas por el hecho de si el ciclo es público o privado. Un ciclo privado es una red con topología de ciclo en la que no se tiene ningún concentrador inteligente conectado a través de un puerto NL. Es como un sistema cerrado que no tiene acceso a otras redes. Un ciclo público es por el contrario un ciclo con un puerto FL y por lo tanto acceso a otras redes. La figura 5-8 muestra las diferencias entre ciclos privados y públicos.

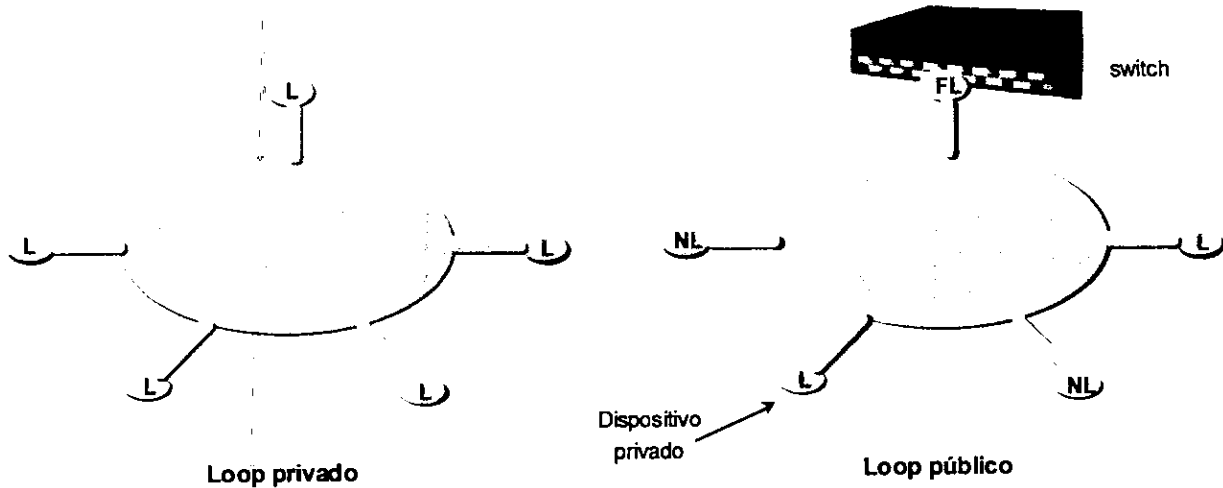


Figura 5-8. Ciclos públicos y privados

Puertos E y G

Los puertos G son puertos generales que pueden ser usados para diferentes funciones en un concentrador de conexiones inteligente tales como puertos F o FL. Por otro lado, un puerto E es un puerto de propósito especial usado para conectar dos concentradores entre sí.

5.4 Topologías

5.4.1 Punto a punto

Una topología punto a punto es simplemente una conexión directa entre dos dispositivos (dos puertos N). Este esquema de conexión como se muestra en la figura 5-9 permite tener todo el ancho de banda dedicado entre los nodos interconectados, típicamente 100 MBps en cada dirección. Antes de que se pueda llevar a cabo cualquier transferencia de información es necesario que los puertos N realicen un proceso de registro (*login*) para que se les pueda asignar una dirección o identificador de puerto.

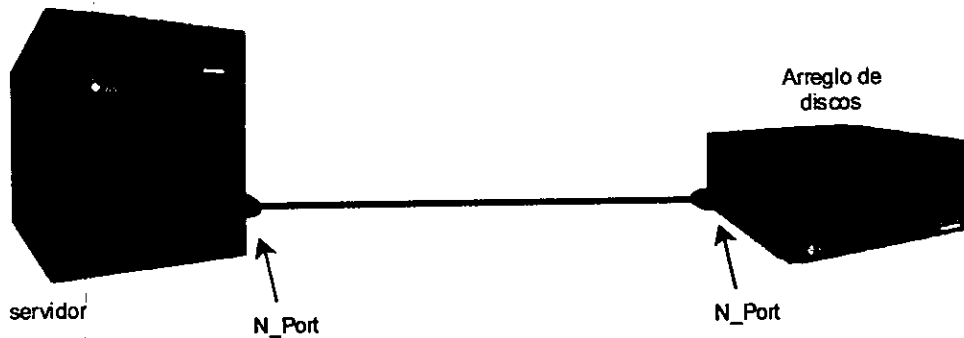


Figura 5-9. Topología punto a punto

Un servidor y un disco en una configuración punto a punto normalmente están haciendo lecturas o escrituras de datos pero nunca ambas operaciones, por eso, aunque es concebible tener una aplicación que requiera transferencias simultáneas en *full duplex*, esto es a 200 MBps, en la práctica solo un lado de la conexión ve el tráfico en un momento dado. El grado de utilización de la conexión en esta topología está determinado por el desempeño de los controladores de canal de fibra y el área de memoria disponible para guardar datos en espera de ser transmitidos o recibidos.

Esta topología tiene la ventaja de ser simple pero no permite la conexión de múltiples dispositivos, sin embargo, para aquellos casos en los que se requiere de una conexión dedicada entre el servidor y el sistema de discos, la conexión punto a punto con fibra óptica ofrece grandes beneficios en términos de desempeño, confiabilidad y distancia comparado con las conexiones por SCSI.

5.4.2 Ciclo Arbitreado

El ciclo arbitreado es la topología de mayor uso en ambientes de redes de almacenamiento en la actualidad ya que permite tener múltiples dispositivos interconectados comparado con las conexiones punto a punto. Además, resulta ser una solución mucho más económica con respecto a los esquemas basados en concentradores de conexiones inteligentes.

Una topología de ciclo arbitreado permite que más de dos dispositivos se puedan comunicar entre sí compartiendo un mismo ancho de banda. Al igual que con las redes *Fast Ethernet* o *Token Ring* el ancho de banda disponible para cualquier dispositivo en el ciclo es determinado por el número total de dispositivos conectados y el nivel de actividad de ellos, entre más puertos activos se tengan, menor ancho de banda disponible.

Aspecto físico de un ciclo

Un ciclo arbitreado es un ciclo físico que se crea haciendo una conexión punto a punto entre todos los dispositivos que formarán parte del ciclo, de tal manera que siempre existe una ruta de datos para todos los puertos permitiendo que cualquier dispositivo pueda acceder a cualquier otro en el ciclo, tal como se muestra en la figura 5-10.

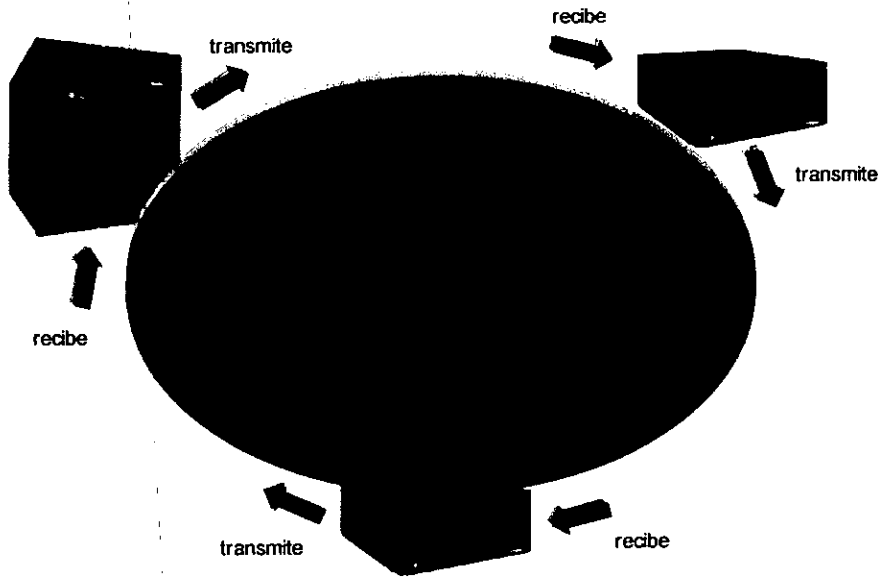


Figura 5-10. Ciclo arbitreado

Los primeros esquemas de ciclo arbitreado fueron creados de esta forma, sin embargo, muy pronto aparecieron desventajas ya que cuando se apagaba o desconectaba cualquier nodo este rompía la cadena y por lo tanto la comunicación del ciclo completo. Este problema de disponibilidad motivó la aparición de concentradores (*hubs*) los cuales permiten una conexión de estrella física en donde cada nodo en el ciclo es conectado a un punto en común. La arquitectura interna de un concentrador completa las conexiones entre transmisores y receptores en un puerto a través de multiplexores y termina el ciclo conectando el transmisor del último puerto del concentrador con el receptor del primero. Una de las características más útiles de los concentradores es la de permitir que el ciclo siga operando aún en el caso de que se presente una falla en cualquiera de los nodos, tal como se muestra en la figura 5-11, esta es la principal razón por la que los concentradores siempre son utilizados aunque se pueda tener una topología de ciclo sin ellos.

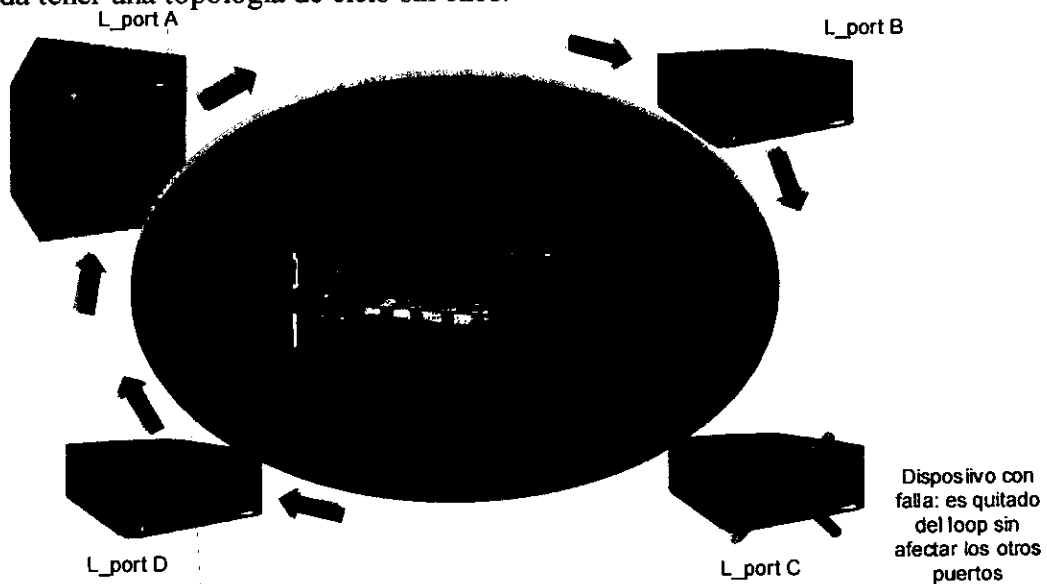


Figura 5-11. Ciclo arbitreado usando concentradores simples (*hubs*)

Direccionamiento de ciclos

A todos los puertos en un ciclo se les asigna una dirección de puerto de 24 bits, esta dirección permite entre otras cosas, definir si los puertos en el ciclo se pueden o no comunicar con otros dispositivos externos, a aquellos puertos que están aislados del mundo exterior y que por lo tanto pertenecen a un ciclo privado se les asignan 0's a sus direcciones en sus dos primeros *bytes*, si los puertos pertenecen a un ciclo público, es decir, si el ciclo tiene comunicación con otros ciclos o concentradores entonces las direcciones de estos puertos tendrán valores de uno en sus dos primeros *bytes*.

Ya sea que se trate de un ciclo privado o público el último *byte* de la dirección de 24 bits es conocido como la dirección física del ciclo arbitreado o AL_PA (*Arbitrared Loop Physical Address*). La dirección AL_PA es determinada durante la inicialización del ciclo y puede, en el caso de los concentradores que soporten conexión a concentradores inteligentes (*switches*) ser modificado por el concentrador. La dirección ALPA proporciona un esquema de direccionamiento bastante compacto que permite asignar a cada dispositivo una identidad única la cual es usada por los conjuntos ordenados durante las transmisiones de información. El número total de direcciones AL_PA disponibles en un ciclo arbitreado es 127, sin embargo, no se recomienda tener 127 dispositivos en el mismo ciclo, por lo menos ahora que el ancho de banda disponible es de 100 MBps.

La topología de ciclo arbitreado tiene un esquema de asignación de direcciones por prioridad para dar oportunidad a ciertos dispositivos como servidores y concentradores de servir al ciclo lo más pronto posible y continuar con otras tareas. La forma de asignar esta prioridad a las direcciones es a través de un esquema de valores numéricos. Entre más pequeño sea el valor numérico, mayor será la prioridad. A un puerto FL, por ejemplo, se le asigna generalmente una dirección de x'00' la cual le da la mayor prioridad sobre cualquier puerto NL que exista en el ciclo, esto permite que el puerto FL siempre gane cuando se tiene que arbitrear el ciclo y asegura que un recurso valioso como un concentrador esté siempre disponible. Durante la selección de direcciones, como se muestra en la figura 5-12, los servidores típicamente toman la más alta prioridad, es decir, los AL_PA's con valores más bajos, mientras que los arreglos de discos toman los valores numéricos mayores, es decir, las prioridades más bajas. Un servidor con un AL_PA de x'01' tendrá mucha mayor posibilidad de ganar el arbitraje comparado con cualquiera de los otros puertos en el ciclo.

La dirección AL_PA de los puertos puede cambiar al agregar, quitar o reiniciar un dispositivo que forme parte del ciclo, este proceso de asignación de direcciones permite que se reduzcan los problemas de administración mientras que se garantiza que cada dispositivo en el ciclo tendrá una dirección única.

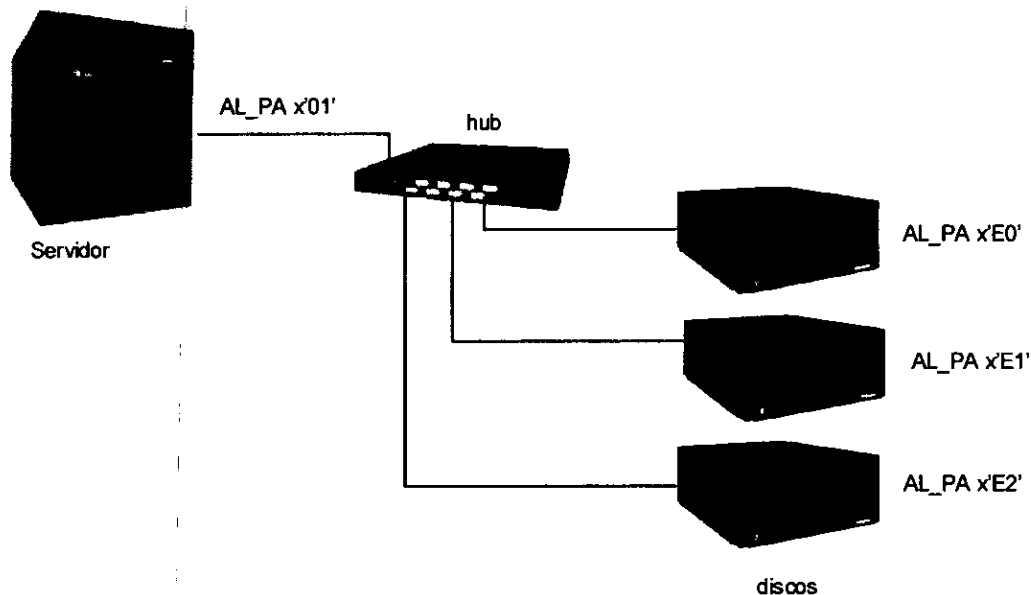


Figura 5-12. Asignación de AL_PA's en un ciclo

Inicialización de ciclos

La inicialización de un ciclo es un proceso esencial para permitir la inclusión de nuevos dispositivos, para la asignación de direcciones a los puertos, para la notificación de cualquier cambio en la configuración y para la recuperación de cualquier falla. Después de que la inicialización se lleva a cabo, el ciclo entra en un modo de monitoreo y a partir de aquí comienza la operación normal. La secuencia de inicialización de un ciclo toma unos cuantos milisegundos pero este tiempo también está en función del número de puertos que tenga el ciclo. Durante el proceso de inicialización se suspende momentáneamente cualquier actividad. La inicialización de un ciclo puede ser provocada por varias razones las más comunes son el agregar, cambiar de posición o quitar un dispositivo del ciclo.

El proceso de inicialización de un ciclo comienza cuando un puerto NL o L envía una secuencia de 12 conjuntos ordenados conocido como Primitivas de Inicialización del ciclo (LIP's, *Loop Initialization Primitives*), debido a la aparición de alguna de las condiciones de inicialización de ciclo. Conforme cada dispositivo recibe estas secuencias, los dispositivos entran en un estado de inicio el cual suspende cualquier operación que se esté llevando a cabo y prepara a los dispositivos para la inicialización del ciclo. Las LIP's son enviadas a lo largo del ciclo a todos los puertos incluyendo el puerto que originó la secuencia de LIP's hasta que todos ellos se encuentren en estado de inicio. En este momento es necesario que uno de los puertos tome el control para convertirse en el maestro del ciclo. A diferencia de *token ring* el maestro del ciclo es asignado de manera temporal y existe un proceso de selección para determinarlo. Una vez seleccionado, el maestro del ciclo, este es el responsable de conducir el resto del proceso de inicialización y regresar el ciclo a su estado de operación normal. La figura 5-13 describe el proceso de inicialización de un ciclo.

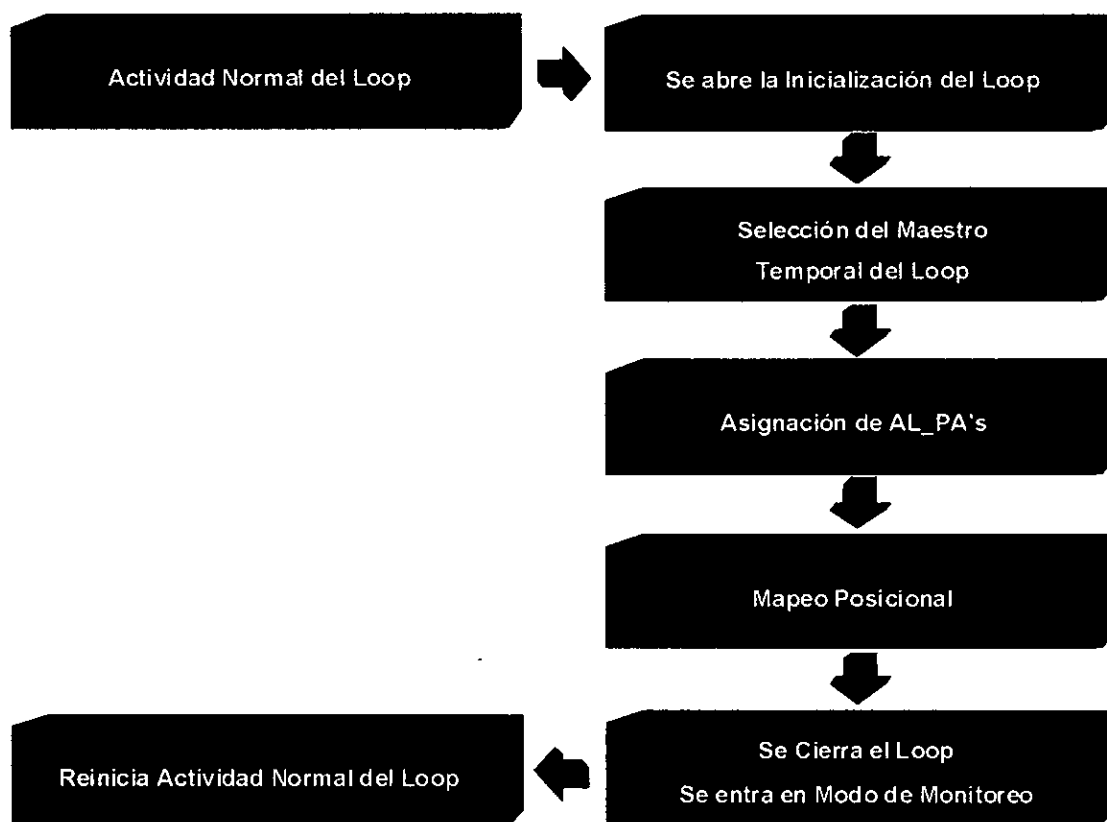


Figura 5-13. Pasos en la sequencia de inicialización de un ciclo

La primera tarea del maestro de un ciclo es mandar una serie de cuatro tramas que permitan a cada participante en el ciclo seleccionar una dirección AL_PA única. El formato de una trama contiene un campo de 128 bits que representa una lista abreviada de todas las AL_PA's disponibles. La posición de cada bit en un mapa de AL_PA's corresponde a la lista secuencial de AL_PA's comenzando con x'00' y terminando con x'EF'. Cuando la primera trama es enviada, todos los *bits* son inicializados a 0, indicando que las AL_PA's no han sido asignadas. Cuando cada dispositivo selecciona su ALPA el *bit* correspondiente es puesto a 1. A continuación de la asignación de AL_PA's se lleva a cabo un proceso de mapeo que tiene como función establecer la forma en la que los dispositivos se encuentran posicionados físicamente en el ciclo. Este mapa posicional es usado para propósitos de optimización y diagnóstico, sin embargo, no es soportado por todos los dispositivos de canal de fibra que se encuentran actualmente en el mercado. Una vez concluido este proceso, el maestro temporal del ciclo termina el proceso de inicialización mandando un conjunto ordenado que indica cerrar el ciclo seguido por señales IDLE. Cada dispositivo recibe la señal de cerrar el ciclo y deja el estado de inicio para regresar a su estado normal de operación. Las señales IDLE continúan circulando en el ciclo hasta que cualquier operación suspendida haya sido reiniciada o comiencen nuevas operaciones.

Registro de puertos

La inicialización de un ciclo permite que cada dispositivo tenga una dirección AL_PA única y por lo tanto que no hayan conflictos de direcciones. Un ciclo que conste de un servidor y 24 discos de canal de fibra por ejemplo, tendrá 25 direcciones AL_PA. Sin embargo, después de la inicialización el servidor no tiene idea de que más hay en el ciclo. Para que un servidor sea capaz de reconocer los discos que se encuentran en el ciclo se requiere un paso adicional, el cual es provisto por un proceso de registro puerto a puerto conocido como PLOGI. Este proceso es levantado por el servidor después de la inicialización del ciclo y consiste en establecer sesiones de registro con todos los puertos NL disponibles enviando tramas PLOGI. Los puertos que acepten los PLOGI's del servidor regresarán una notificación a éste informando su WWN y características adicionales como el máximo tamaño de trama soportado. De esta manera el servidor establece sesiones con todos los participantes activos en el ciclo.

Este proceso de registro es el primero en una serie de intercambios de información entre dispositivos en un ciclo hasta alcanzar la capa FC-4 del estándar del canal de fibra permitiendo la asociación de las direcciones AL_PA en la capa de transmisión con la definición lógica del canal SCSI en la capa de aplicación.

Máquina de estado

Las funciones específicas de un ciclo se encuentran interconstruidas en una máquina de estado de puertos del ciclo (*Loop Port State Machine*) la cual reside entre la capa FC-1 de codificación/decodificación y la capa FC-2 de administración de tramas y control de flujo. Físicamente esta máquina de estado forma parte de los puertos de un concentrador de canal de fibra en forma de silicón o microcódigo.

La máquina de estado de puertos del ciclo es responsable de monitorear y realizar tareas como permitir el acceso del ciclo para que se lleven a cabo las transacciones y asignar el control del ciclo a alguno de los nodos cuando alguna transacción ha sido completada. Estos procesos están definidos en 11 estados y son controlados por una serie de variables de entrada y salida. Conforme las peticiones empiezan a ser manejadas en la capa inmediata superior FC-2 la máquina de estado debe determinar su estado actual así como cualquier acción necesaria para completar cualquier petición. Por ejemplo, si la capa de más arriba tiene tramas que enviar, la máquina de estado debe cambiar de su estado normal de monitoreo a un estado de arbitraje para permitir el acceso al ciclo, una vez que se gana el acceso, la máquina de estado cambia a un estado abierto y notifica al puerto de destino que hay una transacción que realizar, cuando la transmisión se completa se pasa a un estado de recepción y finalmente se regresa al estado de monitoreo. Esta transición lógica de un estado a otro basado en variables de entrada y salida permite tener un mecanismo eficiente para tener un control ordenado del ciclo.

Naturaleza del ciclo arbitreado

El ciclo arbitreado a diferencia de Ethernet o Token Ring es un transporte que no envía notificación a todos los nodos del ciclo, cuando un puerto NL gana el control de un ciclo y empieza a transmitir tramas, estas tramas se envían directamente al puerto de destino, los puertos intermedios entre el puerto de origen y destino verán las tramas pero únicamente los pasarán hacia su destino final. Esta característica de no notificación (*no broadcast*) mejora considerablemente el desempeño al eliminar la carga que significa el manejo de las tramas de por lo menos parte del ciclo, sin embargo, esto también implica que cualquier problema en el ciclo sea más difícil de diagnosticar.

5.4.3 Fábricas

A la conexión de servidores y sistemas de almacenamiento a través de concentradores de conexiones inteligentes (*switches*) de canal de fibra se le conoce como fábrica. Las fábricas proporcionan un ancho de banda de 100 MBps por puerto lo que significa que el ancho de banda se incrementa conforme se agregan nuevos dispositivos al concentrador inteligente a diferencia de los concentradores comunes en donde el ancho de banda está dividido por el número total de puertos y en donde el desempeño se decrementa a medida que hay más dispositivos conectados. La figura 5-14 muestra la manera en que pueden ser conectados servidores y arreglos de discos con un concentrador de canal de fibra.

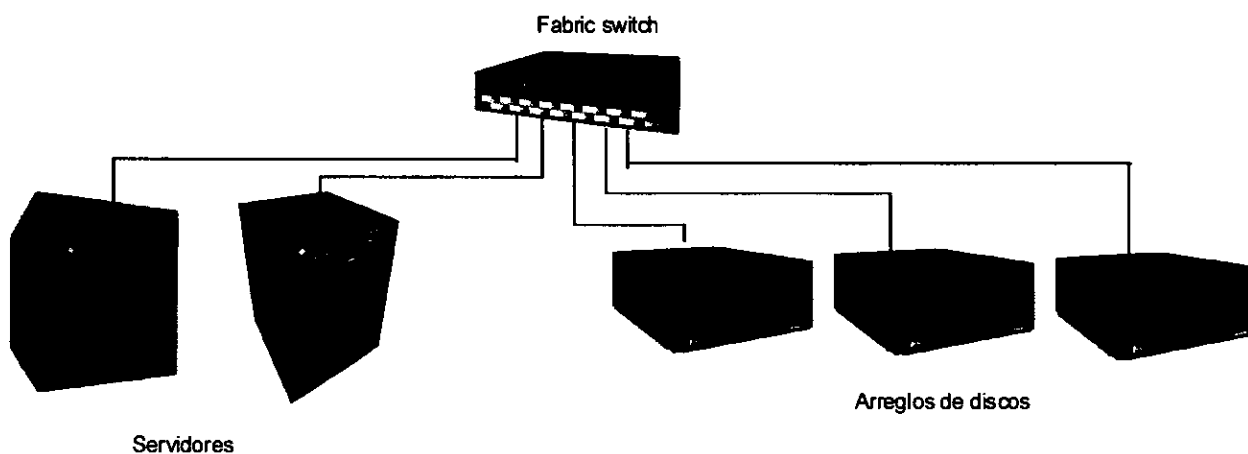


Figura 5-14. Fábrica

Cuando una trama entra a un puerto F desde un puerto N, la lógica del concentrador de conexiones inteligente solo examina el identificador de destino de la trama (dirección de 24 bits) y en base a este se envía la trama al puerto apropiado. Esta tecnología incrementa considerablemente el desempeño y reduce el tiempo requerido para tomar una decisión de ruteo, ya que el identificador de destino reside en los primeros 4 bytes de la cabecera de la trama, una decisión de ruteo se puede hacer en el instante en que la trama entra a un puerto de la fábrica.

Por otro lado el algoritmo de almacenamiento y envío (*store and forward*) ofrecido en algunas tecnologías de concentrador requiere guardar la trama completa antes de que se lleve a cabo una decisión de ruteo. El número de espacios de memoria ofrecidos en cada puerto F es una consideración importante para la selección de un concentrador (*switch*). Durante períodos de congestión, la fábrica puede no ser capaz de transmitir una trama al puerto de destino, especialmente si dos puertos N están enviando datos al mismo puerto de destino. Dependiendo de la clase de servicio, la fábrica puede ser forzada a abandonar las tramas que no pueda procesar. La habilidad de encolar múltiples tramas reduce esta posibilidad y mejora el desempeño.

Esquema de direccionamiento

Las fábricas, a diferencia del ciclo arbitreado no son gobernados por las direcciones AL_PA's. A los puertos N conectados a una fábrica se les asigna una dirección de 24 bits, que teóricamente permite más de 16 millones de posibles direcciones: 2^{24} . Aunque el esquema de direccionamiento de los concentradores reduce este número a aproximadamente 15.5 millones. Esta cantidad de direcciones disponibles resulta bastante buena considerando que en la actualidad la mayoría de los concentradores tienen entre 8 y 64 puertos, lo que da cabida a una gran cantidad de dispositivos adicionales.

El estándar FC-SW (*Fibre Channel Switch Fabric*) divide las direcciones de puerto de 3 bytes en tres campos de un byte: El dominio, que es el más significativo, el área, el siguiente más significativo y el puerto, el de menor importancia. El campo de dominio tiene un número de direcciones reservadas, dejando 239 disponibles para asignárselas a los concentradores. Esto permite que cada concentrador en una fábrica extendida tenga un identificador único pero también limita la fábrica más grande posible a 239 concentradores interconectados. El campo de área proporciona 256 direcciones que pueden ser usadas para identificar puertos FL individuales o grupos de puertos F. Finalmente el último campo provee otras 256 direcciones para identificar puertos N o puertos NL. La figura 5-15 ejemplifica este esquema de direccionamiento.



Figura 5-15. Esquema de direccionamiento en una fábrica

La división del espacio de direcciones de 24 bits en dominio, área y puerto permite una mejora sustancial en el desempeño permitiendo que la decisión de ruteo se haga en base a un solo byte. Si una trama, por ejemplo, tiene una dirección de dominio que pertenece a un concentrador diferente en la fábrica, el motor de ruteo puede avanzar la trama a la interconexión apropiada o puerto E, sin tener que procesar la dirección completa de 24 bits.

Formas de conexión

Ya que las fábricas son extendidas vía conexiones concentrador a concentrador a través de los puertos E, un cuello de botella potencial puede ocurrir durante períodos de gran actividad entre concentradores. Este problema podría ser resuelto proporcionando múltiples puertos E entre dos concentradores interconectados y un algoritmo de balanceo de cargas para distribuir el tráfico. Sin embargo, tal solución puede resultar en una entrega sin orden de las tramas, ya que una secuencia de tramas podría tomar rutas diferentes a través de la fábrica. Consecuentemente, algunas implementaciones permiten solo una conexión activa entre dos concentradores. Si se tiene un puerto E de respaldo este solo es habilitado si el puerto principal falla.

Registro a la fábrica

Cuando un nodo es conectado a un concentrador en una fábrica, el primer proceso que se lleva a cabo es un registro a la fábrica o FLOGI. Como en el registro a puertos, FLOGI es un proceso que establece una sesión entre dos participantes. En el caso de FLOGI, una asociación o sesión de registro es creada entre el puerto N o puerto NL y el concentrador. El proceso de registro comienza cuando un puerto N envía una trama FLOGI conteniendo el nombre del nodo y el puerto (*world wide name*) así como una serie de parámetros de servicio solicitados al concentrador, en el caso de un puerto NL, este primero abre el AL_PA de destino antes de enviar la trama FLOGI. En ambos casos, el manejador de registros en el concentrador acepta el registro retornando una trama ACC al puerto que envió la trama FLOGI. Si el concentrador no soporta algún parámetro de servicio solicitado, pondrá ciertos bits en la trama ACC para indicar esto.

Un puerto N utiliza una dirección de puerto de 24 bits de 'x000000' cuando se registra por primera vez en la fábrica. Esto permite que el concentrador asigne la dirección de puerto apropiada al dispositivo basado en el formato de direccionamiento de dominio/área/puerto. La nueva dirección asignada es contenida en el campo de destino de la trama ACC.

Un proceso similar es usado para un puerto NL, excepto que el *byte* menos significativo es usado para asignar el AL_PA y los 2 *bytes* más significativos contienen el identificador de la fábrica. Ya que el puerto NL puede haber venido de un proceso de inicialización de ciclo es probable que ya tenga una dirección AL_PA asignada, la fábrica decidirá si acepta o no esta dirección. En caso negativo una nueva dirección AL_PA será asignada, lo cual ocasionará que se lleve a cabo otro proceso de inicialización en el ciclo, al cual pertenece dicho puerto NL. Esto asegura que el AL_PA asignada por la fábrica no entre en conflicto con ninguna otra AL_PA en el ciclo.

No se puede asegurar que todas las tarjetas controladoras o discos de canal de fibra soportan registro a fábricas o que diferentes implementaciones de fabricantes de dispositivos de canal de fibra son totalmente compatibles. Una gran base instalada de productos de canal de fibra sin capacidad de fábrica fueron diseñados para configuraciones punto a punto o de ciclo privado. Afortunadamente se están dando pasos firmes buscando que en la actualidad todos los nuevos dispositivos de canal de fibra sean capaces de integrarse a ambientes de fábrica.

Servidor de Nombres

En el caso de un ciclo arbitreado, un servidor o iniciador debe comenzar un proceso de descubrimiento de dispositivos destino realizando registros a los puertos del ciclo para las 126 direcciones AL_PA disponibles. A velocidades de *gigabit*, este proceso se lleva a cabo de manera relativamente rápida y permite que cada servidor encuentre sus discos fácilmente cada vez que hay algún cambio en el ciclo.

Sin embargo, este proceso resulta inadecuado en ambientes de fábrica ya que el espacio potencial de direcciones es de más de 15.5 millones. Incluso a velocidades de *gigabit*, tomaría un tiempo considerable. Las fábricas racionalizan el proceso de descubrimiento de dispositivos a través de un servidor de nombres en cada concentrador (*switch*). El servidor de nombres es una base de datos de objetos que permite que cada dispositivo conectado al concentrador se registre o solicite información útil tal como el nombre, dirección y la clase de servicio de los otros participantes.

El servidor de nombres es accesado realizando un registro a puerto a la dirección 'xFFFFFFC'. Un puerto N o NL se registra en el servidor de nombres después de realizar un PLOGI a esta dirección y terminar la sesión. El dispositivo puede registrar diversos valores para algunos o todos los objetos de la base de datos, pero la información más útil que puede guardar es la dirección de 24 bits del puerto así como su nombre de 64 bits, el nombre del nodo, la clase de servicio y el tipo de puerto, ya sea N o NL. En algunos casos, algunos dispositivos pueden registrar también sus direcciones IPv4 ó IPv6. Ya que el servidor de nombres contiene una tabla para las direcciones de 24 bits y sus correspondientes WWN de 64 bits, un dispositivo puede encontrar a otro fácilmente sin tener que hacer una búsqueda por 15.5 millones de direcciones posibles.

Una área que se encuentra todavía en construcción es el intercambio de información entre servidores de nombres de diferentes concentradores, particularmente entre concentradores de diferentes proveedores. Esta característica será necesaria para implementar grandes redes de almacenamiento inteligente ya que permitirá que cualquier dispositivo en una fábrica extendida descubra y comience operaciones con cualquier otro dispositivo de una manera fácil y rápida.

Notificación de cambio de estado

Ya que las fábricas son ambientes dinámicos y se pueden dar cambios en la topología de conexión conforme nuevos dispositivos son agregados o dispositivos activos son quitados de línea, resulta bastante útil si existe algún mecanismo de notificación para todos los participantes. Esta función es provista por un notificador de cambio de estado (SCN, *State Change Notification*) y una variante de éste denominada notificador de cambio de estado de registro (RSCN, *Registered State Change Notification*). Para que estas funciones estén disponibles, un puerto N o NL deben de especificar su interés de ser notificados de cualquier cambio cuando se registran en el concentrador.

El servicio SCN permite que un puerto N envíe un cambio de notificación directamente a otro puerto N. Mientras que RSCN permite que todos los dispositivos que han sido registrados con la opción de notificación sean informados si ocurre cualquier cambio en la topología. El tipo de cambio que tiene mayor sentido para los participantes es cuando un dispositivo con el que se han estado comunicando desaparece de la fábrica. Los servidores, en particular necesitan ser notificados cuando los recursos ya no están disponibles o cuando nuevos recursos son agregados. Las notificaciones de cambio proporcionan a las fábricas una función que intrínsecamente proporciona la inicialización de un ciclo en un ambiente de ciclo arbitreado.

Soporte de ciclo privado

Las fábricas soportan dispositivos en un ambiente de ciclo arbitreado público a través de los puertos FL. Estos puertos tienen la más alta prioridad, ALPA, 'x00', por lo que siempre ganan el arbitraje en el ciclo, un puerto FL proporciona la interface entre un ciclo y el resto de la fábrica. Sin embargo, para que un dispositivo en un ciclo se pueda comunicar en un ambiente de fábrica, debe ser capaz de realizar un registro en la fábrica seguido de una inicialización de ciclo.

El permitir que un puerto NL en un ciclo privado pueda participar en una topología de concentrador tiene grandes beneficios para las redes de almacenamiento y aunque resulta relativamente sencillo reemplazar una tarjeta controladora que no soporte fábrica por otra que si tenga esa capacidad, el tratar de cambiar un sistema de discos que sea compatible con fábrica en ocasiones puede resultar complicado y sobre todo costoso.

El soporte de un ciclo privado en un concentrador implica inteligencia adicional para monitorear la actividad de los dispositivos en el ciclo privado y asegurar que la existencia de la fábrica es transparente para ellos. Debido a que los dispositivos en un ciclo privado no están concientes de la presencia de la fábrica, el soporte de ciclo privado en un concentrador es a menudo descrito como modo fantasma (*phantom mode*).

La figura 5-16 muestra como una estrategia de segmentación puede permitir que varios dispositivos en un ciclo privado sean dispersos en puertos de fábrica específicos.

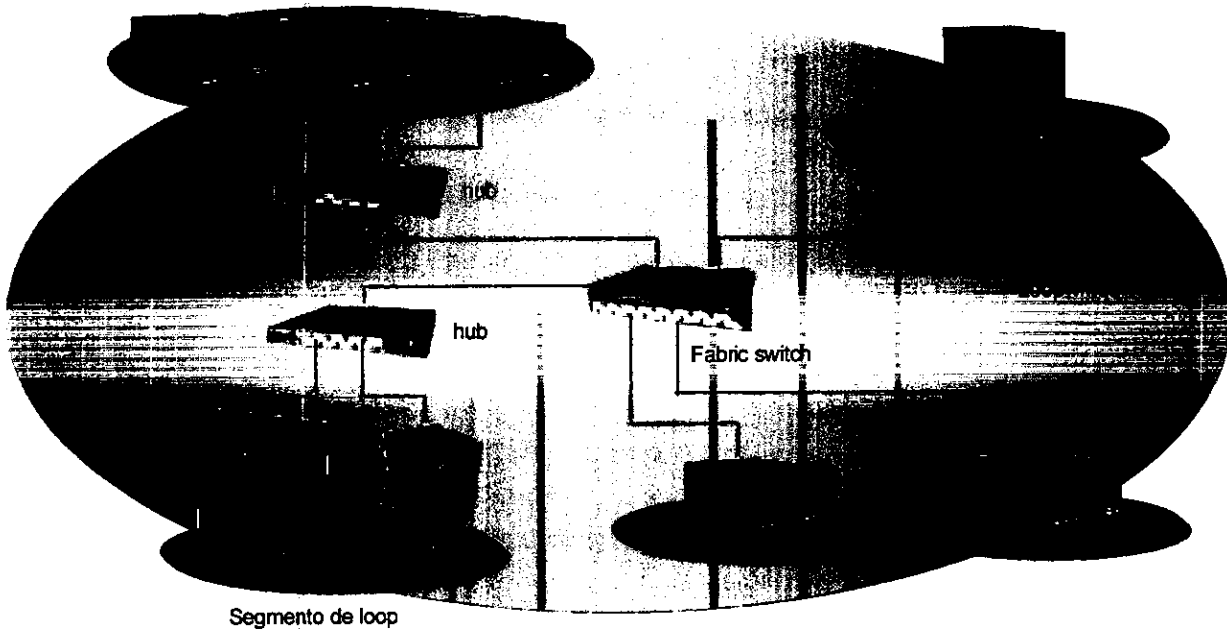


Figura 5-16. Soporte de ciclo privado en un concentrador de conexiones inteligente

Aunque el número total de dispositivos en un ciclo privado en un solo puerto puede estar restringido por el proveedor, generalmente este se encuentra determinado por el ancho de banda y consideraciones de tráfico. En una configuración de fábrica con ciclos privados es posible asignar a cada servidor un puerto de 100 MBps con arreglos de discos representando múltiples AL_PAs, asignadas a otros puertos. Por lo que lo que era un transporte compartido se convierte en una serie de ciclos de 100 MBps, algunos de los cuáles tendrán un solo puerto NL por ciclo. La segmentación transforma lo que previamente era un ciclo físico arbitreado en un ciclo virtual arbitreado.

Ya que los dispositivos en un ciclo privado no saben de la existencia de la fábrica, ellos, arbitrarán, abrirán, y cerrarán transacciones como antes. Un servidor, por ejemplo, arbitrará por el acceso de manera normal pero ya que es el único nodo en ese segmento de ciclo, no tendrá contendientes y ganará inmediatamente. Conforme el servidor envía tramas a una AL_PA de destino, la lógica del concentrador interviene y rutea las tramas directamente al puerto del concentrador que pertenece a la AL_PA correspondiente. Como se puede ver, la segmentación proporciona un método de tener un gran ancho de banda en una configuración de fábrica para todos aquellos dispositivos que no soportan conexión a un concentrador, además de que permite que las transferencias de información ocurran simultáneamente en el mismo ciclo virtual.

Zonificación

La zonificación permite la segmentación o agrupación lógica de múltiples puertos en una fábrica. Sin una manera de separar recursos, el poner por ejemplo información de mercadotecnia y de ingeniería en la misma red de almacenamiento podría ocasionar que se infiltrara información confidencial entre ambos departamentos. La zonificación resuelve estos problemas potenciales asignando puertos o dispositivos a zonas o grupos separados.

La zonificación puede ser implementada en un esquema de puerto a puerto que proporciona la mayor seguridad o en un esquema de direcciones de 24 bits o por WWN. La diferencia entre estos esquemas radica en la necesidad de tener un servidor que se encargue de la asignación y mantenimiento de estas zonas. La zonificación basada en puertos no requiere de este servidor ya que la asignación de puertos a una o más zonas puede ser mantenido por el propio concentrador.

Otra característica que se usa junto con la zonificación es el enmascaramiento de LUN's (*LUN masking*). Los arreglos de discos que utilizan RAID por *hardware* necesitan ser configurados con un nivel de RAID específico antes de que puedan trabajar en conjunto con los servidores, a estas configuraciones de RAID por *hardware* se les conoce como Números de Unidad Lógica (LUN's, Logical Unit Numbers) y consisten de una serie de discos con una configuración de RAID pero que son tratados como una unidad por el servidor, el enmascaramiento de LUN's consiste en dividir un LUN físico de un arreglo de discos en varios LUN's para que puedan ser distribuidos a través de diferentes servidores o equipos de cómputos, algunos sistemas de discos tienen esta funcionalidad integrada mientras que otros necesitan de un componente adicional para lograr esta característica.

Una aplicación clásica de zonificación es el respaldo a cinta de diferentes plataformas y sistemas operativos heterogéneos como se ilustra en la figura 5-17. Como se puede apreciar en la figura los servidores Windows NT con sus respectivos sistemas de discos pueden ser asignados a una zona A, mientras que los servidores Unix y sus arreglos correspondientes pueden ser asignados a una zona B.

Esta separación previene a los servidores NT de acceder la información contenida en los discos asignados a los equipos Unix. Sin embargo, tanto los servidores Sun como NT pueden ser asignados a una zona común C que incluya el sistema de respaldos lo que permite que ambas plataformas puedan usar la misma biblioteca de cintas para respaldar su información.

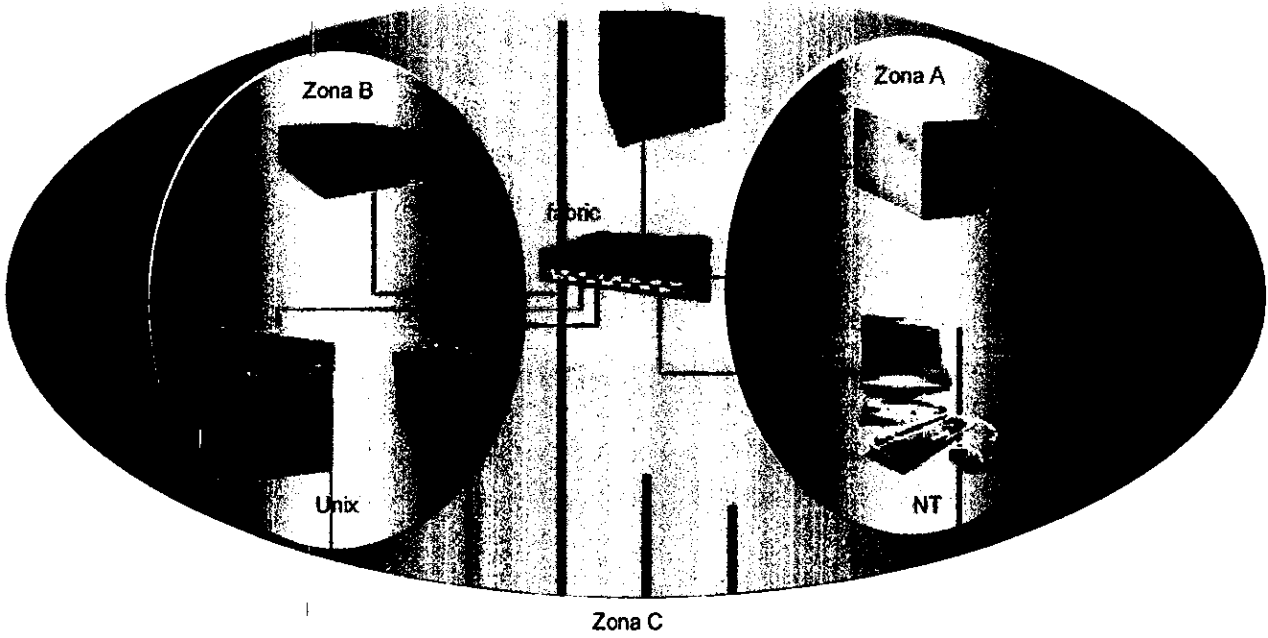


Figura 5-17. Zonificación

VI

Diseño, Implementación y Administración de ISN's

Antes de diseñar una red de almacenamiento inteligente es necesario determinar si realmente se necesita este tipo de solución. Cada empresa tiene una serie de objetivos y necesidades que hacen que lo que puede ser una buena solución de almacenamiento para algunos no lo sea para otros.

Si es posible identificar aplicaciones que sufran de falta de conectividad, de incapacidad para compartir información o recursos de almacenamiento, que no puedan ser respaldadas en los tiempos de los que se dispone o que estén limitadas de tal forma que no puedan proporcionar el servicio que se requiere de ellas, entonces una red de almacenamiento inteligente puede ser la solución. Si los usuarios necesitan mayor espacio en disco y los costos de almacenar y administrar la información se incrementan dramáticamente con el paso del tiempo, entonces una solución basada en ISN puede ser la respuesta. Así mismo, si se piensa adoptar un modelo de negocios basado en internet, proporcionando servicios las 24 horas del día, los siete días de la semana y no es posible tolerar la interrupción de la operación, entonces también una ISN puede ser lo que se está buscando.

El primer paso en el diseño de una red de almacenamiento inteligente consiste pues, en identificar todas las problemáticas asociadas con el manejo de la información, de esta manera se puede cuantificar como una ISN contribuirá a lograr los niveles de operación y servicio demandados por el negocio. Esto puede ser cuantificado en la medida en que una solución como esta traiga beneficios a la empresa al eficientar la habilidad de los administradores y directores de tomar buenas decisiones de negocio gracias a las mejoras en el manejo y disponibilidad de la información. También se puede medir en términos de la seguridad y la confiabilidad de las aplicaciones de misión crítica o en la reducción de costos asociados al *hardware* y a los recursos que se necesitan para su administración y mantenimiento.

6.1 Objetivos técnicos

Una vez que se han entendido los objetivos del negocio, estos marcarán la pauta para evaluar las necesidades técnicas y determinar la manera en que una solución basada en ISN puede satisfacerlas.

Los requerimientos pueden ser medidos en términos de confiabilidad, disponibilidad y servicio así como desempeño, escalabilidad, seguridad, manejabilidad y factibilidad de implementación.

Confiabilidad, Disponibilidad y Servicio

Cuando se habla de estos tres rubros es importante determinar la naturaleza de las aplicaciones que estarán corriendo en un ambiente de ISN. ¿Se trata de aplicaciones de misión crítica? Si no es así, ¿Cuánto tiempo sin servicio puede ser tolerado? ¿Bajo que circunstancias? ¿Cuáles son los costos asociados cuando se interrumpe la operación?

Para algunas empresas la no disponibilidad de las aplicaciones significa la pérdida de miles o millones de pesos por hora, para otros puede no tener tal efecto. La respuesta a estas interrogantes determinará las características de tolerancia a fallas de los equipos, los niveles de soporte y servicio así como los procedimientos y estrategias para prevenir o evitar dichas fallas.

Desempeño

¿Cuáles son las características de desempeño de las aplicaciones que correrán en una ISN? ¿Como se puede medir esto?, ¿En términos de ancho de banda, de IOPS o en tiempo de respuesta? ¿Cuál es la capacidad máxima o el ancho de banda requerido para cargas pico? ¿Qué porcentaje de la capacidad de la ISN será usado en promedio y a que nivel de utilización estará saturada? ¿Qué desempeño es tolerable en caso de una falla en alguno de los componentes?

Escalabilidad

¿Qué tanto crecimiento se espera? ¿La red de almacenamiento inteligente soportará aplicaciones adicionales en un futuro? Si es así, ¿En cuánto tiempo se espera este crecimiento: 6 meses, 1 año, 2 años? ¿Qué tan rápido es el crecimiento de los volúmenes de información que se manejan? ¿Se tiene planeado agregar más servidores o sistemas de almacenamiento a la ISN? Si es así, ¿En cuanto tiempo se tiene pensado hacer esto? ¿Se requiere soportar dispositivos basados en SCSI? ¿Será necesario incluir en el diseño dispositivos ubicados en sitios remotos? ¿Cuáles son las consideraciones de distancia que se deben tomar en cuenta?

Seguridad

¿Se tiene contemplado algún esquema de protección de datos? ¿Qué características de protección contra fallos (*failover*) necesitan ser consideradas en el diseño de una ISN para asegurar la continuidad de la operación en el evento de errores o desastres? ¿Se tendrán múltiples servidores de diferentes plataformas conectados a la ISN?

Este último punto puede resultar crítico en ambientes heterogéneos, especialmente si se tienen servidores NT ya que estos tratan de hacer suyos todos los dispositivos que detecten independientemente de que estén conectados directamente o a través de concentradores. Esta es una de las características que determinará el nivel y tipo de zonificación y enmascaramiento de LUN's apropiados para garantizar así que cada servidor accese únicamente el espacio de disco que le corresponde.

Manejabilidad

Como parte de los requerimientos técnicos es necesario considerar como será administrada la ISN y por lo tanto determinar las herramientas adecuadas para realizar esta función. Estas herramientas de administración deberán cubrir los siguientes aspectos:

Configuración

Facilidades para identificar, operar y controlar los datos así como los dispositivos que conforman la red de almacenamiento inteligente.

Acceso

Herramientas que permitan proteger los datos y asegurar que la información sea accesada solo por los dispositivos autorizados.

Desempeño

Herramientas para analizar el tráfico en la red de almacenamiento así como para analizar el comportamiento de las aplicaciones con el propósito de optimizar la red y hacer planes para enfrentar requerimientos futuros.

Fallas

Habilidad para detectar, aislar, corregir y reportar eventos en una ISN.

Factibilidad de Implantación

Cuando se está considerando una solución basada en ISN's es por que seguramente ya se han determinado los beneficios esperados. Pero el aspecto económico es muchas veces el principal punto a tomar en cuenta, por lo que es fundamental saber cuál es el criterio de decisión más importante en la solución que se está contemplando.

Para una aplicación de misión crítica el requerimiento más importante es la alta disponibilidad mientras que para una aplicación de interconexión remota el tema dominante será la conectividad. Por otro lado para un ambiente departamental, el bajo costo de la solución puede ser lo que se está buscando. Por lo tanto, como en cualquier decisión se tendrá que dar peso a los diferentes requerimientos que se tengan lo que tendrá como consecuencia que algunos aspectos tengan todas las características que se necesitan mientras que otros no.

La tabla 6.1 muestra los aspectos que se tendrán que sacrificar dependiendo del requerimiento principal que se tenga.

Objetivo de diseño	Aspecto Sacrificado
Alta disponibilidad	Mayor equipamiento y mayores costos
Alto desempeño	Mayor equipamiento y mayores costos
Alto nivel de seguridad	Mayor complejidad de uso y mayores costos de monitoreo y administración
Alta escalabilidad	Mayores costos con posibles impactos en alta disponibilidad
Bajo costo	Impacto en disponibilidad, desempeño, seguridad o escalabilidad

Tabla 6-1 Objetivos de diseño

= 6.2 Definiendo la infraestructura de una ISN

Si se está comenzando de cero con una red totalmente nueva entonces se puede hacer el diseño con la selección de la mejor topología y componentes. Pero en la mayoría de los casos es probable que se esté reemplazando una infraestructura de almacenamiento existente o incluso una ISN. Por lo tanto, antes de comenzar con la fase de diseño es importante entender que es lo que se está tratando de mejorar o reemplazar: el sistema de almacenamiento, la infraestructura de comunicaciones, el uso de las aplicaciones, las cargas de tráfico, la disponibilidad, el desempeño o la integridad de los datos.

Toda esta información permitirá determinar qué componentes deberán ser usados en la nueva topología y que se requerirá en el proceso de migración del ambiente actual al nuevo.

6.2.1 Características de tráfico de las aplicaciones

Antes de seleccionar una topología de ISN se necesitará entender la naturaleza del tráfico que se tiene. ¿Que servidores y dispositivos de almacenamiento generarán los movimientos de datos? ¿Cuáles son las fuentes y los destinos finales del flujo de información? ¿Qué cantidad de datos fluirá de un dispositivo de almacenamiento a otro? ¿Qué tipo de información se manejará: bases de datos, información de audio, video o texto simple? ¿Cuáles son los tamaños de los bloques de datos enviados por las diferentes aplicaciones? ¿Qué clase de servicio de canal de fibra se necesitará? ¿Qué departamentos o grupos de trabajo generarán el tráfico en la red?

Estos puntos permitirán encontrar oportunidades para consolidar el almacenamiento, establecer políticas de seguridad y ayudarán también a determinar el número de nodos de canal de fibra así como la topología de la solución.

6.2.2 Plataformas y almacenamiento

¿Qué tipo de servidores y que sistemas operativos formarán parte de la infraestructura de red de almacenamiento? Aunque la mayor parte de las primeras implementaciones de ISNs han sido en ambientes homogéneos, conforme esta tecnología ha ido madurando la tendencia es hacia la integración de redes a gran escala soportando múltiples plataformas de cómputo lo cual tiene implicaciones en seguridad que deben ser tomadas en cuenta.

Para que un servidor sea capaz de integrarse a un ambiente de ISN es necesario que cuente con tarjetas adaptadoras que puedan trabajar en un esquema de fábrica. La elección de una tarjeta adaptadora puede estar ya decidida por el fabricante del servidor o servidores. En la mayoría de los casos se requerirán únicamente dos tarjetas adaptadoras por servidor para poder realizar la conexión hacia una red de almacenamiento inteligente, esto por cuestiones de redundancia y tolerancia a fallas.

Otro aspecto importante a considerar son los sistemas de almacenamiento y bibliotecas de cintas, para lo cual se deberá saber si los arreglos de discos existentes necesitarán integrarse al esquema de ISN para tomar en cuenta los dispositivos necesarios: puentes o compuertas que permitan la integración de sistemas de almacenamiento con interfaces SCSI o SSA al nuevo esquema basado en canal de fibra.

6.2.3 Concentradores de conexiones (hubs)

Similarmente a los concentradores Ethernet o Token Ring un concentrador de canal de fibra es un concentrador de conexiones. Los concentradores fueron diseñados en respuesta a los problemas que aparecieron con las primeras implementaciones de ciclo arbitreado en donde simplemente se interconectaban punto a punto todos los participantes en un ciclo y la falla de cualquier dispositivo ocasionaba que el ciclo completo se viniera abajo. Los concentradores resuelven este problema al permitir que todos los dispositivos sean conectados en un punto central usando una configuración de estrella, permitiendo además que el ciclo pueda ser reconfigurado automáticamente si un dispositivo es agregado o quitado de él.

Los concentradores son dispositivos normalmente pasivos en una red de almacenamiento y su función se restringe principalmente a la verificación de señales. A diferencia de las tarjetas adaptadoras, de los arreglos de discos y de los concentradores inteligentes (*switches*), los concentradores no tienen una dirección de canal de fibra y no participan en ninguna actividad a nivel de protocolos, por lo que la tarea principal de un concentrador es simplemente facilitar la comunicación entre nodos.

Aunque el estándar de ciclo arbitreado permite tener hasta 126 nodos más un puerto de fábrica conectados a un solo ciclo, la realidad es que no se recomienda que una configuración de ciclo arbitreado tenga más de 60 puertos usables. De hecho, la mayoría de las implementaciones que usan topología de ciclo están en el rango de 5 a 30 nodos.

Ya sea que se trate de un solo concentrador o de varios interconectados en cascada estos comparten un solo segmento de red de 100 MBps. El agregar más dispositivos a un solo segmento divide el ancho de banda disponible a cada nodo. En una configuración de red de almacenamiento, los participantes activos son generalmente los servidores, mientras que los arreglos de discos simplemente responden a las peticiones de los servidores, por lo que la consideración más importante cuando se habla del ancho de banda de un ciclo son los nodos activos. De hecho, no se recomienda tener más de 4 nodos activos en una configuración de ciclo. Como se puede ver, los concentradores generalmente son utilizados en implementaciones sencillas con ambientes homogéneos. Algunas de las aplicaciones que se benefician con el uso de concentradores son las soluciones de agrupación de servidores, de consolidación de almacenamiento y de espejeo de discos.

La figura 6.1 muestra una aplicación común usando concentradores. Para asegurar la mayor disponibilidad posible, los servidores están equipados con dobles tarjetas adaptadoras, mientras que los arreglos de discos tienen conexiones a dos concentradores separados. El tráfico normal de datos pasa a través del ciclo primario; si este falla, el *software* en los servidores automáticamente dirigirá el tráfico al otro ciclo. La ruta redundante de datos asegura que la comunicación entre los nodos activos y los arreglos de discos siempre esté disponible.

Este esquema de alta disponibilidad puede ser complementado con *software* en cada servidor, lo cual permite que en caso de que un servidor falle el otro tome el control. Adicionalmente, se pueden implementar técnicas de RAID y espejeo de discos para asegurar la redundancia en los datos. Como se puede apreciar, una configuración compleja como lo es la agrupación de servidores puede ser implementada de una manera relativamente sencilla a través de concentradores.

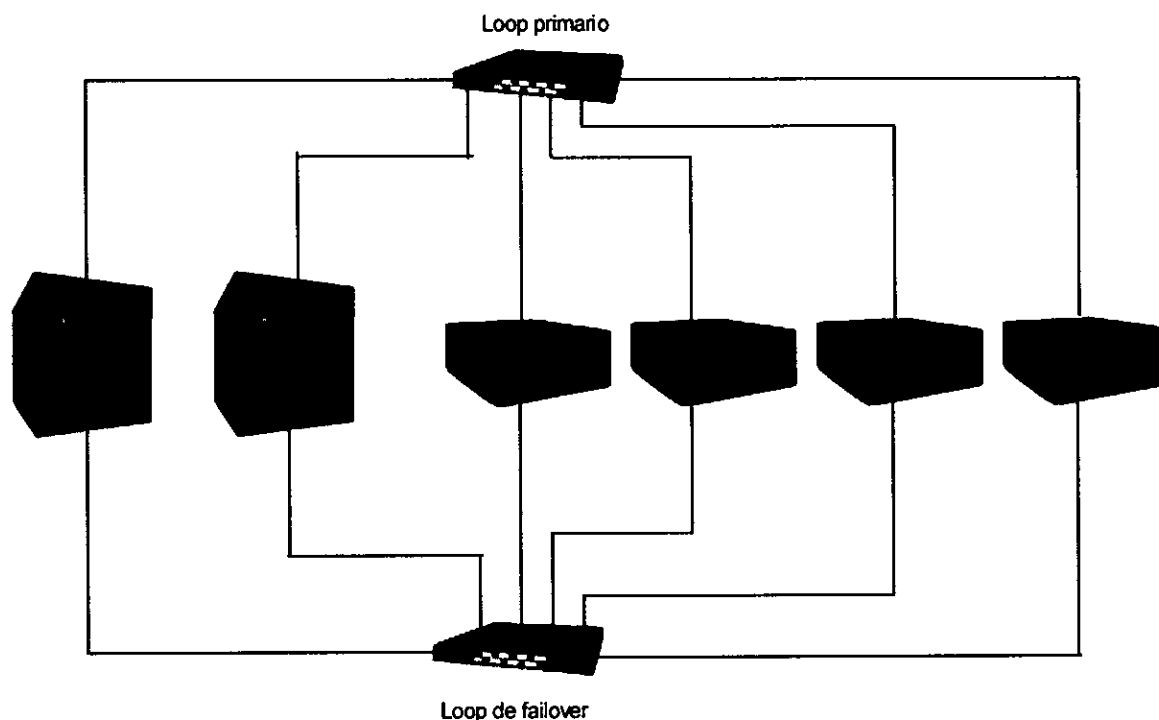


Figura 6-1. Agrupación de servidores (cluster) usando concentradores

6.2.4 Concentradores de conexiones inteligentes (switches)

Los concentradores de conexiones inteligentes son el elemento base para formar una red de almacenamiento inteligente. Estos son considerablemente más complejos que los concentradores simples tanto en diseño como en funcionalidad. Mientras que un concentrador simple es solo un concentrador de conexiones con un segmento compartido de 100 MBps, un concentrador inteligente proporciona un motor de ruteo de alta velocidad y 100 MBps de ancho de banda por cada puerto. Los concentradores inteligentes, a diferencia de los concentradores simples, son elementos altamente activos en las conversaciones de canal de fibra, tanto por los servicios que proporcionan (registro a la fábrica y servicio de nombres entre otros) como por el papel que juegan en la transferencia de datos. Las características de los concentradores inteligentes hacen que sean dispositivos costosos.

Como en el caso de los concentradores simples, los concentradores inteligentes pueden ser conectados en cascada para aumentar el número de puertos disponibles. Teóricamente, una red ISN puede tener hasta 15.5 millones de direcciones disponibles.

La función de un concentrador inteligente no está limitada a descubrir dispositivos sino a interactuar con ellos, los concentradores inteligentes proporcionan una serie de servicios incluyendo un servidor de nombres que es una pequeña base de datos la cual almacena información de los dispositivos que conforman la fábrica así como datos relevantes a los protocolos soportados por los diferentes nodos. Cuando un servidor se registra en un concentrador inteligente puede solicitar a la base de datos de éste información referente a todos los dispositivos que soporten el protocolo SCSI de esta manera el servidor puede establecer sesiones rápidamente con todos los arreglos de discos disponibles. Los concentradores inteligentes también son capaces de notificar a los diferentes nodos de cualquier cambio en la configuración de la red.

El decidir si usar concentradores inteligentes o simples no depende de las características de estos, la decisión radica en la solución de almacenamiento que se esté buscando. Para conectar un par de servidores a un arreglo de discos, por ejemplo, los concentradores simples pueden ser la mejor solución para aplicaciones de tipo transaccional, correo y otras con cargas de tráfico moderadas. De hecho, la mayoría de las implementaciones que usan canal de fibra en la actualidad usan configuraciones de ciclo arbitreado en donde se tienen de uno a cuatro servidores conectados a dos o hasta cuatro arreglos de discos.

Como un reemplazo de SCSI, los concentradores simples simplifican grandemente el esquema de cableado y proporcionan mayor ancho de banda, sin embargo, ya que el ciclo arbitreado es un medio compartido solo un servidor puede estar comunicándose activamente en un momento dado. Si se sabe que se pueden tener varios servidores y que estos no llegarán a saturar el medio compartido de 100 MBps que proporciona una configuración con concentradores simples entonces no tiene mucho sentido usar concentradores inteligentes. Sin embargo, si las aplicaciones demandan grandes anchos de banda entonces una configuración con concentradores inteligentes constituye la mejor opción.

Existen por otro lado algunas aplicaciones que no pueden tolerar ninguna interrupción en el proceso de comunicación por breve que este sea, tales como las aplicaciones de respaldo. Las rutinas de inicialización de ciclo que son un proceso común en el esquema de ciclo arbitreado hacen que el uso de concentradores simples no deba ser siquiera considerado para tales aplicaciones. En este caso el punto a considerar no es el ancho de banda sino la integridad de los datos, cosa que un concentrador inteligente puede garantizar gracias a la característica que tiene de proporcionar anchos de banda dedicados.

En esquemas de conexión a larga distancia se pueden utilizar concentradores simples para hacer las conexiones locales y concentradores inteligentes para las conexiones remotas. Similarmente, en un ambiente distribuido, en un edificio, por ejemplo, los concentradores simples pueden ser usados para conectar servidores y sistemas de discos en cada piso, con conexiones a concentradores inteligentes para lograr la interconexión de los diferentes componentes que se puedan encontrar distribuidos en el edificio.

Los concentradores inteligentes también juegan un papel muy importante en la consolidación de grandes cantidades de almacenamiento. El problema aquí es reducir los costos de administración concentrando múltiples aplicaciones de *software* en un repositorio central de datos. Un gabinete de discos, por ejemplo, puede ser subdividido en múltiples volúmenes, cada uno accesado por un servidor de aplicaciones diferente. Estos gabinetes pueden contener *terabytes* de información y docenas de aplicaciones.

6.3 Topologías de fábrica

El aspecto fundamental en el diseño de una red de almacenamiento inteligente es la selección de la topología más apropiada. La determinación de esta puede estar dada a partir de un diseño basado en una estrategia corporativa que considere la implementación de una ISN con el fin de ayudar a la empresa a cumplir con sus objetivos de negocio. También puede ser abordado desde el punto de vista departamental o de grupos de usuarios que tal vez estén buscando formas de eficientar su trabajo.

La elección de una topología de ISN también puede estar dada en función del riesgo que pueda estar asociado con su implementación. La realidad, es que esta tecnología no deja de ser nueva y que hay muchas cosas alrededor de ella que están todavía por definirse comenzando por los estándares.

La mayoría de las soluciones basadas en ISN que han sido implementadas a la fecha son relativamente pequeñas y fueron diseñadas para resolver problemas específicos. Por ejemplo, muchos usuarios han implementado soluciones de canal de fibra punto a punto para resolver problemas de distancia o desempeño, otros han utilizado la tecnología de canal de fibra para implementar soluciones de agrupación de servidores o de compartición de almacenamiento a través de la topología de ciclo arbitreado y unos más han comenzado a diseñar soluciones de almacenamiento basados en esquemas de fábrica aprovechando todas las ventajas que ofrecen los concentradores de conexiones inteligentes.

La fábrica más pequeña que se puede tener consiste de un concentrador inteligente de canal de fibra, de un servidor y de un sistema de discos, a partir de aquí es posible escalar una red de almacenamiento inteligente hasta el límite soportado por cada una de las topologías de fábrica, las cuáles son:

- Cascada
- Malla
- Anillo
- Arbol

6.3.1 Cascada

Una topología de cascada se puede ver como una cadena de concentradores inteligentes interconectados tal como se muestra en la figura 6.2. Los concentradores son acomodados en un arreglo lineal, cada uno conectado al siguiente concentrador de la línea o acomodados en una cascada vertical con múltiples niveles.

Los diseños de cascada son ideales para aplicaciones en donde el acceso a los datos es local, esto es, en donde los requerimientos de acceso para un servidor o grupo de servidores son típicamente hacia los mismos sistemas de discos. Los grupos de servidores y de sistemas de almacenamiento que están siendo accésados pueden ser conectados a los mismos concentradores proporcionando el mayor desempeño posible. El cascadeo permite que a una ISN se le pueda agregar fácilmente mayor conectividad y que la administración sea centralizada.

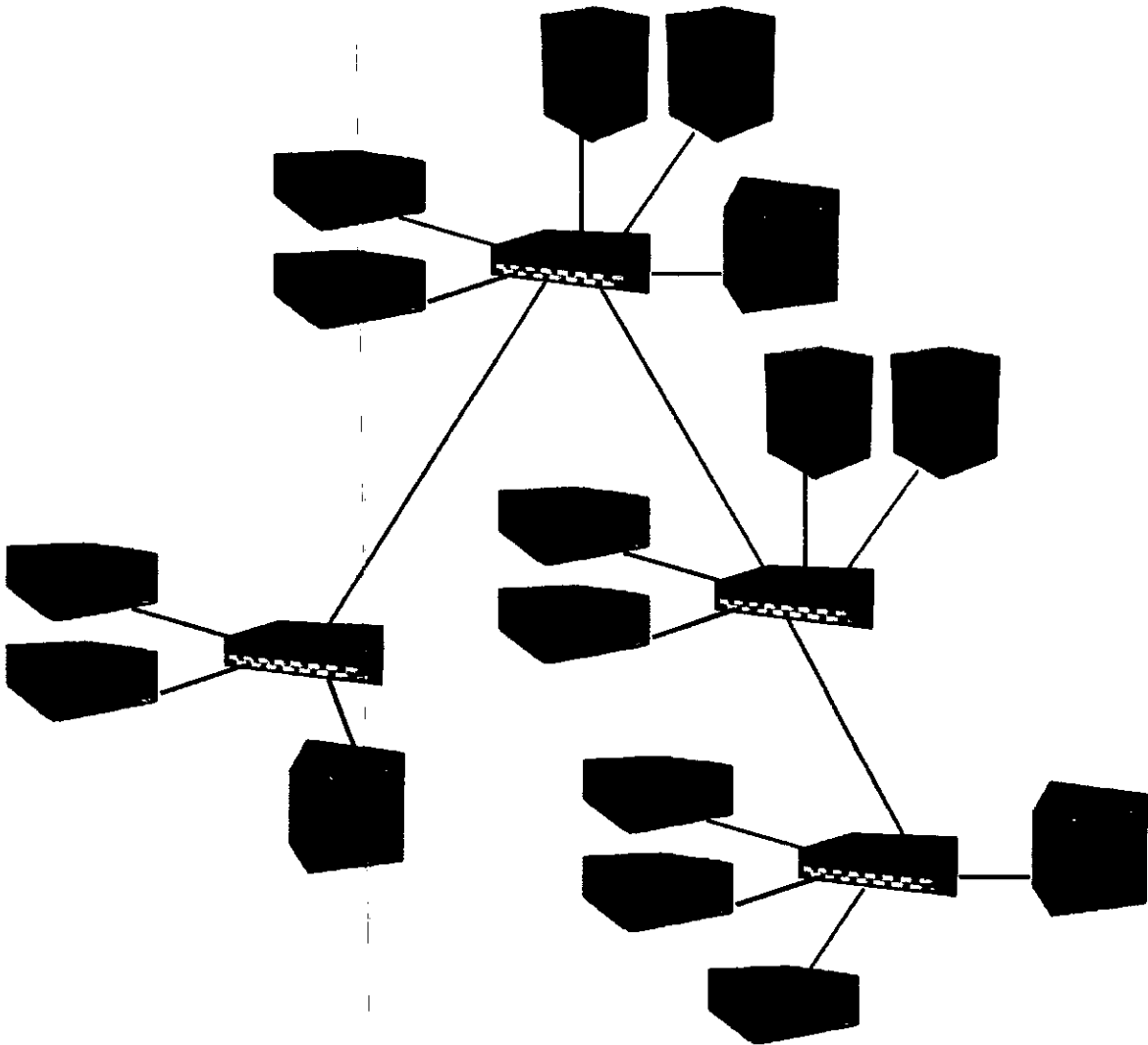


Figura 6-2. Fábrica en cascada

El diseño en cascada también puede ser usado para acceso distribuido, sin embargo, las rutas de tráfico deben ser bien entendidas para asegurarse de que existe el suficiente número de puertos y conexiones para cumplir con los requerimientos de desempeño. El usar más de una conexión entre concentradores inteligentes en cascada es una muy buena idea ya que proporciona rutas redundantes entre un par de concentradores dados en la fábrica

Ventajas del fabric en cascada

- Se adapta a diversas condiciones geográficas
- Escala fácilmente mejorando la conectividad
- Se soporta el respaldo compartido
- Proporciona administración centralizada
- Ideal para accesos locales

6.3.2 Malla

La topología de malla es similar a la de cascada, sin embargo, en este diseño todos los concentradores inteligentes están interconectados de manera que siempre hay por lo menos dos rutas de un concentrador a otro en la fábrica, incluso aunque solo se use una conexión entre concentradores. Este tipo de conectividad proporciona un alto nivel de confiabilidad ya que si falla una conexión o un concentrador, la fábrica puede automáticamente re-rutear los datos hacia la ruta disponible, aún y cuando se tenga que utilizar otro concentrador para ello.

En un diseño de malla con todos los concentradores interconectados, la eficiencia de la fábrica se puede ir decrementando a medida que aumenta el número de concentradores ya que puede llegar un momento en el que el número de puertos disponibles para servidores y sistemas de almacenamiento sea mínimo debido a los puertos ocupados para interconectar los concentradores. Para evitar esto se puede implementar un diseño ligeramente modificado de malla como se muestra en la figura 6-3. En esta variante, conforme los concentradores son agregados, estos son conectados solo al concentrador adyacente, no a todos los concentradores en la fábrica. Este diseño continúa proporcionando los beneficios de una conectividad total sin decrementar la eficiencia.

La topología de malla es ideal para aplicaciones en donde el acceso a los datos es local y distribuido. La completa conectividad de este diseño asegura el acceso de múltiples servidores a diversos sistemas de discos.

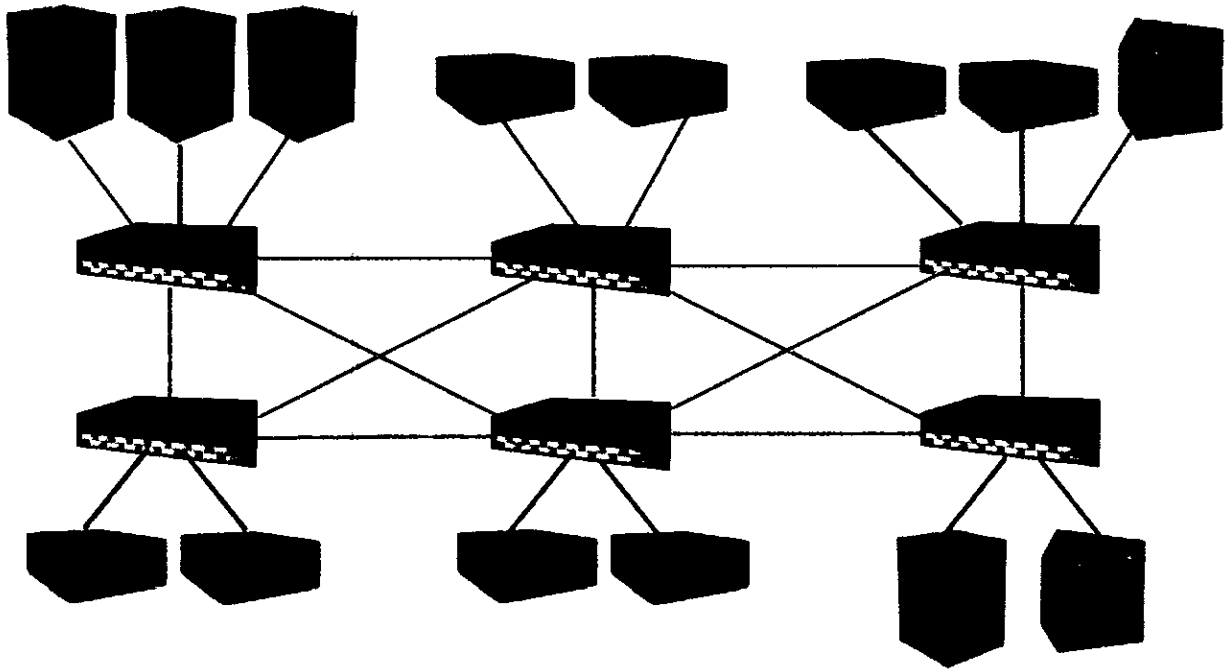


Figura 6-3. Fábrica en malla

Ventajas de la fábrica en malla

- Puede ser configurado para acceso local o distribuido, o ambos.
- Se adapta fácilmente a nuevas o diferentes necesidades de almacenamiento.
- Proporciona protección contra fallas de puertos o concentradores aún y cuando se tenga una sola conexión de un concentrador a otro.
- Escala fácilmente. Este diseño se puede extender a cuatro, seis, ocho o más concentradores sin ocasionar ninguna interrupción a la red de almacenamiento existente.
- Se soporta el respaldo compartido. Una o más bibliotecas de cintas pueden ser agregadas a la fábrica en varios puntos sin impactar el desempeño o la administración.
- Administración centralizada. Las herramientas de administración pueden navegar y administrar la ISN a través de las interconexiones entre concentradores.

6.3.3 Anillo

Esta topología consiste en un anillo de concentradores interconectados en una sola red de almacenamiento. Tal como se muestra en la figura 6-4. Cada concentrador es conectado a los dos concentradores adyacentes, de tal forma que el último y el primer concentrador del anillo se encuentran conectados. Este arreglo proporciona casi el mismo nivel de confiabilidad que el diseño de malla. El máximo número de concentradores soportados en una fábrica en anillo es 14.

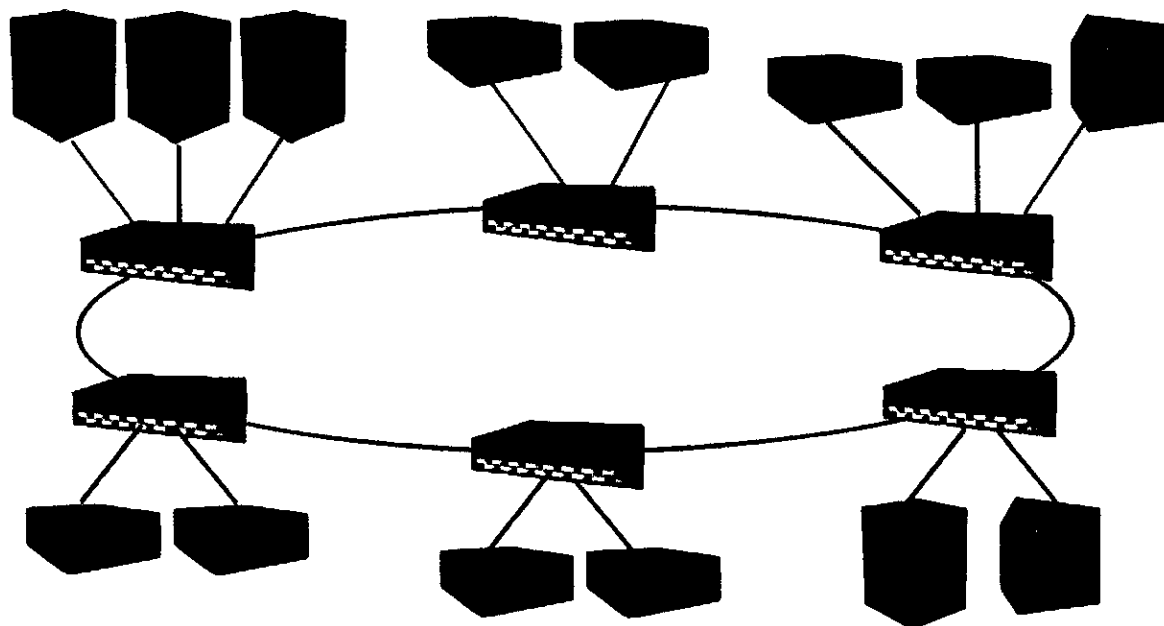


Figura 6-4 Fábrica en anillo

Los diseños de anillo son adecuados para aplicaciones en donde el acceso a los datos siempre es local, en donde los servidores y el almacenamiento que está siendo accedido están en el mismo concentrador y la mayor parte del tráfico es vertical a través del mismo concentrador. Esta implementación proporciona una forma de escalar la fábrica de una manera modular agregando un concentrador y grupos de servidores y de almacenamiento formando una célula, usando un método de bloque para incrementar el tamaño de la ISN.

Una gran ventaja del diseño de anillo es que puede ser preconfigurado e instalado antes de que los requerimientos de las aplicaciones sean conocidos. Este diseño no es recomendado para aplicaciones que requieren una conectividad de muchos a muchos.

Ventajas de la fábrica en anillo

- Fácil de construir. Cada concentrador puede soportar sus propios servidores y almacenamiento, ahorrando tiempo de diseño e implementación.
- Crecer la infraestructura de red es relativamente simple y no ocasiona que se tenga que interrumpir el servicio. Los concentradores pueden ser agregados uno a la vez, conforme se incrementen las necesidades de almacenamiento.
- Se soporta el respaldo compartido. Una o más bibliotecas de cintas pueden ser agregadas al anillo en diferentes puntos sin que haya un efecto negativo en el desempeño o en la administración.
- El diseño inherente de esta topología hace que el acceso local sea óptimo. La mayoría del tráfico de datos se lleva a cabo por un mismo concentrador.
- Diseño modular.

6.3.4 Arbol

Un diseño de árbol tiene uno o más concentradores de canal de fibra que actúan únicamente como medios de comunicación para los demás concentradores en la fábrica. Estos elementos de conexión actúan como ruteadores proporcionando completa conectividad y ancho de banda para los demás concentradores. Este tipo de implementación ofrece la mejor conectividad «muchos a muchos» y ancho de banda distribuido a través de la fábrica.

Este diseño es ideal para implementaciones en donde el tráfico de datos puede variar incluso de manera aleatoria, pero en donde se tiene necesidad de una conectividad de muchos a muchos con grandes demandas de desempeño.

Arbol de tres niveles

Como su nombre lo dice una implementación de este tipo consta de tres niveles en donde los concentradores del nivel medio son usados para propósitos de distribución y comunicación. Los servidores y los sistemas de discos se conectan a los niveles superior o inferior de la infraestructura proporcionando un alto nivel de conectividad y flexibilidad. La figura 6-5 muestra la topología de árbol de tres niveles.

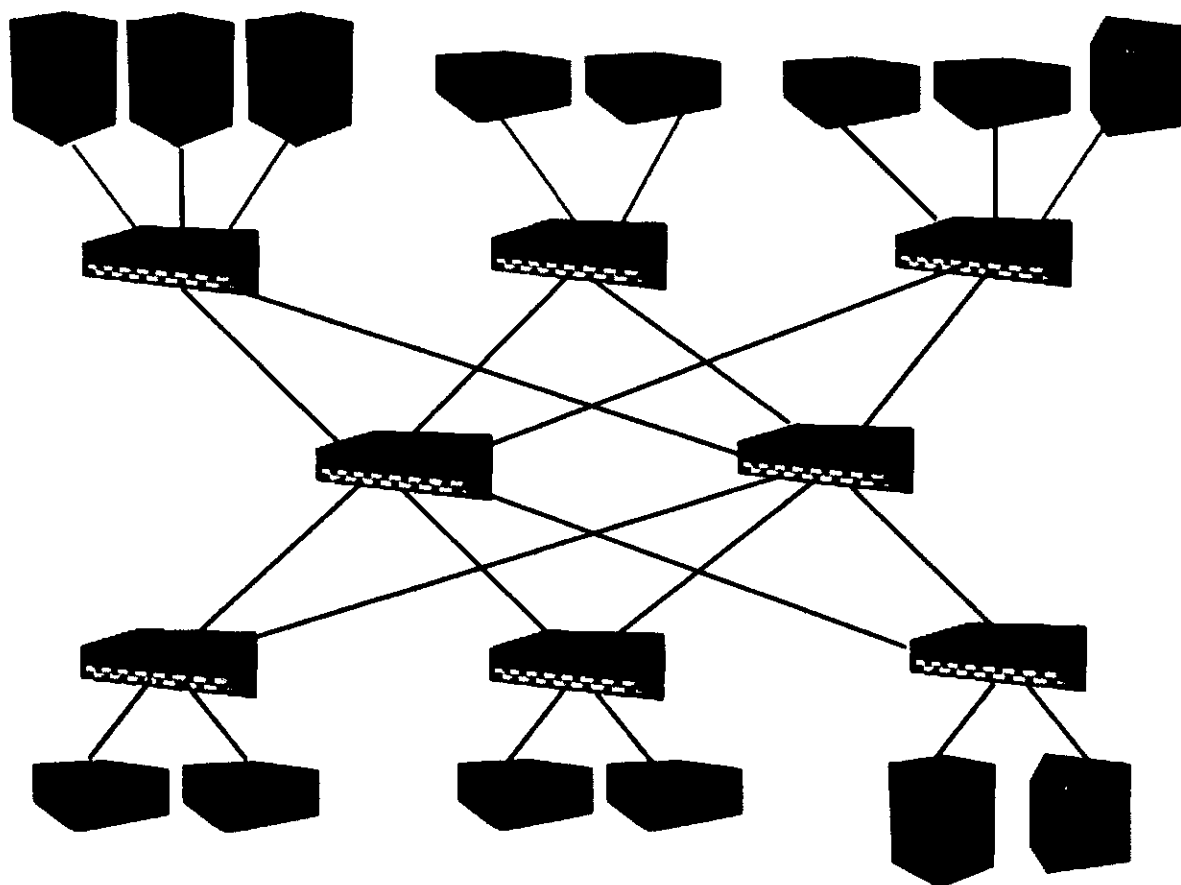


Figura 6-5. Fábrica en árbol de tres niveles

Ventajas de la topología de árbol de tres niveles

- Eficiente expansión de puertos. Los concentradores que se vayan agregando solo necesitan ser conectados a los concentradores de interconexión, esto ahorra tiempo y esfuerzo a la vez que el expandir la red de almacenamiento no provoca ninguna interrupción a la operación.
- Se ajusta fácilmente a las necesidades de alta disponibilidad y redundancia de componentes. El diseño de árbol permite que los concentradores automáticamente re-ruteen el tráfico bajo condiciones de falla.
- La administración es centralizada
- Total conectividad «muchos a muchos» pudiendo manejar un esquema distribuido y altamente redundante.

- Se ajusta a variaciones de conexión y desempeño independientemente del lugar físico en donde se encuentren los servidores y los sistemas de discos.
- Es posible mejorar el desempeño utilizando múltiples conexiones a los concentradores.
- Ofrece la mayor flexibilidad para implementar diferentes tipos de acceso: Local, distribuido o centralizado.
- Puede ser implementado con capacidades de respaldo centralizadas, reduciendo el costo de las operaciones de respaldo y recuperación.
- Es ideal para aprovechar desarrollos futuros tales como la virtualización de almacenamiento.

Cada una de las topologías de fábrica pueden ser caracterizadas por las ventajas que ofrecen dependiendo del tipo de acceso a los datos. La tabla 6-2 muestra el grado de eficiencia de las diferentes topologías de ISN dependiendo de la forma en que se va acceder la información: local, centralizada o distribuida.

Topología	Localidad de los datos		
	Local	Centralizado	Distribuido
	Uno a uno	Muchos a uno	Muchos a muchos
Cascada	La más alta	No recomendado	No recomendado
Malla	Alta	Media	Alta
Anillo	La más alta	Media	No recomendado
Arbol	Media	La más alta	La más alta

Tabla 6-2. Comportamiento de las topologías de ISN de acuerdo al tipo de acceso a los datos

Algunas de estas topologías pueden ser modificadas para cumplir los requerimientos específicos de las aplicaciones. Por ejemplo, la topología de árbol de tres niveles proporciona la mejor conectividad «muchos a muchos» y permite que algunas porciones de la red puedan ser configuradas para acceso local, esto se logra conectando los servidores y los sistemas de disco al mismo concentrador.

La tabla 6-3 muestra el número máximo de puertos y dispositivos soportado por cada una de las topologías.

Topología	Número total de puertos		Número de puertos para dispositivos	
	Una fábrica	Dos fábricas	Una fábrica	Dos fábricas
<i>Cascada, Malla Anillo</i>				
Una fábrica: 2 a 20	320	640	200	400
Dos fábricas: 4 a 40				
<i>Arbol</i>				
Un fábrica: 5 a 20	320	640	192	384
Dos fábricas: 10 a 40				

Tabla 6-3. Número máximo de puertos y dispositivos por topología

6.4 Disponibilidad de la información en una ISN

La disponibilidad de los datos en una fábrica puede estar influenciada por muchos factores, incluyendo la arquitectura, el número de concentradores y el número de puertos disponibles. La disponibilidad de una fábrica puede ser clasificada en cuatro niveles.

- Rutas únicas hacia el servidor y el sistema de discos, usando un solo concentrador.
- Rutas únicas hacia el servidor y el sistema de discos, usando topología de malla.
- Múltiples rutas hacia el servidor y el sistema de discos, usando topología de malla.
- Múltiples rutas hacia el servidor y el sistema de discos con múltiples fábricas.

Nivel 1

En este diseño los servidores y arreglos de discos son conectados al concentrador usando conexiones no redundantes tal como se muestra en la figura 6-6. Este esquema no proporciona ningún nivel de seguridad en la fábrica pero ofrece la mayor conectividad. Esta implementación debe ser usada cuando el desempeño está por encima de la disponibilidad de la información.

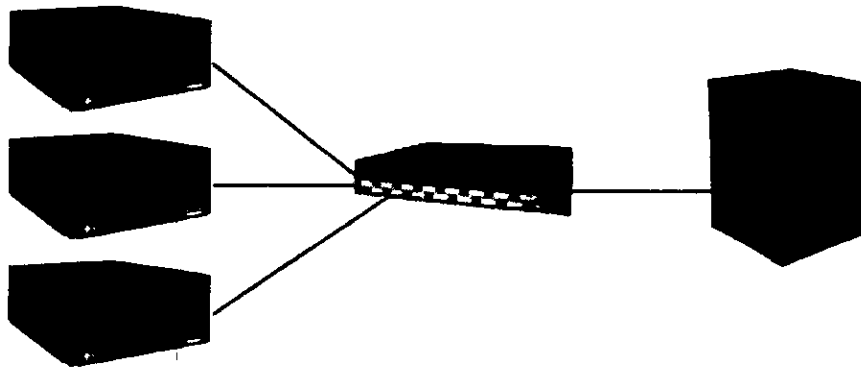


Figura 6-6. Nivel 1 de disponibilidad en una fábrica

Nivel 2

En este diseño se requieren por lo menos dos concentradores los cuáles son interconectados entre sí con conexiones redundantes. Los servidores y los sistemas de discos se conectan a la fábrica usando conexiones simples. La figura 6-7 ilustra este nivel de disponibilidad.

Este esquema asegura que en el evento de que una conexión entre los dos concentradores falle, los datos sean ruteados a través de la otra conexión. Sin embargo, la falla de una ruta de conexión entre los discos o los servidores hacia el concentrador impide que estos puedan continuar accedendo la información. La falla de un concentrador también impide que algunos de los dispositivos en la ISN puedan tener acceso a los datos.

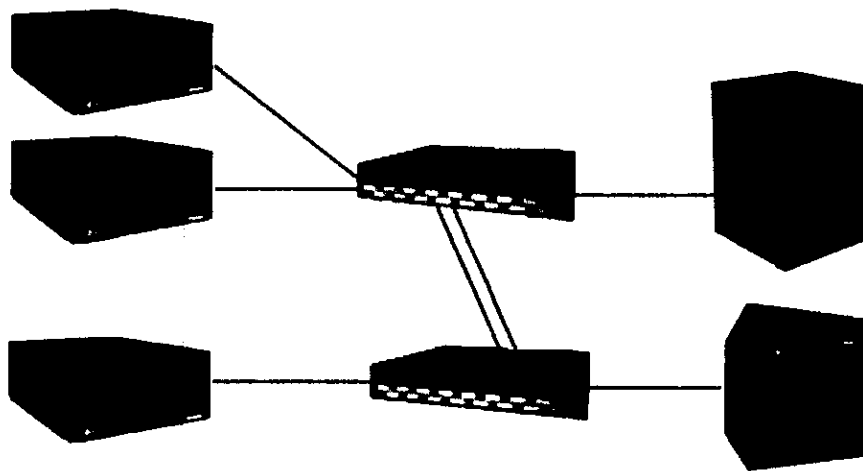


Figura 6-7. Nivel 2 de disponibilidad en una fábrica.

Nivel 3

Este diseño de conexión está basado en el nivel dos, pero a diferencia de este, los servidores y sistemas de discos tienen múltiples conexiones hacia los concentradores. Este diseño permite que aunque haya una falla en un concentrador, tarjeta adaptadora o ruta, el acceso a la información de cualquiera de los servidores a cualquier de los sistemas de discos esté garantizado. Este nivel requiere el uso de zonificación para definir dos rutas separadas a la fábrica. La figura 6-8 muestra un ejemplo de este nivel de disponibilidad.

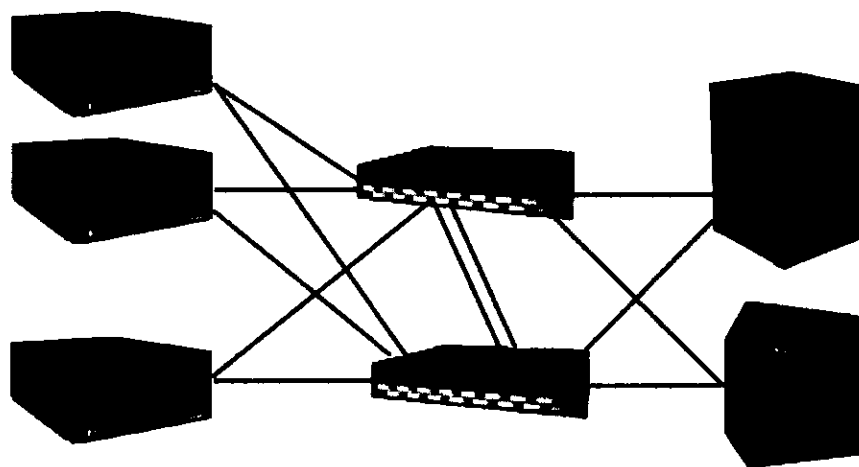


Figura 6-8. Nivel 3 de disponibilidad en una fábrica.

Nivel 4

Este esquema proporciona múltiples rutas hacia servidores y sistemas de discos como en el nivel tres, pero en este caso se utilizan múltiples fábricas en la implementación. Este diseño proporciona el mayor nivel de disponibilidad ya que no existe un solo punto de falla. Cualquier actividad que pueda afectar el desempeño o la disponibilidad de la fábrica será solucionado ruteando los datos a la otra fábrica. Los servidores no ven ninguna interrupción en el flujo de I/O.

Este diseño elimina cualquier vulnerabilidad que pueda existir por fallas en la fábrica, por ejemplo, errores humanos al reemplazar concentradores, errores en configuración o falla de alguno de los componentes a la vez que proporciona el mayor desempeño así como un gran número de puertos disponibles.

Este tipo de implementación está disponible para todas las topologías de ISN, basta con duplicar la infraestructura de comunicaciones entre los nodos participantes. Sin embargo, esto incrementa al doble el costo de la implementación, pero los beneficios obtenidos son también multiplicados al doble.

Por ejemplo, el implementar una fábrica utilizando cuatro concentradores en una topología de malla con múltiples rutas a servidores y arreglos de discos proporciona 52 puertos para servidores y discos. Implementando la misma topología, pero usando dos fábricas hace que se tengan hasta 104 puertos disponibles para conectividad de arreglos y discos. El utilizar dos fábricas también permite que se puedan hacer transiciones de una topología a otra.

La figura 6-9 muestra la forma de implementar un nivel de disponibilidad de nivel cuatro con múltiples fábricas.

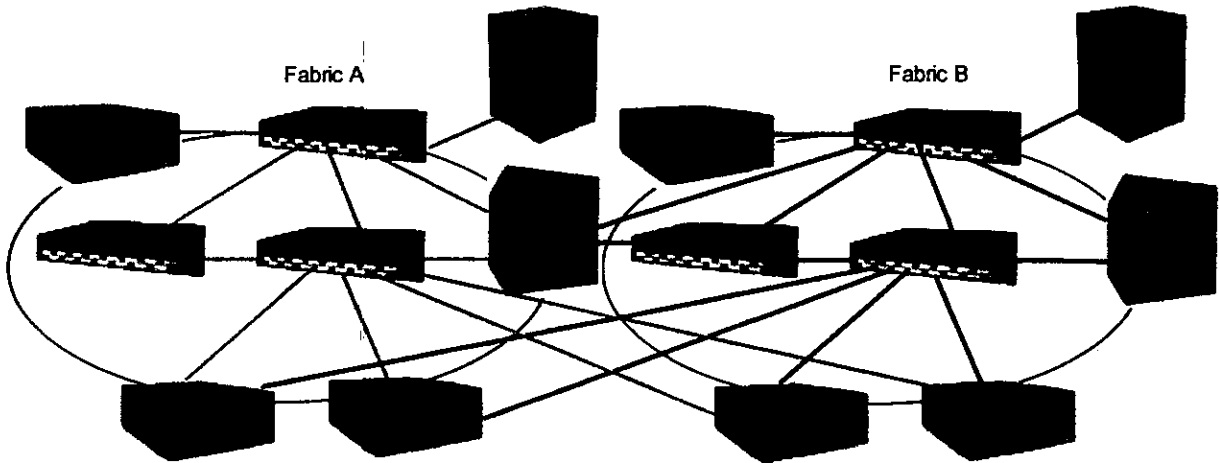


Figura 6-9. Nivel 4 de disponibilidad en una fábrica

La tabla 6-4 muestra la relación entre las topologías de fábrica y los niveles de disponibilidad de datos.

Diseño de fábrica	Nivel de disponibilidad	Topología
Una fábrica	(1) Sin redundancia	Uno o varios concentradores con conexiones simples
Una fábrica de malla	(2) Medio	Cascada, malla, anillo, árbol
Múltiples fábricas	(3) Alto	Todas
Una fábrica de malla	(3) Alto	Todas
Múltiples fábricas	(4) El más alto	Todas
Múltiples conexiones	(4) El más alto	Todas
Dos o más fábricas	(4) El más alto	Todas
Múltiples conexiones	(4) El más alto	Todas

Tabla 6-4. Niveles de disponibilidad en fábricas

Cuando se está pensando en seleccionar un nivel de disponibilidad el costo y la continuidad en el acceso a los datos son los criterios principales a considerar. Para aplicaciones de misión crítica se debe considerar siempre diseños con múltiples fábricas y conexiones redundantes, el costo adicional de tal implementación puede ser justificado fácilmente si se toman en cuenta los costos asociados cuando la operación es interrumpida.

La tabla 6-5 muestra los factores que se deben tomar en cuenta para los diferentes niveles de disponibilidad. En la tabla la x representa el costo de una fábrica.

Diseño de fábrica	Nivel	Costo de hardware
Una fábrica	1	x
Una fábrica de malla	2	x + puertos adicionales
Múltiples fábricas		
Una fábrica de malla	3	x + puertos adicionales
Múltiples fábricas		- puertos de HBAs
Múltiples conexiones		
Dos o más fábricas	4	x + puertos adicionales
Múltiples conexiones		+ HBAs adicionales + swiches adicionales

Tabla 6-5. Factores de disponibilidad

— 6.5 Aspectos a considerar al implementar ISN's

Reglas de Topología

- Aunque teóricamente se pueden tener hasta 15.5 millones de puertos disponibles en una fábrica, no se recomienda tener más de 20 concentradores en las topologías de cascada, malla o árbol.
- Hasta 14 concentradores inteligentes en una topología de anillo.
- No se recomienda tener más de siete niveles de concentradores entre dos dispositivos.

Recomendaciones para desempeño

- Cuando sea posible, los dispositivos que intercambien grandes cantidades de datos deben estar conectados al mismo concentrador.
- Cuando los dispositivos que intercambian datos se encuentren en diferentes dispositivos tomar en cuenta lo siguiente:
 - Reducir al mínimo el número de niveles de concentradores entre dispositivos.
 - Para aplicaciones que demandan gran ancho de banda, configurar un máximo de dos sistemas de discos activos por puerto.
 - Para aplicaciones con mucha demanda de I/O, configurar un máximo de seis dispositivos de almacenamiento por puerto.
 - Para aplicaciones con cargas mixtas, configurar un máximo de cuatro arreglos de discos por puerto.

Reglas de zonificación

Como ya se mencionó la zonificación puede ser usada para separar lógicamente dispositivos, plataformas de *hardware* y sistemas operativos en una misma ISN física. El uso de la zonificación es opcional en varios casos, pero su empleo es requerido bajo las siguientes condiciones:

- Cuando se mezclen diferentes plataformas de *hardware*, sistemas operativos y sistemas de discos que están soportados solo en ambientes de ISN homogéneos.
- Cuando se sabe que hay problemas en la interacción de diferentes plataformas de *hardware*, sistemas operativos y arreglos de discos, como cuando se tienen que integrar equipos con sistemas operativos UNIX y NT en la misma ISN.

Reglas de Interconexión

- Usar los puertos E para conectividad de concentrador a concentrador
- Emplear los puertos F para conectar iniciadores de dispositivo con soporte de fábrica (tarjetas adaptadoras)
- Utilizar puertos FL cuando se quieren conectar dispositivos de ciclo arbitreado a un ambiente de fábrica.

Consideraciones de distancia en fábricas

- La longitud mínima de un cable de fibra óptica debe ser de dos metros entre dispositivos de canal de fibra.
- Hasta 150 metros de distancia entre dispositivos y concentradores o entre concentradores usando cables de fibra óptica multimodo de 62.5/125 micrones y GBICs o GLMs de onda corta.
- Hasta un máximo de 500 metros de distancia entre dispositivos y concentradores o entre concentradores usando cables de fibra óptica multimodo de 50/125 micrones y GBICs o GLMs de onda corta.
- Hasta 10 kilómetros de distancia entre dispositivos y/o concentradores usando cables de fibra óptica monomodo de 9/125 micrones y GBICs de onda larga.
- Hasta un máximo de 70 kilómetros de distancia entre dispositivos usando siete segmentos de 10 kilómetros interconectando concentradores. Esta regla está basada en un máximo de separación de 10 Km entre concentradores.
- Hasta 100 kilómetros de distancia entre dos concentradores de canal de fibra usando cables de fibra monomodo de 9/125 micrones y GBICS de onda muy larga.

Reglas de configuración de ISN's heterogéneas

- Las siguientes plataformas y sistemas operativos están soportados en un ambiente de ISN: Compaq OpenVMS 7.2-1, 7.2-1H1, Tru64 UNIX 4.0F/G, Tru64 UNIX 5.0A/5.1, HP-UX 10.20/11.0, IBM AIX 4.2.1/4.3.0/4.3.1/4.3.2/4.3.3, Microsoft Windows NT 4.0 Windows 2000 SP1, Novell Netware 4.2/5.x, RedHat Linux 6.1/6.2, SUSE Linux 6.3, SGI IRIX 6.5.7/6.5.8, Sun Solaris 2.6/7/8
- En algunas situaciones se pueden necesitar múltiples arreglos de discos para acomodarse a los requerimientos de las diferentes plataformas y sistemas operativos.
- Todas las plataformas de *hardware* y sistemas operativos soportados en ambientes homogéneos están también soportados en un ambiente de ISN heterogéneo siempre cuando cada plataforma esté configurada en una zona de fábrica separada.
- Los servidores y arreglos de discos pueden estar conectados a múltiples fábricas.

6.6 Prácticas recomendadas en ISN's

6.6.1 Planeando una ISN

Una planeación apropiada debe considerar tanto los requerimientos actuales como los futuros. Por lo que es importante, que cuando se diseñe una ISN no solo se considere la mejor opción en términos de desempeño y necesidades actuales sino también se tome en cuenta que la ISN será expandida con el tiempo.

En la medida que se le dedique más tiempo a la fase de planeación y diseño mayores serán los beneficios en las etapas de configuración e implementación. Las siguientes son recomendaciones que se deben tener en cuenta en la fase de planeación y diseño:

Estrategia de desarrollo. En muchos casos es una buena idea pensar en desarrollar varias ISN's pequeñas con la idea de incrementar la capacidad de éstas independientemente o interconectándolas en un futuro. Las redes de almacenamiento pequeñas son más fáciles de construir y también permiten que el riesgo sea mínimo mientras se va ganando experiencia.

Diseño de topología. Es importante considerar una topología que dé respuesta a las necesidades presentes y futuras sin necesidad de cambiarla y que permita la migración a otros esquemas de una manera relativamente sencilla.

Avances en tecnología. Un buen diseño debe considerar los avances tecnológicos y debe ser capaz de adaptarse a ellos sin necesidad de hacer muchos cambios. Es necesario planear pensando en la flexibilidad como uno de los aspectos más importantes.

Documentar el diseño. Este es uno de los puntos más importantes y que a veces se llega a pasar por alto. Este permite que se pueda revisar y reevaluar el diseño, hacer cambios y comunicarlos a los grupos involucrados. Otro beneficio de documentar el diseño es que se puede utilizar como plan de implementación. Dentro de este punto por lo menos se deberían documentar los siguientes aspectos:

- **Mapa de la topología.** En donde se muestre el esquema lógico de la ISN.
- **Diseño de la configuración.** En donde se especifique el diseño físico de toda la implementación. Este debe tener información más detallada que la encontrada en el mapa de la topología.

- *Mapa de almacenamiento.* En donde se defina el arreglo, la configuración de los sistemas de discos y se incluyan detalles como los niveles de RAID que se vayan a implementar. Este mapa define como es configurado el almacenamiento en la ISN.
- *Mapa de zonificación.* Define el acceso de los servidores a los sistemas de discos, así como cuáles son los nodos o dispositivos que tienen permiso para comunicarse entre sí dentro de la ISN.

6.6.2 Configurando una ISN

Registrar toda la información pertinente a las conexiones en el diagrama de configuración. Registrar los WWN de todos los nodos y dispositivos para identificar en donde se encuentran físicamente. Se recomienda también el usar etiquetas en los diferentes dispositivos con información tal como direcciones IP y WWN's.

Es importante también etiquetar los diferentes cables para saber que dispositivos están siendo conectados así como para su identificación posterior.

Es una buena práctica el usar simetría en las conexiones, esto significa el tratar de ser consistentes en la conexión de servidores, en la colocación de tarjetas adaptadoras y en la conexión a los concentradores. Por ejemplo, si se están configurando dos fábricas con múltiples rutas, conectar la tarjeta adaptadora 1 a la fábrica 1 y la tarjeta adaptadora 2 a la fábrica 2. Esta forma de configuración permite detectar fallas en la conexión o en los componentes de una manera más sencilla.

Verificar que todos los concentradores en una fábrica tengan los mismos parámetros de configuración y que cada uno tenga un identificador único.

Asegurarse de que los diferentes servidores tienen las versiones apropiadas de sistema operativo, firmware y parches.

Usar el mapa de almacenamiento para configurar cada uno de los sistemas de discos, con los correspondientes parámetros y niveles de RAID.

Al configurar zonas, considerar primero las más pequeñas para que puedan ser probadas y luego continuar con las más complejas. Es muy importante guardar las configuraciones de zonificación antes de crear otras, esto por seguridad.

Se puede hacer zonificación por sistema operativo o por arreglo de discos. Por ejemplo, se pueden configurar los servidores NT1, NT2 y NT3 para acceder el sistema de discos ARRAY1 y configurar los servidores VMS1, VMS2 y VMS3 para acceder el ARRAY2 tal como se muestra en la tabla 6-6.

ZONA_NT	ZONA_VMS
NT1	VMS1
NT2	VMS2
NT3	VMS3
ARRAY1	ARRAY2

Tabla 6-6 Zonificación por sistema operativo

También es posible configurar una zona en donde varios servidores accesen al mismo arreglo de discos, si este lo permite, tal como se muestra en la tabla 6-7.

ZONA_ARRAY1	ZONA_ARRAY2
NT1	NT3
NT2	NT4
VMS2	VMS1
VMS3	NT5
SUN1	SUN2

Tabla 6-7. Zonificación por arreglo de discos

Al establecer las zonas, es importante usar nombres que tengan sentido y ser consistente con la convención de nombres. Por ejemplo se puede usar el alias NT1_HBA1 para el servidor NT1 conectado a la tarjeta adaptadora1, en general se puede utilizar cualquier convención siempre y cuando tenga significado para el administrador.

6.6.3 Escalando una ISN

Estas son las consideraciones que se deben tener en cuenta al escalar una red de almacenamiento inteligente.

- Reemplazar concentradores de 8 puertos por concentradores de 16 puertos.
- Agregar concentradores hasta los límites establecidos por las diferentes topologías.
- Agregar una segunda fábrica para mayor disponibilidad.
- Desarrollar múltiples ISN's.

Cada vez que se vaya a expandir una infraestructura de redes de almacenamiento se debe estar seguro de que los nuevos concentradores y/o dispositivos que se están agregando son consistentes con la topología original de la ISN en términos de diseño. En lo posible se debe evitar el hacer cambios en la topología, cada vez que se vaya a hacer un cambio en la ISN se debe respaldar toda la información y de ser posible suspender la actividad de I/O, esto con el propósito de asegurar la integridad de los datos. A continuación se describen los procedimientos que se deben seguir para expandir una fábrica en topología de cascada, malla, anillo y árbol.

Fábrica en cascada

Para expandir una fábrica en cascada se debe conectar el nuevo concentrador a un puerto disponible de un concentrador existente. Si no hay puertos disponibles, entonces se deben desconectar los dispositivos del concentrador que ya existe para poder conectar el nuevo concentrador y después conectar en este último los dispositivos que se hayan desconectado.

Fábrica en malla

En este caso se debe seguir el mismo procedimiento que para la fábrica en cascada pero aquí se debe asegurar que existen múltiples rutas que conecten el nuevo concentrador a la fábrica en malla.

Fábrica en anillo

Para expandir una fábrica en esta topología se debe romper el anillo para insertar el nuevo concentrador. Cuando se está expandiendo fuera del anillo, se debe asegurar que no hayan más de siete niveles entre dos dispositivos que necesiten comunicarse.

Fábrica en árbol

Este esquema es muy simple basta con conectar los nuevos concentradores a los concentradores de interconexión, dependiendo del diseño y del número de puertos disponibles se podrá requerir algún concentrador adicional.

6.7 Administración de ISN's

La administración de una ISN tiene un objetivo central: asegurarse de que los datos sean transportados sin ningún problema a través de la red. La administración de los recursos de datos, incluyendo los componentes que forman la ISN, el acceso a los archivos, los administradores de volúmenes y las tareas de respaldo son esenciales en un ambiente de redes de almacenamiento inteligente.

Mientras que los servidores y los arreglos de discos estuvieron aislados de la red, las aplicaciones de administración de disco se enfocaban en el servidor - el dueño del almacenamiento. Las ISN's cambian también el esquema de administración de los recursos de almacenamiento haciéndola independiente de la plataforma, sistema operativo o arreglo de discos.

Como se muestra en la figura 6-10 la administración de una ISN está conformada por una jerarquía de funciones que pueden existir como aplicaciones separadas o como un sistema de administración integrado. Las capas inferiores mueven la información de eventos y estado de los elementos de la red hacia las capas superiores, mientras que las capas de arriba envían comandos y peticiones solicitando información a los agentes de las capas de abajo. La estructura de administración de una ISN tiene su base en los dispositivos de interconexión tales como las tarjetas adaptadoras, concentradores y puentes. Estos dispositivos se comunican a sus respectivas aplicaciones de administración usando una serie de protocolos dentro de los que destaca el protocolo de administración de redes simples (SNMP, *Simple Network Management Protocol*).

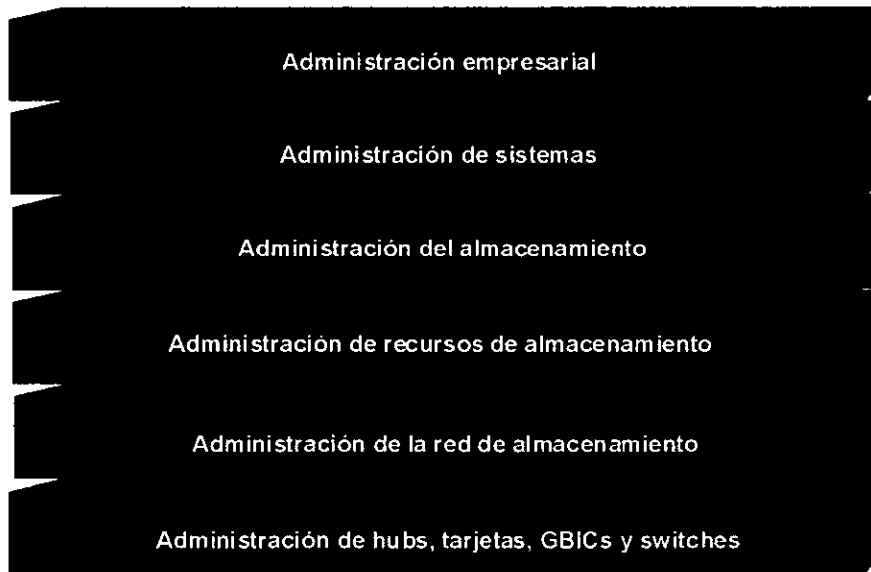


Figura 6-10. Jerarquía de administración en una ISN

Administración de Redes de Almacenamiento

Después de la capa de dispositivos viene la capa de administración de redes de almacenamiento, la cual es responsable de interactuar directamente con el equipo de interconexión. Todos los proveedores de dispositivos de ISN proporcionan una consola o interface gráfica que por lo menos permite la configuración de puertos y el manejo de mensajes para reportar el estado de los componentes, dicha interfase se conoce como administrador de dispositivos, sin embargo, en la mayoría de los casos ésta es única al componente, por ejemplo, una herramienta de administración para una tarjeta adaptadora dará información exclusivamente para el HBA pero no permitirá ver información de cualquier otro dispositivo dentro de la red de almacenamiento.

Para evitar el tener que usar múltiples administradores de dispositivos en una red de almacenamiento inteligente, algunos proveedores utilizan una interfase común que integra a los diferentes agentes de administración, mientras que otros proporcionan interfases para que administradores de dispositivos de terceros interactúen con sus productos. La idea es tener un esquema de administración centralizado.

Un administrador de dispositivos tiene componentes de *hardware*, *software* y *firmware* tal como se muestra en la figura 6-11.



Figura 6-11. Arquitectura de administración de dispositivos

La comunicación entre un dispositivo y la consola de administración se puede llevar a cabo a través del medio de transporte de fibra, a dicha administración se le conoce como en banda (*in-band*). Alternativamente, la administración puede ser realizada por otro medio distinto al de canal de fibra utilizando, por ejemplo, una interface Ethernet o RS-232, dicha administración es conocida como fuera de banda (*out-band*). La administración fuera de banda a través de Ethernet utiliza generalmente SNMP, aunque también es posible utilizar HTTP en las implementaciones en donde la administración es basada en navegador.

Administración en banda

Aunque es posible correr SNMP sobre canal de fibra, el protocolo para administración en banda más usado es el (SES, *SCSI Enclosure Service*). Para que un dispositivo pueda soportar comandos SES debe tener una dirección de canal de fibra. La administración en banda tiene funciones que son difíciles de implementar en un esquema fuera de banda. Por ejemplo, si un concentrador puede enviar peticiones SES a las tarjetas adaptadoras y discos, este puede proporcionar un mapa de la topología de la ISN detallando la identidad, estado y localización de los diferentes dispositivos que conforman la ISN. La administración en banda también simplifica la infraestructura de red ya que esta se realiza usando la misma red de almacenamiento.

La desventaja de este esquema es cuando se presenta algún problema en el transporte de canal de fibra ya que automáticamente se interrumpe la administración. Este tipo de administración se recomienda solo en ambientes de redes de almacenamiento pequeñas.

Administración fuera de banda

Este tipo de administración utiliza típicamente Ethernet. Para la comunicación entre los dispositivos se usa SNMP aunque no es rara la administración a través de telnet y Web.

Ya que esta implementación no depende del medio de fibra, la administración de los diferentes componentes se puede seguir llevando a cabo incluso en el evento de una falla en el medio, además de que se puede tener integración con cualquier otro dispositivo que no forme parte de la red de almacenamiento a través de SNMP. La principal desventaja de este método es que no es posible tener información tan detallada de la topología de red como la que se puede lograr a través de la administración en banda.

Los protocolos utilizados en la administración fuera de banda corren sobre IP, lo que permite que las peticiones y respuestas de los diferentes dispositivos puedan ser manejados desde cualquier lugar en la red.

SNMP

SNMP es el protocolo predominante en las redes LAN / WAN y es ampliamente soportado por ruteadores y concentradores. SNMP proporciona un conjunto de comandos solicitando el estado de los diferentes dispositivos. La plataforma de administración SNMP corre típicamente como una interfase gráfica en ambientes UNIX o NT y es capaz de administrar cientos de dispositivos en una red. La plataforma de administración contiene un manejador de SNMP que se comunica con los agentes que deben ser instalados en cada uno de los dispositivos a monitorear.

La información que se puede recabar con este esquema incluye: número serial, identificador del proveedor, estado, tipo de puerto, estado operacional del puerto, volúmenes de tráfico y condiciones de error. Esta información es organizada en una Base de Información de Administración (MIB, *Management Information Base*), que es mantenida por una estación de trabajo de administración y los diferentes agentes SNMP en cada uno de los dispositivos.

HTTP

La administración a través del protocolo HTTP permite que esta se puede llevar a cabo mediante un navegador. Esta implementación tiene dos componentes: Un servidor de HTTP que reside en el dispositivo administrado y un navegador de Web que actúa como una interfase gráfica de administración. Cada vez que se quiere verificar el estado de un componente o cambiar algunos de sus parámetros basta con especificar su dirección IP y navegar a través de un par de pantallas.

Debido a que los dispositivos solo pueden ser accedidos a través de sus direcciones IP la administración se puede volver complicada a medida que aumenta el número de componentes, es por eso que la administración basada en HTTP se recomienda para ISN's pequeñas.

Otra de las desventajas de este método de administración es el alto riesgo de seguridad que se tiene ya que a través de la dirección IP es posible que cualquiera pueda acceder a la red de almacenamiento aún y cuando el acceso esté restringido. Por lo que se recomienda que la solución sea complementada con un *firewall* en grandes implementaciones.

La administración basada en navegador también puede ser implementada en un esquema cliente servidor tal como se muestra en la figura 6-12, esto permite que se pueda monitorear y administrar la red de almacenamiento incluso desde puntos remotos.

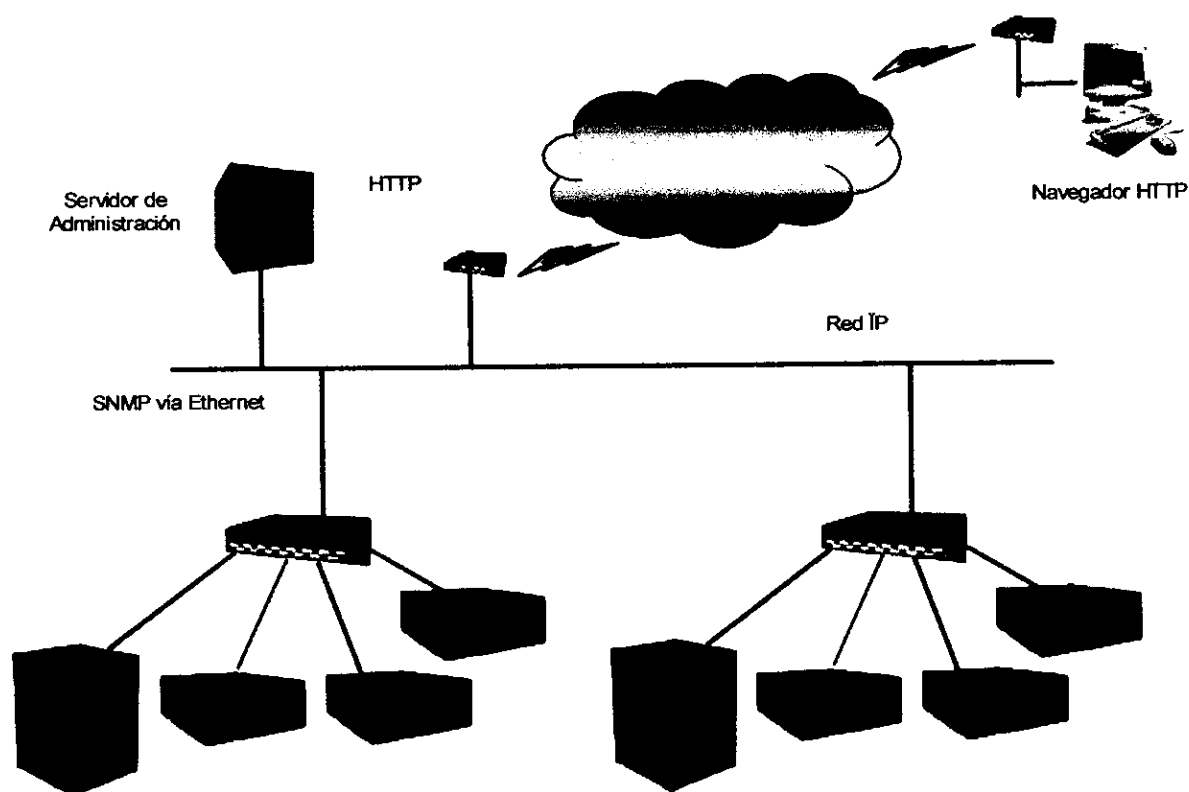


Figura 6-12. Administración cliente/servidor mediante HTTP.

TELNET

A diferencia de las soluciones basadas en SNMP o HTTP. Una implementación de TELNET solo proporciona una interfase basada en texto. Cuando se quiere ver el estado de un dispositivo es necesario hacer un telnet a su dirección IP y establecer una sesión. Esta forma de administración es muy popular en ambientes UNIX pero su uso puede ser ineficiente en esquemas grandes de ISN.

Administración de recursos de almacenamiento

Las aplicaciones de administración de recursos de almacenamiento (SRM, Storage Resource Management) están disponibles como programas independientes o módulos *plug-in*. Una de las principales ventajas de estas aplicaciones es la habilidad que tienen de hacer visibles todos los discos (independientemente de la forma como estén distribuidos) desde una sola consola. Estas aplicaciones no son únicas para ambientes de ISN, sin embargo, representan una herramienta valiosa en este tipo de esquemas.

Sin un mecanismo de administración de recursos de almacenamiento, la administración estaría limitada a servidores individuales, lo cual sería una verdadera pesadilla en ambientes grandes de cómputo y lo que provocaría la duplicidad de aplicaciones en múltiples servidores y la mala utilización del espacio en disco.

La forma en la que los recursos de almacenamiento son administrados es a través de un cliente de SRM que corre en cada servidor y que periódicamente actualiza la información de sus volúmenes asignados enviando dicha información al administrador de SRM. Este último, consolida la información que recibe de los diferentes clientes en una base de datos relacional y puede a través del uso de políticas de almacenamiento enviar alertas si algún servidor se está quedando sin disco.

Administración del almacenamiento

Esta involucra la administración de sistemas de discos, bibliotecas de cintas, políticas de almacenamiento y respaldo, administración de volúmenes así como políticas de archivos. Como en el caso de las aplicaciones de SRM, la administración del almacenamiento es facilitada por las ISN's.

Las aplicaciones de administración de almacenamiento pueden monitorear el acceso a recursos compartidos tales como sistemas de discos y bibliotecas de cintas, a través de ellas también es posible ejecutar rutinas de defragmentación de disco, controlar la asignación de recursos y la integridad de los sistemas de archivos.

Conforme más elementos se van agregando a la ISN estas herramientas de administración se vuelve aún más útiles al acceder los recursos directamente y de una manera centralizada sin tener que pasar por los servidores.

VII

Productos y Soluciones Disponibles

El ámbito de las redes de almacenamiento inteligente está conformado por una gran variedad de productos de *hardware* y *software* que se complementan o compiten entre sí. Algunos fabricantes tienen solo una porción de una solución de ISN, mientras que otros son capaces de ofrecer soluciones completas integrando productos propios y de otras compañías ya que ningún proveedor fabrica todos los elementos que se requieren para construir una ISN.

Para lograr la interoperabilidad entre productos de diferentes marcas, algunos proveedores que no compiten entre sí han establecido alianzas, mientras que otros han optado por invertir en laboratorios para verificar la compatibilidad de sus productos con los de otros fabricantes.

El espectro de productos de *hardware* abarca desde los transceptores, tarjetas adaptadoras, concentradores de conexiones, arreglos de discos hasta dispositivos más sofisticados que permiten múltiples conexiones a grandes distancias utilizando diferentes protocolos. Por otro lado, los componentes de *software* van desde los manejadores de dispositivos que vienen con las tarjetas adaptadoras, hasta *software* de administración.

7.1 Productos de hardware

Existen algunos componentes de *hardware* que son indispensables en una red de almacenamiento inteligente, mientras que otros son productos más especializados que pueden ser requeridos dependiendo de las características particulares de la solución. La tabla 7-1 muestra los elementos esenciales y especializados encontrados en una ISN. Más adelante se explican con detalle cada uno de ellos.

Redes de Almacenamiento Inteligente

Productos esenciales	Productos especializados
Transceptores <i>GBICS, GLMs, SFFs</i>	Puentes
Tarjetas adaptadoras	Extensores (DWDMs)
Concentradores simples	Compuertas
Concentradores inteligentes	
Arreglos de discos <i>JBODs, RAIDs</i>	

Figura 7-1. Componentes de hardware encontrados en una ISN

7.1.1 Transceptores

En la tecnología de canal de fibra, las tramas (*frames*) son movidas del dispositivo fuente al de destino usando velocidades de *Gigabit*. Para comunicarse en un transporte de altas velocidades, ambos lados deben tener soporte a este tipo de comunicación. Esto se logra utilizando dispositivos que manejen estas características de medio y velocidad o usando interfases especialmente diseñadas llamadas transceptores para convertir un transporte de comunicación determinado al de fibra. Los transceptores más populares en el mercado son:

- Convertidor de Interface a Gigabit (GBIC, *Gigabit Interface Converter*)
- Módulo de Conexión de Gigabit (GLM, *Gigabit Link Module*)
- Transceptor de Pequeño Factor de Forma (SFF, *Small Factor Form*)
- Adaptador de Interface al Medio (MIA, *Media Interface Adapter*)

Convertidor de Interface a Gigabit

Los GBIC's son transceptores que convierten los datos entrantes de un lado de un dispositivo a señales ópticas del otro lado. Los GBIC's pueden ser usados para aplicaciones que manejen fibra óptica y *Gigabit Ethernet*. Estos dispositivos reciben una señal de datos serial como entrada y el GBIC lo convierte en una señal que pueda ser manejada en un medio óptico o de cobre. El medio de fibra es preferido para alcanzar mayores distancias así como por su inmunidad al ruido.

Cuando el GBIC recibe una señal óptica, ésta es convertida a una señal eléctrica que después es manejada por el servidor como una señal de datos serial. Hay dos tipos de conectores del lado del servidor o arreglo de discos: DB-9 y HSSDC.

Existen dos modos de operación en el medio óptico. La diferencia de ellos es la longitud de onda del láser y el modo en como se transmiten los pulsos ópticos. Estos modos son:

- Modo de onda corta SW
- Modo de onda larga LW

Las distancias de conexión que se pueden tener con cada uno de estos modos se muestran en la tabla 7-2.

Tipo de fibra	SW	LW
fibra óptica de 9 micrones	n/a	10 km
fibra óptica de 50 micrones	2 - 500 m	2 - 550 m
fibra óptica de 62.5 micrones	2 - 175 m	2 - 550 m

Tabla 7-2. Distancias de conexión usando fibra óptica

Los GBIC's son dispositivos de reemplazo en caliente (*hot swap*) que pueden ser configurados y reemplazados sin tener que dar de baja el equipo o sistema de discos, pero son usados también en concentradores de conexiones constituyendo la interfase de conexión más popular de fibra óptica en nuestros días. Las velocidades de transferencia son de 1063 MB/s. La figura 7-1 muestra un GBIC.



Figura 7-1. Convertidor de Interfase a Gigabit (GBIC)

La selección de un GBIC en una ISN es un aspecto muy importante en donde se deben tomar en cuenta los aspectos de distancia, simplicidad en el cableado, desempeño e inmunidad al ruido que se requieren en la red de almacenamiento. La tabla 7-3 muestra las características de los GBICs.

Convertidor de Interface a Gigabit

Características	Aplicaciones
- Láser clase 1	- Canal de fibra gigabit
- Velocidades de 1063 Mb/s a 1250 Mb/s	- Gigabit ethernet
- Compatible con el estándar de canal de fibra ANSI	- Ambientes cliente/servidor
- Compatible con Gigabit Ethernet (IEEE 802.3z/d5.0)	- Multiprocesamiento distribuido
- Opciones de onda corta (distancia <= 550 m) y onda larga (distancia <= 10 km)	- Aplicaciones tolerante a fallas
- Interfase serial eléctrica de Gigabit	- Visualización y video
- Conversión de señales eléctricas a luz	- Extensores de canal, almacenamiento de datos

Tabla 7-3. Características y aplicaciones de los GBIC's

Módulos de Conexión de Gigabit

Aunque todavía es posible encontrarlos en el mercado los GLMs fueron los primeros transceptores usados en aplicaciones de canal de fibra. Estos son mucho más baratos que los GBIC's pero no tienen las características de reemplazo en caliente encontradas en los GBICs, por lo que es necesario dar de baja el equipo para instalar, mantener o reparar un GLM.

Los módulos de Conexión de Gigabit utilizan también el medio de fibra para la transmisión, manejando dos tipos de láser: Onda corta y onda larga, por lo que es posible alcanzar las mismas distancias que usando GBICs. Las velocidades de transferencia disponibles son de 266 MB/s y 1063 MB/s en modo *full duplex*. En una transmisión, un GLM convierte los datos codificados serializándolos en pulsos de luz láser para transmisión en fibra óptica tal como lo hacen los GBIC's. La figura 7-2 ilustra un GLM.

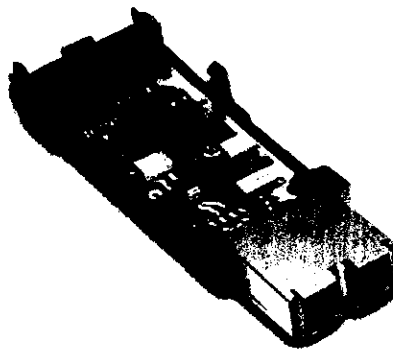


Figura 7-2. Módulo de Conexión de Gigabit

Transceptor de pequeño factor de forma

Los SFF tienen un tamaño mucho más pequeño y disipan menos energía que los GBICs o los GLMs, por lo que estos transceptores son ideales para aplicaciones que requieren de alto desempeño y poco consumo de energía. Los transceptores de pequeño factor de forma tienen las mismas funciones que los GBICs permitiendo transmitir o recibir datos a velocidades de 1 ó 2 Gbps, proporcionando una forma de poder actualizar la infraestructura de una red de almacenamiento inteligente de una manera transparente. La figura 7-3 muestra un SFF.

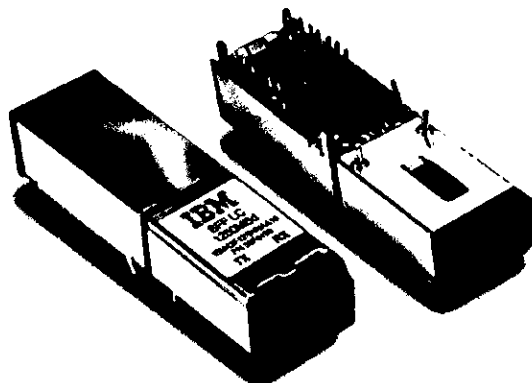


Figura 7-3. Transceptor de pequeño factor de forma

Adaptadores de Interface al Medio

Los Adaptadores de Interface del medio (MIA, *Media Interface Adapters*) son utilizados para facilitar la conversión entre conexiones con interfaces ópticas y de cobre. Típicamente los MIA's son conectados a las interfases de las tarjetas adaptadoras, pero también pueden ser usados con arreglos de discos y concentradores de conexiones.

Algunos fabricantes de sistemas de discos, por ejemplo, prefieren utilizar interfases de cobre porque les permite que la circuitería sea menos complicada además de que se reducen costos, sin embargo, cuando se trata de conectar los arreglos a los servidores se prefiere utilizar el medio óptico ya que la distancia de conexión es mucho mayor además de que no se tienen problemas de ruido. Para este tipo de situaciones los MIA's son el recurso utilizado ya que permiten una conversión de interfase del medio de cobre al óptico de manera transparente. La figura 7-4 muestra la imagen de un MIA.

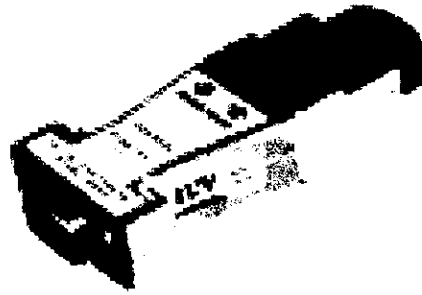


Figura 7-4. Adaptador de interfase al medio

7.1.2 Tarjetas adaptadoras

Las tarjetas adaptadoras (HBAs, *host bus adapters*) son utilizadas para permitir la conexión de servidores, arreglos de discos y bibliotecas de cintas a la red de almacenamiento. Las HBA's están disponibles para diferentes tipos de canales y conexiones físicas, siendo las más comunes aquellas con interface de canal PCI y transceptores de onda corta. Cada tarjeta viene con manejadores de *software* para el sistema operativo en donde será instalada. Las primeras tarjetas de fibra óptica que aparecieron en el mercado solo soportaban las topologías de conexión punto a punto y de ciclo arbitreado, sin embargo, en la actualidad la mayoría de las tarjetas soportan las tres topologías de conexión de fibra. La tabla 7-4 muestra las características de tres diferentes HBAs de canal de fibra.

HBA FC - SBus

- Interfase SBus de 64 bits
- Velocidad de transmisión de 1 ó 2 Gb/s con autonegociación
- Soporte de todas las topologías de FC
- Soporte de protocolos FCP, IP y FC-tape
- Soporte de clases de servicio 2 y 3
- Transceptor de onda corta
 - 50/125 micrones (hasta 300 m)
 - 62.5/125 micrones (hasta 90 m)
- Detección automática de topología
- Protección completa de paridad en todas las rutas de datos paralelas
- LEDs externos de estado
- Ambiente operativo Solaris

HBA FC - PCI 64 bits

- Interfase PCI de 64 bits @ 33 Mhz
- Velocidad de transmisión de 1 Gb/s
- Soporte de todas las topologías de FC
- Conector PCI universal
- Soporte de clase de servicio 3
- Transceptor de onda corta
 - 59/125 micrones (hasta 500 m)
 - 62.5/125 micrones (hasta 300 m)
- Muy baja latencia
- Protección completa de paridad en todas las rutas de datos paralelas
- LEDs externos de estado
- Ambientes operativos Solaris, Windows NT, 2000, AIX, HP-UX, Mac OS

HBA FC - PCI Dual Port

- Dos interfaces PCI de 64 bits totalmente independientes
- Velocidad de transmisión: 1 GB/s
- Soporte de todas las topologías de FC
- Utilización eficiente del bus PCI
- Soporte de clase de servicio 3
- Transceptor de onda corta
 - 50/125 micrones (hasta 500 m)
 - 62.5/125 micrones (hasta 300 m)
- Muy baja latencia
- Protección completa de paridad en todas las rutas de datos paralelas
- LEDs externos de estado
- Ambientes operativos Solaris, Windows NT, 2000, AIX, HP-UX, Redhat Linux, Mac OS, Novell Netware

Tabla 7-4. Características de tres tarjetas adaptadoras de canal de fibra

Aunque la mayoría de las HBA's tienen un solo transceptor, no es difícil encontrar tarjetas con dobles e incluso con cuatro conexiones de fibra, en un ambiente de fábrica cada uno de estos puertos aparece como un nodo de canal de fibra, cada uno con su nombre de puerto WWN y su dirección de 24 bits. Las tarjetas con múltiples puertos constituyen un ahorro en términos de ranuras de expansión pero también representan un solo punto de falla. La figura 7-5 ilustra algunas de las HBA's disponibles en el mercado.

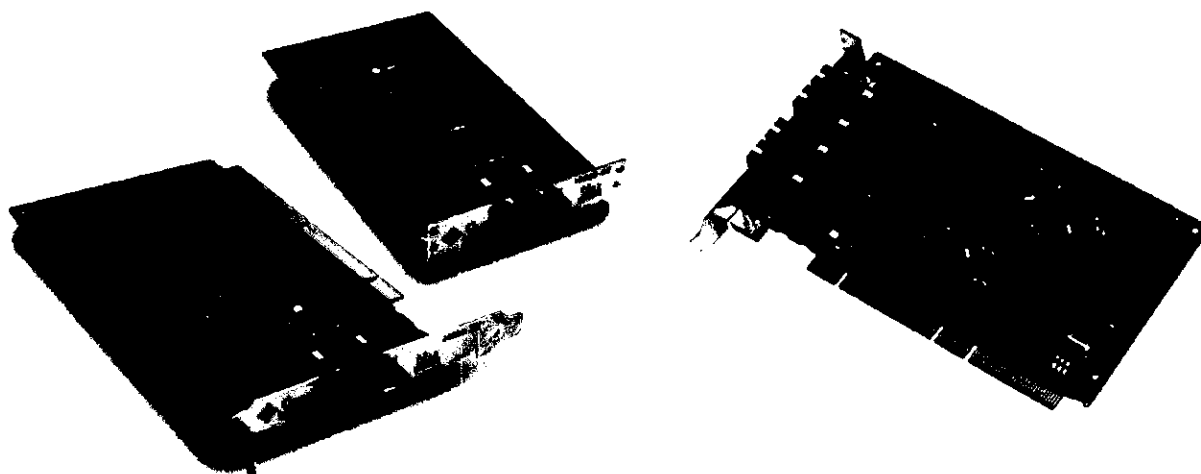


Figura 7-5 Tarjetas adaptadoras de canal de fibra

Las tarjetas adaptadoras tienen que ver con las cuatro capas del canal de fibra. En el nivel 0, una HBA tiene funciones de transmisión y recepción para la conexión física, usando como interfaces GBICs, GLMs o SFFs con conexiones SC mientras que para el medio de cobre se utilizan conectores DB-9. En la capa FC-1 se incorpora la lógica de codificación 8b/10b para datos salientes y una lógica de decodificación para datos entrantes así como funciones de administración de conexión de bajo nivel, de control de flujo y de monitoreo. A continuación la capa FC-2 proporciona el protocolo de señalización para la segmentación y ensamblaje de tramas, la clase de servicio así como servicios para fábrica y registro de puertos. En la capa FC-4 la mayor parte de las HBA's proporcionan el mapeo de los protocolos SCSI-3 e IP para los diferentes sistemas operativos.

La figura 7-6 muestra un diagrama funcional de un tarjeta adaptadora y su relación con las capas del canal de fibra.

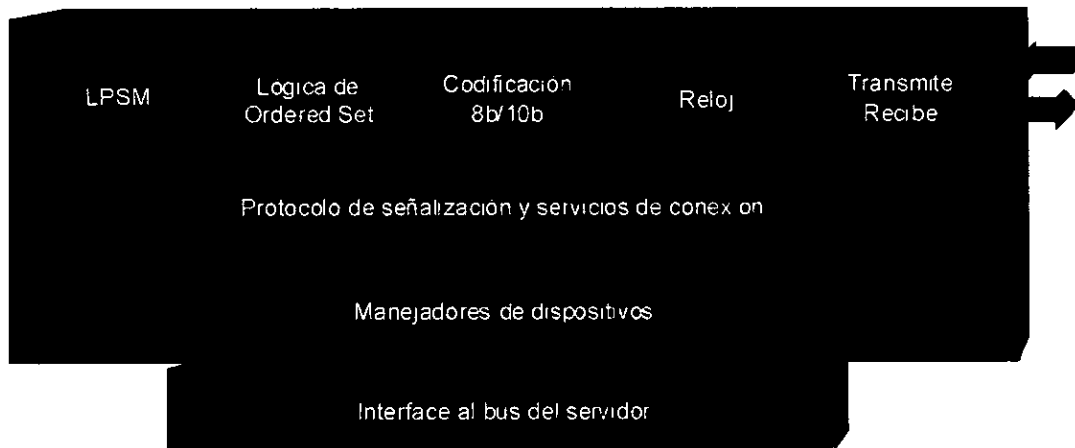


Figura 7-6. Diagrama funcional de una tarjeta adaptadora de canal de fibra

7.1.3 Concentradores de conexiones simples

Los concentradores simples de canal de fibra son muy populares en implementaciones de ISN, debido fundamentalmente a su facilidad de instalación y al bajo costo por puerto. Los proveedores de productos de canal de fibra ofrecen diferentes modelos de concentradores que pueden variar dependiendo de la configuración de los puertos, del tipo de interfases y del nivel de administración. La figura 7-7 ilustra dos modelos de concentradores.

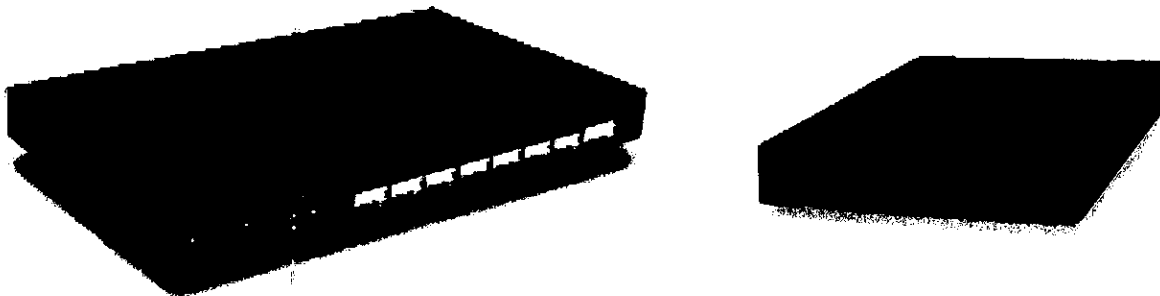


Figura 7-7. Concentrador de conexiones simple

Arquitectura de un concentrador simple

El diseño de un concentrador varía de un proveedor a otro, pero todos incorporan ciertas características básicas. Un concentrador intrínsecamente tiene una topología de ciclo la cual se logra interconectando cada uno de los puertos del concentrador y uniendo el último puerto con el primero. Estos dispositivos utilizan una circuitería que permite que el ciclo se autoconfigure en el evento de que un dispositivo sea agregado o quitado del concentrador, tal como se muestra en la figura 7-8.

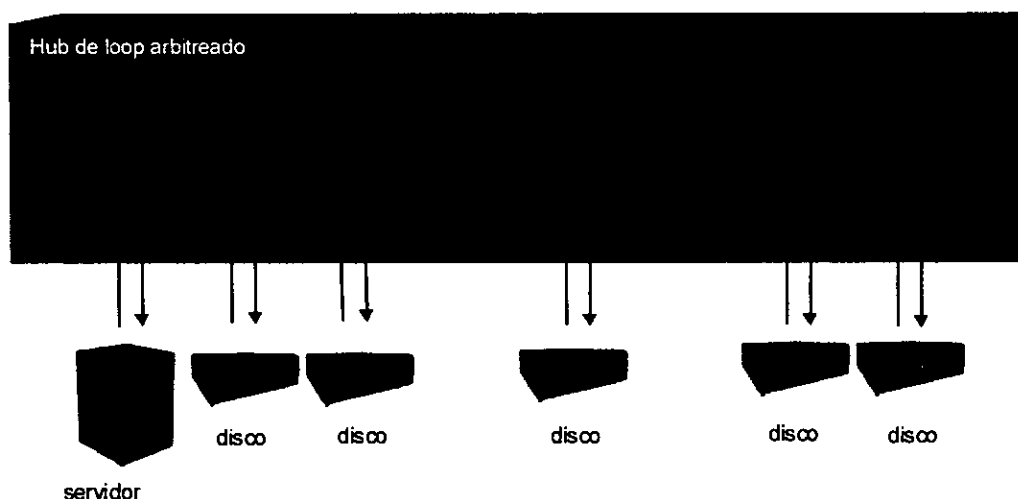


Figura 7-8. Arquitectura interna de un concentrador

Conforme los servidores o arreglos son quitados de línea o habilitados la circunferencia de la circuitería del ciclo se contrae o expande. A nivel de conexión, la inicialización del ciclo maneja la reconfiguración automática de las direcciones de puerto para dar cabida a los nuevos dispositivos conectados. La combinación de la circuitería y la inicialización del ciclo facilitan cualquier cambio en la topología con mínima interrupción del tráfico de datos. Generalmente los concentradores incluyen uno o más LED's para desplegar el estado del puerto. En la mayoría de los casos se tienen dos LED's: uno verde que indica el estado de conexión y otro ambar que indica el modo de operación. Dependiendo de la especificación del proveedor, la combinación de los estados de LED verde y ámbar puede ser usado para desplegar diferentes condiciones de puerto tal como se muestra en la tabla 7-5.

Verde	Ambar	Estado del puerto
Apagado	Apagado	No hay dispositivo conectado
Encendido	Apagado	Dispositivo conectado y ciclo insertado
Encendido	Encendido	Dispositivo conectado y modo de paso
Apagado	Encendido	GBIC con falla, puerto en modo de paso
Parpadeando	Parpadeando	Modo mantenimiento

Tabla 7-5. Diferentes estados de un puerto en un concentrador

La arquitectura de un concentrador simple está regida por los estándares definidos para FC-AL, FC-AL-2 y la capa física FC-0. Mientras no se violen estos estándares, los fabricantes de concentradores son libres de hacer variaciones en la densidad de puertos, en el tipo de puertos, en el procesamiento de señales y en las características de administración.

La densidad de puertos en una unidad puede variar de 5 a 32. La densidad más común es de entre 6 y 12 puertos por unidad. Este rango es adecuado para la mayoría de las implementaciones de ciclo y se pueden tener muchos más puertos cascadeando los concentradores. Estos ocupan una unidad de gabinete (1U), lo cual permite que se pueda tener un gran número de puertos si se tienen varios concentradores configurados en un gabinete estándar de 19".

El tipo de puerto también puede variar dependiendo del fabricante, es posible encontrar concentradores con interfases fijas de cobre, los cuáles son bastante accesibles económicamente hablando, sin embargo, estos han caído en desuso debido a los problemas de distancia y de emisiones de ruido inherentes al medio de cobre. En la actualidad, la mayoría de los concentradores simples ofrecen puertos basados en GBICs lo que permite que incluso se puedan tener mezclas de medios de cobre y fibra en onda corta y larga. Este diseño permite que se puedan ir agregando puertos conforme se requiera y que sea posible cambiar de un tipo de medio a otro sin tener que cambiar el concentrador. Las tres implementaciones de concentrador más comunes en el mercado son:

- Concentradores sin administración
- Concentradores con administración
- Concentradores con cierta inteligencia

Concentradores sin administración

Este tipo de concentradores es usado para ambientes pequeños y homogéneos, como es el caso de esquemas en donde un solo fabricante provee el servidor, las tarjetas adaptadoras, concentradores, arreglos, GBICs y cableado y por tanto no se requiere el nivel de administración necesario en un esquema de ISN complejo y heterogéneo.

Los concentradores sin administración son una solución de interconexión simple y de bajo costo. Típicamente este tipo de dispositivos ofrecen circuitería de paso basado en una sola señal así como LED's para indicar el estado de los puertos. Si un dispositivo es desconectado o apagado, un concentrador sin administración hará que la señal se vaya al siguiente puerto activo. Sin embargo, debido a la falta de inteligencia cualquier situación que signifique una violación de protocolo hará que el ciclo se vaya abajo. Por lo tanto el uso de este tipo de puertos puede constituir un riesgo.

Los concentradores sin administración son una opción lógica para soluciones de red de almacenamiento de muy bajo costo en donde el aspecto central de diseño está basado en lo económico.

Concentradores con administración

Estos son concentradores que tienen cierto grado de inteligencia, el cual puede variar de un producto a otro junto con capacidades mejoradas de monitoreo y administración, debido a su complejidad en el diseño estos concentradores son más caros que los concentradores sin administración.

Los concentradores con administración más sencillos ofrecen capacidades de monitoreo y control de puertos desde un navegador, TELNET o a través de SNMP, mientras que los concentradores más sofisticados ofrecen diagnósticos avanzados y características de administración proactivas en una amplia variedad de plataformas de administración incluyendo *HP Open View* y aplicaciones basadas en *Java*.

La administración en concentradores de ciclo arbitreado tiene dos componentes básicos: A nivel de *hardware* se usa una circuitería adicional que es necesaria para monitorear el estado de la fuente de poder, ventilador, temperatura y puertos del concentrador, existen algunos concentradores con capacidades avanzadas que proporcionan diagnósticos más extensivos para monitorear el estado del ciclo, la actividad del protocolo y dar información más detallada acerca del estado del concentrador y de sus puertos. Por otro lado, para que esta circuitería sea útil, el concentrador debe ser capaz de reportar y aceptar comandos de una estación de administración externa. Esto se realiza típicamente a través de un puerto *Ethernet* en el concentrador, usando comandos y peticiones de SNMP que son enviados a y desde una consola NT o UNIX. El *software* de aplicación usado para manejar un concentrador es provisto por el proveedor ya sea como una aplicación separada o como una interfase integrada a una aplicación de administración de ISNs.

Los concentradores con administración que soportan SNMP o navegadores pueden ser manejados desde cualquier lugar en una red IP. Esto constituye una ventaja sobre todo para aquellas organizaciones en las que la infraestructura de las redes de almacenamiento inteligente va más allá de una sola localidad física. En general se prefiere usar SNMP ya que múltiples concentradores pueden ser administrados desde una sola estación de trabajo a diferencia de la administración basada en navegador que típicamente permite ver un solo concentrador a la vez.

Concentradores con cierta inteligencia

Los concentradores con cierta inteligencia son una solución híbrida de ISN que es intermedia entre los concentradores simples (*hubs*) y los concentradores inteligentes (*switches*). De manera similar a la funcionalidad de ciclo privado o modo fantasma de algunas fábricas, los concentradores con cierta inteligencia proporcionan la simplicidad del ciclo privado con el mayor ancho de banda ofrecido por los concentradores inteligentes. Los concentradores con cierta inteligencia son más caros que los concentradores estándar pero son razonablemente más accesibles que los concentradores inteligentes (*switches*) Para configuraciones que requieren de un gran ancho de banda pero en donde no se requieren más de 126 dispositivos, esta alternativa constituyen una buena solución en términos de precio y desempeño.

Estos dispositivos tienen generalmente de seis a ocho puertos, cada uno soportando 100 MBps. Los nodos conectados al ciclo son configurados en un ciclo virtual conformado por múltiples segmentos de ciclo de 100 MBps esto permite que el desempeño no se vea degradado a medida que se agregan más dispositivos al ciclo.

La segmentación de un ciclo lógico en varios segmentos también permite que el comportamiento del ciclo pueda ser modificado. La inicialización de un ciclo, puede ser por ejemplo, restringida a un solo puerto, lo cual elimina problemas potenciales para aplicaciones sensibles, como las de manejo de video y respaldos. Además, se pueden tener conversaciones concurrentes entre varios dispositivos en un mismo ciclo virtual al estar conformado por varios ciclos físicos. Un servidor en el puerto 3, por ejemplo, puede estar accediendo información en el puerto 7 mientras un servidor en el puerto 1 podría acceder datos en el puerto 6.

Este tipo de concentradores incluye también capacidades de administración vía SNMP o SES y dependiendo del diseño de cada fabricante algunos productos incluyen características de diagnóstico avanzadas como reruteo de tráfico.

La tabla 7-6 muestra las características más comunes encontradas en los concentradores.

Hub sin administración

- Soporte de configuraciones en alta disponibilidad
- Velocidad de transferencia de 100 MBps en total para todos los puertos
- Permite conexiones extendidas de hasta 10 Km
- Siete puertos FC-AL de reemplazo en caliente de los puertos no usados
- 4 GBICs de onda corta
- Opción de 3 GBICs adicionales SW o LW
- Capacidad de cascadeo

Hub administrado

- Soporte de aplicaciones de cluster, respaldo sin LAN consolidación de servidores y copia remota de disco
- Velocidad de transferencia de 100 MBps para todos los puertos
- Mejora la disponibilidad mediante el aislamiento de fallas
- 8 puertos FC-AL hot swap con siete puertos fijos SW y una ranura de expansión disponible
- Expansión de ciclo mediante cascadeo o conexión a otros concentradores.
- Software de administración para manejar múltiples concentradores

Tabla 7-6 Características de concentradores simples

7.1.4 Concentradores inteligentes (switches)

Debido a la gran popularidad que han ido teniendo las soluciones de ISN y a la reducción de costos, los concentradores inteligentes (*switches*) se han convertido en el componente principal en la infraestructura de una ISN. Todos los concentradores inteligentes soportan 100 MBps por puerto y todos ellos tienen un motor de ruteo que les permite dirigir tramas a otra conexión en caso de que la principal no esté disponible. Afortunadamente para los usuarios finales, la competencia entre proveedores ha originado la cooperación de estos para establecer estándares que al final permitirán la integración total de componentes de diferentes fabricantes, - una de las mayores promesas de las redes de almacenamiento inteligente. Sin embargo, siempre habrán diferencias de un producto a otro, dichas diferencias se centran en cuestiones como la densidad de puertos, desempeño y funcionalidad adicional para facilitar la instalación y la administración.

Los concentradores inteligentes pueden tener 8 ó 16 puertos para aplicaciones departamentales ó de 32 a 128 puertos para aplicaciones empresariales, la figura 7-9 muestra los modelos de concentradores inteligentes de Brocade y QLogic las dos compañías que liderean el mercado de concentradores de canal de fibra en la actualidad.

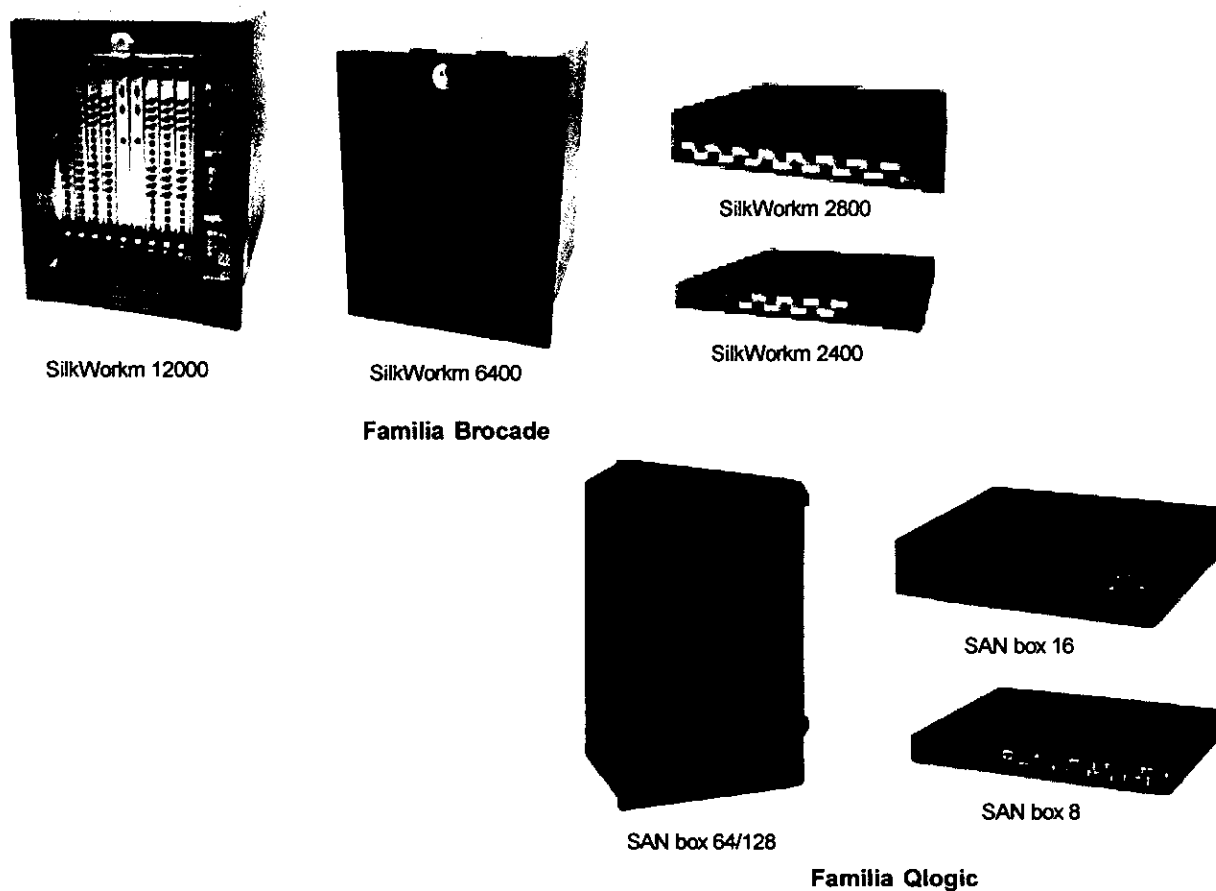


Figura 7-9 Concentradores inteligentes de Brocade y Qlogic

El cascadeo de concentradores inteligentes via puertos de expansión (puertos E) permite que configuraciones pequeñas y medianas se puedan expandir conforme las necesidades de almacenamiento aumenten, con la desventaja de que se requieren puertos adicionales para la interconexión entre los concentradores lo cual también se puede convertir en un potencial cuello de botella. Algunos productos permiten múltiples conexiones entre dos concentradores con balanceo de cargas. Esto resuelve el problema de la congestión pero puede causar problemas en la recepción de la información si las tramas no son entregadas en el orden en que fueron enviadas. Una mejor alternativa para resolver estos problemas es el usar concentradores intermedios dedicados exclusivamente a interconectar los otros concentradores en donde los servidores y los discos estarán conectados, es decir, empleando una topología de árbol, tal como se muestra en la figura 7-10.

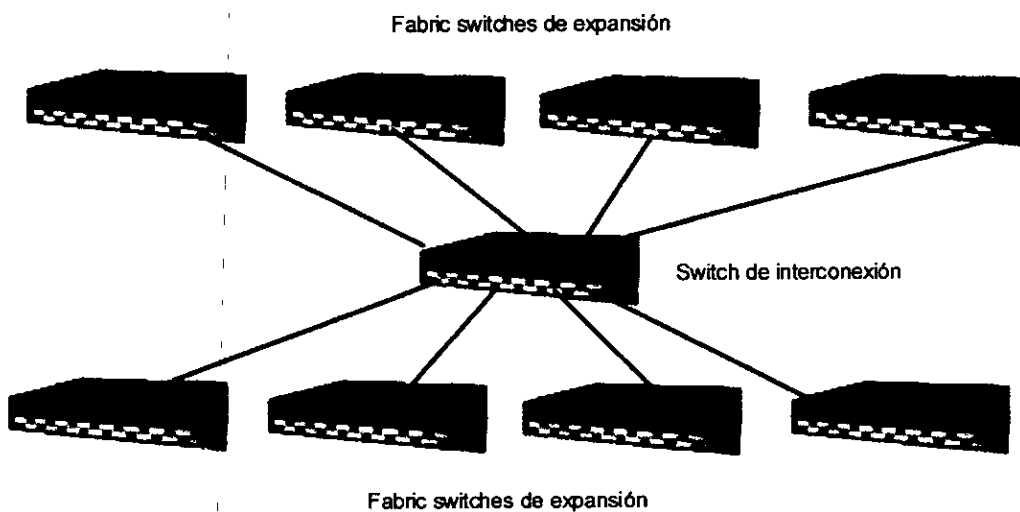


Figura 7-10. Uso de la topología de árbol para mejorar el desempeño

Los estándares de canal de fibra definen los puertos F para conexión de nodos, los puertos E para expansión y los puertos G para soportar otros puertos N o dispositivos FC-AL. Sin embargo, estos estándares no definen como estos tipos de puertos pueden ser implementados en *hardware* por lo que los diseños de cada fabricante pueden variar.

Algunos productos ofrecen un diseño modular que permite agregar puertos a la fábrica con el propósito de incrementar el número de puertos disponibles. Existen también diseños que permiten que un puerto pueda ser configurado *via software* para soportar cualquier tipo de puerto.

En relación al número de puertos FL soportados, estos requieren de muchos recursos del concentrador y aunque algunos fabricantes no ponen restricciones a este respecto, otros limitan el número de puertos FL a cuatro.

Con respecto al desempeño, la diferencia en latencia entre concentradores de diversos proveedores puede tener una variante de hasta 2 microsegundos. Sin embargo, la latencia también se puede ver beneficiada o afectada por la capacidad de memoria de cada puerto. Algunos productos proporcionan suficiente memoria para encolar 2 tramas, otros pueden encolar 16 o más. Adicionalmente, estos espacios de memoria permiten que el concentrador pueda manejar más eficientemente las congestiones de tráfico sin tener que tirar tramas lo cual incrementa la velocidad de comunicación.

Los concentradores también incluyen funcionalidad adicional no contenida por los estándares de canal de fibra, tal como soporte para dispositivos de ciclo privado, zonificación y administración.

Los ciclos privados en concentradores inteligentes pueden ser implementados de diferentes maneras, sin embargo, la forma más eficiente es usando una estrategia de ciclo virtual que permite que los dispositivos de ciclo privado puedan comunicarse con otros como si estuviera en el mismo segmento de ciclo. La configuración física, divide el ciclo en segmentos más pequeños por concentrador lo que permite el switcheo entre segmentos de ciclo privado. El hecho de tener conversaciones concurrentes en un mismo ciclo virtual así como la habilidad de restringir la propagación de LIP's hace que las configuraciones de ciclo privado virtual sean sumamente atractivas para instalaciones que deben soportar dispositivos que no tienen capacidades de fábrica. Adicionalmente, la traslación de ciclos privados extiende estas capacidades permitiendo que puertos N públicos se comuniquen con dispositivos de ciclo privado, de esta manera un servidor con capacidades de fábrica se puede comunicar con arreglos de discos con ciclo privado.

La zonificación permite que un puerto pueda ser asignado a un grupo exclusivo de puertos. En la mayoría de las implementaciones, un puerto puede ser asignado a varios grupos, dependiendo de los requerimientos de las aplicaciones. La zonificación generalmente es una característica que no implica un costo adicional y es la mejor manera de segregar servidores con su respectivo disco de otros servidores, también es una forma de aislar información entre departamentos. Algunas opciones incluyen zonificación por WWN, técnica que ofrece mayor flexibilidad pero también que implica mayor costo ya que se requiere de un servidor externo y de *software* para manejar las zonas.

De manera similar a los concentradores simples, la administración de los concentradores inteligentes es normalmente realizada a través de *Ethernet* usando SNMP o Telnet, aunque también son comunes las implementaciones que utilizan canal de fibra y peticiones SES. La mayoría de las aplicaciones de administración de concentradores son manejadores de dispositivos, esto significa que pueden manejar el concentrador pero no tienen visibilidad de ningún otro componente dentro de la ISN. Las interfaces gráficas de administración de concentradores pueden incluir mapeo de topologías, estadísticas de los puertos y de los concentradores, información de ruteo y gráficas de desempeño de puertos. Estas características permiten a los administradores tener en cualquier momento un reporte completo del estado de la fábrica.

Las tablas 7-7 y 7-8 muestran un comparativo de las características de los principales modelos de concentradores inteligentes de *Brocade* y *Qlogic*.

Concentradores inteligentes de 8 y 16 puertos

Brocade Silkworm 2400 / 2800

- 8 ó 16 puertos universales de canal de fibra
- Soporte de hasta 239 switches interconectados
- Modos de operación: fábrica, ciclo público, ciclo privado
- Soporte de todas las topologías de conexión: cascada, malla, anillo y árbol
- Soporte de estándares de canal de fibra: FC-AL, FC-AL-2, FC-FLA, FC-GS-2, FC-FG FC-PH, FC-PH-2, FC-PH-3, FC-PLDA, FC-SW
- Velocidad de transferencia de 1.0625 Gb/s por puerto
- Latencia < 2 microsegundos
- 224 memoria para tramas (modelo 2400)
- 448 memoria para tramas (modelo 2800)
- Clases de servicio: 2, 3 y F
- Tipos de puertos: FL, F y E
- Transceptores reemplazables en caliente GBIC SW (500 m), GBIC LW (10 km) Cobre (13 m)
- Opciones para GBICs, fuente de poder redundante y kit de montaje
- Administración basada en Telnet, SNMP Brocade WEB Tools y Brocade SES
- Acceso a la administración a través de 100 BASE T o canal de fibra
- Servicios de fábrica: Servidor de nombres simple, Notificación de cambio de estado, Zonificación Administración remota

Qlogic SANbox 8 / 16

- 8 ó 16 puertos universales de canal de fibra
- Soporte de hasta 239 switches interconectados
- Modos de operación: fábrica, ciclo público, ciclo privado
- Soporte de todas las topologías de conexión: cascada, malla, anillo y árbol
- Soporte de estándares de canal de fibra: FC-AL, FC-AL-2, FC-FLA, FC-GS-2, FC-FG FC-PH, FC-PH-2, FC-PH-3, FC-PLDA, FC-SW
- Velocidad de transferencia de 1.0625 GB/s por puerto
- Latencia menor a 0.6 microsegundos
- Tamaño máximo de tramas: 2148 bytes
- Clases de servicio: 2, 3 y sin conexión
- Tipos de puertos: F, FL, E, TL y SL
- Transceptores reemplazables en caliente GBIC SW (500 m), GBIC LW (10 km) Cobre DB-9 (13 m), Cobre HSSDC (13 m)
- Opciones para GBICs y kit de montaje
- Administración basada en SNMP, TFTP, Telnet y SES
- Acceso a la administración a través de Ethernet o canal de fibra
- Servicios de fábrica: Servidor de nombres simple, Notificación de cambio de estado, Zonificación basada en hardware, WWN y puertos, Selección automática de Rutas

Tabla 7-7. Características de concentradores inteligentes de 8 y 16 puertos

Concentradores inteligentes de 64 puertos

Brocade Silkworm 6400

- 64 puertos universales de canal de fibra
- Soporte de hasta 239 switches interconectados
- Rutas de datos internas redundantes
- Núcleo de switcheo redundante
- Detección y aislamiento de fallas automático
- Modos de operación: fábrica, ciclo público, ciclo privado
- Soporte de todas las topologías de conexión: cascada, malla, anillo y árbol
- Soporte de estándares de canal de fibra: FC-AL, FC-AL-2, FC-FLA, FC-GS-2, FC-FG FC-PH, FC-PH-2, FC-PH-3, FC-PLDA, FC-SW
- Velocidad de transferencia de 1.0625 Gb/s por puerto
- Latencia < 2 microsegundos en cada módulo
- 448 frame buffers
- Clases de servicio: 2, 3 y F
- Tipos de puertos: FL, F y E
- Transceptores de reemplazo en caliente GBIC SW (500 m), GBIC LW (10 km) Cobre (13 m)
- Opciones para GBICs y kit de montaje
- Administración basada en Telnet, SNMP Brocade WEB Tools y Brocade SES
- Acceso a la administración a través de 100 BASE T o canal de fibra
- Servicios de fábrica: Servidor de nombres simple, Notificación de cambio de estado, Zonificación Administración remota

Qlogic SANbox 64

- 8 ó 64 puertos universales con incrementos de 8
- Soporte de hasta 239 switches interconectados
- Rutas de datos internas redundantes
- Núcleo de switcheo redundante
- Detección y aislamiento de fallas automático
- Modos de operación: fábrica, ciclo público, ciclo privado
- Soporte de todas las topologías de conexión: cascada, malla, anillo y árbol
- Soporte de estándares de canal de fibra: FC-AL, FC-AL-2, FC-FLA, FC-GS-2, FC-FG FC-PH, FC-PH-2, FC-PH-3, FC-PLDA, FC-SW
- Velocidad de transferencia de 1.0625 GB/s por puerto
- Latencia entre 0.6 y 3 microsegundos
- Tamaño máximo de frames: 2148 bytes
- Clases de servicio: 2, 3 y sin conexión
- Tipos de puertos: F, FL, E, TL y SL
- Transceptores de reemplazo en caliente GBIC SW (500 m), GBIC LW (10 km) Cobre DB-9 (13 m), Cobre HSSDC (13 m)
- Opciones para GBICs fuentes adicionales y kit de montaje
- Administración basada en SNMP, TFTP, Telnet y SES
- Acceso a la administración a través de Ethernet o canal de fibra
- Servicios de fábrica: Servidor de nombres simple, Notificación de cambio de estado, Zonificación basada en hardware, WWN y puertos, Selección automática de Rutas

Tabla 7-8. Características de concentradores inteligentes de 64 puertos

7.1.5 Puentes

Los puentes (*bridges*) permiten que dispositivos SCSI puedan integrarse a un esquema de redes de almacenamiento inteligente. Estos normalmente proporcionan una o más interfases de canal de fibra para conexión a la ISN y de dos a cuatro puertos SCSI para arreglos de discos y bibliotecas de cintas basados en SCSI. Además de esta conversión física y de transporte, los puentes hacen translaciones del protocolo SCSI-3 al protocolo SCSI apropiado para los dispositivos de legado. La figura 7-11 muestra algunos de los modelos de puentes que actualmente se encuentran en el mercado.

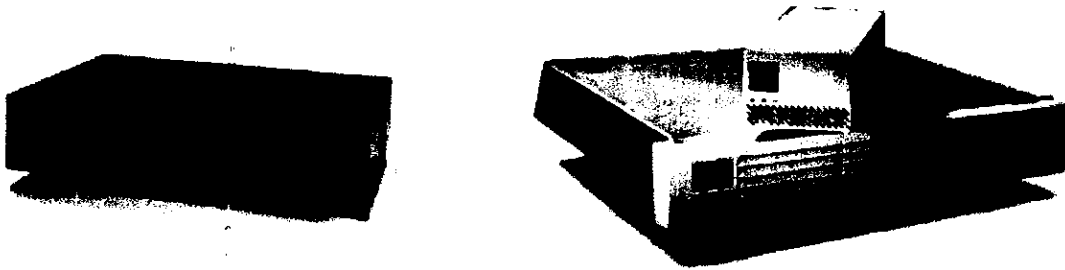


Figura 7-11. Puentes

La aplicación más común de los puentes de SCSI a canal de fibra es el soporte a modelos anteriores de bibliotecas de cintas tal como se muestra en la figura 7-12.

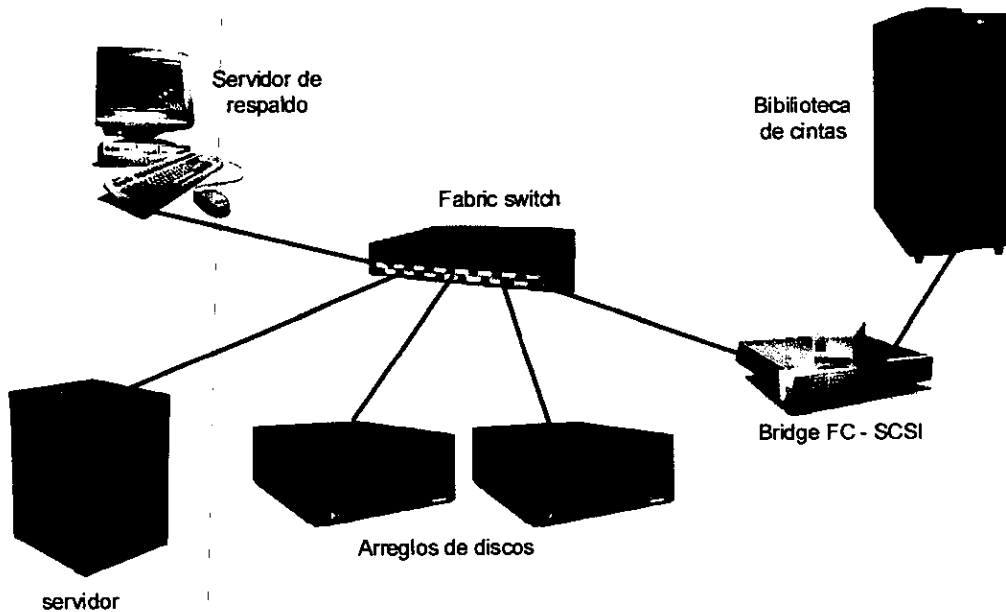


Figura 7-12. Uso de los puentes FC-SCSI en soluciones de respaldo

Las bibliotecas de cintas representan una inversión sustancial y no es fácil rediseñarlas para soportar canal de fibra. Sin embargo, el respaldo a cinta, es un requerimiento esencial en cualquier red de almacenamiento inteligente. Sin una manera de acceder la cinta a través del medio de canal de fibra, se requiere del uso de una red LAN para realizar esta función lo cual puede agregar considerable tráfico a la red de producción. Por tanto, los puentes permiten preservar la inversión hecha de *hardware* a la vez que se optimiza el proceso de respaldos.

El colocar un sistema de cintas atrás de un puente también elimina el hecho de que un solo servidor sea el dueño del dispositivo permitiendo que este pueda ser controlado por cualquier servidor en la red de almacenamiento.

El papel que juegan los puentes en una ISN irá decayendo conforme se difundan más los dispositivos basados en canal de fibra. Sin embargo, dada la gran inversión que las compañías han hecho en dispositivos basados en SCSI los puentes seguirán siendo un elemento importante en las redes de almacenamiento inteligente por un tiempo considerable.

7.1.6 Arreglos de discos

RAID

Los sistemas RAID de canal de fibra usan una controladora encargada de manejar las operaciones de I/O así como los cálculos de paridad, lo que libera al servidor de realizar estas funciones. Esta controladora se encuentra entre la interfase de conexión y el arreglo de discos y recibe peticiones de SCSI-3 a través de un ciclo o fábrica para almacenar u obtener los datos del arreglo usando el nivel apropiado de RAID.

Desde el punto de vista de interconexión, la interfase física entre la controladora de RAID y sus discos es irrelevante. La controladora aparece como un puerto N o NL al mundo exterior y puede usar un canal propietario, SCSI o cualquier otra arquitectura para hablar con sus unidades. Con respecto al desempeño y confiabilidad, los fabricantes de discos RAID también pueden incorporar tecnología de canal de fibra detrás de la controladora. Esto permite el uso de unidades con dobles conexiones.

Existe un gran número de proveedores de sistemas de discos en la industria, pero los más reconocidos son EMC, Sun, IBM, HP, Hitachi y Compaq. La figura 7-13 ilustra algunos de los modelos de RAID de estos proveedores.

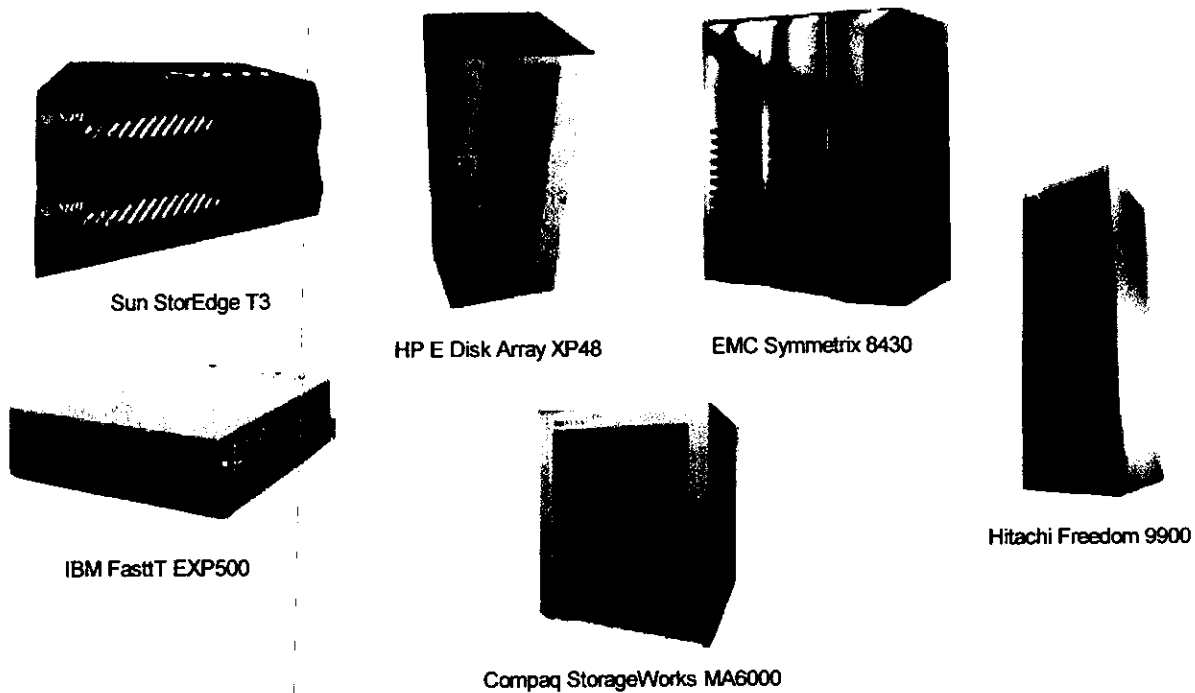


Figura 7-13 Sistemas de discos RAID

Para redes de almacenamiento inteligente en esquemas de misión crítica, los sistemas RAID ofrecen un gran número de ventajas. El soporte de diagnósticos avanzados, las características de administración, la alta confiabilidad y escalabilidad inherentes a estos productos los hacen la elección preferida. Los servidores también se benefician de las bondades de estos sistemas de discos descargando las tareas de manejo de paridad y de I/O en la controladora de RAID. Adicionalmente, cuando el servidor es liberado de las tareas de manejo de RAID, también se libera de la pertenencia exclusiva del dispositivo de almacenamiento. Por este motivo, cualquier servidor puede hacerse propietario de una porción de los datos del arreglo de discos lo que permite que sean una de las mejores opciones para soluciones de agrupación de servidores y otras aplicaciones en donde se requiere que dos o más servidores compartan los recursos de almacenamiento.

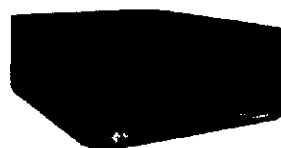
Desde el punto de vista de diseño, los RAID's de canal de fibra simplifican las configuraciones ya que la controladora de RAID aparece como una sola dirección de puerto al ciclo o fábrica. Agregar unidades de disco al sistema de RAID no incrementa el número de dispositivos en la fábrica aunque internamente utilicen topología de ciclo arbitreado, de esta manera es posible tener terabytes de información en ciclo arbitreado sin consumir todas las AL_PA's disponibles. Para esquemas de fábrica, la controladora de RAID y todo su espacio en disco aparecen como un solo puerto N en su propio segmento de 100 MBps mientras que un sistema JBOD aparece como múltiples puertos NL compartiendo un mismo ancho de banda. La tabla 7-9 muestra un comparativo de algunos sistemas de RAID de algunos proveedores.

Características	EMC Symmetrix 8430	Hitachi Freedom 9900	HP E Disk Array XP48
Proveedor	EMC	Hitachi Data Systems	Hewlett Packard
Tipo de producto	Arreglo de discos	Arreglo de discos	Arreglo de discos
Capacidad Máxima (en el mismo gabinete)	17,376 GB	3,504 GB	3,504 GB
Interfases de conexión al host	ESCON, FC-AL, Ultra SCSI 2	ESCON, FC-AL	ESCON, FC-AL, Wide SCSI
Interfase de los discos	Ultra SCSI	FC-AL	FC-AL
Máximo número de discos	96	48	48
Número de interfaces	24	24	24
Plataformas soportadas	Compaq, HP, IBM, servidores Intel, Silicon Graphics, Sun Unisys, Windows NT/2000 Linux, NCR, Sequent	MVS, Sun, HP-UX Tru64 SGI IRIX, Linux, Windows NT/2000, Novell Netware Sequent	HP-UX, IBM-AIX, Sun Tru64, SGI IRIX, Linux Windows NT/2000, Novell Netware
Discos soportados	36 GB, 73 GB, 181 GB	18.4 GB, 47 GB, 73 GB	18 GB, 73 GB
Controladores de RAID	Procesadores Power PC	ACP	ACP
RAID soportado	1, S	0+1, 5	0+1, 5
Tamaño del caché	32,768 MB	16,384 MB	16,000 MB

Tabla 7-9. Comparativo de sistemas RAID

JBOD

Un arreglo JBOD (*Just a Bunch of Disks*) es un sistema con múltiples unidades de disco de canal de fibra conectados a una tarjeta madre común. La tarjeta madre proporciona las conexiones de transmisión y recepción que permiten la configuración de los discos en un esquema de ciclo arbitreado con electrónica de paso lo que permite que los discos puedan ser insertados o quitados sin que el ciclo se vea afectado. La figura 7-14 muestra algunos de los modelos de arreglos JBOD.



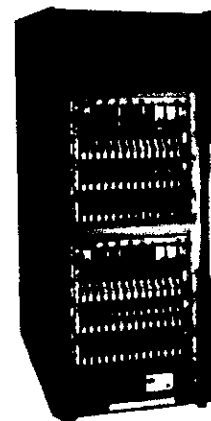
Sun StorEdge A5200



EMC FC4500



HP Storage 12FC



IBM 2106

Figura 7-14 Sistemas de discos JBOD

Cuando un JBOD es conectado a un concentrador de fibra óptica, la conexión no es hecha a un solo dispositivo de canal de fibra sino a múltiples dispositivos de ciclo independientes dentro del sistema de discos. Un sistema JBOD con ocho discos, por ejemplo, aparece como ocho AL_PA's al mundo exterior. Si la conexión se hace a un concentrador inteligente, entonces el concentrador debe tener un puerto FL ya que el sistema JBOD por si mismo es un segmento de un ciclo. Si la conexión se hace a un concentrador de ciclo arbitreado, el número de dispositivos en el ciclo se incrementa con el número de unidades en el JBOD por lo que estos dispositivos, a diferencia de los sistemas RAID tienen un impacto directo en la topología de la ISN a la que estén conectados.

Algunos sistemas de discos basados en JBOD incorporan una lógica adicional para propósitos de administración, usualmente una controladora de canal de fibra que soporta comandos SES. Esto agrega una AL_PA adicional a la configuración la cual puede ser manejada por una estación de trabajo desde la cual se puede solicitar información del arreglo de discos. Los JBODs también pueden incluir opciones para configurar la tarjeta madre y soportar conexiones duales a un conjunto de discos o dividir los discos en conjuntos más pequeños con una sola conexión de ciclo tal como se muestra en la figura 7-15.

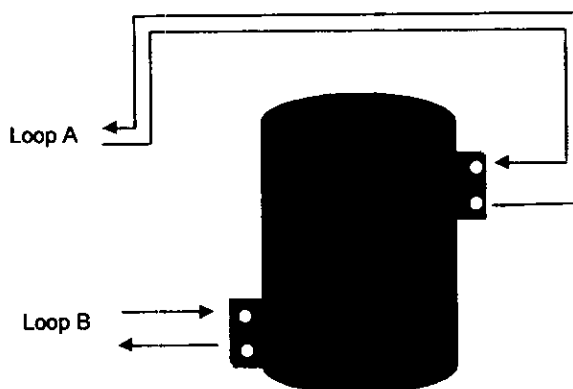


Figura 7-15. Dividiendo la tarjeta madre de un JBOD en ciclos separados

Los sistemas JBOD normalmente están conformados por bahías de 8 a 11 discos y en algunos casos pueden estar configurados para soportar fallas en las rutas de datos. Los arreglos de discos más grandes pueden venir montados en gabinetes de 19" en donde también se pueden incluir concentradores para interconectar los diferentes módulos JBOD hacia los servidores permitiendo tanto conexiones simples como duales.

Adicionalmente, fuentes de poder, ventiladores y conexiones redundantes permiten a los JBOD's ser usados en ambientes en donde se requiere de alta disponibilidad. Algunos proveedores también pueden ofrecer tarjetas controladoras de RAID con el propósito de mejorar el desempeño de los sistemas de discos a la vez que se protege la inversión hecha. La tabla 7-10 muestra un comparativo de las características de algunos sistemas JBOD que se pueden encontrar en el mercado.

Características	Sun StorEdge A5200	EMC CLARIION FC4500	IBM 2106 Storage
Proveedor	Sun	EMC	IBM
Tipo de producto	Arreglo de discos	Arreglo de discos	Arreglo de discos
Capacidad Máxima (en el mismo gabinete)	4,805GB	7,300 GB	4,300 GB
Interfases de conexión al host	FC-AL	FC-AL	FC-Switch, FC-AL
Interfase de los discos	FC-AL	FC-AL	Ultra SCSI
Máximo número de discos	22	100	72
Número de interfaces	4	2	2
Plataformas soportadas	Sun Solaris, Windows NT	HP-UX, IBM AIX, Sun Solaris, Windows NT, Windows 2000	Open VMS, IBM AIX Linux, Windows NT/2000 Windows NT/2000, Novell Netware
Discos soportados	18.2 GB 36.4 GB	18.4 GB, 36 GB, 73 GB	18 GB, 36 GB
RAID soportado	JBOD	JBOD	JBOD

Figura 7-10. Comparativo de sistemas JBOD

7.1.7 Multiplexores

Los multiplexores permiten que múltiples cadenas de datos puedan ser transmitidas a través de un solo cable de fibra óptica, cada uno a una longitud de onda específica, lo que permite un incremento en el ancho de banda, aumentando la cantidad de datos que puede ser manejada por una red de almacenamiento inteligente y permitiendo conectar varias ISN's separadas varios kilómetros entre sí. Esto elimina la necesidad de rentar o instalar múltiples cables de fibra óptica lo que representa una reducción significativa en los costos. Una ventaja clave de los multiplexores es que son independientes de la velocidad de transmisión así como del protocolo por lo que pueden operar sobre Canal de Fibra, ESCON, Gigabit Ethernet, Fast Ethernet o SONET/SDH y manejar velocidades de transmisión de datos de entre 10 Mb/s y 2.5 Gb/s. La figura 7-16 muestra un diagrama de la operación de un multiplexor.

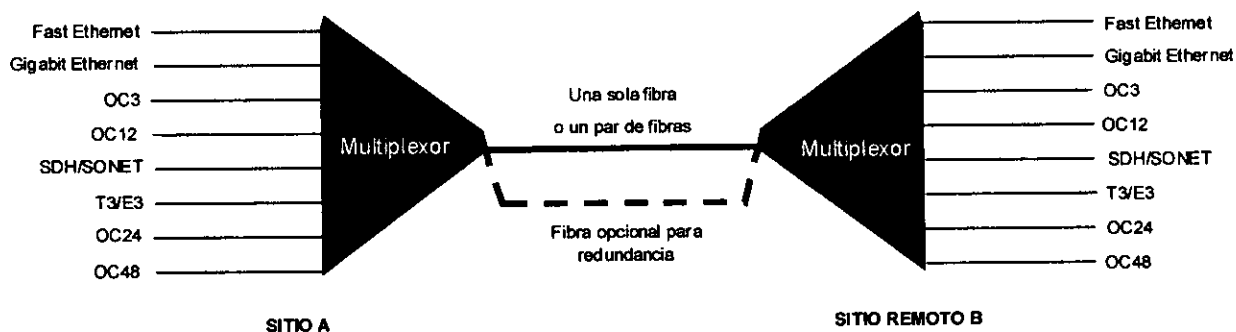


Figura 7-16. Diagrama de operación de multiplexores

Estas características permiten que una fábrica pueda ser extendida sobre grandes distancias sin realizar ninguna traslación de protocolos, además junto con las capacidades de memoria de los puertos E, se puede garantizar la velocidad de transmisión haciendo que estos dispositivos puedan ser usados en aplicaciones tales como replicación de datos síncrona lo que facilita la continuidad de operación del negocio y una recuperación rápida en caso de una contingencia. La figura 7-17 muestra la forma en la que se puede expandir una fábrica a través de multiplexores.

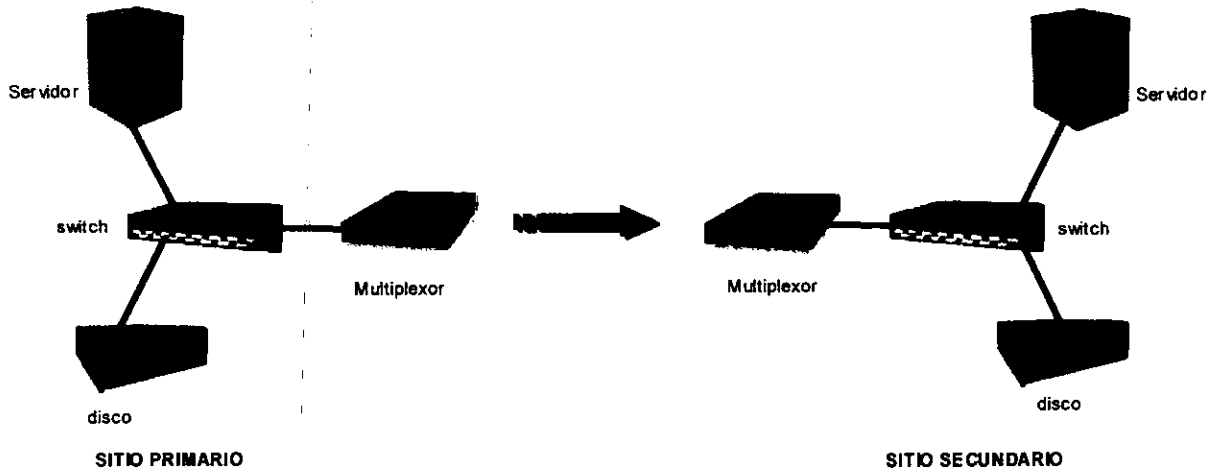


Figura 7-17. Configuración de fábrica extendida con multiplexores

Algunos multiplexores son capaces de concentrar hasta 32 canales de entrada en uno o dos cables de fibra, con cada canal de entrada manejando una velocidad de transmisión de hasta 2.5 GB/s para una velocidad de transmisión combinada de 80 Gb/s. Estos dispositivos constan de dos secciones funcionales: las tarjetas y el módulo de multiplexaje óptico. Cada tarjeta convierte sus señales de entrada a una longitud de onda específica. La administración de estos dispositivos es realizada a través de agentes SNMP permitiendo que puedan ser manejados de manera local o remota. La tabla 7-11 muestra las características típicas de un multiplexor de fibra óptica.

Características

- Arquitectura modular con expansión de canales sin interrupción de la operación.
- Independiente del protocolo y velocidad de transmisión
- Distancias de conexión de hasta 63 y 83 kilómetros
- Componentes redundantes
- Protocolos soportados: Canal de fibra, ESCON, ATM, FE, GE, FDDI, T3/E3
- Administración a través de SNMP

Especificaciones de interfase

- Entrada: 10 Mb/s, 100 Mb/s ó 1.25 GB/s multimodo o monomodo
- Salida: Hasta 83 kilómetros usando un par de fibras
- Hasta 63 kilómetros usando una sola fibra

Tabla 7-11. Características de un multiplexor

7.1.8 Compuertas

Las compuertas (*gateways*) son dispositivos que permiten conectividad sobre distancias globales permitiendo configuraciones de fábrica remota. Las compuertas aprovechan la infraestructura de WAN existente así como protocolos tales como ATM, IP y SONET para aplicaciones tales como replicación y respaldo remoto. La figura 7-18 muestra la manera de implementar una red de almacenamiento remota a través de compuertas.

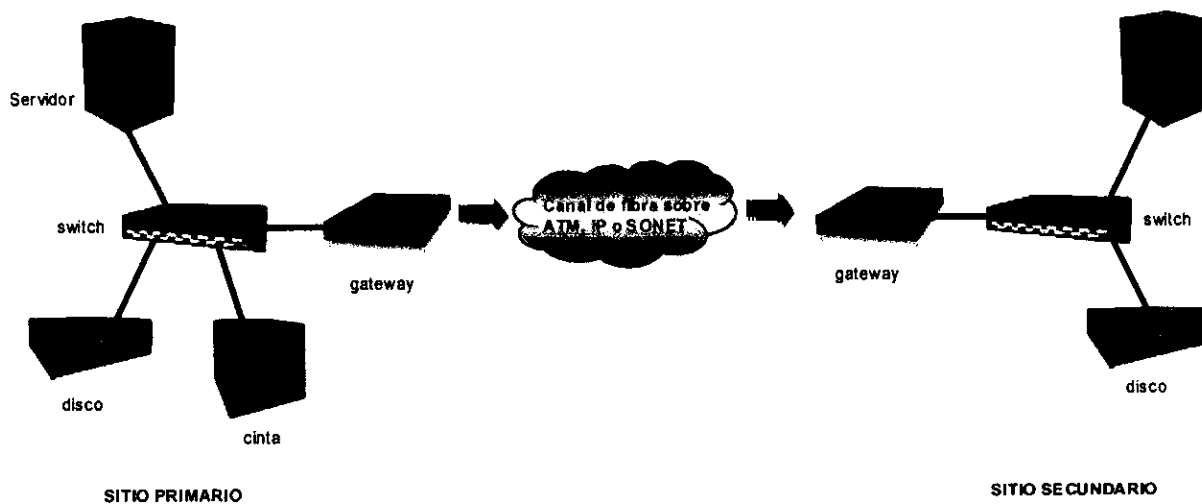


Figura 7-18. Configuración de fábrica extendida usando compuertas

Una de las funciones principales de las compuertas consiste en hacer traducciones de protocolos tal como canal de fibra sobre ATM o IP. De esta manera un sistema de discos se conecta a una SAN local y después a través de la compuerta el arreglo aparece como un dispositivo SCSI local en el site remoto.

Estos dispositivos permiten que las redes de almacenamiento inteligente puedan ubicarse en diferentes continentes con el mejor desempeño posible para mover la información. Dentro de las características incluidas en las compuertas se encuentran:

- Detección y Recuperación de Errores.
- Conversión de protocolos y enmascaramiento de LUN's.

La tabla 7-12 muestra las características principales de las compuertas.

Características	Soporte de copia remota de discos, respaldo remoto y compartición de disco Soporte de topologías de fábrica y ciclo así como conexión de dispositivos SCSI Soporte de ATM / SDH / Sonet (OC3) T3/E3 Conversión de protocolos y enmascaramiento de LUN's Administración a través de SNMP Detección de fallas y fuentes de poder redundantes
------------------------	--

Tabla 7-11. Características de una compuerta

7.2 Productos de software

Para que los productos de *hardware* que conforman una red de almacenamiento inteligente puedan operar es necesario que exista un elemento de *software* que los guíe en sus funciones. Es importante mencionar que estos son los productos que tienen que ver directamente con la red de almacenamiento ya que arriba de ellos se encuentran las aplicaciones que propiamente manejarán la información como pueden ser bases de datos, sistemas de ERP, soluciones de agrupación de servidores, respaldo y replicación por mencionar algunos. A continuación se mencionan los productos de *software* necesarios para la operación, manejo y administración de una ISN.

- Sistema operativo para fábricas
- Traductor de ciclos
- Zonificación
- SES
- Administración basada en Web

7.2.1 Sistema operativo para fábricas

Gran parte de la inteligencia que tienen las fábricas se encuentra en el sistema operativo que reside en estos, el cual permite el soporte de cientos de dispositivos interconectados, a la vez que se asegura la integridad de la información transmitida y la velocidad de transmisión. El sistema operativo conocido también como *Fabric OS* es el responsable de manejar diferentes configuraciones y topologías permitiendo que se vaya agregando mayor funcionalidad a la red de almacenamiento con el transcurso del tiempo. A continuación se mencionan las características y funciones primordiales del sistema operativo para fábricas.

- Identifica el tipo de puerto y lo inicializa dependiendo del tipo de dispositivo conectado: concentrador, servidor o dispositivo de ciclo privado, de esta manera cada puerto del concentrador puede ser de uso universal.
- Automáticamente descubre y registra nuevos dispositivos conforme estos son conectados a la ISN.
- Deshabilita la transferencia de datos a cualquier puerto que tenga una falla, así mismo una vez que se ha corregido el problema el puerto es automáticamente rehabilitado.
- Proporciona el servidor de nombres simple así como notificación de cambios de estado.
- Selecciona la ruta más eficiente para transferir datos en un ambiente de múltiples concentradores
- Proporciona balanceo de cargas.
- Automáticamente reconfigura rutas alternas cuando una conexión falla.
- Garantiza que las tramas son entregadas en el orden en el que fueron enviadas.
- Proporciona servicios de administración de SNMP.
- Proporciona un sistema de administración en línea de comandos.

La figura 7-19 ilustra las funciones y servicios ofrecidos por el sistema operativo y su relación con otros productos de *software* utilizados en un ambiente de ISN.



Figura 7-19. Funciones y servicios ofrecidos por el sistema operativo para fábricas

7.2.2 Traductor de ciclos

El traductor de ciclos ofrece la posibilidad de que dispositivos de canal de fibra en esquemas de ciclo privado puedan ser incorporados a un ambiente de ISN sin requerir ninguna modificación en los manejadores de los servidores. La manera en la que se forma un ciclo privado es configurando lógicamente los puertos para transferir datos de manera independiente lo que incrementa considerablemente el ancho de banda disponible ya que este no es compartido, soportando un máximo de 126 dispositivos por cada ciclo privado.

Es posible también tener en un mismo concentrador configuraciones de ciclo privado y fábrica, esto permite que nuevos dispositivos sean incorporados a la configuración mientras que se protege la inversión hecha con los dispositivos que no soportan fábrica.

El traductor de ciclos también tiene la capacidad de detectar y aislar cualquier puerto con falla en el ciclo y monitorearlo constantemente hasta que se haya recuperado para después regresarlo a su estado de operación normal.

7.2.3 Zonificación

La zonificación permite que un administrador pueda crear segmentos o zonas dentro de una fábrica, asegurando que la información será accesada solo por los dispositivos definidos en una zona.

Las zonas pueden ser configuradas dinámicamente variando en tamaño y forma siendo el número de estas ilimitado. Los dispositivos pueden ser miembros de más de una zona y es posible crear zonas temporales para uso específico como el caso de respaldos. La configuración de una zona involucra usar un conjunto de comandos para crear, borrar y desplegar las zonas, para agregar o quitar miembros de una zona y para configurar conjuntos de zonas, una vez definidas, la restricción de acceso es automática.

Una zona es referenciada por un nombre mientras que un miembro de una zona es especificado por su correspondiente número de puerto físico en la fábrica, por el nombre del nodo o por el nombre del puerto. Es posible también usar alias para facilitar el trabajo del administrador. La tabla 7-13 muestra las características principales de la zonificación.

Características	Descripción
Especificación	Creación, borrado y despliegue de zonas Adición y borrado de miembros de zonas Despliegue de miembros por zona o configuración Creación de alias. Creación de zonas temporales
Aseguramiento	Servidor de Nombres Simple
Administración	Telnet
Respaldo	La configuración de una zona es guardada en la memoria flash del concentrador inteligente

Tabla 7-13. Características de la zonificación

7.2.4 SES

Este producto permite que cualquier servidor conectado a una fábrica que use el protocolo SES pueda manejar todos los concentradores dentro de una ISN. La administración se realiza a través del medio de fibra por lo que este *software* constituye la opción ideal para manejar ambientes en donde existen arreglos de discos que no tienen una interfase de red disponible para propósitos de monitoreo.

Para que un servidor pueda manejar una ISN a través de SES, el servidor necesita realizar un registro normal a la fábrica y después iniciar la solicitud apropiada de comandos de SES. Basado en la información obtenida via SES, el servidor puede realizar una operación en el concentrador inteligente, por ejemplo, puede habilitar o deshabilitar un puerto, tomar las lecturas del sensor de temperatura del concentrador o monitorear el desempeño o errores encontrados en la transmisión de datos. La tabla 7-14 muestra las características principales de SES.

Características	Descripción
Estándares	SCSI-3 FCP, SCC, SES y SPC
Comandos soportados	Recepción de resultados de diagnóstico, reporte de LUN's Petición de sensado, información del estado de los componentes
Interfase	Administración en banda sobre canal de fibra
Requerimientos	Servidor con manejo del protocolo SES

Tabla 7-14. Características principales de SES

7.2.5 Administración basada en navegador

La administración a través de un navegador permite que una red de almacenamiento inteligente pueda ser monitoreada y manejada desde una interfase intuitiva y fácil de usar. Basta con tener una estación de trabajo o computadora personal con un navegador que soporte *Java* para simplemente escribir la dirección IP de cualquier concentrador en la fábrica y a partir de ahí recabar la información que se necesite o realizar la tarea de administración que se requiera. Existen herramientas de administración que permiten que esta se pueda llevar a cabo a través de conexiones de canal de fibra (en banda) o a través de una red Ethernet (fuera de banda).

Es posible tener diferentes vistas de una ISN, de esta manera se puede monitorear la red de almacenamiento completa o un componente específico tal como se menciona a continuación:

Vista de la fábrica. Despliega todos los concentradores que conforman la ISN desde una sola pantalla. Este despliegue gráfico muestra todos los concentradores configurados en la fábrica y proporciona un punto central a partir del cual se puede monitorear y administrar cualquier fábrica en la ISN, tal como se muestra en la figura 7-20.

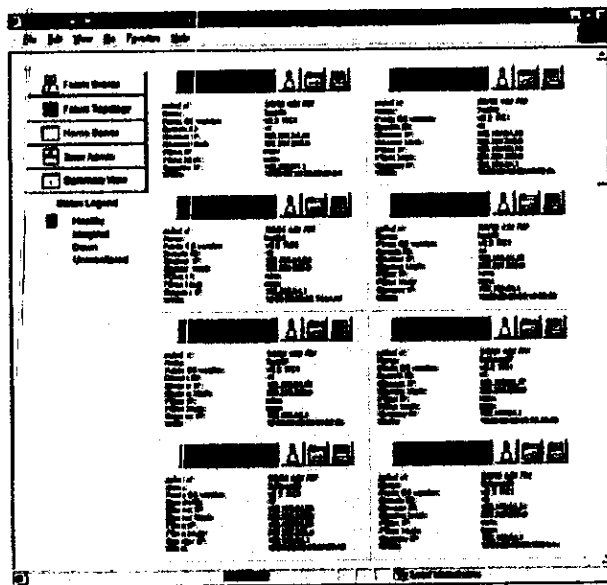


Figura 7-20. Vista de una fábrica a partir de un navegador

Vista de Eventos en la fábrica. Despliega sucesos de la fábrica recolectados del sistema de mensajes de cada concentrador

Vista de la Topología. Lista la configuración física incluyendo dominios activos, rutas e información de ruteo para todos los concentradores en la ISN.

Vista del Servidor de Nombres. Despliega información acerca de los dispositivos registrados en la fábrica. Esta información es automáticamente actualizada cuando se agregan nuevos dispositivos a la ISN.

Vista del Switch. Permite obtener información de un concentrador tal como estado de las fuentes de poder ventiladores, puertos y temperatura. La figura 7-21 muestra la vista de un concentrador.

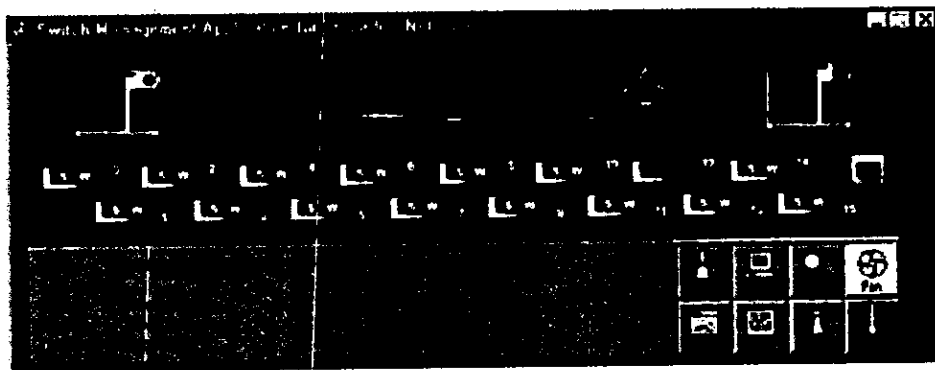


Figura 7-21. Vista de un concentrador de conexiones inteligente

Vista de Desempeño. Muestra información del desempeño de cada puerto o de un concentrador completo.

Vista Detallada de Puertos. Despliega estadística, información y estado de los puertos de un concentrador.

VIII

Casos de Estudio

Aunque las redes de almacenamiento inteligente comparten componentes comunes en la forma de servidores, arreglos de discos, dispositivos de interconexión y *software*, la configuración de una ISN está determinada por las aplicaciones que correrán en ella. Los requerimientos de aplicaciones de respaldo local o remoto varían de una solución de consolidación de disco, así como los elementos necesarios para implementar una solución de replicación de datos remota pueden ser diferentes de aquellos indispensables para una configuración de agrupación de servidores. Sin embargo, es posible satisfacer las necesidades de múltiples aplicaciones dentro de una misma configuración de ISN si ésta es diseñada apropiadamente, tal como una red LAN puede dar servicio a diferentes aplicaciones.

En algunos casos, el desarrollo de una nueva infraestructura de red provee oportunidades adicionales de resolver problemas que no están relacionados directamente con el problema que se está tratando de solucionar. Un diseño de ISN para una aplicación que requiere de grandes anchos de banda, por ejemplo, puede dar cabida a una solución más eficiente de respaldo, aunque este último no sea el objetivo principal del diseño.

Aunque como ya se comentó las redes de almacenamiento inteligente pueden no ser la respuesta para todos los problemas de manejo de información si pueden brindar una amplia gama de soluciones. A continuación se examinan las características de cuatro aplicaciones de las ISN's así como ejemplos prácticos de implementación de ellos

8.1 Aplicaciones

- Respaldos
- Agrupaciones de servidores
- Redes de Almacenamiento Remotas
- Replicación de datos

8.1.1 Respaldos

La necesidad de respaldar la información a un medio secundario como son las cintas magnéticas involucra muchos retos, los cuales no son resueltos de manera eficiente utilizando los métodos de respaldo tradicionales basados en la conexión directa de las bibliotecas de cinta a los servidores o llevando a cabo el respaldo a través de la red LAN. En la figura 8-1 cuatro servidores departamentales comparten un recurso de respaldo común a través de una red de producción LAN.

Incluso si se tuviera una red conmutada Ethernet de 100 Mbps y disponibilidad total de la red para llevar a cabo los respaldos, la velocidad máxima de transferencia de datos entre los servidores y la cinta sería de aproximadamente 25 GB/h. Suponiendo que cada servidor almacenara unos 100 GB de datos el llevar a cabo un respaldo completo de toda la información se tomaría unas 16 horas. Estas restricciones de tiempo obligan a que se implementen políticas de respaldo en donde se combinen respaldos totales e incrementales para eficientar el uso de recursos y acomodarse a las ventanas de respaldo disponibles.

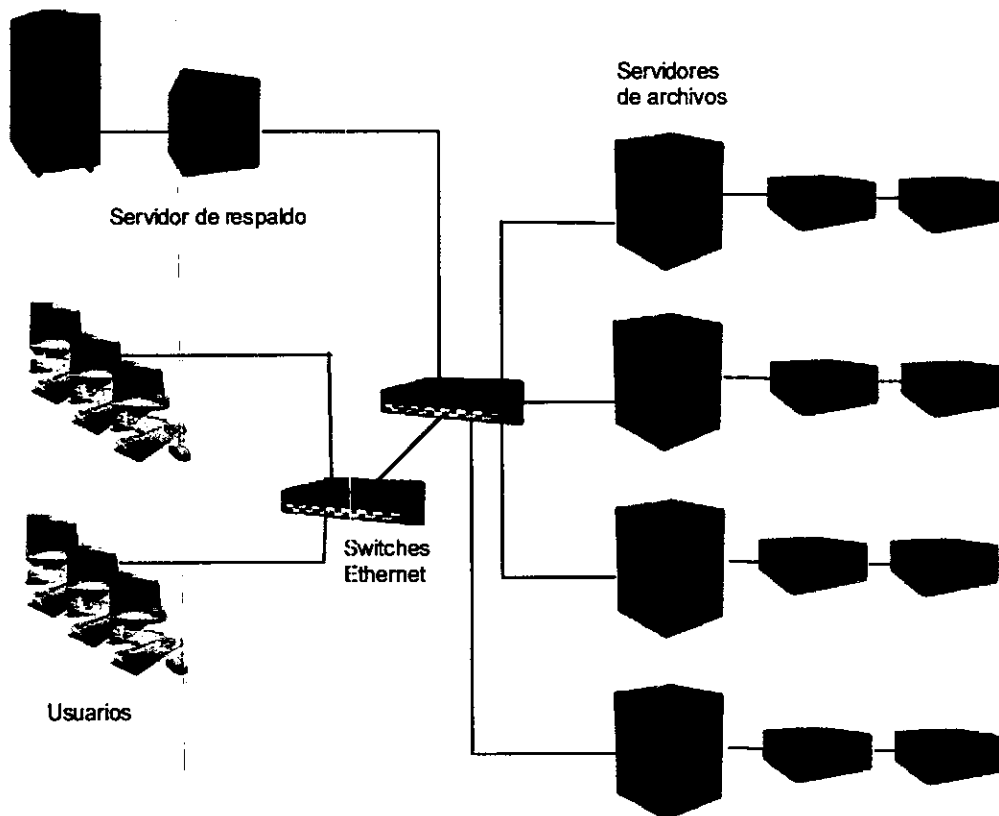


Figura 8-1. Respaldo tradicional a través de una red LAN de producción

Por otro lado, conforme el volumen de datos a respaldar se incrementa, el tiempo para hacer el respaldo también será mayor lo que implica cambiar la infraestructura de red para tener mayor ancho de banda o utilizar una red alterna. El cambiar a *Gigabit Ethernet* por ejemplo, puede liberar a la red de producción, sin embargo, la relación servidor/arreglo de discos/biblioteca de cintas continúa sin ningún cambio. Por lo que la mejor manera de resolver el conflicto entre el tráfico de datos y los requerimientos de almacenamiento y respaldo es utilizar una red separada del ambiente de producción.

De esta manera el esquema representando en la figura 8-1 se puede llevar a un ambiente de ISN como se ilustra en la figura 8-2. Como se puede apreciar la biblioteca de cintas SCSI se queda intacta, pero se incluyen nuevos componentes a la configuración tales como tarjetas adaptadoras de canal de fibra, concentradores y puentes para permitir la conversión de SCSI a canal de fibra de la biblioteca de cintas.

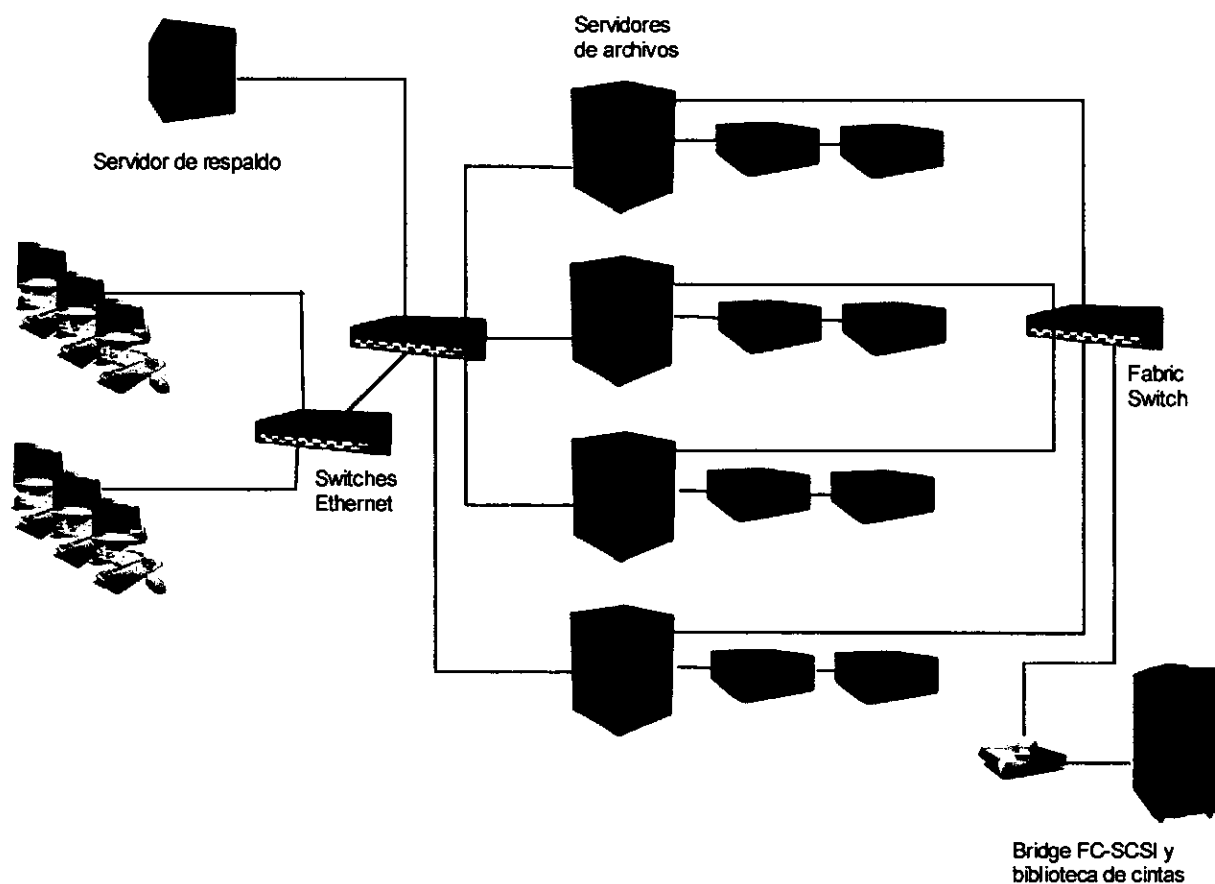


Figura 8-2. Transición a una implementación de respaldo sin usar la LAN

Ya que el sistema de cintas aparece como cualquier otro dispositivo SCSI en un canal separado para cada uno de los servidores, este es accesible desde cualquier servidor. El *software* administrador de respaldos instruye a cada servidor sobre cuando y como hacer los respaldos.

El ancho de banda de 100 MBps del canal de fibra y la flexibilidad de mover datos en su propio medio de transporte resuelven el problema del tráfico en la red de producción pero poco pueden hacer por mejorar los tiempos de respuesta si no se cuenta con una biblioteca de cintas que proporcione la velocidad de transferencia de datos necesaria, además con este tipo de implementación se siguen usando ciclos de procesador de cada servidor en el proceso de respaldo. Una alternativa a esto consiste en conectar los sistemas de discos directamente a la red de almacenamiento para quitar a los servidores de la ruta de datos de respaldo tal como se muestra en la figura 8-3, de esta manera los servidores siempre están disponibles para el acceso a los usuarios y debido a que el respaldo no interfiere más con la operación diaria, la ventana de respaldo pasa a segundo plano ya que los respaldos pueden llevarse a cabo en cualquier momento mediante el control del *software* de respaldo que también se encarga de manejar los permisos y actualizaciones de los archivos.

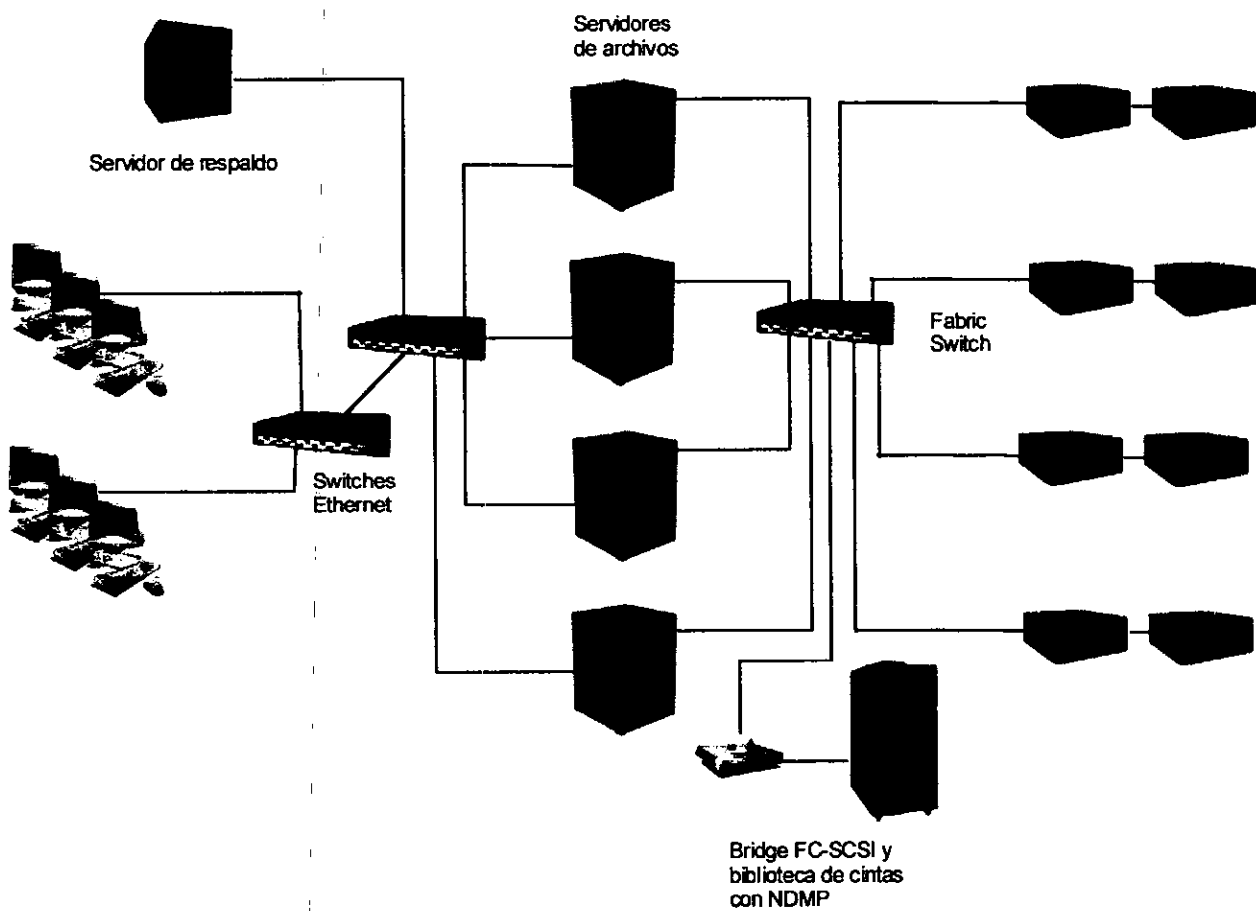


Figura 8-3. Implementación de un esquema de respaldos sin usar la LAN y los servidores

Para que se pueda enviar información directamente desde el arreglo de discos a la biblioteca de cintas se necesita de un agente de respaldo que soporte el Protocolo de Administración de Datos sobre la red (NDMP, *Network Data Management Protocol*) o de algún otro protocolo de copia de datos sobre la red, este agente puede residir en un servidor de respaldos conectado a la ISN, en un puente SCSI-Canal de fibra o en algunos casos en la misma biblioteca de cintas.

En esta configuración los datos a respaldar son leídos directamente de los arreglos de discos a través del agente NDMP y son escritos a cinta, sin pasar por los servidores y aunque la ISN proporciona el vehículo para mover los datos, es el *software* el encargado de determinar cuando y como se hacen los respaldos. El acceso a los datos de parte de los usuarios y el respaldo se pueden hacer al mismo tiempo siempre y cuando el *software* de respaldo lleve un control de los cambios que hayan en los archivos.

8.1.2 Agrupaciones de servidores (clusters)

Conforme las aplicaciones empresariales basadas en Internet y en sistemas de cómputo se han hecho cada vez más críticas el acceso continuo a la información se ha vuelto más importante que nunca. Por esta razón, los fabricantes de equipos de cómputo han respondido con diseños más sofisticados que incluyen componentes redundantes tales como fuentes de poder, interfases de conexión, procesadores y otras características encaminadas a mejorar la disponibilidad y confiabilidad de los sistemas. La falla potencial de un componente es resuelta a través de esta redundancia. Pero esta característica puede ser extendida al *software* usando múltiples servidores corriendo las mismas aplicaciones. De esta manera, la falla de un módulo de *hardware* o *software* que implique la incapacidad de un servidor de seguir prestando servicio es solucionada moviendo a los usuarios conectados del servidor con falla a uno o más servidores dentro de la agrupación.

El *software* requerido para reasignar usuarios de un servidor a otro con interrupción mínima es muy complejo, ya que en el proceso de recuperación se debe preservar la dirección de red para acceso a los usuarios, la información de registro así como el estado y configuración de la máquina: archivos abiertos, aplicaciones, etc. Adicionalmente, algunas soluciones de agrupación de servidores pueden incluir características de balanceo de carga que además de incrementar los niveles de disponibilidad también mejoran los tiempos de respuesta.

Pequeñas agrupaciones de servidores pueden ser implementadas a través de conexiones SCSI para compartir la información pero generalmente están limitados a dos servidores, lo cual es una restricción importante. El canal de fibra, por otro lado, permite que se puedan tener configuraciones del orden de decenas de servidores compartiendo decenas de terabytes de información.

Ya que el objetivo principal de una agrupación de servidores es mantener los más altos niveles de disponibilidad, el tener una agrupación de servidores en un ambiente de ISN requiere del uso de múltiples conexiones a los arreglos de discos y de un componente de *software* en cada uno de los servidores que monitoree el estado de los componentes de *hardware* y de las aplicaciones proporcionando información de cualquier condición de error a los demás servidores dentro de la agrupación. Esta información de monitoreo es generalmente enviada a través de una conexión LAN que también es redundante.

En la figura 8-4 está representada una configuración de agrupación con seis servidores compartiendo dos arreglos de discos. Como se puede observar se tienen dos rutas de acceso a los discos a través de dos concentradores, mientras que la señal de estado se encuentra configurada de manera redundante a través de conexiones *Ethernet*. Esta configuración puede ser escalada agregando servidores o arreglos adicionales usando los puertos disponibles o agregando más concentradores en cascada.

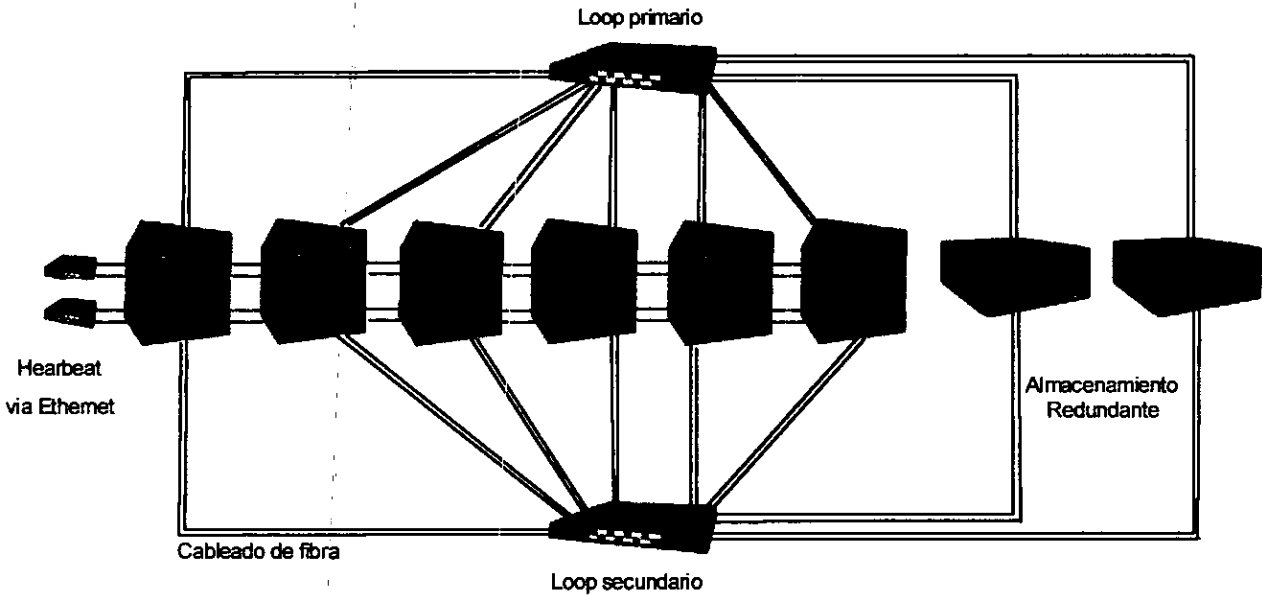


Figura 8-4. Agrupación de servidores usando ciclo arbitrado

8.1.3 Redes de Almacenamiento Remotas

La administración de un ambiente de cómputo en ocasiones puede presentar varias contradicciones. Por ejemplo, el descentralizar los servidores y los sistemas de discos proporciona la ventaja de mejorar las velocidades de transmisión de la información y del tiempo de acceso de los usuarios pero el trabajo de administrar múltiples sitios es mayor comparado con el de administrar uno solo. Por otro lado, el centralizar los servidores y la infraestructura de cómputo en general reduce los requerimientos de administración y permite la consolidación de recursos pero restringe el ancho de banda disponible para los usuarios remotos.

La tecnología de canal de fibra proporciona el medio para poder tener un ambiente centralizado y distribuido que se adapte a las necesidades de consolidación de recursos pero que también permita buenos tiempos de respuesta a los usuarios que están accediendo la información de manera remota. Usando láseres de onda larga y dispositivos para extender las señales junto con cableado monomodo, múltiples sitios en una red de área metropolitana pueden ser unidos en una red de almacenamiento inteligente extendida.

Como se puede apreciar en la figura 8-5, cada edificio en una corporación puede tener una ISN local basada en ciclo arbitreado o fábrica proporcionando acceso de alta velocidad a todos los usuarios de ese sitio. Por otro lado, al unir cada uno de estos puntos al sitio principal los usuarios en cada centro de datos remoto siguen teniendo la misma velocidad de acceso del centro de datos principal.

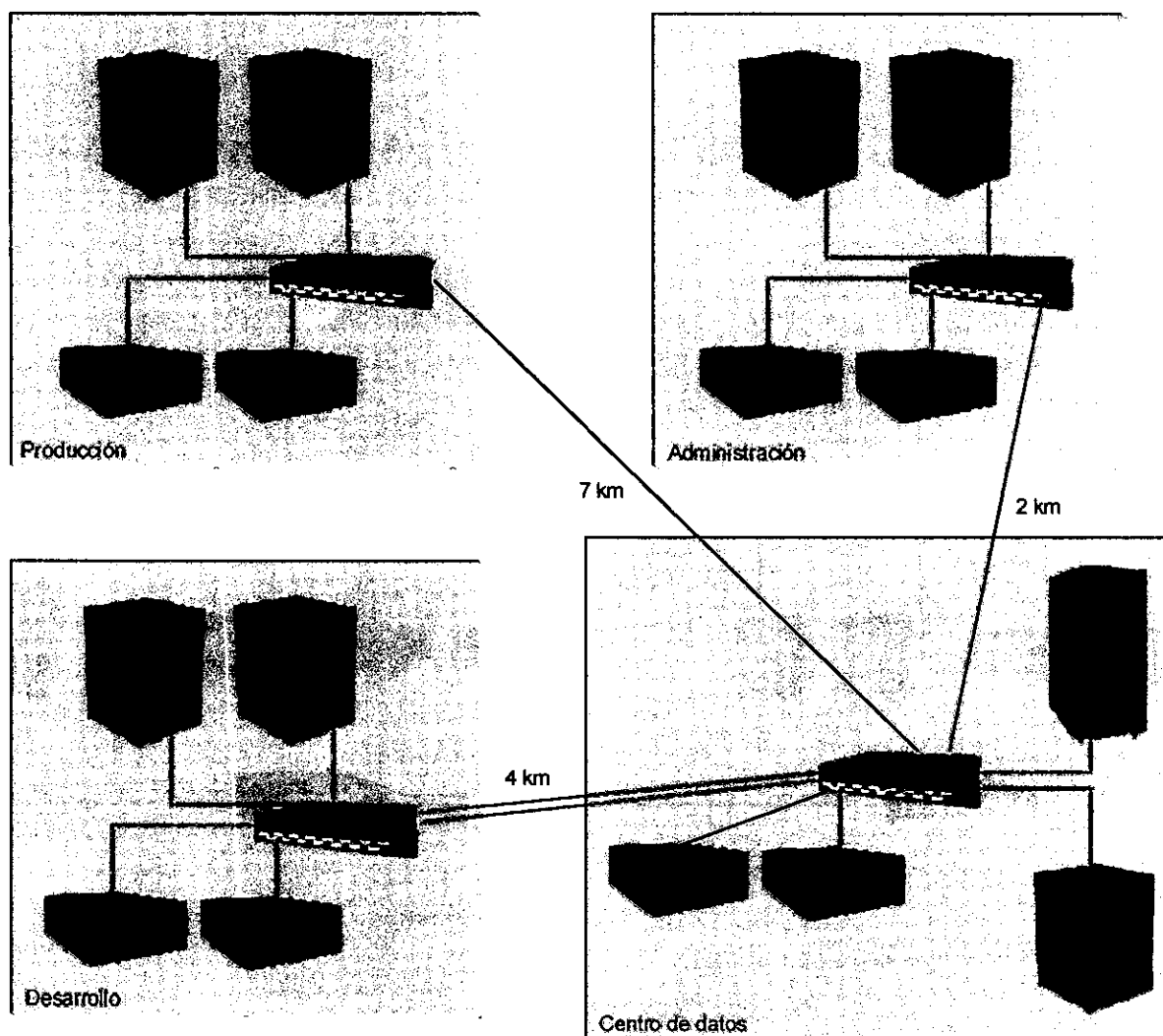


Figura 8-5. Red de almacenamiento inteligente metropolitana

En este ejemplo, el edificio de desarrollo tiene dos conexiones hacia el centro de datos para el acceso a archivos grandes, esta doble conexión permite que haya un balanceo de cargas en el acceso a la información. Este esquema de ISN extendida resuelve muchos problemas de integridad en los datos a través del respaldo y la compartición de recursos de almacenamiento centralizados desde múltiples sitios remotos pero para lograr esto se requiere de *software* para el control de la asignación de volúmenes y de métodos de bloqueo de archivos en caso de que los datos sean compartidos entre varios servidores. De la misma manera, el acceso a una biblioteca de cintas compartida requiere de un proceso de calendarización y de asignación dinámica del recurso a través de zonificación y de *software* de respaldo.

8.1.4 Replicación de datos

La replicación remota de la información es uno de los métodos más utilizados como parte de un esquema de recuperación de desastres. La figura 8-6 ilustra una solución de replicación de datos en donde se utiliza fibra monomodo de onda larga para la conexión entre el sitio de producción y el sitio secundario. Como se puede apreciar estas conexiones son redundantes.

Otra característica importante en esta configuración es que para evitar retrasos en la propagación de la información se utilizan concentradores inteligentes en lugar de concentradores simples. En este ejemplo, la aplicación principal en el sitio primario es una base de datos relacional. Para mantener actualizada la información en el sitio remoto, solo los registros que presentan cambios son transmitidos lo que reduce considerablemente el uso del ancho de banda disponible. Periódicamente también se pueden realizar respaldos a partir de los discos en el sitio remoto lo que asegura la integridad de la información sin agregar carga adicional a la ISN.

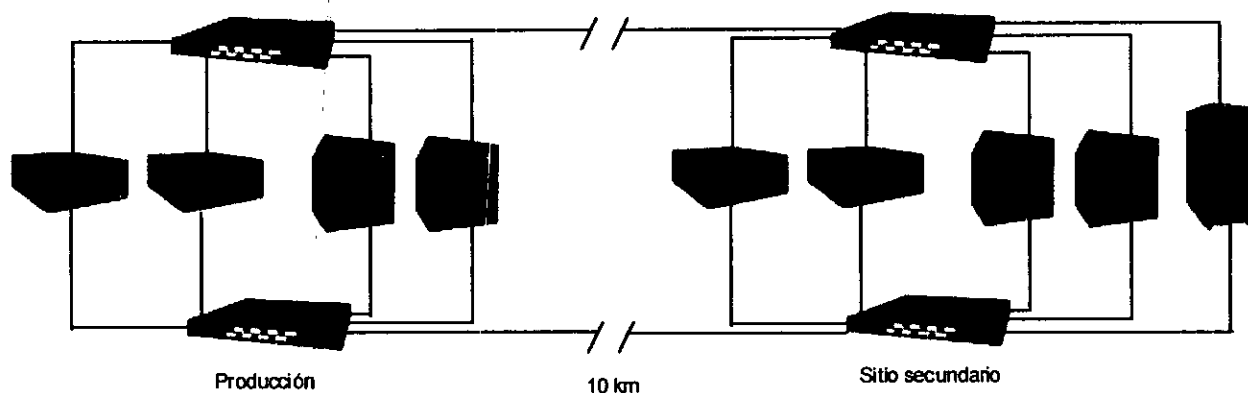


Figura 8-6. Esquema de replicación remota.

8.2 Ejemplos de Implementación

A continuación se explican tres ejemplos prácticos de implementación de redes de almacenamiento inteligente:

- Consolidación de almacenamiento
- Respaldo remoto
- Copia remota

8.2.1 Consolidación de almacenamiento

La consolidación de almacenamiento en un ambiente de ISN resuelve el problema de utilización de espacio en disco así como los problemas asociados con la administración ya sea en ambientes de múltiples centros de datos o de redes de almacenamiento metropolitanas.

Se ha comprobado que la utilización en disco se mejora del 10 al 40% usando consolidación de almacenamiento. Una ISN proporciona una plataforma que permite que los arreglos de discos y servidores puedan ser agregados con mínima interrupción de la operación.

En pocas palabras la consolidación de almacenamiento proporciona gran conectividad, mejor utilización de recursos, mayor escalabilidad y administración simplificada.

Configuración del Sistema

En este ejemplo de implementación se tienen siete sitios remotos localizados a diferentes distancias del sitio en donde se tienen los arreglos de discos. Los servidores ubicados en cada sitio son conectados al centro de datos a través de DWDM's.

En esta configuración, cada sitio está conformado por una fábrica que consta de un solo concentrador Brocade 2400 conectado a un equipo Sun Solaris o Microsoft Windows NT. Como se puede apreciar en la figura 8-7 algunos servidores tienen conectado un JBOD de canal de fibra. En el centro de datos se utilizaron 2 concentradores 2800 y como servidor de respaldo se tiene un servidor Windows NT.

Con respecto a los arreglos de discos se tiene un ADTX y un Hitachi 9900, para la parte de respaldos se tiene conectado a la ISN del centro de datos una biblioteca de cintas ATL P1000 con 4 unidades de cinta DLT 7000.

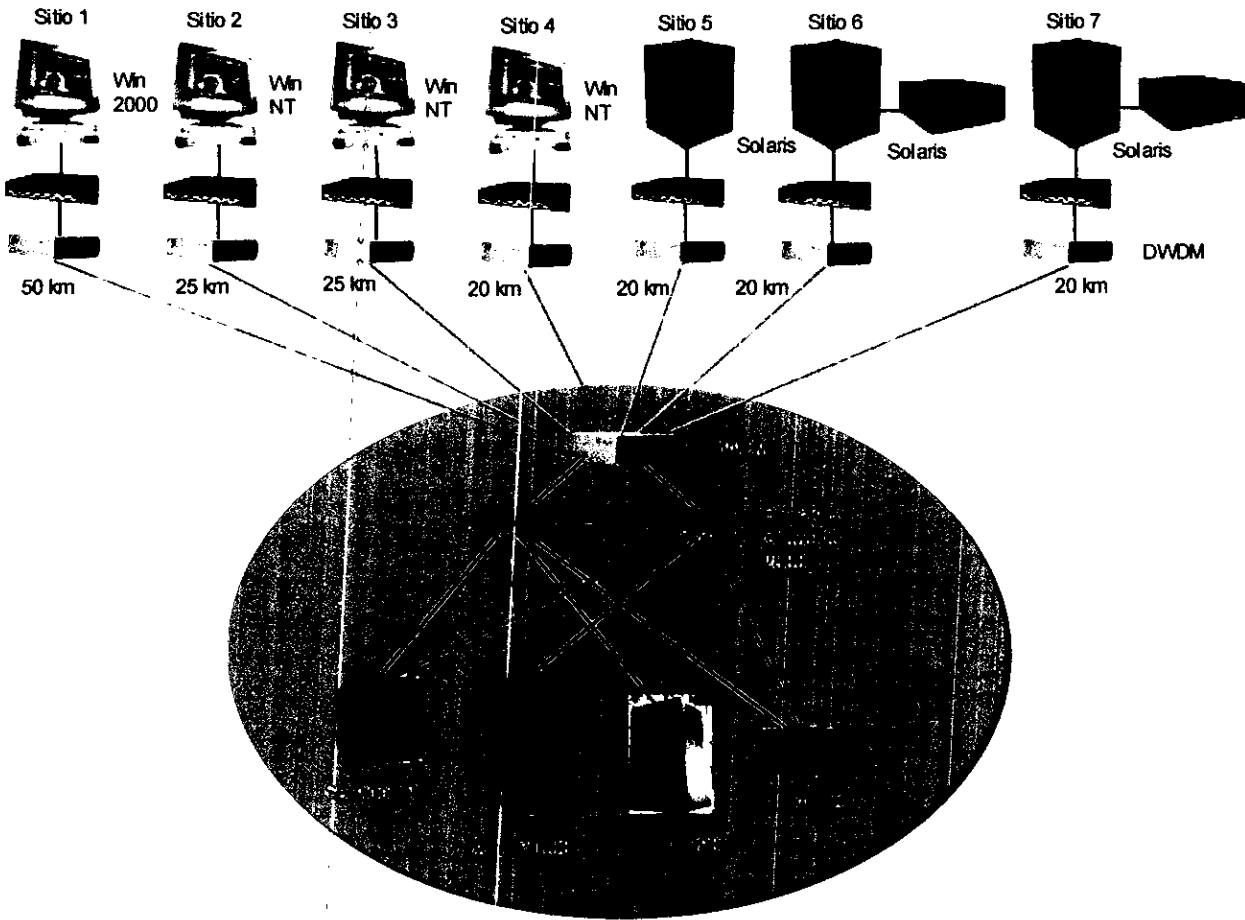


Figura 8-7. Esquema de una solución de consolidación de almacenamiento

La tabla 8-1 muestra con detalle los componentes de *hardware* y *software* utilizados en la configuración.

DETALLE DE LOS COMPONENTES USADOS EN LA CONFIGURACION

SERVIDORES

Sun E3500	Solaris 8	Sin parches
Tarjeta SBUS JNI FC64-1063	v2.5	
Sun Ultra 30	Solaris 8	
Tarjeta PCI Emulex LP8000	V4.1.0g	
Servidor IBM Netfinity 4500R	Windows NT, 2000 Enterprise Edition	
Emulex LP8000	3.80a3	

SOFTWARE

Veritas	Netbackup Datacenter 3.4 Volume Manager 3.1 File System 3.4	Sin parches
---------	---	-------------

SWITCH

Brocade	Silk Worm 2400, 2800	Firmware 2.3
---------	----------------------	--------------

ARREGLOS DE DISCOS

ADTX	ArrayMaStor RAID Array	
Hitachi Data Systems	Freedom Storage 9900	
Xyratex FC JBOD	12 X Seagate ST318304FC	Firmware 005

BIBLIOTECA DE CINTAS

ATL P1000	4 DLT 7000s	
-----------	-------------	--

INFRAESTRUCTRA DE DWDM

LuxN WavSystem	Wav Station DWDM	
DWDM	WavFarer DWDM	

Tabla 8-1. Detalle de la configuración utilizada en el esquema de consolidación de almacenamiento remoto.

Configuración de los concentradores inteligentes

A los concentradores inteligentes se les debe asignar una dirección IP única para que puedan ser configurados, administrados y monitoreados. Para lograr esto por primera vez se utiliza el puerto serial. Una vez hecho esto, la configuración se puede continuar a través de una sesión de telnet desde una PC o desde un sistema Unix.

Adicionalmente en el caso de los concentradores Brocade se puede utilizar WEB TOOLS para realizar la administración vía un navegador.

Configuración de la ISN

Todas las conexiones de larga distancia en la fábrica deben ser configuradas como conexiones de fábrica extendidas para lograr el óptimo desempeño de I/O. La tabla 8-2 muestra el nivel de fábrica extendido que debe ser configurado de acuerdo a la distancia.

Distancia	Nivel de Fábrica Extendido
0 - 10	Nivel 0
11 - 50	Nivel 1
51 - 100	Nivel 2

Tabla 8-2. Niveles de fábrica extendidos

Para configurar un concentrador en modo extendido se deben seguir los siguientes pasos:

- Desabilitar el concentrador
- Correr la utilería “*configure*”
- Poner el parámetro de *Long Distance Fabric* a 1
- Asegurarse de que cada concentrador tiene asignado un dominio único
- Configurar cada puerto que lo requiera en modo extendido

Configuración de los DWDM's

Se tiene un DWDM en cada uno de los sitios remotos, los DWDM's multiplexan dos canales, uno para Canal de Fibra y el otro para Gigabit Ethernet, para propósitos de la ISN solo se utiliza el canal de fibra. En el centro de datos se tiene otro DWDM que actúa como una compuerta a y desde el repositorio de disco. La conexión en cada sitio se lleva a cabo con cables multimodo los cuales se conectan a los concentradores y de ahí al DWDM usando la capacidad extendida de conexión de los concentradores Brocade, de ahí cada uno de los siete DWDMs se conecta al DWDM del centro de datos.

8.2.2 Respaldo remoto

Las redes de almacenamiento inteligente proporcionan una infraestructura que permite nuevas soluciones de recuperación de desastres protegiendo la información a larga distancia. Una de estas soluciones la constituyen los respaldos remotos. Esta opción permite que en caso de una contingencia en el sitio principal se pueda seguir prestando servicio a partir de la información respaldada en los sitios remotos. Esta solución reduce los costos de operación asociados con la identificación de las cintas de respaldo, inventario y envío de la información respaldada de un sitio a otro.

Configuración del sistema

En este ejemplo de implementación se tiene una solución de respaldo remota sobre una ISN metropolitana usando Veritas Net backup. El servidor maestro de netbackup es instalado en el sitio principal y es aquí a donde se lleva a cabo la consolidación de los archivos de índices de todos los clientes de los que se respalda su información. En la configuración se utiliza un servidor IBM Netfinity 4500R corriendo Microsoft Windows 2000 que funge como el servidor maestro de respaldos y que puede ser incluso un sistema más pequeño ya que el verdadero procesamiento se lleva a cabo en los servidores esclavos instalados en los sitios remotos. Uno de los sitios utiliza un servidor Sun E3500 corriendo Solaris 8, mientras que el otro utiliza también un equipo IBM Netfinity corriendo Windows NT. Como se puede ver esta mezcla de plataformas prueba que ambientes heterogéneos pueden coexistir perfectamente en una solución de ISN extendida. La figura 8-8 muestra el esquema de configuración.

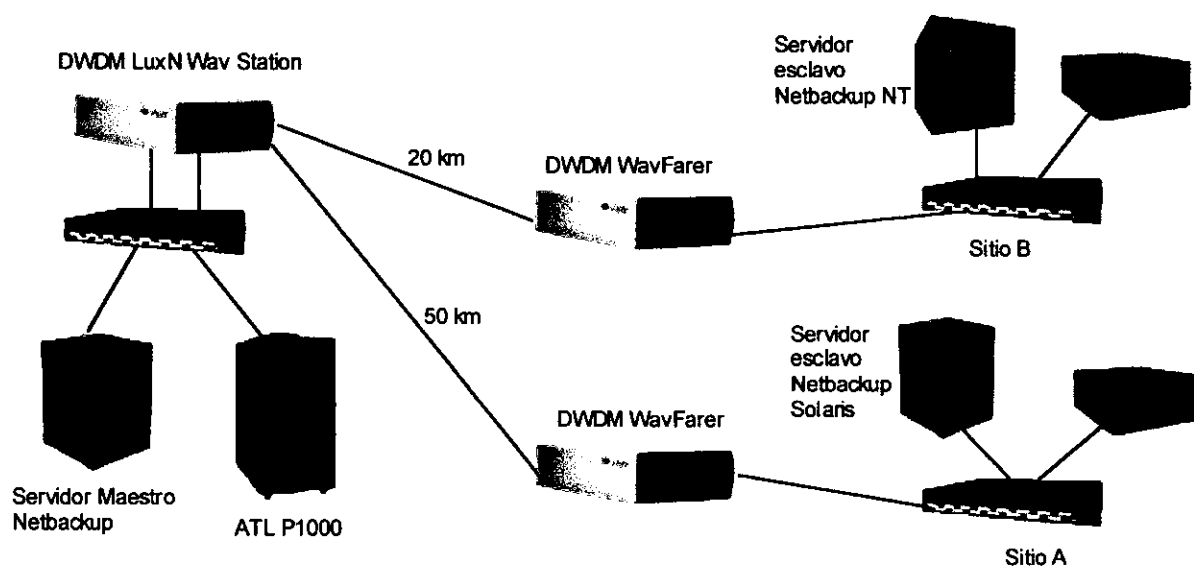


Figura 8-8. Esquema de una solución de respaldo remoto

Para asegurarse de que la biblioteca de cintas será accesada solo por los servidores y clientes designados se utiliza una configuración de zonificación. Cada servidor esclavo en el sitio remoto es configurado dentro de una zona de respaldo, incluyendo a la biblioteca de cintas. Esta característica proporciona seguridad adicional a los datos almacenados en las cintas ya que ningún otro servidor en la fábrica puede tener acceso a estos datos. El *software* de cliente de netbackup puede ser instalado en múltiples clientes en los diferentes sitios sin que estos se encuentren en el ambiente de ISN. Estos pueden acceder sus propios datos y enviarlos al servidor esclavo para finalmente ser respaldos en la biblioteca de cintas remota.

Ya que múltiples servidores esclavos accedan a una misma biblioteca de cintas, es necesario utilizar una opción adicional del *software* de Netbackup conocida como Opción de Almacenamiento Compartido (*Shared Storage Option*). Esta característica permite que múltiples servidores esclavos puedan acceder una biblioteca de cintas compartida, en la configuración se utiliza una biblioteca de cintas ATL P1000 con cuatro unidades de cinta DLT 7000. Los respaldos para cada uno de los clientes en los sitios remotos se pueden realizar de manera concurrente siempre y cuando haya una unidad de cinta disponible para esto.

8.2.3 Copia Remota

Otra solución interesante que pueden brindar las redes de almacenamiento inteligente es la copia síncrona de datos. Esto permite que múltiples clientes en un sitio remoto puedan a través de una ISN espejear datos críticos a un sitio alterno. La información se mantiene en el otro sitio para que en caso de que haya una contingencia los datos puedan ser accedidos a partir del sitio remoto, una vez que los discos que han fallado son reemplazados, los sistemas de archivos pueden ser resincronizados al sitio principal. Este esquema proporciona una solución que ofrece redundancia en la información e incrementa el tiempo de servicio de las aplicaciones.

Configuración del sistema

En este ejemplo de implementación se utiliza Veritas Volume Manager y Veritas File System corriendo en Solaris. En la configuración se utilizó un servidor Sun E3500 en un sitio y un servidor Ultra 30 en el otro con arreglos de discos JBOD, asignando seis discos a cada servidor y doce discos en el sitio principal. Para evitar que algún servidor accese información que no le corresponde se utilizó una configuración de zonificación la cual contiene un servidor Solaris con su disco local y un JBOD en el sitio remoto, de esta manera se puede copiar la información del disco local del sitio B al mirror del centro de datos B sin interferir la copia del disco local del sitio A al mirror del centro de datos B, tal como se muestra en la figura 8-9.

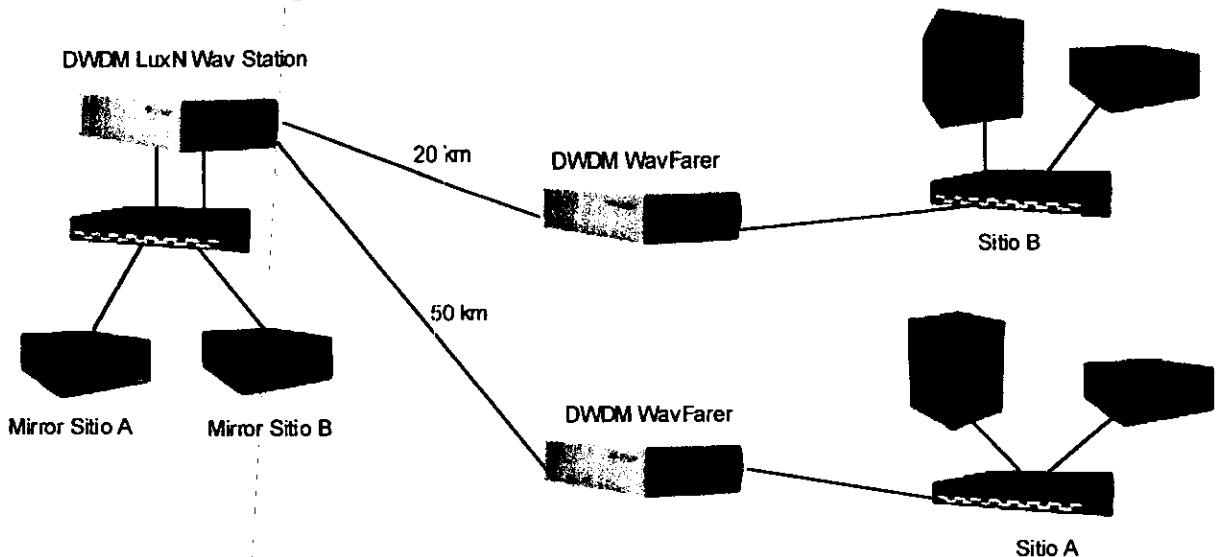


Figura 8-9. Esquema de configuración de copia remota

Conclusiones

Después de varios años de desarrollo, la tecnología de las redes de almacenamiento inteligente ha pasado de ser una promesa a una realidad. Sin embargo, la complejidad inherente a esta tecnología y sobre todo la falta de conocimiento que existe alrededor de ella implica que se debe ser muy cuidadoso desde el momento en que se está considerando una ISN como una posible solución de almacenamiento, así como en el proceso de configuración, implementación y mantenimiento.

Hasta la fecha, la mayor parte de las implementaciones de ISN han sido confinadas a resolver problemas muy específicos, esto de ninguna manera es malo ya que por un lado permite que se pueda ir ganando experiencia y después mejorar la infraestructura de la red de almacenamiento, pero la realidad es que ahora se cuenta con la tecnología y conocimiento para implementar soluciones relativamente más complejas, pero en donde el proceso de implementación puede ser largo y riesgoso, lo cual debe ser analizado con mucho cuidado y en donde se debe evaluar la relación costo/beneficio para determinar la factibilidad de implementación.

Lo más importante antes de analizar una solución basada en ISN es decidir claramente cuales son los problemas que se piensan resolver y los beneficios esperados, tanto en términos informáticos como de negocio. No en todos los casos una solución basada en ISN será la respuesta a lo que se está buscando, también es importante considerar que si realmente no existen problemas que solucionar entonces justificar una ISN puede ser muy difícil. Adicionalmente, no se deben pasar por alto las alternativas de solución que actualmente existen en el mercado y que pueden ofrecer en algunos casos un mejor costo beneficio que las ISNs como es el caso de las soluciones de almacenamiento conectado directamente a la red o incluso de conexión directa al servidor. De hecho, en algunos casos la mejor solución de almacenamiento puede ser una mezcla de las tres.

En el caso de que un centro de datos se beneficie de varias de las características de una ISN se debe considerar cuidadosamente si estas funciones deben ser configuradas en una o varias redes de almacenamiento. Esto da paso al tipo de topología que se debe usar, que es probablemente el aspecto más importante a tomar en cuenta en el diseño de una ISN por cuestiones de funcionalidad, desempeño y de crecimiento futuro.

En general las topologías complejas pueden presentar problemas en administración y desarrollo y una alternativa es considerar el segmentar o tener varias redes de almacenamiento pequeñas que después una vez que se haya ganado experiencia se pueden expandir de manera independiente o unirse en una sola ISN. Sin embargo, el tener una sola ISN puede ser riesgoso si no se considera un esquema de alta disponibilidad a nivel de múltiples fábricas.

Es fundamental entender las limitantes y beneficios reales de las ISNs y de las soluciones de almacenamiento en general. Prácticamente todos los proveedores ofrecen las mismas funcionalidades y características. Lamentablemente también algunos de ellos pueden exagerar en las capacidades de sus productos con tal de no perder terreno en el mercado. Por lo que es importante hacer un estudio cuidadoso de las soluciones disponibles y trabajar alrededor de estas consideraciones conforme la tecnología continúe desarrollándose.

Finalmente, después de haber hecho este trabajo de tesis, se puede concluir que las redes de almacenamiento inteligente constituyen una alternativa de solución real para muchos de los retos de manejo de datos que enfrentan las corporaciones de hoy en día ofreciendo grandes beneficios comparados con los que ofrecen las otras soluciones de almacenamiento existentes en el mercado.

Glosario

Adaptador de Interface al Medio, *Media Interface Adapter (MIA)*

Adaptador utilizado para cambiar la interfase de conexión de cobre a fibra en sistemas de almacenamiento y servidores.

Administración de Recursos de Almacenamiento, *Storage Resource Management (SRM)*

Es el conjunto de programas encaminado a administrar los recursos de hardware y software que integran una red de almacenamiento inteligente.

Almacenamiento Conectado Directamente al Servidor, *Direct Attached Storage (DAS)*

Método de conectar un arreglo de discos directamente al servidor, típicamente a través de un canal SCSI.

Almacenamiento Conectado a la Red, *Network Attached Storage (NAS)*

Método de conectar un arreglo de discos a la red, este método utiliza protocolos de compartición de archivos por red para que los equipos puedan acceder la información.

Almacenamiento de Datos Digitales, *Digital Data Storage (DDS)*

Formato de grabación usado por las unidades DAT

Ancho de banda

Cantidad de datos que puede ser transmitida en un momento dado a través de un canal de comunicaciones.

Ancho de columna, *Stripe Width*

Es el número de discos físicos usados en un *stripe* de RAID sin considerar la información de paridad.

Arquitectura de Almacenamiento Serial, *Serial Storage Architecture (SSA)*

Método de alta velocidad para conectar discos, cintas, unidades de CD-ROM, impresoras y otros dispositivos a un sistema de cómputo.

Arquitectura Estándar de la Industria, *Industry Standard Architecture (ISA)*

Canal estándar de 16 bits para el sistema PC/AT de IBM. ISA fue el único canal estándar para PC's hasta la aparición de EISA y de la arquitectura de Microcanal.

Arquitectura Estándar de la Industria Mejorada, *Enhanced ISA (EISA)*

Estándar para canal de expansión diseñado para competir con la Arquitectura de Microcanal de IBM.

Arquitectura de Microcanal, *Microchannel Architecture (MCA)*

Canal de expansión de 32 bits diseñado por IBM e introducido con sus equipos PS/2 de alto rendimiento. Este canal no es compatible con otras arquitecturas de canal de I/O.

Arreglo Redundante de Discos Independientes, *Redundant Array of Independent Disks (RAID)*

Método que consiste en combinar varios discos duros en una unidad lógica de almacenamiento que ofrece tolerancia a fallas y mejor desempeño que un solo disco duro.

Asociación de Estándares Electrónicos para Video, *Video Electronics Standard Association (VESA)*

Canal de I/O diseñado para trabajar con computadoras personales basadas en el procesador Intel 486

ATA

Define las señales y el protocolo para los periféricos IDE.

Autocargador, *Autoloader*

Dispositivo de respaldo de cinta con una sola unidad que cuenta con varios cartuchos. Está diseñado para automatizar el proceso de respaldo a través del uso de un brazo mecánico.

Base de Información de Administración, *Management Information Base (MIB)*

Base de datos mantenida por una estación de trabajo y por agentes en un esquema de administración basado en SNMP. Es de este repositorio de donde se tiene la información que se necesita de los dispositivos en un momento dado.

Biblioteca de Cintas Automatizada, *Automated Tape Library (ATL)*

Sistema de respaldos con varias unidades de cintas y varios cartuchos usado para esquemas de respaldo automatizados en donde se necesitan respaldar grandes volúmenes de información. Estos sistemas cuentan con varias características avanzadas como lectores de códigos de barras.

Campus cluster

Es una agrupación entre equipos que se encuentran en sitios remotos.

Canal, *Bus*

Conjunto de conductores que forman una ruta de interconexión entre el procesador o procesadores de un sistema y sus subsistemas periféricos. Un canal (*bus*) típicamente lleva datos, energía eléctrica y otras señales relacionadas.

Canal de fibra

Interface de alta velocidad que ofrece gran desempeño y capacidad para redes y sistemas de almacenamiento, sus características permiten tener un alto grado de flexibilidad y escalabilidad a la vez que se simplifica el cableado. Se puede pensar en el canal de fibra como un vehículo de transporte.



Chip

Circuito integrado hecho de silicón o de algún otro material semiconductor. Un *chip* puede ser un dispositivo de memoria, un microprocesador o un dispositivo lógico digital.

Ciclo, *loop*

Es una estructura de interconexión en la que cada punto tiene conexiones a dos puntos adyacentes. En una topología de ciclo el ancho de banda disponible es compartido.

Ciclo arbitreado, *loop arbitreado*

Tecnología de interconexión de canal de fibra que permite hasta 126 nodos y un nodo activo de comunicación.

Cinta de Audio Digital, *Digital Audio Tape (DAT)*

Tecnología de almacenamiento que usa una cinta de 4 mm para almacenar datos. DAT es similar a una cinta de audio pero en lugar de almacenar los datos linealmente a lo largo de la cinta, la información es guardada con un ángulo de inclinación.

Cinta Lineal Abierta, *Linear Tape Open (LTO)*

Formato de cinta desarrollado por HP, IBM y Seagate Technology.

Cinta Lineal Digital, *Digital Linear Tape (DLT)*

Tecnología de grabación a cinta usada en soluciones de respaldo en donde se manejan grandes volúmenes de información

Cinta magnética

Dispositivo de memoria de acceso secuencial cuyo medio magnético es una cinta en un cassette.

Cliente/Servidor

Ambiente de cómputo en donde ciertos equipos llamados servidores proporcionan servicios o aplicaciones a otros equipos conocidos como clientes los cuales, para operar apropiadamente necesitan de estos servicios.

Cluster

Método de configurar múltiples computadoras conectándolas entre sí para que puedan operar como una sola entidad. Esta característica ofrece grandes beneficios en términos de desempeño y disponibilidad.

Colisión

Problema de comunicación provocado por transmisiones simultáneas de dos o más nodos en una red.

Columna, *Stripe*

Es el bloque de datos más pequeño que puede ser leído o escrito a un disco físico antes de moverse al siguiente disco en una configuración de RAID 0.

Compuerta, *Gateway*

Dispositivo que conecta dos redes de computadora que usan diferentes protocolos.

Concentrador Inteligente, *Switch*

Dispositivo de monitoreo de tráfico que controla la comunicación entre múltiples nodos en una red.

Convertidor de Interfase al Medio, *Media Interface Converter (MIC)*

Convertidor para cables de fibra óptica usado en redes FDDI

Conexión de Sistemas Empresariales, *Enterprise System Connection (ESCON)*

Protocolo de conexión desarrollado por IBM para conectar dispositivos periféricos a mainframes.

Control de Acceso al Medio, *Media Access Control (MAC)*

Dirección física de un dispositivo conectado a una red, la cual es expresada con un número hexadecimal de 48 bits.

Convertidor de Interfase Gigabit, *Gigabit Interface Converter (GBIC)*

Es el transceptor de mayor uso en la actualidad para la conversión de señales eléctricas a ópticas y viceversa.

Disco Compacto, *Compact Disk (CD-ROM)*

Dispositivo de almacenamiento de solo lectura con capacidad de hasta 600 MB, el cual utiliza un disco láser intercambiable similar a un disco compacto de audio.

Disco duro

Dispositivo electromagnético que usa varios discos rígidos junto con cabezas de grabación y que es usado como un medio de almacenamiento primario. En la actualidad es posible encontrar discos de hasta 73 GB de capacidad y 15000 rpm.

Disco magneto óptico

Medio de almacenamiento que utiliza una combinación de magneto y láser para alterar las direcciones de flujo magnético en una superficie de grabación de un disco y de esta manera almacenar datos en él.



Disco de Video Digital, *Digital Video Disk (DVD)*

Disco parecido a un disco compacto estándar pero con capacidad de almacenar hasta siete veces más información. Actualmente un DVD puede almacenar hasta 4.7 GB lo que lo hace atractivo como un medio de almacenamiento de películas y animaciones.

Dispositivo crudo

Forma de almacenar la información en donde las referencias a los datos se hacen a través de dispositivos de bloque.

Electrónica de Unidad Integrada, *Integrated Drive Electronics (IDE)*

Interfase de unidad de disco usada para computadoras personales.

Enmascaramiento de LUN's, *LUN masking*

Es una técnica utilizada para especificar como los dispositivos estarán disponibles a determinados servidores.

Entrada/Salida, *Input/Output (I/O)*

Uno de los componentes principales de la arquitectura de un sistema de cómputo, constituye el enlace entre el microprocesador y los componentes que lo rodean.

Espacio de Memoria, *Buffer*

Area de almacenamiento temporal usada para el procesamiento de información.

Ethernet

Estándar de comunicación muy común para redes de área local. Es el estándar de comunicación más usado en computadoras y maneja velocidades de 10, 100 y 1000 Megabits por segundo.

Fábrica

Uno o más switches de canal de fibra en una topología de red.

Failover

Transferencia de la operación de un componente con falla a otro similar y cuya finalidad es asegurar la continuidad de la operación.

FDDI

Estándar para transmitir datos en cables de fibra óptica a una velocidad de 100 Mbps.

Firewall

Esquema de seguridad utilizado para prevenir que usuarios externos puedan tener acceso a información de una empresa a través del Internet o a través de otra red.



Firmware

Instrucciones permanentes y datos programados directamente en un circuito integrado para controlar la operación de una computadora.

Full duplex

Método de comunicación en la que el envío y recepción de datos entre dos dispositivos se puede llevar a cabo al mismo tiempo.

Gabinete, Rack

Nombre con el que también se le conoce a los gabinetes usados para montar arreglos de discos y servidores. El ancho estándar es de 19".

Gigabit

Unidad que representa 1000 Megabits

Grupo de Expertos de Fotografía Unidos, Joint Photographic Experts Group (JPEG)

Método para almacenar fotografías de color de alta calidad en un formato de bitmap.

Hacker

Experto en computación que se dedica a violar sistemas de seguridad.

Heartbeat

Mecanismo utilizado en ambientes de agrupación de servidores (*cluster*) para determinar el estado de los servidores y activar la función de *failover* en caso de que alguno de los servidores tengan algún problema.

Hot spare

Unidad de disco duro que tiene como función tomar el lugar de un disco con falla en una configuración de RAID.

Hot swap

Capacidad de que un componente pueda ser quitado, reemplazado y configurado sin necesidad de apagar el sistema al que pertenece.

Hub

Dispositivo que divide una conexión de red en varias conexiones, cada una conectada a un equipo de cómputo o sistema de almacenamiento de fibra diferente.



Interconexión de Componentes Periféricos, *Peripheral Component Interconnect (PCI)*

Canal estándar de la industria usado en servidores, estaciones de trabajo y PC's

Interfase Paralela de Alto Desempeño, *High Performance Parallel Interface (HIPPI)*

Canal de comunicaciones de alta velocidad usado para cubrir distancias cortas.

Instituto de Ingenieros Eléctricos y Electrónicos, *Institute of Electrical and Electronic Engineers (IEEE)*

Organización de ingenieros, científicos y estudiantes destinada a definir estándares para sistemas de cómputo y equipos de comunicaciones.

IOPS

Número de señales de computadora de recepción y transmisión que pueden ser manejados a lo largo de la ruta de conexión entre el canal del procesador y las unidades de disco.

Java

Lenguaje de programación orientado a objetos desarrollado por Sun Microsystems y cuya principal característica es su independencia de plataforma.

JBOD

Arreglo de discos que no tiene ninguna inteligencia incorporada para el manejo de las operaciones de I/O o de RAID y que es visto por el servidor solo como un conjunto de discos.

Jiro

Es una tecnología desarrollada por Sun Microsystems que tiene como objetivo el tener un solo ambiente de administración inteligente de recursos de cómputo.

Jukebox

También conocido como biblioteca de discos ópticos, es un gabinete que contiene múltiples unidades de discos ópticos y cartuchos. Usado para almacenar grandes volúmenes de información, principalmente imágenes y videos.

Mainframe

Computadora con gran capacidad de procesamiento encontrada en centros de procesamiento de datos.

Manejador, *Driver*

Programa o conjunto de programas usado para operar un periférico.

Memoria de Acceso Aleatorio, *Random Access Memory (RAM)*

Memoria de circuito integrado que permite el almacenamiento de información y su recuperación por un controlador o microprocesador. La información puede ser almacenada o accesada en cualquier orden y todas las localidades de almacenamiento son igualmente accesibles.

Memoria virtual

Método para extender el tamaño aparente de la memoria de acceso aleatorio de un sistema de cómputo mediante el uso de uno o varios discos duros.

Metal Evaporado Avanzado, *Advanced Metal Evaporated (AME)*

Es un método de elaboración de cintas que consiste en evaporar el material magnético utilizado en su elaboración, creando una capa magnética libre de lubricantes o impurezas.

Micron

Millonésima de metro, aproximadamente 0.0000394 pulgadas

Mid range

Este término se utiliza para hacer referencia a equipos de cómputo con capacidad de procesamiento de rango medio en donde se pueden manejar hasta centenares de usuarios.

Modo de Transferencia Asíncrona, *Asynchronous Transfer Mode (ATM)*

Método rápido de switcheo de paquetes de información que utiliza un paquete de tamaño fijo llamado celda. Esta técnica hace posible la transmisión de datos a alta velocidad, dentro de los tipos de datos que se benefician con esta tecnología se encuentran: video, audio y videoconferencia

Módulo de Conexión de Gigabaud, *Gigabaud Link Module (GLM)*

Transceptor utilizado para convertir señales eléctricas a ópticas, usado tanto para sistemas de almacenamiento como para servidores.

Módulo *Plug-in*

Es un pedazo de software que se agrega a una aplicación para darle mayor funcionalidad.

Multimedia

Combinación de varios formatos usados para la entrega de información. Estos formatos pueden ser texto, fotos, audio, animación y video.

Opción de Almacenamiento Compartido, *Shared Storage Option (SSO)*

Software de la compañía Veritas que permite que varios servidores puedan usar de manera dinámica una biblioteca de cintas en un ambiente de ISN.



Parche

Programa o conjunto de programas utilizado para corregir algún problema de *hardware* o *software* o para agregar funcionalidad a un sistema de cómputo.

Portadora

Tono continuo establecido por dos sistemas de comunicación.

Procesamiento de Transacciones en Línea, *Online Transaction Processing (OLTP)*

Aplicaciones orientadas a operaciones que se traducen en transacciones tales como agregar, añadir y eliminar registros.

Protocolo

Conjunto de convenciones que gobiernan el formato de los mensajes que se intercambian en un sistema de comunicaciones.

Protocolo de Administración de Datos sobre la Red, *Network Data Management Protocol (NDMP)*

Protocolo utilizado para transmitir información desde un arreglo de discos directamente a la biblioteca de cintas, sin pasar por los servidores.

Protocolo de Administración de Red Simple, *Simple Network Management Protocol (SNMP)*

Protocolo estándar que corre sobre IP para proporcionar administración de dispositivos sobre la red.

Protocolo de Configuración de Host Dinámico, *Dynamic Host Configuration Protocol (DHCP)*

Protocolo utilizado para asignar dinámicamente direcciones IP a equipos de cómputo conectados a una red IP.

Protocolo de Datagrama de Usuario, *User Datagram Protocol (UDP)*

Protocolo de comunicaciones para la capas de red, transporte y sesión de internet que hace posible enviar un mensaje de una computadora a una aplicación corriendo en otra computadora.

Protocolo de Internet, *Internet Protocol (IP)*

Estándar que describe la forma en que una computadora conectada al Internet debe descomponer los datos en paquetes para su transmisión en la red y la forma en la que deben dirigirse estos paquetes para que lleguen a su destino.

Protocolo de Transferencia de Hiper Texto, *Hyper Text Transfer Protocol (HTTP)*

Es el estándar de internet utilizado para soportar el intercambio de información.

Prueba de Desempeño, *Benchmark*

Prueba que se corre en un equipo de cómputo para medir su desempeño.

Puente, *Bridge*

Dispositivo que permite el intercambio de datos entre dos redes con diferentes protocolos y/o topologías.

Respaldo sin usar la LAN, *LAN free backup*

Técnica de respaldo en donde no se utiliza la LAN de producción para este proceso sino una red alterna como una ISN.

Respaldo sin usar el Servidor, *Server Free Backup*

Técnica de respaldo en donde se libera al servidor de realizar cualquier tipo de procesamiento que tenga que ver con la tarea de hacer respaldos.

Ruteador

Dispositivo electrónico que conecta dos o más redes y avanza los paquetes de entrada a la red apropiada.

S-Bus

Canal de I/O diseñado por Sun Microsystems para sus equipos SPARC. Originalmente est canal proporcionaba mayor desempeño que PCI pero nunca llegó a ser lo suficientemente popular para convertirse en un estándar y fue alcanzado por PCI en funcionalidad y desempeño.

SCSI Enclosure Services (SES)

Estándar utilizado para definir un conjunto de comandos con el propósito de obtener información del estado de un sistema de discos.

Sistema de Archivos


Método de almacenamiento de información en donde las referencias a los datos se hacen directamente hacia los caracteres.

Sistema de Interface para Computadoras Pequeñas, *Small Computer System Interface (SCSI)*

Interface estándar de la industria usada para conectar periféricos tales como impresoras, arreglos de discos y unidades de cinta a un equipo de cómputo. SCSI abarca estándares de *hardware* y *software* permitiendo a las computadoras y periféricos comunicarse entre sí.

Sistema de Planeación de Recursos Empresariales, *Enterprise Resource Planning (ERP)*

Consiste en la planeación y administración de todos los recursos de una empresa a través de una solución de cómputo.



Sistema de Soporte para la toma de Decisiones, *Datawarehousing*

Sistema que contiene una base de datos centralizada diseñada para mantener y manejar la información de una empresa por un período largo de tiempo. Estos datos son usados para hacer predicciones, para determinar las tendencias del negocio y para monitorear resultados.

Stream

Flujo de datos

Tarjeta Adaptadora, *Host Bus Adapter (HBA)*

Es una unidad de *hardware* que agrega unidades de I/O a un sistema de cómputo tal como son dispositivos de comunicación o periféricos.

Tarjeta madre, *Backplane*

Tarjeta principal que se utiliza para interconectar los diferentes componentes electrónicos de un arreglo de discos o servidor.

Tecnología de paso (*Autobypass*)

Método que consiste en ignorar un puerto o dispositivo con falla en un ciclo, de tal manera que la falla de un componente no afecte al ciclo completo.

Tiempo entre Fallas, *Mean Time Between Failures (MTBF)*

Tiempo promedio entre una falla y otra. Esta es una medida de fallas de componentes y no de un sistema de cómputo completo.

Token Ring

Red de área local en la que los equipos de cómputo están configurados en un anillo y un mensaje llamado *token* es pasado de estación a estación

Tramas (*Frames*)

Porción de datos transmitidos por un dispositivo de comunicación.

Transceptor, *Transceiver*

Dispositivo que se utiliza para cambiar señales eléctricas a otro tipo de señal así como para cambiar una interface de conexión.

Tunneling

Proceso que consiste en hacer traducción de protocolos



Unidad Central de Procesamiento, *Central Processing Unit (CPU)*

Término con el que también se le conoce al sistema principal de un equipo de cómputo conformado por el procesador, memoria y sistema de I/O.

Verificador de Redundancia Cíclica

Método para verificar la integridad de los datos dentro de un *frame* en una comunicación de dos dispositivos de fibra.

Zonificación

Es la agrupación de múltiples puertos de un *switch* para formar una red de almacenamiento privada virtual.

Bibliografía

Building Storage Networks

Marc Farley
Osborne / McGraw Hill, 2000

Storage Area Networks

Briang Wong
Prentice Hall, 2000

Backup and Restore Practices for Sun Enterprise Servers

Stan Stringfellow / Miroslav Klivansky / Michael Barto
Prentice Hall, 2000

Unix Backup and Recovery

W. Curtis Preston
O'Reilly, 2000

Configuration and Capacity Planning for Unix Servers

Briang Wong
Prentice Hall, 1998

Diccionario para Usuarios de Computadoras e Internet

Bryan Pfaffenberger
Prentice Hall, 1997

Página de Web de Brocade

<http://www.brocade.com>

Página de Web de Qlogic

<http://www.qlogic.com>

Página de Web de Jaycor

<http://www.jni.com>

Página de Web de Sun Microsystems

<http://www.sun.com>

Página de Web de IBM

<http://www.ibm.com>

Página de Web de HP

<http://www.hp.com>

Página de Web de Compaq

<http://www.compaq.com>

Página de Web de Brocade

<http://www.brocade.com>

Página de Web de EMC

<http://www.emc.com>

Página de Web de Vixel

<http://www.vixel.com>

Página de Web de Hitachi

<http://www.hitachi.com>