



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES ARAGÓN

“ALMACENAMIENTO DE VOLÚMENES DE DATOS
UTILIZANDO VERITAS VOLUME MANAGER PARA
SERVIDORES UNIX CON SISTEMA OPERATIVO
SOLARIS Y LA SOLUCIÓN DE
ALMACENAMIENTO PARA UNA EMPRESA DEL
SECTOR ELÉCTRICO”

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

INGENIERO EN COMPUTACIÓN

P R E S E N T A :

JOSÉ GABRIEL PÉREZ RODRÍGUEZ

ASESOR: HUGO PORTILLA VAZQUEZ

BOSQUES DE ARAGÓN ESTADO DE MÉXICO, 2009





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

A mi Padre Celestial por permitirme vivir la vida en este tiempo, por darme los talentos que tengo, por conocer los principios y formas para lograr ser un hombre recto y honesto, con todo mi amor para ti padre.

A mi señor Jesucristo por el amor que me tiene, por el evangelio que nos dio a la humanidad, y por estar a mi lado en todo momento de mi vida.

A mi familia, Carolina por ser la esposa ejemplar que me ha apoyado y aconsejado en todas las decisiones importantes de mi vida, por incitar a mi superación personal y profesional pero sobre todo gracias por tu amor. A mi hijos Gabriel y Leonardo, por ser el motor de mi vida, por sus talentos individuales que contagian energía y vitalidad.

A mi madre Margarita Rodríguez Guadalajara, por darme sus cuidados, consejos, y apoyo en todas las etapas de mi vida, a ella todos mis logros. A mi padre José Pérez Fernández (qepd), el tiempo que estuvimos juntos fue suficiente para enseñarme como debe ser un buen padre.

Con todo cariño a mis hermanos, Gerardo por el apoyo y amor que me ha dado y porque sus éxitos a base de trabajo y dedicación son un ejemplo para mí. A Erick por todo el apoyo brindado durante mi etapa de estudiante. A David por su apoyo incondicional en todo momento. A Norma por el amor que demuestra a mí y a mi familia. A todos ustedes y sus familias gracias totales.

A mis sobrinos Junior, Gerardo, Priscila, que este trabajo sea un estímulo para alcanzar las metas que se impongan.

A mi familia Esquivel. Sergio, Martha, Lourdes, Sheyla, con mi total agradecimiento a ustedes quienes son parte de este logro.

Gracias a mis amigos Pablo Cepeda e Israel Ortiz, por su compañía, apoyo y por compartir nuestros éxitos, éste es tan solo uno de ellos.

A mi Alma Máter, por ser el orgullo más grande de mi país. Es un honor mi formación profesional en esta institución. Con todo orgullo menciono que soy parte de la UNAM.

Finalmente quiero agradecer a todas las personas que me han acompañado a lo largo y ancho de mi carrera profesional, a los cercanos que están lejos y a los lejanos que están cerca.

José Gabriel Pérez Rodríguez.

Contenido

Introducción	1
Dispositivos de almacenamiento	1
I.1 Unidades de discos flexibles (floppy).....	1
I.2 Unidades Zip y Jaz.....	2
I.3 CD-ROM.....	2
I.4 DVD-ROM.....	3
I.5 Unidades de cinta magnética.....	4
I.6 Memoria USB.....	5
I.7 Reproductores de audio y video digital.....	5
I.8. Librerías de cintas.....	6
I.9. Discos Duros.....	6
1.10 Arreglos de Discos Duros.....	7
Discos Duros	8
2.1 Estructura física de un disco duro.....	9
2.1.1 Los discos.....	9
2.1.2 Las cabezas.....	9
2.1.3 El eje.....	9
2.1.4 Actuador.....	9
2.2 Estructura lógica de un disco duro.....	10
2.2.1 Cilindros.....	10
2.2.2 Pistas.....	10
2.2.3 Sectores.....	10
2.3 Funcionamiento de la unidad de disco duro.....	11
2.4 Medidas que describen el desempeño de un Disco Duro.....	12
2.4.1 Tiempo de Búsqueda (seek time).....	12

2.4.2 Latencia (latency).....	12
2.4.3 Command Overhead.....	12
2.4.4 Transferencia.....	13
2.4.5 Buffer y Caché.....	13
2.5 Interfaces y controladoras.....	13
2.5.1 ST506, MFM y RLL.....	13
2.5.2 ESDI: Enhanced Small Devices Interface.....	14
2.5.3 IDE: Integrated Drive Electronics.....	14
2.5.4 Enhanced IDE.....	16
2.5.5 SCSI: Small Computer System Interface.....	17
2.5.6 Differential o High Voltage Differential (HVD).....	21
2.5.7 Low Voltage Differential (LVD).....	21
2.5.8 Fiber Channel Arbitrated Loop (FCAL).....	21
2.5.9 Fiber Channel SCSI.....	21
2.5.10 Serial ATA.....	22
2.5.11 SATA Externo.....	23
2.5.12 9-USB.....	23
2.5.13 10-Firewire o IEEE 1394.....	25
2.6 Compañías líderes fabricantes de discos duros.....	28
2.7 Futuro de los discos duros.....	29
Administración de Discos Duros utilizando el Sistema Operativo Solaris.....	30
3.1 Particionamiento de Discos Duros.....	32
3.2 Nombre de las Particiones.....	34
3.3 Introducción a las convenciones de nombres en Solaris.....	36
3.3.1 Nombre Lógico.....	36
3.3.2 Nombres físicos.....	36
3.3.3 Nombres de instancia.....	37

3.4 Partición del disco.....	37
3.4.1 Fundamentos de la partición de discos.....	37
3.4.2 Espacio en disco y condiciones indeseables.....	37
3.4.3 Desperdicio de espacio en disco.....	38
3.4.4 Superposición o traslape de particiones.....	39
3.5 Tablas de partición de discos.....	40
3.6 Usos del comando format.....	41
3.7 Guardar la tabla de particiones.....	51
3.8 Uso de una tabla de particiones personalizada.....	52
3.9 Manejo de las etiquetas de disco.....	54
3.9.1 Uso del comando verify.....	54
3.10 Reetiquetando un disco.....	56
3.11 Manejo del sistema de archivos (file systems) en Solaris.....	57
3.12 File Systems Basados en Disco.....	57
3.13 File Systems Distribuidos.....	58
3.14 Pseudo file systems.....	58
3.15 Creación de un file system ufs.....	59
3.16 Etiqueta del disco (VTOC).....	61
3.17 Bloque de arranque.....	61
3.18 Super bloque primario.....	61
3.19 Super bloque de respaldo.....	61
3.20 Grupo de cilindros.....	61
3.21 Bloques de grupos de cilindros.....	62
3.22 El Inodo ufs.....	63
3.23 Apuntadores Directos.....	65
3.24 Apuntadores Indirectos.....	65
3.25 Bloques de datos.....	65

3.26 Fragmentación.....	66
Tecnología RAID.....	68
4.1 ¿Qué es RAID?.....	69
4.2 Técnica del sistema RAID.....	69
4.3 Sistemas RAID como una solución a la operación de los Discos Duros.....	70
4.4 Ventajas del uso de RAID.....	71
4.5 Niveles de arreglos.....	71
4.5.1 RAID 0. Striping sin tolerancia a fallos.....	71
4.5.2 RAID 1. Espejo (mirroring) y duplicación.....	73
4.5.3 RAID 2 Código de Corrección de Errores con código de Hamming.....	74
4.5.4 RAID 3 Transferencia en paralelo con paridad.....	76
4.5.5 RAID 4 Arreglo de discos independientes.....	77
4.5.6 RAID 5 Arreglo de discos independientes.....	78
4.5.7 RAID 6 Arreglo de discos independientes.....	80
4.5.8 RAID 10 o 1+0.....	82
4.6 Como seleccionar un disco RAID.....	83
Veritas Storage Foundation.....	86
5.1 Almacenamiento Virtualizado.....	86
5.1.1 Almacenamiento Virtualizado basado en el Almacenamiento.....	88
5.1.2 Almacenamiento Virtualizado basado en Host.....	88
5.1.3 Almacenamiento Virtualizado Basado en la Red.....	88
5.2 ¿Qué es Veritas Volume Manager (VxVM)?.....	89
5.3 ¿Qué es VERITAS File System?.....	89
5.4 Beneficios de VERITAS Storage Foundation.....	90
5.4.1 Manejabilidad.....	90
5.4.2 Disponibilidad.....	91
5.4.3 Rendimiento.....	91

5.4.4 Escalabilidad.....	91
5.5 Veritas Storage Foundation.....	92
5.6 Importancia.....	92
5.6.1 Almacenamiento Físico de Datos.....	92
5.6.2 Estructura Física de Datos.....	93
5.6.3 Nombres para los discos físicos.....	93
5.6.4 Arreglos de Discos.....	94
5.7 Almacenamiento Virtual de Datos.....	94
5.7.1 ¿Qué es un Volumen?.....	94
5.7.2 ¿Cómo se accesa a un Volumen?.....	95
5.8 ¿Por qué usar Volume Manager?.....	95
5.9 Controladora de discos de Volume Manager.....	96
5.10 Cuando un disco es puesto bajo el control de Volume Manager.....	96
5.11 Objetos de Almacenamiento de Volume Manager.....	97
Solución de Almacenamiento en una empresa del sector eléctrico.....	103
6.1 Descripción de la Compañía.....	103
6.2 Situación Actual.....	104
6.3 Análisis de la problemática.....	104
6.4 Propuesta.....	109
6.5 Implementación.....	111
6.6 Resultados.....	124
Conclusiones.....	125

Introducción

En la actualidad la información se ha convertido en lo más importante y valioso para las empresas, las cuales necesitan proteger los datos y recuperarlos rápidamente en caso de que se produzcan catástrofes, naturales o provocadas por el hombre, errores del operador o fallos tecnológicos o de aplicación. Sin una estrategia de protección de datos efectiva, las operaciones pueden quedar en punto muerto, con las consiguientes pérdidas millonarias. Cualquier pérdida o falta de disponibilidad de la información pone en peligro la propia supervivencia de negocios y empresas.

El almacenamiento apropiado de esta información es esencial para las operaciones de dichas empresas. Debido a esto, cada día se están tomando mayores medidas de seguridad.

Actualmente, se utilizan sistemas informáticos basados en la combinación de hardware y software para implementar soluciones automatizadas de respaldo, aunque estos sistemas son indispensables y necesarios para una buena administración y, no siempre es el método más rápido en el caso de una restauración de datos.

La elección de una solución correcta de almacenamiento depende de las necesidades específicas y los objetivos a largo plazo de cada empresa.

Existen varios criterios clave a considerar para la elección de una solución; los cuales son:

- **La capacidad.** Considerando la cantidad y el tipo de datos que serán almacenados.
- **El funcionamiento.** Refiriéndonos a la entrada, salida y exigencias de rendimiento.
- **Adaptabilidad.** Hablamos del crecimiento de datos a largo plazo.
- **Disponibilidad y fiabilidad.** ¿Las aplicaciones manejadas por la empresa son de misión crítica?
- **Protección de datos.** Reserva y exigencias de recuperación.
- **Personal capacitado y recursos humanos disponibles.**
- **Presupuesto.**

Las empresas y negocios requieren una solución de almacenamiento que pueda administrarse fácilmente y que proporcione la más alta disponibilidad del sistema.

Aunque esta necesidad es evidente, no siempre es fácil decidir cual solución es la adecuada para la organización. Existen una variedad de opciones las más utilizadas son el modelo de almacenamiento directo

(Direct Attached Storage: DAS), el sistema de almacenamiento conectado a la red (Network Attached Storage: NAS) y las redes de área de almacenamiento (Storage Area Network: SAN).

Las soluciones anteriores requieren de un manejador de volúmenes de datos empleando la tecnología RAID (Redundant Array of Inexpensive Disks) basada ya sea en hardware o software que permite mayor disponibilidad y seguridad de la información así como un mejor desempeño de las unidades de almacenamiento como lo son los discos duros.

El presente trabajo tiene el objetivo de mostrar la forma de almacenar datos de forma segura utilizando un software manejador de volúmenes (VERITAS Volume Manager® by Symantec). Se describirán las características de dicho programa, el cual es ideal para entornos informáticos empresariales y de organizaciones que dependen de datos y que requieren acceso constante y coherente a datos de misión crítica.

Comenzaremos por conocer los diferentes tipos de dispositivos de almacenamiento, detallando las principales características de cada uno de ellos; dando importancia al más utilizado en la actualidad “el disco duro”, en el cual empleamos un capítulo para tener un conocimiento completo de su funcionamiento.

Se estudiará la forma en la que el sistema operativo Solaris maneja los discos duros, la forma de organización de los datos, particiones y sistemas de archivos.

Daremos una introducción al concepto RAID explicando su significado. Se describirán cada uno de los diferentes tipos de niveles de RAID utilizados en la actualidad, mencionando características, aplicaciones, ventajas y desventajas de cada uno. Se aclarará la diferencia entre RAID por Software y RAID por hardware. Conoceremos las diferentes compañías líderes dedicadas al desarrollo de esta tecnología.

Por último abundaremos en el empleo de un programa manejador de volúmenes dando a conocer sus principales ventajas y beneficios. Ejemplificaremos su utilización en la puesta a punto de la solución de almacenamiento de datos ofrecida a una empresa del sector eléctrico.

1

Dispositivos de almacenamiento

Existen multitud de dispositivos diferentes donde almacenar la información o copias de seguridad, desde un simple disco flexible, unidades de cinta de última generación hasta arreglos inteligentes de discos duros. Evidentemente, cada uno tiene sus ventajas y sus inconvenientes,

Aquí vamos a comentar algunos de los dispositivos de almacenamiento de datos más utilizados hoy en día; de todos ellos (o de otros, no listados aquí) cada administrador de sistema de información ha de elegir el que más se adapte a sus necesidades.

I.1 Unidades de discos flexibles (floppy).



Los discos flexibles o más bien conocidos como disquetes fueron un medio de almacenamiento de tipo magnético y decimos fueron porque en la actualidad se encuentran obsoletos. Existieron dos versiones para PC que fueron primero de 5 ¼ pulgadas y de 3 ½ por el tamaño del cartucho plástico.

La información que éste tipo de discos contenía podía perderse o afectarse de forma fácil ya que el disco magnético se veía afectado por el tiempo, el polvo, la humedad, el magnetismo y hasta el calor.

Las capacidades de almacenamiento eran de 1.2 MB para la versión 5 ¼ y de 1.44 MB para la de 3 ½. La unidad encargada de procesar los datos en éste tipo de discos es llamada disquetera o también y comúnmente floppy. Esta unidad ya no forma parte de las máquinas actuales sin embargo es posible conectarlas a través de puertos IDE.

I.2 Unidades Zip y Jaz.



Son dispositivos magnéticos al igual que los discos flexibles pero con mayores capacidades de almacenamiento hasta 100 MB la Zip y 1 GB la Jaz, Son más robustos y fiables, Al igual que sus antecesores están en un proceso de desaparición si no es que ya han llegado a ser obsoletos también su capacidad no es suficiente para hacer respaldos de grandes cantidades de información.

Algunas computadoras adicionaron éste tipo de dispositivos de manera interna por una interfaz IDE o SCSI, pero también podían hacerse conexiones externas a través de los puertos SCSI o hasta paralelo, sin tener prescindir de la impresora.

Sin duda en su momento fueron dispositivos de gran uso especialmente para Pc's ya que se podían tener almacenados varios documentos en un mismo disco.

I.3 CD-ROM



Los diskettes y los discos duros almacenan la información en un medio magnético. Otra forma de almacenamiento secundario de forma óptica y es en un disco compacto o CD, en donde podemos llevar a cabo el proceso. Un disco compacto almacena la información de forma binaria. Miden aproximadamente 12 cm. De diámetro y pesan sólo unos pocos gramos tienen capacidad hasta de 640 MB, pero puede llegar a ser un poco más. La información contenida en el disco puede ser leída por un haz de luz láser desde una lectora a lo que le llamamos CD-ROM.

Los discos compactos son de dos tipos CD-R (Recordable, grabable una única vez), y CD-RW (ReWritable, regrabable múltiples veces). Los CD's tiene una buena distribución en los datos por estar extendidos, la velocidad de las lectoras grabadoras no es muy alta hablamos en la práctica de no mas de 1.8 MB/s los 12x. La grabación se hace más o menos a 4x 600 Bb/s algo que si se requiere de una gran cantidad de información no resulta del todo favorable.

La tecnología de CD-ROM se usa en proyectos de multimedia, que presentan la información utilizando una combinación de medios como sonido, gráficos, animación, video y texto. Otros usos para discos de CD-ROM son en grandes bases de datos, tablas financieras y otros cuerpos de datos extremadamente grandes. En la actualidad ya se encuentran disponibles las unidades de CD-WR (Compact Disk Write Read), que al igual que los CD-ROM disponen de formatos estándares de 700 MB por

disco, pero también disponen de capacidad para leer y escribir en formatos no estándares y lo más impactante es que pueden re-escribir el disco óptico previo borrado, lo que permite actualizaciones e incluso dan lugar a una nueva tecnología para el almacenamiento de información en forma segura y transportable a voluntad del usuario. Estas unidades funcionan correctamente con los discos existentes en el mercado de solo lectura de escritura una vez y lectura muchas veces, pero también existen ya discos especiales que permiten ser borrados y re-escritos varias veces.

I.4 DVD-ROM.



El disco versátil digital es un dispositivo de almacenamiento masivo de datos tienen el mismo tamaño que el CD o sea, 12 cm de diámetro u 8 cm para la versión mini; contienen hasta 25 veces más información y puede transmitirla al ordenador o computadora unas 20 veces más rápido que un CD-ROM. Los DVD guardan los datos utilizando un sistema de archivos denominado UDF.

Un DVD simple puede guardar hasta 4.7 GB, los hay de una cara y doble capa éste tipo almacena hasta 8.5 GB, los de dos caras y capa simple en ambas tiene una capacidad de 9.4 GB y los de dos caras, capa doble en ambas hasta 17.1 GB.

Las unidades lectoras de DVD permiten leer la mayoría de los CD's, ya que ambos son discos ópticos; no obstante, los lectores de CD no permiten leer DVD's.

El uso de este tipo de discos ópticos son muy variados se puede almacenar video, audio, datos, pueden ser de solo lectura, grabables una sola vez, regrabables, grabables una vez de doble capa o regrabables de doble capa.

En 1999 aparecieron los DVD-Audio, que emplean un formato de almacenamiento de sonido digital de segunda generación con el que se pueden recoger zonas del espectro sonoro que eran inaccesibles al CD-Audio.

Por su parte, los lectores de disco compacto, CD, y las unidades de DVD, disponen de un láser, ya que la lectura de la información se hace por procedimientos ópticos. En algunos casos, estas unidades son de sólo lectura y en otros, de lectura y escritura.

La velocidad de transferencia de datos de una unidad DVD está dada en múltiplos de 1.350 KB/s, esto es que la unidad lectora de 16x, transmite a $16 \times 1.350 = 21.600$ K/B/s. En términos de rotación física (revoluciones por minuto), un múltiplo de velocidad en DVD equivale a nueve múltiplos de velocidad en CD. Así podemos decir que una unidad de

DVD es tres veces más rápida que una de CD-ROM.

I.5 Unidades de cinta magnética.



La cinta o banda de material magnético (como óxido de hierro o algún cromato). Es un medio de almacenamiento externo, generalmente legible y escribible, almacena datos en la forma de cargas electromagnéticas de tal manera que pueden ser leídas y también ser borradas.

Puede grabarse cualquier tipo de información de forma digital o analógica. Los primeros sistemas utilizaban cintas tipo riel abierto, los nuevos como los de las imágenes suelen ser cartuchos tipo casetes.

En la actualidad, la cinta es el medio más utilizado para almacenar sistemas de respaldo.

Una ventaja del drive es que las cintas tienen una capacidad grande para almacenar datos y son económicas cuando están comparadas al costo de almacenaje del disco duro. Una desventaja es que los drives almacenan datos secuencialmente. El tipo de acceso secuencial quiere decir que, si se requiere acceder al n -ésimo bloque de la cinta será necesario leer antes los $n-1$ bloques que le preceden.

La capacidad de datos en una cinta magnética es medida en BPI (bits por pulgada), que puede ir desde 800 bpi hasta 6250 bpi. A mayor densidad en la cinta más datos se guardan por pulgada, depende del formato de la misma (DDS "Digital Data Storage", DLT "Digital Linear Tape", LTO "Linear Tape Open", etc).



DDS



DLT



LTO

I.6 Memoria USB.



Universal Serial Bus es un puerto de gran velocidad soporta dispositivos plug&play y conexiones en caliente. Las transferencias de datos son a 12 MBps, El puerto permite ser usado por más de 127 dispositivos periféricos como ratones, módems, teclados y medios de almacenamiento como lo son las llamadas memorias USB.

Una memoria USB son también llamadas memorias flash. Es un tipo de memoria no volátil como las utilizadas en los celulares, las cámaras digitales, PDAs, reproductores portátiles, etc, pueden borrarse y reescribirse. Puede decirse que son una evolución de las memorias EEPROM que permiten que múltiples posiciones de memoria sean escritas o borradas en una misma operación mediante impulsos eléctricos.

Al principio tenían una capacidad de almacenamiento de 8 MB, actualmente están revasando los 64 GB con velocidades de transmisión de 20 MB/s.

Estas memorias son resistentes a los rasguños y al polvo que han afectado a las formas previas de almacenamiento portátil, como los CD y los disquetes.

Estas memorias se han convertido en el sistema de almacenamiento y transporte personal de datos más utilizado, Su gran popularidad le ha supuesto infinidad de denominaciones populares relacionadas con su pequeño tamaño y las diversas formas de presentación

Los sistemas operativos actuales pueden leer y escribir en las memorias sin más que conectarlas al puerto USB del equipo encendido, recibiendo la energía de alimentación a través del propio conector.

I.7 Reproductores de audio y video digital.



Son dispositivos que almacenan, organizan y reproducen archivos de video y/o música digital. Comúnmente se les denomina reproductores MP3 (debido a la ubicuidad de este formato), pero estos tipos de dispositivos de almacenamiento son capaces de almacenar diferentes tipos de archivos, algunos de los formatos son propios del fabricante y otros son de tipo libre de patente o como se les llama comúnmente abiertos.

Principalmente existen tres tipos de reproductores de audio y video digital: los reproductores de CD's MP3, reproductores basados en flash; son dispositivos que almacenan archivos en memorias internas o externas,

normalmente son de poca capacidad de almacenamiento que pueden ser ampliados con memoria adicional, y los reproductores basados en disco duro; dispositivos que leen archivos desde un diminuto disco duro, estos dispositivos tienen cantidades de almacenamiento más grandes que los anteriores, dependiendo de la tecnología de disco que contengan.

I.8. Librerías de cintas.



óptico.

Una librería de cintas, también conocido como robots de respaldo, es un conjunto de cintas magnéticas y sus unidades lectoras también conocidas como drives.

Estas librerías cuentan con un dispositivo robótico (brazo) ubicado en el interior, el cual tiene la función de seleccionar las cintas, montarlas y desmontarlas en sus respectivos drives; el brazo sabe distinguir entre una cinta y otra ya que cada cinta cuenta con un código de barras y el brazo cuenta con un lector

Estos sistemas son administrados por un software manejador de respaldos que permite crear políticas de respaldos automatizados y programados, facilitando así, las tareas del administrador.



I.9 Discos Duros.



Los discos duros son muy importantes ya que son el medio de almacenamiento más seguro, confiable y rápido. Pueden almacenar una gran cantidad de datos y se puede recuperar más rápidamente que en cualquier otro dispositivo de almacenamiento. Los discos duros contienen una o más superficies rígidas para grabar datos, cada una cubierta con un óxido metálico que registra cargas magnéticas. La cabeza de escritura /lectura se encuentra típicamente a unos catorce millonésimos de pulgada por encima de la superficie del disco.

Los platos giran sobre un eje desde 7200 hasta 15000 rpm (revoluciones por minuto) dependiendo del modelo de disco (cuanto más rápido gire un disco, más rápido la cabeza de lectura/escritura puede encontrar datos).

En la actualidad existen discos duros de tecnologías que le permiten un tiempo de acceso muy corto; entre los líderes de este tipo de tecnologías se cuenta con empresas como *Quantum*, *Seagate*, *Maxtor*, *Hitachi Data Systems*, quienes fabrican discos duros con superficies aprovechables (libre de errores) en casi un 100%.

1.10 Arreglos de Discos Duros.



Los arreglos de discos duros son un conjunto de discos duros dentro de una “caja” la cual es conectada al host (servidor) para ofrecer un gran espacio de almacenamiento.

Pueden ser inteligentes o simples; Al decir arreglos simples significa que el arreglo de discos simplemente ofrece al host la cantidad total de discos sin manejar tecnologías RAID; al mencionar inteligentes significa que dentro de esa “caja” de discos existe una o más tarjetas controladoras de los discos, estas tarjetas controladoras contienen componentes electrónicos de tal forma que pueden ofrecer una solución de almacenamiento empleando tecnologías RAID, a esto se le conoce como Raid por Hardware.

Estos arreglos inteligentes tienen la gran ventaja que manejan memoria caché, es decir los datos a primera instancia son almacenados en memoria y posteriormente a disco, esto hace que la escritura sea más eficiente y mucho más rápida.

Los arreglos hoy en día están disponibles con interfaces SCSI o fibra y pueden almacenar varios terabytes de información.

A través de este capítulo hemos visto un poco del desarrollo de los dispositivos de almacenamiento, a este tiempo algunos de ellos han casi desaparecido o son obsoletos, ya que han sido reemplazados por otros de mayor tecnología. El siguiente capítulo se ha dedicado al dispositivo que nos compete para el desarrollo de éste trabajo que es el disco duro, conceptos de gran importancia para el resto de los capítulos.



2

Discos Duros

Los discos duros pertenecen a la llamada memoria secundaria o almacenamiento secundario de una computadora. Fueron inventados en los años 50s y se les dio el nombre de discos duros para distinguirlos de los discos flexibles o “floppys” (descritos en el capítulo anterior). También se les conoce como discos rígidos por su fabricación a base de una capa rígida de aluminio.

Han sufrido muchos cambios desde su invención; en un principio su tamaño físico era de las dimensiones de una caja de zapatos y con capacidades de almacenamiento muy pequeñas. Al día de hoy los fabricantes de discos (de los que hablaremos más adelante) los fabrican en tamaños compactos y reducidos con capacidades medidas hasta en terabytes (millones de bytes).

Para que un disco forme parte de una computadora es necesario conectarlo a través de una tarjeta electrónica la cuál recibe el nombre de controladora; durante el desarrollo de éste capítulo entraremos a detalle sobre los diferentes tipos de controladoras describiendo sus características y usos.

Para entender mejor el uso de los discos duros comenzaremos por describir su construcción física así como funcionamiento y conceptos involucrados en el mismo.

2.1 Estructura física de un disco duro.

Los componentes físicos de un disco duro son:

2.1.1 Los discos.

Están elaborados de compuestos de vidrio, cerámica o aluminio finamente pulidos y revestidos por ambos lados con una capa muy delgada de una aleación metálica (habitualmente óxido de hierro). Los discos están unidos a un eje y un motor que los hace girar a una velocidad constante. Convencionalmente los discos duros están compuestos por varios platos, es decir varios discos de material magnético montados sobre un eje central. Estos discos normalmente tienen dos caras que pueden usarse para el almacenamiento de datos, si bien suele reservarse una para almacenar información de control.

2.1.2 Las cabezas.

Están ensambladas en pila y son las responsables de la lectura y la escritura de los datos en los discos. La mayoría de los discos duros incluyen una cabeza de lectura/escritura a cada lado del disco, sin embargo algunos discos de alto desempeño tienen dos o más cabezas sobre cada superficie, de manera que cada cabeza atiende a la mitad del disco reduciendo la distancia del desplazamiento radial. Las cabezas de lectura/escritura no tocan el disco cuando este está girando a toda velocidad; por el contrario, flotan sobre una capa de aire extremadamente delgada (100 millonésimas de pulgada). Esto reduce el desgaste en la superficie del disco durante la operación normal, cualquier polvo o impureza en el aire puede dañar suavemente las cabezas o el medio. Su funcionamiento consiste en una bobina de hilo que se acciona según el campo magnético que detecte sobre el soporte magnético, produciendo una pequeña corriente que es detectada y amplificada por la electrónica de la unidad de disco.

2.1.3 El eje.

Es la parte del disco duro que actúa como soporte, sobre el cual están montados y giran los platos del disco.

2.1.4 Actuador .

Es un motor que mueve la estructura que contiene las cabezas de lectura entre el centro y el borde externo de los discos. Un “actuador” usa la fuerza de un electromagneto empujado contra magnetos fijos para mover las cabezas a través del disco. La controladora manda más corriente a través del electromagneto para mover las cabezas cerca del borde del disco. En caso de una pérdida de energía eléctrica, un resorte mueve la cabeza nuevamente hacia el centro del disco sobre una zona

donde no se guardan datos. Dado que todas las cabezas están unidas al mismo rotor ellas se mueven al unísono.

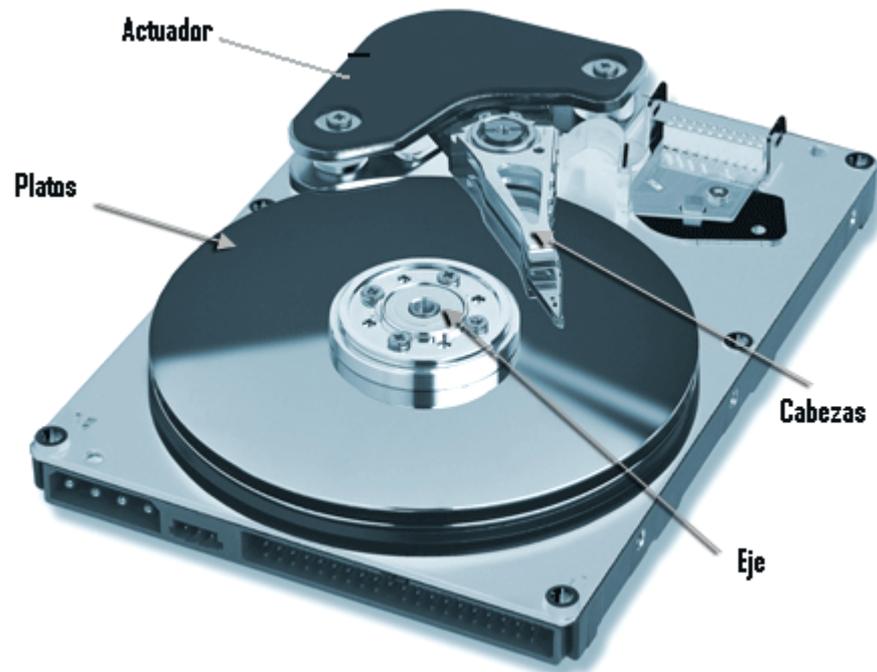


Figura 2.1 Estructura física de un disco duro.

2.2 Estructura lógica de un disco duro.

2.2.1 Cilindros.

Se llama cilindro al par de pistas en lados opuestos del disco o plato. Si el disco duro contiene múltiples platos, un cilindro incluye todos los pares de pistas directamente una encima de la otra. Los discos duros normalmente tienen una cabeza a cada lado del disco. Dado que las cabezas de lectura/escritura están alineadas una con otra, la controladora puede escribir en todas las pistas del cilindro sin mover el rotor. Como resultado los discos duros de múltiples platos se desempeñan levemente más rápido que los discos duros de un solo plato.

2.2.2 Pistas

Un disco está dividido en delgados círculos concéntricos llamados pistas. Las cabezas se mueven entre la pista más externa ó pista cero a la más interna. Es la trayectoria circular trazada a través de la superficie circular del plato de un disco por cabeza de lectura/escritura.

2.2.3 Sectores.

Un byte es la unidad útil más pequeña en términos de memoria. Los discos duros almacenan datos en pedazos gruesos llamados sectores. La

mayoría de discos usan sectores de 512 bytes. La controladora del disco duro determina el tamaño de un sector en el momento en que el disco es formateado. Algunos modelos permiten especificar el tamaño del sector. Cada pista del disco está dividida en 1 ó 2 sectores dado que las pistas exteriores son más grandes que las interiores, las exteriores contienen más sectores.

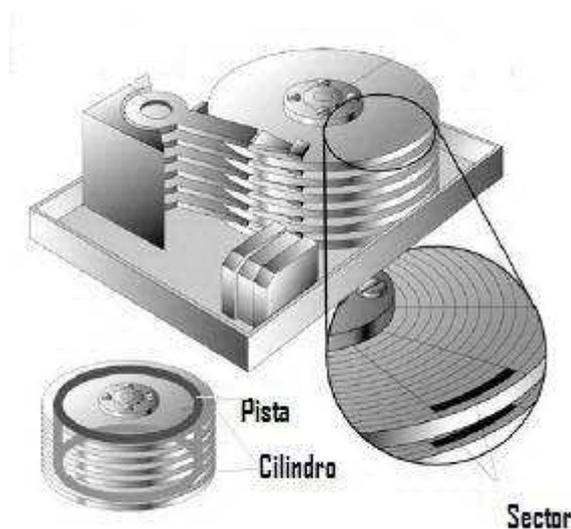


Figura 2.2 Estructura lógica de un Disco Duro.

2.3 Funcionamiento de la unidad de disco duro.

A continuación describiremos el conjunto de mecanismos necesarios para el acceso a los datos almacenados en los discos duros.

En primer lugar, el número de cabezas de lectura/escritura que tiene un disco es igual al número de platos que contenga multiplicado por dos ya que cada cara requiere de un lector. El conjunto de cabezas de lectura/escritura pueden desplazarse del exterior al interior de la pila de platos mediante un brazo mecánico que los soporta, para que las cabezas puedan acceder al total de los datos es necesario que los discos giren, este movimiento debe ser a velocidad constante y permanecerá mientras la computadora esté conectada a la corriente eléctrica; a diferencia de los discos flexibles, los cuales sólo giran si se tiene una orden de lectura o escritura.

Para escribir, la cabeza se sitúa sobre la celda a grabar y se hace pasar por ella un pulso de corriente, lo cual crea un campo magnético en la superficie. Dependiendo del sentido de la corriente, así será la polaridad de la celda. Para leer, se mide la corriente introducida por el campo magnético de la celda, es decir, que al pasar por la zona detectará un

campo magnético que según se encuentre magnetizado en un sentido u otro, indicará si en esa posición hay almacenado un 0 o un 1. En el caso de la escritura el proceso es inverso, la cabeza recibe una corriente que provoca un campo magnético, el cual pone la posición sobre la que se encuentre la cabeza en 0 o en 1 dependiendo del valor del campo magnético provocado por dicha corriente.

2.4 Medidas que describen el desempeño de un Disco Duro.

Los fabricantes de discos duros miden la velocidad en términos de tiempo de acceso, tiempo de búsqueda, latencia, transferencia, buffer y caché. Estas medidas también aparecen en las advertencias, comparaciones y en las especificaciones. Tiempo de acceso (access time) termino frecuentemente usado en discusiones de desempeño, en el intervalo de tiempo entre el momento en el que el drive recibe un requerimiento por datos, y el momento en el que el drive empieza a despachar el dato. El tiempo de acceso de un disco duro es una combinación de tres factores:

2.4.1 Tiempo de Búsqueda (seek time)

Es el tiempo que les toma a las cabezas de lectura/escritura moverse desde su posición actual hasta la pista donde está localizada la información deseada. Como la pista deseada puede estar localizada en otro lado del disco o en una pista adyacente, el tiempo de búsqueda variara en cada búsqueda.

2.4.2 Latencia (latency)

Cada pista en un disco duro contiene múltiples sectores una vez que la cabeza de lectura/escritura encuentra la pista correcta, las cabezas permanecen en el lugar e inactivas hasta que el sector pasa por debajo de ellas. Este tiempo de espera se llama latencia. La latencia promedio es igual al tiempo que le toma al disco hacer media revolución y es igual en aquellos drivers que giran a la misma velocidad.

2.4.3 Command Overhead.

Tiempo que le toma a la controladora procesar un requerimiento de datos. Este incluye determinar la localización física del dato en el disco correcto, direccional al "actuador" para mover el rotor a la pista correcta, leer el dato, redireccionarlo a la instancia que solicitó el dato.

2.4.4 Transferencia.

Los discos duros también son evaluados por su transferencia, la cual generalmente se refiere al tiempo en el cual los datos pueden ser leídos o escritos en el drive, el cual es afectado por la velocidad de los discos, la densidad de los bits de datos y el tiempo de acceso. La mayoría de los discos duros actuales incluyen una memoria RAM que es usada como caché o almacenamiento temporal. Dado que los procesadores y los discos duros se comunican a través de un bus de entrada/salida, el tiempo de transferencia actual entre ellos está limitado por el máximo tiempo de transferencia del bus, el cual en la mayoría de los casos es mucho más lento que el tiempo de transferencia del drive.

2.4.5 Buffer y Caché.

Prácticamente todos los discos incluyen una memoria buffer, en la que almacenan los últimos sectores leídos; ésta, es importantísima de cara al rendimiento, e incluso imprescindible para poder mantener altas cuotas de transferencia. Se la denomina caché cuando incluye ciertas características de velocidad,; concretamente, los procesos se optimizan cuando el sistema viene de una operación de copiado de datos a la unidad sin esperar a que ésta haya finalizado. También utilizan otra técnica diferente consistente en que la unidad informa de la finalización de una operación de escritura en el momento de recibir los datos, antes de comenzar a grabarlos en el disco.

2.5 Interfaces y controladoras.

La interfaz es la conexión entre el mecanismo de la unidad de disco y el bus del sistema. Define la forma en la que las señales pasan entre el bus del sistema y el disco duro. En el caso del disco, se denomina controladora o tarjeta controladora, y se encarga no sólo de transmitir y transformar la información que parte de y llega al disco, sino también de seleccionar la unidad a la que se quiere acceder, del formato, y de todas las ordenes de bajo nivel en general. La controladora a veces se encuentra dentro de la tarjeta madre o mother board.

2.5.1 ST506, MFM y RLL.

Los primeros discos duros eran manejados por controladoras ST506, un estándar creado por la empresa Seagate®. Dentro de esta norma se implementaron los modos MFM (Modified Frequency Modulation) y RLL (Run Length Limited), dos sistemas para el almacenamiento de datos que, si bien diferentes en su funcionamiento, a nivel físico externo del disco presentaban la misma apariencia, siendo conocidos de forma genérica en el mundo como discos "MFM". Estas unidades incluían externamente tres conectores: el primero, y común a cualquier disco duro, es el de alimentación. En los restantes se conectaba un cable de control y un cable de datos, desde el disco a la controladora; el cable de control gestionaba la posición de las cabezas de lectura/ escritura y el de datos transmitía el

flujo de información desde y hasta la controladora.

La diferencia entre MFM y RLL es a nivel interno; MFM y RLL son dos métodos de la información binaria. RLL permite almacenar un 50% más de datos que el MFM, al aumentar la densidad de almacenamiento. También la tasa de transferencia es superior en RLL, debido al más eficiente método de grabación usado, sin embargo, la velocidad de rotación era la misma en ambos casos: 3600 rpm.

2.5.2 ESDI: Enhanced Small Devices Interface.

Con esta interfaz (interfaz mejorada para dispositivos pequeños), se daba un paso adelante. Para empezar, una parte de la lógica decodificadora de la controladora se implementó en la propia unidad, lo que permitió elevar el ratio de transferencia a 10 Mbits por segundo. Así mismo, se incluyó un pequeño buffer de sectores que permitía transferir pistas completas en un único giro o revolución de disco.

No obstante, estas unidades no se extendieron demasiado, y únicamente compañías como IBM ®, fueron las que más lo emplearon en sus máquinas. Estas unidades no solían tener una capacidad superior a 630 Megas y en cualquier caso se trató más bien de una tecnología de transición, ya que un tiempo después tuvo lugar el salto cuantitativo y cualitativo con la interfaz que detallaremos a continuación.

2.5.3 IDE: Integrated Drive Electronics

El sistema IDE , "Dispositivo con electrónica integrada") o ATA (Advanced Technology Attachment,) controla los dispositivos de almacenamiento masivo de datos, como los discos duros y ATAPI (Advanced Technology Attachment Packet Interface), además añade dispositivos como las unidades CD-ROM.

Fue creado por la firma Western Digital, curiosamente por encargo de Compaq® para una nueva gama de computadoras personales. Su característica más representativa era la implementación de la controladora en el propio disco duro, de ahí su denominación. Desde ese momento, únicamente se necesita una conexión entre el cable IDE y el bus del sistema, siendo posible implementarla en la tarjeta madre o mother board. Igualmente se eliminó la necesidad de disponer de dos cables separados para control y datos, basado en un cable de 40 hilos desde el bus al disco duro. Se estableció también el termino ATA (AT Attachment) que define una serie de normas a las que deben de acogerse los fabricantes de unidades de este tipo.

IDE permite una transferencia de 4 Megas por segundo, aunque dispone de varios métodos para realizar estos movimientos de datos. La interfaz IDE supuso la simplificación en un proceso de instalación y configuración de discos duros, y estuvo durante un tiempo a la altura de

las exigencias del mercado.

Paralell ATA (algunos están utilizando la sigla PATA)

ATA-1

ATA-2, soporta transferencias rápidas en bloque y multiword DMA.

ATA-3, es el ATA2 revisado.

ATA-4, conocido como Ultra-DMA o ATA-33 que soporta transferencias en

33 Mbps

ATA-5 o Ultra ATA/66, originalmente propuesta por Quantum para transferencias en 66 Mbps

ATA-6 o Ultra ATA/100, soporte para velocidades de 100MBps.

ATA-7 o Ultra ATA/133, soporte para velocidades de 133MBps.

* Serial ATA, remodelación de ATA con nuevos conectores (alimentación y datos), cables, tensión de alimentación y conocida por algunos como SATA.

Las controladoras IDE casi siempre están incluidas en la mother board, normalmente dos conectores para dos dispositivos cada uno. De los dos discos duros, uno tiene que estar como esclavo y el otro como maestro para que la controladora sepa a qué dispositivo mandar los datos y también de cual serán recibidos. La configuración se realiza mediante jumpers. Habitualmente, un disco duro puede estar configurado de una de estas tres formas:

a) Como maestro ('master'). Si es el único dispositivo en el cable, debe tener esta configuración, aunque a veces también funciona si está como esclavo. Si hay otro dispositivo, el otro debe estar como esclavo.

b) Como esclavo ('slave'). Debe haber otro dispositivo que sea maestro.

c) Selección por cable (cable select). El dispositivo será maestro o esclavo en función de su posición en el cable. Si hay otro dispositivo, también debe estar configurado como cable select. Si el dispositivo es el único en el cable, debe estar situado en la posición de maestro.

Para distinguir el conector en el que se conectará el primer bus IDE se utilizan colores distintos.

Este diseño (dos dispositivos a un bus) tiene el inconveniente de que

mientras se accede a un dispositivo el otro dispositivo del mismo conector IDE no se puede usar. En algunos chipset no se podría usar siquiera el otro IDE a la vez.



Figura 2.3 Disco Duro IDE y Cable IDE 25 pines

No obstante, no tardaron en ponerse en manifiesto modificaciones en su diseño. Dos importantes eran de capacidad de almacenamiento, de conexión y de ratios de transferencia; en efecto la tasa de transferencia se iba quedando atrás ante la demanda cada vez mayor de prestaciones por parte del software. Así mismo, sólo podían coexistir dos unidades IDE en el sistema, y su capacidad no solía exceder de los 528 Megas.

2.5.4 Enhanced IDE.

La interfaz EIDE o IDE mejorado, propuesto también por Western Digital, logra una mejora de flexibilidad y prestaciones. Para empezar, aumenta su capacidad 8.4 Gigas, y la tasa de transferencia empieza a subir a partir de los 10 Megas por segundo, según el modo de transferencia usado. Además, se implementaron dos sistemas de traducción de los parámetros físicos de la unidad de forma que se pudiera acceder a superiores capacidades. Estos sistemas, denominados CHS y LBA aportaron ventajas innegables, ya que con mínimas modificaciones se podían hacer a las máximas capacidades permitidas.

Otra mejora del EIDE se reflejó en el número de unidades que podían ser instaladas al mismo tiempo que se aumentó a cuatro. Para ello se obligó a fabricantes de sistemas y de BIOS a soportar los controladores secundarios (dirección 170h, IRQ 15) siempre presentes en el diseño de PC pero nunca usados hasta el momento, de forma que se pudiera montar una unidad y otra esclava, configuradas como secundarias. Más aún se habilitó la posibilidad de instalar unidades de CD-ROM y de cinta, coexistiendo pacíficamente en el sistema. A nivel externo, no existen prácticamente diferencias con el anterior IDE, en todo caso un menor tamaño o más bien una superior integración de un mayor número de componentes en el mismo espacio.

2.5.5 SCSI: Small Computer System Interface

Acrónimo inglés Small Computer System Interface, es un interfaz estándar para la transferencia de datos entre distintos dispositivos del bus de la computadora.

Para montar un dispositivo SCSI en una computadora es necesario que tanto el dispositivo como la mother board dispongan de un controlador SCSI. Es habitual que el dispositivo venga con un controlador de este tipo, pero no siempre es así, sobre todo en los primeros dispositivos. Se utiliza habitualmente en los discos duros y los dispositivos de almacenamiento sobre cintas, pero también interconecta una amplia gama de dispositivos, incluyendo scanners, unidades CD-ROM, grabadoras de CD, y unidades DVD. De hecho, el estándar SCSI entero promueve la independencia de dispositivos, lo que significa que teóricamente cualquier cosa puede ser hecha SCSI (incluso existen impresoras que utilizan SCSI).

En el pasado, era muy popular entre todas las clases de computadoras. Actualmente sigue siendo popular en lugares de trabajo de alto rendimiento, servidores, y periféricos de gama alta. Las computadoras personales y las portátiles utilizan habitualmente las interfaces más lentas de IDE para los discos duros y USB (el USB emplea un conjunto de comandos SCSI para algunas operaciones) así como Firewire a causa de la diferencia de costo entre estos dispositivos.

Tipos de SCSI :

SCSI 1

Bus de 8 bits. Velocidad de transmisión de datos a 5 Mbps. Su conector genérico es de 50 pin's (Figura 2.4 conector Centronics) y baja densidad. La longitud máxima del cable es de seis metros. Permite hasta 8 dispositivos (incluida la controladora), identificados por las direcciones 0 a

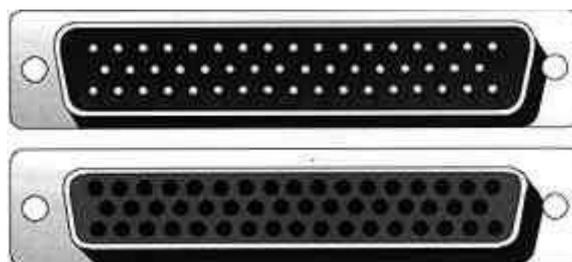


Figura 2.4 Conector SCSI 1

SCSI-2

Este término describe la norma ANSI publicada (X3.131-1994). SCSI-2 fue una actualización de la Interfaz original SCSI. Los cambios incluyeron tasa de datos mas rápida, mensajes de mandato y estructura de comandos para mejorar la compatibilidad. La tasa de transferencia sincrónica de datos para SCSI-2 es de 2.5 a 10 Mbytes/seg. para un bus de datos de 8 bits, y 5 a 20 Mbytes/seg. para un bus de datos de 16 bits.

SCSI-3

Este término describe un grupo de normas ANSI que están actualmente siendo desarrolladas para el bus SCSI. El documento SCSI-2 es muy grande (más de 400 páginas) y cubre todo el rango de tópicos. SCSI-3 separa este documento en una serie de documentos menores, cubriendo cada uno, una "capa" de la definición de la Interfaz.

Las capas básicas son:

- física (conectores, asignación de pines, especificaciones eléctricas).
- protocolo (la actividad de la capa física es organizada en fases de bus, paquetes,etc.)
- arquitectura (una descripción de como los requerimientos de comando son organizados, encolados y respondidos por algún protocolo)
- comandos primarios (descripción de los comandos que deben ser soportados por todos los dispositivos SCSI)
- comandos específicos de dispositivo (comandos que son específicos para una clase de dispositivos en particular; lectoras de CD-ROM o lecto-grabadoras WORM.

El conjunto de normas necesarias para lograr una implementaron de Interfaz paralela SCSI-3 para disco es:
SPI (SCSI Parallel Interface) para la capa física.
SIP (SCSI Interlocked Protocol) para la capa de protocolo.
SAM (SCSI Architecture Model) para la arquitectura.
SPC (SCSI Primary Commands) para el conjunto de comandos primario.
SBC (SCSI Block Commands) para el conjunto de comandos especifico de disco.

Las normas SCSI-3 están divididas en capas de esta manera para permitir la sustitución de partes de la estructura a medida que emerja nueva tecnología. Por ejemplo, un conjunto comparable de normas para una Interfaz SCSI Fiber Channel para disco reemplaza las capas física y protocolar con nuevos documentos pero usa los mismos documentos para las otras tres capas. El principal punto a recordar aquí es que los términos SCSI-2 o SCSI-3 no implican alguna performance particular per se, mas allá de referirse a la generación de documentos a la cual conforma un

producto. Partiendo de que, las mas nuevas características están solo en SCSI-3 y tienden a ser de performance superior, los dispositivos SCSI-3 deberían demostrar mejor performance que los SCSI-2 en la mayoría de los casos.

SCSI FAST

Se refiere a las relaciones de tiempos definidas en SCSI-2 para una tasa de 10 MegaTransfer/sec. Un "MegaTransfer" (MT) es una unidad de medida referente a la tasa de señales en la interface sin importar el ancho del bus. Por ejemplo, una tasa de 10 MT/sec. sobre un bus de 1 byte de ancho (narrow), resulta en una tasa de transferencia de 10 Mbytes/sec., pero en un bus de 2 bytes (wide), esta resulta en una tasa de transferencia de 20 Mbyte/sec.

SCSI FAST-20

Se refiere a las relaciones de tiempos definidas en SPI SCSI-3 para una tasa de transferencia de 20MT/sec, la cual alcanza tasas de datos el doble de rápidas que SCSI FAST. Por ejemplo 20 Mbyte/sec. para narrow y 40 Mbytes/sec. para wide.

SCSI FAST-40

Se refiere a las relaciones de tiempos que están siendo definidas para una futura revisión del SPI SCSI-3 que alcanzan 40 MT/sec., lo cual es el doble SCSI FAST-20. Por ejemplo 40 Mbyte/sec. para narrow y 80 Mbytes/sec. para wide.

SCSI FAST-80

Se refiere a las relaciones de tiempos que están siendo definidas para una futura revisión del SPI SCSI-3 que alcanzan 80 MT/sec., lo cual es el doble SCSI FAST-20. Por ejemplo 80 Mbyte/sec. para narrow y 160 Mbytes/sec. para wide.

Ultra SCSI

Este término describe la última norma ANSI publicada (X3T10/1071D rev. 6), comúnmente conocida como Fast-20. Al igual que en todas las transferencias sincrónicas, Ultra SCSI lo hace a una velocidad de reloj negociada.

Ultra SCSI-2

Este término describe la última norma ANSI publicada (X3T10/1071D rev. 6), comúnmente conocida como Fast-40. Al igual que en todas las transferencias sincrónicas, Ultra SCSI lo hace a una velocidad de reloj negociada.

SCSI Narrow

Este término se refiere al bus de datos de 1 byte de ancho sobre una Interfaz paralela de 50 pines, tal como esta definida en la norma ANSI (X3.131-1986). El bus narrow consiste en 8 líneas de datos con paridad, una serie de líneas de protocolo y sus correspondientes líneas de masa asociadas.

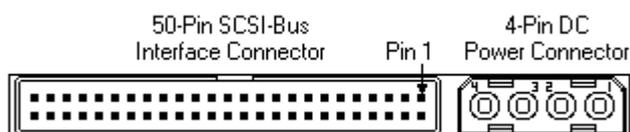


Figura 2.5 Conector SCSI Narrow

SCSI Wide

Este término usualmente se refiere al bus de datos de 2 bytes de ancho sobre una Interfaz paralela de 68 pines, tal como esta definida en el documento SCSI-3 SPI. El término puede ser aplicado genéricamente a cualquier implementaron más ancha que 1 byte, pero, al momento de escribir esto, no existen implementaciones mayores que 2 bytes. Las futuras implementaciones pueden incluir más bytes de datos porque las tasas de transferencia FAST estarán a pleno con las transferencias de 2 bytes hasta que la interfaces seriales (como Fiber Channel o FireWire) se tornen más populares.

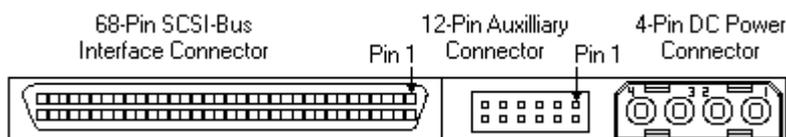


Figura 2.6 Conector SCSI Wide

2.5.6 Differential o High Voltage Differential (HVD)

Differential es un sistema de señales lógicas usado en algunos dispositivos SCSI. Este usa un nivel de señales más y menos apareadas, para reducir el efecto del ruido sobre el bus SCSI. Cualquier ruido inyectado sobre la señal estaría igualmente presente sobre el mas y el menos y, por lo tanto, sería anulado. Debido a cambios en las definiciones, Differential es ahora conocido como High Voltage Differential (HVD).

2.5.7 Low Voltage Differential (LVD)

LVD es un esquema lógico diferencial pero usando niveles de Voltage más bajos que HVD.

2.5.8 Fiber Channel Arbitrated Loop (FCAL)

Este es el nombre formal para el sistema Fiber Channel usado por SCSI. Es más comúnmente conocido como Fiber Channel SCSI. La parte Loop del nombre se refiere a la forma en que el sistema está conectado a un gran anillo. Debido a las características de este anillo, esta interface tiene más en común con LAN que con SCSI paralelo.

2.5.9 Fiber Channel SCSI

Se refiere a productos con capas físicas y protocolares Fiber Channel usando el set de comandos SCSI. La interface Fiber Channel es completamente diferente de la SCSI paralela, ya que esta es una interface serial, donde la información de comando y los datos son transmitidos sobre una corriente de señal organizada en paquetes. La fibra puede ser un cable de tipo cobre coaxial o fibra óptica. La señal sobre la primer implementaron utiliza una tasa de 1 GHz, alcanzando los 100 Mbytes/sec. sobre el cable. FC también implementa un incrementado software de control y configuración y eleva el número máximo de dispositivos a 126 ID's, frente a los 8 o 16 del bus paralelo.

SCA-2

Este es un conector tipo D, miniatura de 80 pines, usado en dispositivos SCSI para insertar en conectores de panel. El SCA-2 provee las líneas de masa, Voltage y control necesarias para permitir la conexión en caliente de dispositivos de interfaz SCSI paralela.

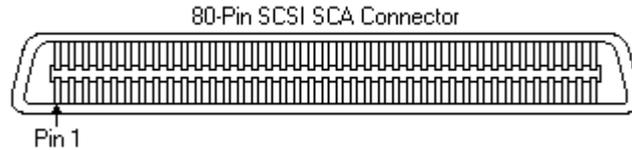


Figura 2.7 Conector SCSI SCI

2.5.10 Serial ATA

Serial ATA o S-ATA es una interfaz para discos que sustituye a la tradicional Parallel ATA o P-ATA (estándar que también se conoce como IDE o ATA). El S-ATA proporciona mayores velocidades, mejor aprovechamiento cuando hay varios discos, mayor longitud del cable de transmisión de datos y capacidad para conectar discos en caliente (con la computadora encendida).

Mientras que la especificación SATA1 llega como máximo a unos 150 MB/s, SATA2 incrementa el límite a 300 MB/s. Actualmente es una interfaz ampliamente aceptada y estandarizada en las tarjetas madre de PC.

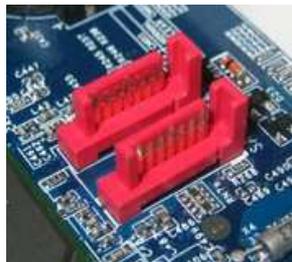


Figura 2.8 Puerto SATA en una mother board o tarjeta madre.

Cable y Conector



El cable se compone de dos pares apantallados a los que se suministra una impedancia de 100 Ohmios

Pin	Nombre	Descripción
1	GND	Tierra
2	A+	Transmisión +
3	A-	Transmisión -
4	GND	Tierras
5	B-	Recepción -
6	B+	Recepción +
7	GND	Tierra

2.5.11 SATA Externo

Fue estandarizado a mediados de 2004, con definiciones específicas de cables, conectores, y requisitos de la señal para unidades eSATA externas. eSATA se caracteriza por:

Velocidad de SATA en los discos externos (se han medido 115 MB/s con RAID externos) sin conversión de protocolos de PATA/SATA a USB/Firewire, todas las características del disco están disponibles para el anfitrión.

La longitud de cable se restringe a 2 metros; USB y Firewire permiten mayores distancias.
Se aumentó el voltaje de transmisión mínimo y máximo a 500mV - 600mV (de 400mV - 600mV)
Voltaje recibido disminuido a 240mV - 600mV (de 325 mV - 600 mV).

Actualmente, la mayoría de las tarjetas madre no tienen un conector para eSATA, pero es posible usar adaptadores de bus o tarjetas PC-Card y CardBus.

La interfaz eSATAII es la especificación para SATAII externo. eSATAII permite disfrutar de las funciones SATAII ofreciendo velocidades de transferencia superiores a 3.0Gb/s.

Alternativas

También en SCSIW se está preparando un sistema en serie, que además es compatible con SATA, esto es, se podrán conectar discos SATA en una controladora SAS (Serial Attached SCSI). Estos conectores de 7 contactos también permiten una mayor circulación de aire: disminuyendo así la temperatura dentro del gabinete.

2.5.12 9-USB

El Bus de Serie Universal (USB, de sus siglas en inglés Universal Serial Bus) es una interfaz que provee un estándar de bus serie para conectar dispositivos a un ordenador personal (generalmente a un PC). Un sistema USB tiene un diseño asimétrico, que consiste en un solo servidor y múltiples dispositivos conectados en serie para ampliar la gama de conexión, en una estructura de árbol utilizando concentradores especiales. Se pueden conectar hasta 127 dispositivos a un sólo servidor, pero la suma debe incluir a los concentradores también, así que el total de dispositivos realmente usables es algo menor.

Fue creado en 1996 por siete empresas: IBM, Intel, Northern Telecom, Compaq, Microsoft, Digital Equipment Corporation y NEC.

El estándar incluye la transmisión de energía eléctrica al dispositivo conectado. Algunos dispositivos requieren una potencia mínima, así que se pueden conectar varios sin necesitar fuentes de alimentación extra. La mayoría de los concentradores incluyen fuentes de alimentación que brindan energía a los dispositivos conectados a ellos, pero algunos dispositivos consumen tanta energía que necesitan su propia fuente de alimentación. Los concentradores con fuente de alimentación pueden proporcionarle corriente eléctrica a otros dispositivos sin quitarle corriente al resto de la conexión (dentro de ciertos límites).

El diseño del USB tenía en mente eliminar la necesidad de adquirir tarjetas separadas para poner en los puertos bus ISA o PCI, y mejorar las capacidades plug-and-play permitiendo a esos dispositivos ser conectados o desconectados al sistema sin necesidad de reiniciar. Cuando se conecta un nuevo dispositivo, el servidor lo enumera y agrega el software necesario para que pueda funcionar.

El USB puede conectar periféricos como ratones, teclados, escáneres, cámaras digitales, impresoras, discos duros, tarjetas de sonido y componentes de red. Para dispositivos multimedia como escáneres y cámaras digitales, el USB se ha convertido en el método estándar de conexión. Para impresoras, el USB ha crecido tanto en popularidad que ha empezado a desplazar a los puertos paralelos porque el USB hace sencillo el poder agregar más de una impresora a un ordenador personal.

En el caso de los discos duros, el USB es poco probable que reemplace completamente a los buses como el ATA (IDE) y el SCSI porque el USB tiene un rendimiento un poco más lento que esos otros estándares. El nuevo estándar Serial ATA permite tasas de transferencia de hasta aproximadamente 150 MB por segundo. Sin embargo, el USB tiene una importante ventaja en su habilidad de poder instalar y desinstalar dispositivos sin tener que abrir el sistema, lo cual es útil para dispositivos de almacenamiento desinstalables. Hoy en día, una gran parte de los fabricantes ofrece dispositivos USB portátiles que ofrecen un rendimiento casi indistinguible en comparación con los ATA (IDE).

El USB no ha reemplazado completamente a los teclados AT y ratones PS/2, pero virtualmente todas las tarjetas madre o mother board de PC traen uno o más puertos USB. En el momento de escribir éste documento, la mayoría de las tarjetas madre traen múltiples conexiones:

USB 2.0.

Tarjeta PCI-USB 2.0.

Tarjeta PCI-USB 2.0.

El estándar USB 1.1 tenía dos velocidades de transferencia: 1.5 Mbit/s para teclados, ratón, joysticks, etc., y velocidad completa a 12 Mbit/s. La mayor ventaja del estándar USB 2.0 es añadir un modo de alta velocidad

de 480 Mbit/s. En su velocidad más alta, el USB compite directamente con FireWire.

Las especificaciones USB 1.0, 1.1 y 2.0 definen dos tipos de conectores para conectar dispositivos al servidor: A y B. Sin embargo, la capa mecánica ha cambiado en algunos conectores. Por ejemplo, el IBM UltraPort es un conector USB privado localizado en la parte superior del LCD de los computadoras portátiles de IBM. Utiliza un conector mecánico diferente mientras mantiene las señales y protocolos característicos del USB. Otros fabricantes de artículos pequeños han desarrollado también sus medios de conexión pequeños, y una gran variedad de ellos han aparecido. Algunos de baja calidad.

Una extensión del USB llamada "USB-On-The-Go" permite a un puerto actuar como servidor o como dispositivo - esto se determina por qué lado del cable está conectado al aparato. Incluso después de que el cable está conectado y las unidades se están comunicando, las 2 unidades pueden "cambiar de papel" bajo el control de un programa. Esta facilidad está específicamente diseñada para dispositivos como PDA, donde el enlace USB podría conectarse a un PC como un dispositivo, y conectarse como servidor a un teclado o ratón. El "USB-On-The-Go" también ha diseñado 2 conectores pequeños, el mini-A y el mini-B, así que esto debería detener la proliferación de conectores miniaturizados de entrada.



Clavija Macho USB tipo A



Cable USB tipo B



Tarjeta PCI-USB 2.0.

2.5.13 10-Firewire o IEEE 1394

El IEEE 1394 o FireWire o i.Link es un estándar multiplataforma para entrada/salida de datos en serie a gran velocidad. Suele utilizarse para la interconexión de dispositivos digitales como cámaras digitales y videocámaras a computadoras.

Historia

El FireWire fue inventado por Apple Computer a mediados de los 90, para luego convertirse en el estándar multiplataforma IEEE 1394. A principios de este siglo fue adoptado por los fabricantes de periféricos digitales hasta convertirse en un estándar establecido. Sony utiliza el estándar IEEE 1394 bajo la denominación i.Link, que sigue los mismos

estándares pero solo utiliza 4 conexiones, de las 6 disponibles en la norma IEEE 1394, suprimiendo las dos conexiones encargadas de proporcionar energía al dispositivo, que tendrá que proveerse de ella mediante una toma separada.

Ventajas de Firewire

Alcanzan una velocidad de 400 megabits por segundo.

Es hasta cuatro veces más rápido que una red Ethernet 100Base-T y 40 veces más rápido que una red Ethernet 10Base-T.

Soporta la conexión de hasta 63 dispositivos con cables de una longitud máxima de 425 cm.

No es necesario apagar un escáner o una unidad de CD antes de conectarlo o desconectarlo, y tampoco requiere reiniciar el ordenador.

Los cables FireWire se conectan muy fácilmente: no requieren números de identificación de dispositivos, conmutadores DIP, tornillos, cierres de seguridad ni terminadores.

FireWire funciona tanto con Macintosh como con PC.

FireWire 400 envía los datos por cables de hasta 4,5 metros de longitud. Mediante fibra óptica profesional, FireWire 800 puede distribuir información por cables de hasta 100 metros, lo que significa que podrías disparar ese CD hasta la otra punta de un campo de fútbol cada diez segundos. Ni siquiera necesitas ordenador o dispositivos nuevos para alcanzar estas distancias. Siempre que los dispositivos se conecten a un concentrador FireWire 800, puedes enlazarlos mediante un cable de fibra óptica supereficiente.

Características:

- Elevada velocidad de transferencia de información.
- Flexibilidad de la conexión.
- Capacidad de conectar un máximo de 63 dispositivos.
- Su velocidad hace que sea la interfaz más utilizada para audio y vídeo digital. Así, se usa mucho en cámaras de vídeo, discos duros, impresoras, reproductores de vídeo digital, sistemas domésticos para el ocio, sintetizadores de música y escáneres.

Existen dos versiones:

a) FireWire 400: tiene un ancho de banda 30 veces mayor que el USB 1.1.

b) IEEE 1394b, FireWire 800 o FireWire 2: duplica la velocidad del FireWire 400.

Así, para usos que requieran la transferencia de grandes volúmenes de información, resulta muy superior al USB.

- Arquitectura altamente eficiente. IEEE 1394b reduce los retrasos en la negociación, gracias a 8B10B (código que codifica 8 bits en 10 bits, que fue desarrollado por IBM y permite suficientes transiciones de reloj, la codificación de señales de control y detección de errores. El código 8B10B es similar a 4B5B de FDDI, el que no fue adoptado debido al pobre equilibrio de corriente continua), que reduce la distorsión de señal y aumenta la velocidad de transferencia. Proporciona, por tanto, una mejor vivencia como usuario.
- Da igual cómo conectes tus dispositivos entre ellos, FireWire 800 funciona a la perfección. Por ejemplo, puedes incluso enlazar a tu Mac la cadena de dispositivos FireWire 800 por los dos extremos para mayor seguridad durante acontecimientos en directo.
- Compatibilidad retroactiva. Los fabricantes han adoptado el FireWire para una amplia gama de dispositivos, como videocámaras digitales, discos duros, cámaras fotográficas digitales, audio profesional, impresoras, escáneres y electrodomésticos para el ocio. Los cables adaptadores para el conector de 9 contactos del FireWire 800 te permiten utilizar productos FireWire 400 en el puerto FireWire 800. FireWire 800 comparte las revolucionarias prestaciones del FireWire 400.
- Flexibles opciones de conexión. Conecta hasta 63 computadoras y dispositivos a un único bus: puedes incluso compartir una cámara entre dos Macs o PCs.
- Distribución en el momento. Fundamental para aplicaciones de audio y vídeo, donde un fotograma que se retrasa o pierde la sincronización arruina un trabajo, el FireWire puede garantizar una distribución de los datos en perfecta sincronía.
- Alimentación por el bus. Mientras el USB 2.0 permite la alimentación de dispositivos sencillos y lentos que consumen un máximo de 2,5 W, como un ratón, los dispositivos con FireWire pueden proporcionar o consumir hasta 45 W, más que suficiente para discos duros de alto rendimiento y baterías de carga rápida.
- Conexiones de enchufar y listo. No tienes más que enchufar un dispositivo para que funcione.

2.6 Compañías líderes fabricantes de discos duros.

TOSHIBA

Toshiba, compañía líder en la innovación y desarrollo de alta tecnología ofrece al mercado mundial productos eléctricos y electrónicos, Tiene reconocimiento internacional en la fabricación de discos duros para equipos de cómputo portátiles.

FUJITSU

Fujitsu, uno de los mayores fabricantes de discos duros a nivel mundial, ofreciendo soluciones de Discos Duros para las necesidades del mercado en informática móvil y equipos profesionales: Discos SCSI para servidores, Discos para portátiles UDMA 100. Gracias a que Fujitsu realiza fuertes inversiones en desarrollos tecnológicos y mejora de sus productos, podemos ofrecer a nuestros clientes un alto grado de seguridad en cuanto a la calidad de los productos se refiere, y una excelente relación Precio/Prestaciones.

Seagate

Seagate, Uno de los líderes a nivel mundial en el diseño, fabricación y marketing de unidades de disco duro, y ofrece productos para una amplia gama de aplicaciones empresariales, ordenadores de sobremesa, equipos informáticos móviles y productos de electrónica de consumo.

HITACHI DATA SYSTEMS

Hitachi Data Systems Empresa líder en soluciones de infraestructura de almacenamiento, software de administración de almacenamiento y servicios de consultoría de almacenamiento. Uno de los líderes a nivel mundial del sector de electrónica que ofrece una amplia variedad de sistemas, productos y servicios en los sectores del mercado tales como, sistemas de información, dispositivos electrónicos, sistemas eléctricos e industriales, productos para el consumidor, materiales y servicios financieros.

WD Western Digital

Western Digital es pionera en el almacenamiento de datos y es también uno de los líderes en la industria del disco duro. La empresa brinda soluciones rentables de almacenamiento para personas y organizaciones que recopilan, administran y utilizan información digital.



almacenamiento.

IBM además de ser líder en equipos de alto procesamiento es uno de los líderes y pioneros en la fabricación de discos duros y soluciones de



Hewlett Packard Otra empresa altamente reconocida por su diversidad de productos electrónicos y las soluciones que ofrece en procesamiento y almacenamiento.



Sun Microsystems Además de desarrollar tecnología de hardware y software también es líder en soluciones de almacenamiento central.



EMC, empresa altamente distinguida en el almacenamiento y protección de la información, ofrece soluciones competitivas de almacenamiento

2.7 Futuro de los discos duros.

Seagate acaba de ser el último fabricante de discos duros en anunciar que entra al mercado de los discos duros en estado sólido (SSD), en el 2008.

Seagate ya ofrece discos híbridos que tienen un disco duro tradicional junto a memoria flash que es utilizada para *cachear* datos frecuentemente utilizados, por lo que ya tiene algo de experiencia en el proceso, pero según la empresa discos SSD están en todas las vertientes de sus planes futuros, desde laptops hasta discos empresariales.

Lo cierto es que la única cosa que mantiene a los SSD fuera del mercado de los discos duros tradicionales es el precio, pues aparte de eso ofrecen ventaja en todos los aspectos posibles, desde bajo consumo de energía y cero ruido sin partes móviles, hasta mayor velocidad y mas confiabilidad.

En este capítulo hemos descrito a los discos en forma completa desde sus estructuras física y lógica hasta las formas a través de las cuales se conectan a una tarjeta madre que hace posible tener acceso al recurso de almacenamiento masivo que ofrece este dispositivo, en los capítulos posteriores seguiremos abordando algunos de estos conceptos así como más de las características, el manejo y la administración de los discos a través del sistema operativo Solaris.

3

Administración de Discos Duros utilizando el Sistema Operativo Solaris

En 1969 Ken Thompson y David Retchie desarrollaron un sistema operativo para la computadora llamada PDP-7 lo llamaron “UNICS” (Uniplexed Information and Computing System), eventualmente para modernizarlo el nombre fue cambiado a UNIX.

UNIX fue desarrollado principalmente en lenguajes de alto nivel solo una parte relativamente pequeña está escrita en código ensamblador (a esto se le llama kernel) lo cual lo hizo fácilmente transportable ya que solo se necesita reescribir el kernel para instalarlo en una nueva plataforma.

Después de que el sistema se dio a conocer oficialmente se distribuyó de forma gratuita a las universidades con propósitos educativos. Dos de las universidades que recibieron el código fuente fueron: la universidad de California en Berkeley (UCB) y la universidad de New South Wales (UNSW) en Sídney, Australia.

En 1974 dos estudiantes graduados de Berkeley: Hill Joy y Check Haley, comenzaron a realizar mejoras al sistema; un editor visual (vi), el compilador conocido como Pascal y la interface llamada lenguaje C. Le realizaron algunas correcciones y mejoraron el sistema.

En 1977, Bill Joy junto con Berkeley Software Distribution (BSD) distribuyeron la nueva versión mejorada, para 1978 ya existía una segunda versión llamada 2BSD la cual incorporaba modificaciones y mejoras. La tercera distribución (3bsd) fue instalada en una computadora de 32 bits VAX e incluía la implementación de memoria virtual. Después de esta tercera distribución la organización americana llamada United Defense Advanced Research Projects Agency (DARPA) comenzó la distribución de la primera parte de Internet.

En 1980, debutó 4BSD con una mejora del compilador Pascal. Para las próximas versiones y por cuestiones comerciales la BSD acordó en solo incrementar un número decimal al nombre y así la 5ª. Versión fue llamada 4.1BSD, para 1983 la versión 4.2BSD estaba lista la cual fue un éxito atribuido a la inclusión de el protocolo TCP/IP.

Hasta ahora las modificaciones y las versiones de BSD no se han detenido.

A principios de 1982, Hill Joy anunció que dejaría Berkeley para trabajar en una pequeña compañía en San Francisco llamada Sun Microsystems. Su trabajo comenzó sobre un sistema operativo llamado Sun Operating System (SunOS), la versión comercial de UNIX de Sun basada en la versión 4.2BSD. La cual tuvo una mejora en 1983 dando como resultado SunOS 1.0

En 1992, Sun migro su sistema basado en BSD SunOS a Solaris, El cual está basado en UNIX System V Release 4 (SVR4) de AT&T, empresa con la cual Sun colaboraba en el desarrollo de sistemas en concreto con el SVR4, contribuyendo con tecnologías como el NFS (Network File System).

Solaris usa una base de código común para las arquitecturas que soporta: SPARC y X86 (incluyendo AMD64/EM64T).

Solaris tiene una reputación de ser muy adecuado para el multiprocesamiento simétrico (SMP), soportando un gran número de CPUs. También ha incluido soporte para aplicaciones de 64 bits SPARC desde Solaris 7. Históricamente Solaris ha estado firmemente integrado con la plataforma hardware de Sun, SPARC, con la cual fue diseñado y promocionado como un paquete combinado. Esto proporcionaba frecuentemente unos sistemas más fiables pero con un coste más elevado que el del hardware de PC. De todas formas, también ha soportado sistemas x86 desde la versión Solaris 2.1 y la última versión, Solaris 10, ha sido diseñada con AMD64 en mente, permitiendo a Sun capitalizar en la disponibilidad de CPUs de 64 bits basadas en la arquitectura AMD64. Sun ha promocionado intensamente Solaris con sus estaciones de trabajo de nivel de entrada basadas en AMD64, así como con servidores que en 2006 varían desde modelos dual-core hasta modelos a 16 cores.

Los puntos siguientes explicarán la forma en la que el sistema operativo Solaris administra a los discos duros en forma física y lógica.

3.1 Particionamiento de Discos Duros.

Los discos duros están divididos lógicamente en porciones individuales conocidas como particiones. Las particiones en un disco son agrupaciones de cilindros que son comúnmente usados para organizar los datos según el uso de éstos.

Por ejemplo; una partición puede almacenar los archivos de sistema y programas, mientras que otra partición, en el mismo disco, almacena los archivos creados por los usuarios. El tener agrupados los cilindros en particiones facilita el respaldo y la organización de los datos así como también provee espacio de swap (memoria virtual).

En Solaris un disco puede ser dividido sólo en ocho particiones que son etiquetadas de la partición 0 a la partición 7.

Por convención, la partición 2 representa el disco entero, la cual contiene información importante acerca del disco, como el tamaño y el número total de cilindros disponibles para el almacenamiento de archivos y directorios.

Cada partición está definida por un cilindro de inicio y un cilindro final. Estos cilindros determinan el tamaño de la partición.

La figura 3.1 muestra un ejemplo de las particiones en un disco.

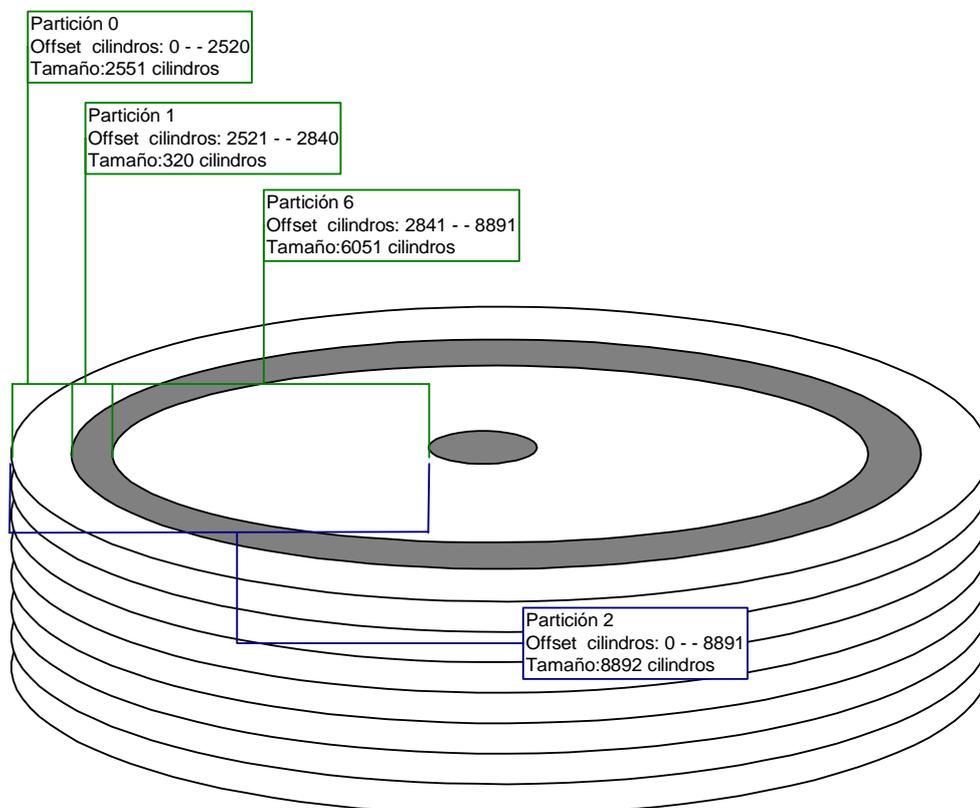


Figura 3.1 Particiones de un disco.

La tabla 3.1 muestra las particiones de disco y los diferentes sistemas de archivos que pueden contener.

Partición	Nombre	Función
0	/	Directorio raíz ,archivos de sistema
1	Swap	Area de swap
2	**	Disco total
5	/opt	Software Opcional
6	/usr	Programas y Sistemas ejecutables
7	/export/home	Directorios y archivos de usuario

Tabla 3-1 Particiones de Disco.

La figura 3.2 muestra una posible configuración para organizar datos. En el ejemplo el disco es dividido en particiones que organizan lógicamente los datos en el disco de arranque.

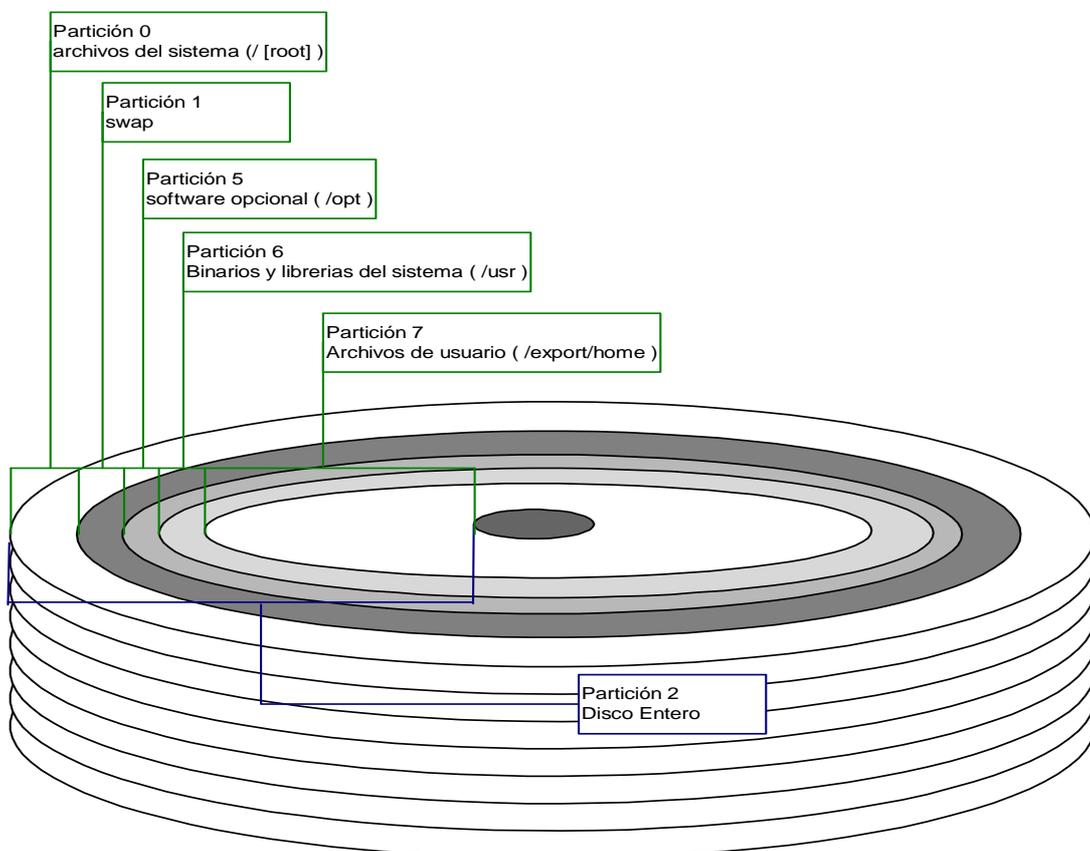


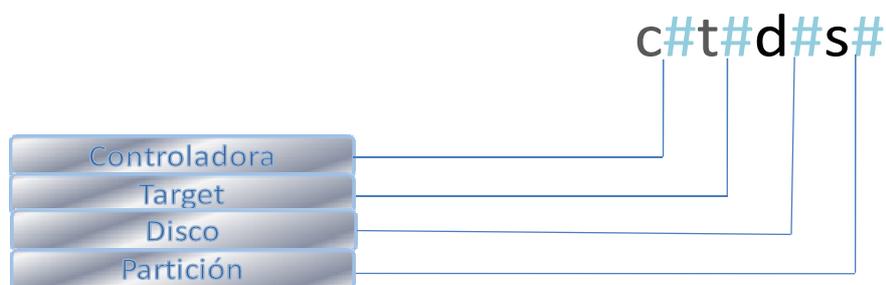
Figura 3.2

3.2 Nombre de las Particiones.

Típicamente una cadena de 8 caracteres representa el nombre completo de una partición. La cadena incluye el número de controladora, el número de target, el número de disco y el número de la partición.

Número de controladora	Identifica a la tarjeta (HBA) que controla la comunicación entre el sistema y la unidad de disco. La HBA se encarga del envío y la recepción de datos e instrucciones del dispositivo. El número de controladora se asigna en orden secuencial, por ejemplo: c0, c1, c2, ...,cn.
Identificador Target	El número de target, como por ejemplo t0, t1, t2,...,tn, corresponden a una dirección única de hardware que es asignada para cada disco, cinta o CD-ROM. Algunas unidades de discos externos tienen un switch de direcciones y algunos discos internos tienen pins que pueden ser colocados de tal forma que asignan el número de target del disco.
Número de disco	El número de disco es también conocido como el número de unidad lógica (LUN). Este número refleja el número de disco de acuerdo a la localidad del target.
Número de partición	Un número de partición entre el 0 y el 7. (como ya se mencionó Solaris solo permite realizar 8 particiones en el disco)

La figura 3.3 muestra la cadena de ocho caracteres que representa el nombre completo de una partición de disco.



La figura 3.4 muestra la configuración en una arquitectura SCSI.

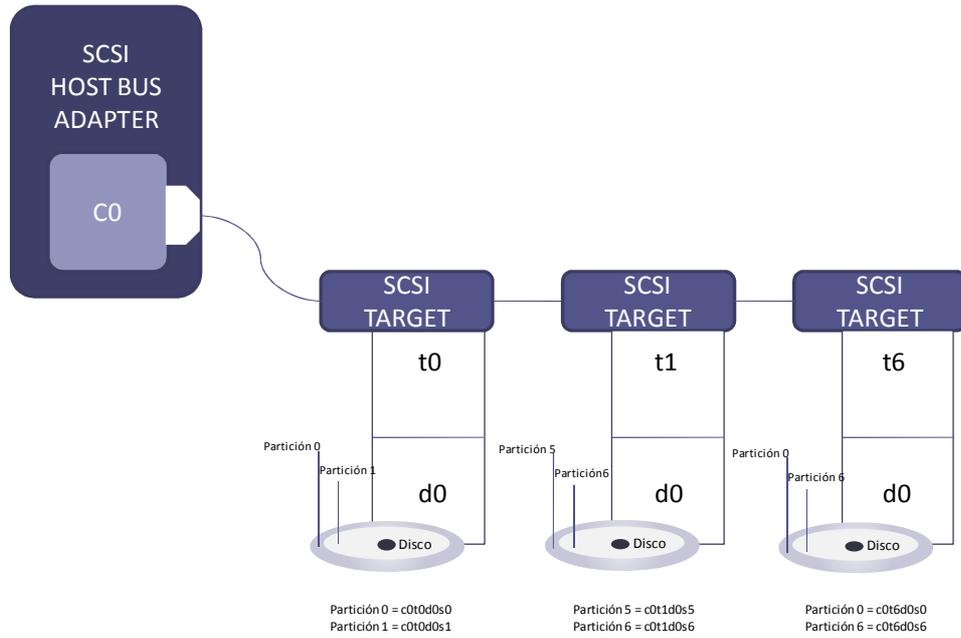


Figura 3.4

La figura 3.5 muestra la configuración en una arquitectura IDE.

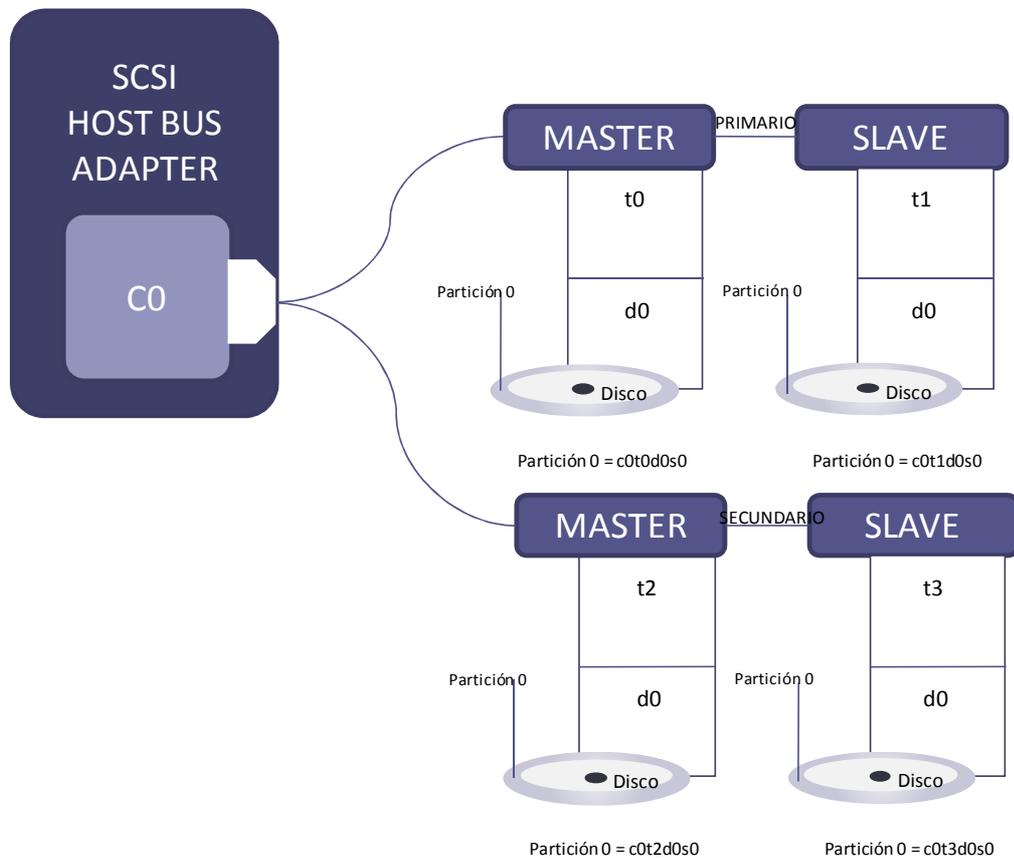


Figura 3.5

3.3 Introducción a las convenciones de nombres en Solaris.

En Solaris, todos los dispositivos están representados por tres nombres, dependiendo de cómo están referenciados:

- Nombre lógico
- Nombre físico
- Nombre de instancia

3.3.1 Nombre Lógico.

En el sistema operativo Solaris los nombres lógicos de los discos son ligas simbólicas a el nombre físico del dispositivo el cual está almacenado en el directorio `/devices`. Los nombres lógicos son usados para referirse al dispositivo cuando se dan instrucciones en la línea de comandos. Todos los nombres lógicos se guardan en el directorio `/dev`. Este tipo de nombre contiene el número de controladora, el número de target, el número de disco y el número de partición (ya antes definidos).

Cada unidad de disco tiene una entrada en ambos directorios en `/dev/dsk` y `/dev/rdisk`.

Por ejemplo:

El nombre **c0t0d0s0**. Identifica a la partición 0 del disco 0 conectado a con el target 0 en la controladora 0.

3.3.2 Nombres físicos.

Los nombres físicos únicamente identifican la posición física del dispositivo en el sistema y están almacenados en el directorio `/devices`.

Estos nombres contienen la información del hardware, representados por una serie de nombres, separados por guiones, que indican la ruta del dispositivo.

Por ejemplo:

.../ ../devices/pci@1f,0/pci@1, 1/ide@3/dad@0, 0:a

Utilizamos el comando `# ls -l /dev/dsk/c0t0d0s0` como ejemplo. Nótese que se utilizó el nombre lógico para ejecutar el comando.

El manejador del dispositivo (driver) es el software que permite la comunicación con el dispositivo. Este software debe estar disponible en el kernel del para que el sistema pueda usar el dispositivo.

Durante el inicio del sistema el Kernel identifica la localización del dispositivo y lo asocia a un nodo con una dirección de la forma `nombredelnodo@dirección` y este es el nombre físico del dispositivo.

3.3.3 Nombres de instancia.

Los nombres de instancia son nombres abreviados asignados por el kernel para cada dispositivo en el sistema.

Por ejemplo:

sdn

Donde **sd** es el nombre del disco y **n** es el número, como **sd0** para la primera unidad de disco SCSI.

3.4 Partición del disco duro.

El comando `format` es una herramienta de administración del sistema usada principalmente para preparar los discos.

Aunque también realiza una variedad de actividades de manejo de disco. La función principal de la utilería es dividir el disco en particiones.

3.4.1 Fundamentos de la partición de discos.

Para dividir un disco en particiones es necesario:

1. Identificar el disco correcto.
2. Planificar la disposición del disco.
3. Usar la utilería `format` para dividir el disco en particiones.
4. Etiquetar el disco con la información de la nueva partición.

Solo el usuario `root` puede usar la utilería o comando `format`. Si un usuario sin permisos intenta utilizarla recibirá un mensaje de error.

3.4.2 Espacio en disco y condiciones indeseables.

Las particiones de disco están definidas por `offset` y un tamaño de cilindros. El `offset` es la distancia al cilindro 0. La figura 3.6 muestra un ejemplo de los tamaños de particiones de disco y su `offset`.

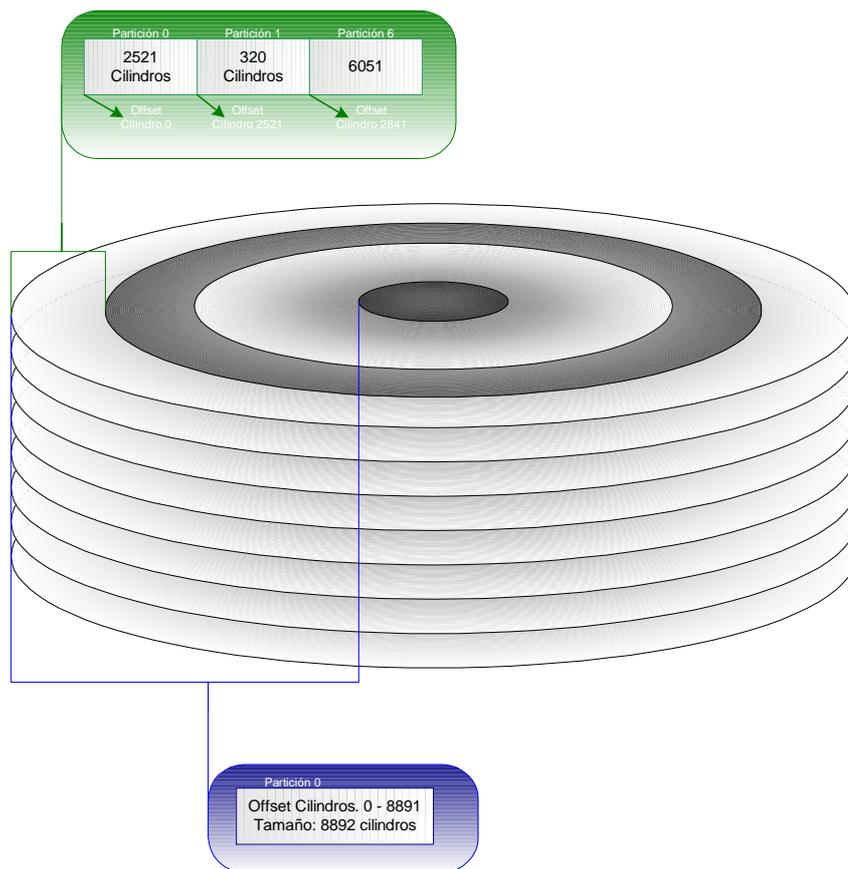


Figura 3.6

El offset para la partición 0 es 0 cilindros, y su tamaño es 2521 cilindros. La partición 0 comienza en el cilindro 0 y termina en el cilindro 2520.

El offset para la partición 1 es 2521 cilindros, y su tamaño es de 320 cilindros. La partición 1 comienza en el cilindro 2521 y termina en el cilindro 2840.

El offset para la partición 6 es 2841 cilindros, y su tamaño es 6051 cilindros. La partición 6 comienza en el cilindro 2841 y termina en el último cilindro disponible, el cual es el cilindro 8891.

3.4.3 Desperdicio de espacio en disco.

El desperdicio en espacio ocurre cuando uno o mas cilindros no están asignados a una partición. La figura 3.7 muestra un cilindro que no está asignado.

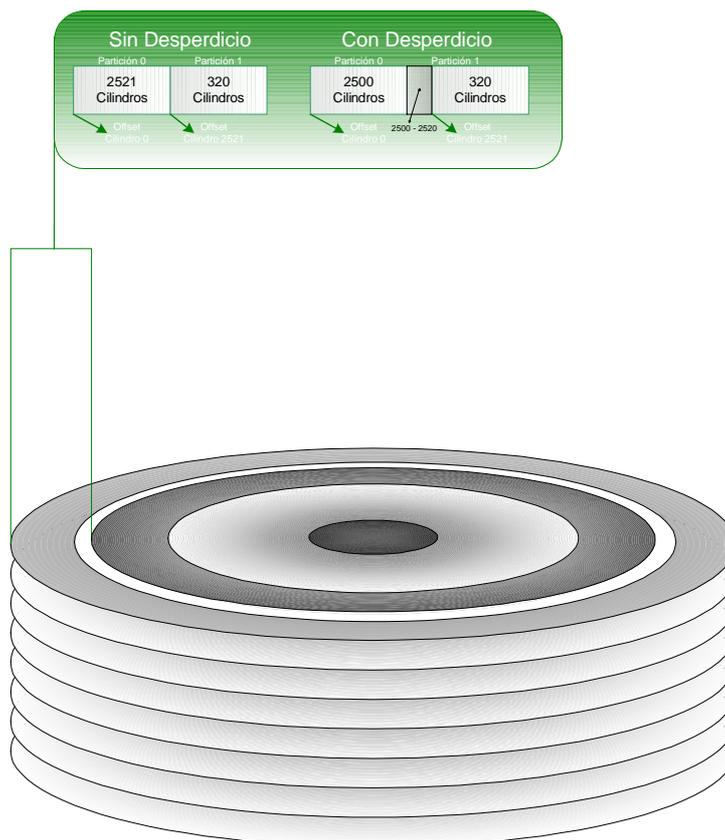


Figura 3.7

Los cilindros 2500 al 2520 no se están utilizando porque no están asignados a una partición.

El desperdicio de espacio ocurre durante la partición de los discos cuando uno o más cilindros no se asignan a una partición. Esto puede ocurrir intencionalmente o por accidente. Los cilindros que no están dentro de una partición pueden ser asignados a una después.

3.4.4 Superposición o traslape de particiones.

La superposición o traslape de particiones es cuando una cantidad de cilindros que puede ser desde uno hasta n están asignados a más de una partición a la vez. La figura 3.8 muestra un ejemplo de la superposición de particiones.

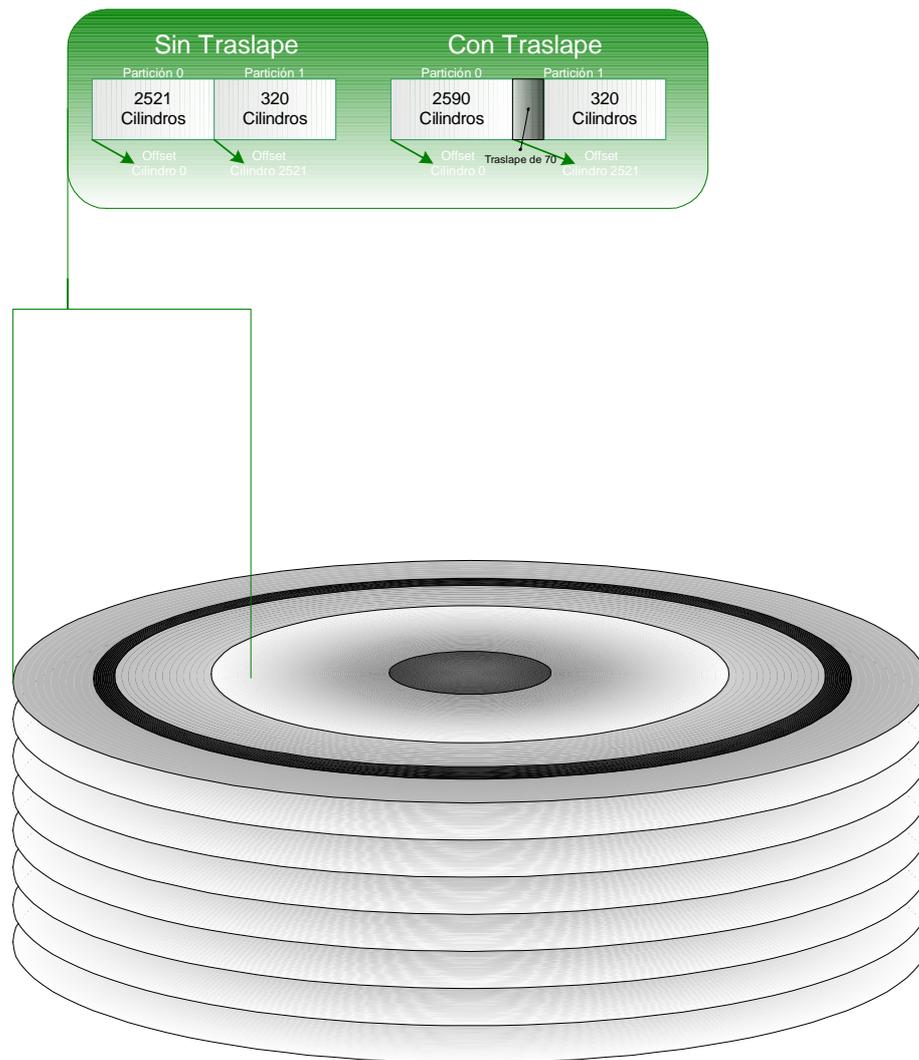


Figura 3.8

En la figura se muestra cómo se sobreponen los cilindros 2521 al 2590 en dos particiones.

El comando `format` ofrece una advertencia si se quiere utilizar un cilindro de alguna otra partición. Ya que esto ocurre si se incrementa el tamaño de laguna partición y no se modifica la siguiente.

3.5 Tablas de partición de discos.

Etiquetas de los discos.

La etiqueta del disco es una parte en donde se guarda información acerca de la controladora del disco, la distribución y las particiones del mismo. A esta información también se le llama tabla de contenidos del volumen o VTOC (Volume Table of Contents).

Etiquetar discos es escribir información de las particiones sobre el propio disco. Si se comete un error al etiquetar un disco una vez que se han definido las particiones se pierde la información de tales particiones.

Una parte importante de la etiqueta del disco es la tabla de particiones la cual contiene el total del tamaño y los límites de las particiones en cilindros.

La figura 3.9 muestra la relación entre la etiqueta en el disco, la etiqueta en memoria y la etiqueta predefinida en el archivo /etc/format.dat (este archivo es leído por el sistema cuando se utiliza el comando format).

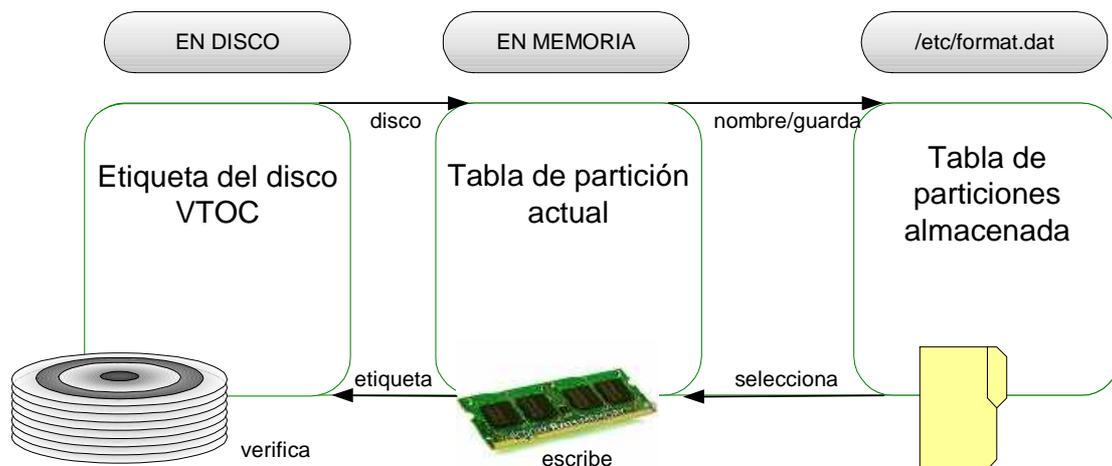


Figura 3.9

3.6 Usos del comando format.

La utilidad `format` tiene organizados sus comandos en dos partes si se introduce `format` en la línea de comandos aparece la lista de los comandos que permiten entre otra funciones seleccionar un disco, seleccionar una partición, salvar un nuevo disco y definir sus particiones, y escribir la etiqueta.

La segunda parte de comandos aparece al introducir `partition` en seguida del prompt `format>`. Este grupo de comandos permite definir las características de cada partición, desplegar la tabla de particiones actual y escribir el mapa de la partición y la etiqueta del disco, entre otras funciones.

La tabla 3.2 describe la terminología de particionamiento de discos.

	Término	Descripción
Part		Indica el numero de la partición (0 - 7)
Tag		Un valor que indica cómo será usada la partición 0= unassigned 1=boot 2=root 3=swap 4=usr 5=backup 6=stand 8=home
Flag		00 wm = La partición del disco es de escritura y montable 01 wu = La partición del disco es de escritura y no montable. <i>Este es el estado por default para una partición dedicada a un área de swap</i> 10 rm = La partición del disco es de lectura y montable 11 ru = La partición del disco es de lectura y no montable.
Cylinders		El cilindro de inicio y cilindro final de una partición del disco
Size		El tamaño de la partición Mbytes (MB), Gbytes (GB), bloques (b) o cilindros (c).
Blocks		El número total de cilindros y el número total de sectores por partición.

Tabla 3.2

A continuación se muestra un ejemplo de partición de discos utilizando el comando format.

Los siguientes pasos muestran cómo dividir un disco en particiones.

1. Como el usuario root escribir el comando format en el prompt, y presionar Enter.

```
#format
Searching for disks...done
```

AVAILABLE DISK SELECTIONS:

```
0. c0t0d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
  /pci@50,600000/pci@0/scsi@1/sd@0,0
1. c0t1d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
  /pci@50,600000/pci@0/scsi@1/sd@1,0
2. c1t0d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
  /pci@54,600000/pci@0/scsi@1/sd@0,0
3. c1t1d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
  /pci@54,600000/pci@0/scsi@1/sd@1,0
Specify disk (enter its number): Specify disk (enter its number):
```

El comando format busca todos los discos conectados al sistema. Por cada disco que encuentra despliega el nombre lógico del dispositivo, el nombre asignado por SUN, los parámetros físicos y el nombre físico.

Se selecciona el disco a particionar, en este caso se selecciona el primer disco (disco 0)

Posteriormente aparece el siguiente menú utilizado para ver, cambiar o modificar particiones, incluye las siguientes instrucciones.

```
Specify disk (enter its number): 0
selecting c1t1d0
[disk formatted]
Warning: Current Disk has mounted partitions.
```

FORMAT MENU:

```
disk      - select a disk
type      - select (define) a disk type
partition - select (define) a partition table
current   - describe the current disk
format    - format and analyze the disk
repair    - repair a defective sector
label     - write label to the disk
analyze   - surface analysis
defect    - defect list management
backup    - search for backup labels
verify    - read and display labels
save      - save new disk/partition definitions
inquiry   - show vendor, product and revision
```

```
volname - set 8-character volume name
!<cmd> - execute <cmd>, then return
quit
format>
```

Donde:

Partition	Despliega el menú de la partición
Label	Escribe como está definida la partición actual en la etiqueta del disco
Verify	Lee y despliega la etiqueta del disco
Quit	Sale del comando format

Si se escribe la opción partition en el prompt de format, aparece el menú Partition.

```
format> partition
```

PARTITION MENU:

```
0 - change `0' partition
1 - change `1' partition
2 - change `2' partition
3 - change `3' partition
4 - change `4' partition
5 - change `5' partition
6 - change `6' partition
7 - change `7' partition
select - select a predefined table
modify - modify a predefined partition table
name - name the current table
print - display the current table
```

```
label - write partition map and label to the disk
!<cmd> - execute <cmd>, then return
quit
partition>
```

El menú partition permite realizar las siguientes funciones:

0-7	Especifica el offset y el tamaño de las 8 particiones
Select	Selecciona una tabla de particiones predefinida en el archivo /etc/format.dat (archivo leído por el sistema cuando se trabaja con el comando format)
Modify	Modifica la tabla de particionamiento
Name	Da un nombre para identificar a la tabla de particiones en el archivo /etc/format.dat
Print	Despliega la tabla de particiones actual que se tiene en memoria.
Label	Escribe la tabla de particiones actual en la etiqueta del disco
!<cmd>	Salte del menú y ejecuta un comando del shell.
Quit	Salte al menú anterior

0. Si se escribe la opción print después del prompt del directorio partition se despliega la etiqueta del disco que se copio a RAM cuando se utilizó el comando format.

```
partition> print
```

El nombre de la tabla aparece entre paréntesis en la primera línea de la tabla.

```
partition> print
```

```
Current partition table (original):
```

```
Total disk cylinders available: 14087 + 2 (reserved cylinders)
```

Part	Tag	Flag	Cylinders	Size	Blocks
0	root	wm	0 - 3091	15.00GB	(3092/0/0) 31464192
1	swap	wu	3092 - 3916	4.00GB	(825/0/0) 8395200
2	backup	wm	0 - 14086	68.35GB	(14087/0/0) 143349312
3	var	wm	3917 - 4947	5.00GB	(1031/0/0) 10491456
4	unassigned	wm	4948 - 6184	6.00GB	(1237/0/0) 12587712
5	unassigned	wm	6185 - 7421	6.00GB	(1237/0/0) 12587712
6	unassigned	wm	7422 - 14086	32.34GB	(6665/0/0) 67823040
7	unassigned	wm	0	0	(0/0/0) 0

```
partition>
```

Las columnas en la tabla tienen los siguientes significados:

Part	El número de la partición
Tag	La etiqueta dada
Flag	La bandera predefinida
Cylinders	El primer y último cilindro de la partición
Size	El tamaño de la partición en bloques (b), cilindros (c), Mega Bytes (MB) o Giga Bytes (GB)
Blocks	El total de cilindros y sectores por partición

5. Siguiendo con el ejemplo si seleccionamos la partición 0 tecleando 0. Así:

```
partition> 0
```

Part	Tag	Flag	Cylinders	Size	Blocks
0	root	wm	0 - 3091	15.00GB	(3092/0/0) 31464192

6. Si deseamos modificar el identificador Tag, debemos teclear un símbolo ? seguido de un enter, cuando aparezca un enunciado como el siguiente:

```
Enter partition id tag[root]: ?
```

```
Expecting one of the following: (abbreviations ok):
```

unassigned	boot	root	swap
usr	backup	stand	var
home	alternates	reserved	

7. Para el ejemplo que tenemos elegimos alternates seguido de un enter.

```
Enter partition id tag [unassigned] : alternates
```

Para modificar los permisos y queremos ver antes las opciones disponibles,

Es necesario teclear el símbolo ? cuando un enunciado como el siguiente aparece.

```
Enter partition permission flags [wm] : ?
```

Aparece una lista de posibles permisos para la partición.

```
Expecting one of the following : (abbreviations ok):
```

wm	-read.write, mountable
wu	-read-write, unmountable
rm	-read-only, mountable
ru	-read-only, unmountable

Enter partition permission flags [wm] :

9. Si damos un enter se acepta la bandera por default que es wm.

10. Se presentan otras opciones como en la que se debe elegir el cilindro en el que iniciará la partición. Si damos un enter la elección será por default el cilindro 0.

Enter new starting cyl [0] : <enter>

11. Otra opción es la de especificar el tamaño para la partición. Para el ejemplo elegimos 400 mega bytes de la siguiente manera.

Enter partition size [0b, 0c, 0e, 0.00mb, 0.00gb]: 400mb

12. Si imprimimos en pantalla nuevamente la tabla (tecleando la palabra print seguido de un enter en el prompt de la herramienta partition), tendremos la información modificada.

```
partition> print
```

```
current partition table (unnamed):
```

```
Total disk cilindres available: 1965 + 2 (reserved cylinders)
```

Part	Tag	Flag	Cylinders	Size	Blocks
0	alternates	wm	0-602	400.43MB	(603/0/0) 820080
1	unassigned	wm	0	0	(0/0/0) 0
2	backup	wu	0-1964	1.27GB	(1965/0/0) 2672400
3	unassigned	wm	0	0	(0/0/0) 0
4	unassigned	wm	0	0	(0/0/0) 0
5	unassigned	wm	0	0	(0/0/0) 0
6	unassigned	wm	0	0	(0/0/0) 0
7	unassigned	wm	0	0	(0/0/0) 0

La tabla muestra los cambios realizados a la partición 0.

Ahora para seguir con el ejemplo mostramos como es posible modificar el cilindro de inicio para la partición 1 así podemos ver los alcances de la herramienta partition.

13. Como ya lo vimos vasta con teclear el número de la partición para acceder a sus opciones.

partition > 1

Part	Tag	Flag	Cylinders	Size	Blocks
1	unassigned	wm	0	0 (0/0/0)	0

14. Debemos elegir el tag que le asignaremos para el ejemplo seleccionamos swap.

Enter partition id tag [unassigned] : swap

15. Los permisos se los asignaremos con la abreviatura wu.

Enter partition permission flags [wm] : wu

16. En el siguiente enunciado es donde modificamos el cilindro de inicio para la partición que elegimos.

Enter new starting cyl [0] : 603

17. Modificamos tambien su tamaño.

Enter partition size [0b, 0c, 0e, 0.00mb, 0.00gb]: 60mb

18. Nuevamente imprimimos la tabla para revisar los cambios. (print en el prompt de partition).

partition> print

current partition table (unnamed):

Total disk cylinders available: 1965 + 2 (reserved cylinders)

Part	Tag	Flag	Cylinders	Size	Blocks
0	alternates	wm	0-602	400.43MB (603/0/0)	820080
1	swap	wu	603-693	60.43MB (91/0/0)	123760
2	backup	wu	0-1964	1.27GB (1965/0/0)	2672400
3	unassigned	wm	0	0 (0/0/0)	0
4	unassigned	wm	0	0 (0/0/0)	0
5	unassigned	wm	0	0 (0/0/0)	0
6	unassigned	wm	0	0 (0/0/0)	0
7	unassigned	wm	0	0 (0/0/0)	0

La tabla muestra los cambios realizados a la partición 1.

Notemos que el cilindro de inicio de la partición 1 es el número inmediato superior al del final de la partición 0 con esto evitamos la superposición o el traslape del que ya habíamos hablado anteriormente.

Ahora vamos a modificar el cilindro de inicio para la partición 7 así vemos otras opciones de la utilería (partition).

19. Seleccionamos la partición.

partition > 1

Part	Tag	Flag	Cylinders	Size	Blocks
7	unassigned	wm	0	0	(0/0/0) 0

20. Debemos elegir el tag que le asignaremos para el ejemplo seleccionamos home.

Enter partition id tag [unassigned] : home

21. Los permisos se los asignaremos por default.

Enter partition permission flags [wm] : <enter>

22. En el siguiente enunciado es donde modificamos el cilindro de inicio para la partición que elegimos.

Enter new starting cyl [0] : 694

23. Modificamos también su tamaño. Tecleamos el símbolo \$ esto hará que se la asigne el espacio restante a la partición.

Enter partition size [0b, 0c, 0e, 0.00mb, 0.00gb]: \$

24. Nuevamente imprimimos la tabla para revisar los cambios. (print en el prompt de partition).

```
partition> print
```

```
current partition table (unnamed):
```

```
Total disk cylinders available: 1965 + 2 (reserved cylinders)
```

Part	Tag	Flag	Cylinders	Size		Blocks
0	alternates	wm	0-602	400.43MB	(603/0/0)	820080
1	swap	wu	603-693	60.43MB	(91/0/0)	123760
2	backup	wu	0-1964	1.27GB	(1965/0/0)	2672400
3	unassigned	wm	0	0	(0/0/0)	0
4	unassigned	wm	0	0	(0/0/0)	0
5	unassigned	wm	0	0	(0/0/0)	0
6	unassigned	wm	0	0	(0/0/0)	0
7	home	wm	694-1964	844.02MB	(1271/0/0)	1728560

Si hacemos la suma de los cilindros en la columna Blocks de las particiones 0, 1 y 7 tenemos el total de cilindros que contiene la partición 2.

25. Una vez que hemos revisado la tabla y que estamos seguros que no existen errores podemos etiquetar el disco escribiendo la palabra label.

```
partition> label
```

```
Ready to label disk, continue? Y
```

```
partition>
```

3.7 Guardar la tabla de particiones.

Ya vimos como crear una tabla de particiones ahora es necesario saber cómo almacenarla ya que Solaris nos da la opción de guardar las tablas en un archivo (/etc/format.dat) para después utilizarlas en otros discos del mismo tipo sin necesidad de volver a crearlas.

Para salvar o guardar una tabla de particiones personalizada, podemos seguir los siguientes pasos:

En el menú de la utilidad partition:

1. Escribimos la palabra name para poder asignar el nombre a la tabla el cual debe ser único a menudo se le asigna el nombre del fabricante del disco.

```
partition> name
```

```
Enter table name (remember quotes): SUN1.3G
```

Si el nombre se compone de dos o más palabras debe de ir entre comillas.

2. Salimos del menú con la palabra quit.

```
partition> quit
```

3. Escribimos la palabra save para guardar la tabla.

```
format> save
```

```
Saving new partition definition
```

Es necesario indicar el nombre de la ruta completo.

```
Enter file name ["/.format.dat"] : /etc/format/dat
```

3.8 Uso de una tabla de particiones personalizada.

Como lo mencionamos Solaris nos permite utilizar una tabla de particiones creada con opciones personales en otros discos. Vamos a ejemplificar éste uso de la tabla, para ello es necesario entrar a la herramienta format y realizar lo siguiente:

1. Escribir la palabra partition para entrar dentro de la utilería.

```
format> partition
```

2. Escribimos select para desplegar la lista de tablas creadas, elegimos alguna seleccionando el número asignado.

```
partition> select
```

```
0. SUN1.3G
```

```
1. original
```

```
2. SUN4.2
```

Supongamos que queremos la tabla SUN1.3G entonces:

```
Specify table (enter its number)[3] : 0
```

3. Etiquetamos el disco con la tabla de particiones seleccionada.

```
partition> label
```

```
Ready to label disk, continue? yes
```

4. Salimos del menu partition.

```
partition> quit
```

5. Leemos la nueva etiqueta del disco.

```
format> verify
```

```
Primary label contents:
```

```
Volume name = <      >
```

```
ascii name   = <SUN1.3G cyl 1965 alt 2 hd 17 sec 80>
```

```
pcyl        = 3500
```

```
ncyl        = 1965
```

```
acyl        =    2
```

```
nhead       =   17
```

```
nsec        =   80
```

Part	Tag	Flag	Cylinders	Size		Blocks
0	alternates	wm	0-602	400.43MB	(603/0/0)	820080
1	swap	wu	603-693	60.43MB	(91/0/0)	123760
2	backup	wu	0-1964	1.27GB	(1965/0/0)	2672400
3	unassigned	wm	0	0	(0/0/0)	0
4	unassigned	wm	0	0	(0/0/0)	0
5	unassigned	wm	0	0	(0/0/0)	0
6	unassigned	wm	0	0	(0/0/0)	0
7	home	wm	694-1964	844.02MB	(1271/0/0)	1728560

6. Salimos del comando format.

```
format > quit
```

3.9 Manejo de las etiquetas de disco.

Como ya lo hemos mencionado cada disco tiene asociada una etiqueta en la que se almacena información sobre la controladora de disco, las particiones y la organización del disco.

Existen dos métodos para localizar y ver la etiqueta del disco o VTOC.

- Usando el comando verify dentro de la utilidad format.
- Desde la línea de comandos utilizamos la instrucción prtvtoc.

3.9.1 Uso del comando verify.

El comando verify permite ver la etiqueta del disco cuando se está dentro de la utilidad format. Vamos a hacer un ejemplo de cómo utilizarla.

1. Estando en el prompt de format escribimos la palabra verify.

```
format> verify
```

```
Primary label contents:
```

```
Volume name      = <          >
```

```
ascii name       = < SUN1.3G cyl 1965 alt 2 hd 17 sec 80 >
```

```
pcyl             = 3500
```

```
ncyl            = 1965
```

```
acyl            = 2
```

```
nhead           = 17
```

```
nsect           = 80
```

Part	Tag	Flag	Cylinders	Size		Blocks
0	root	wm	0-602	400.43MB	(603/0/0)	820080
1	swap	wu	603-693	60.43MB	(91/0/0)	123760
2	backup	wu	0-1964	1.27GB	(1965/0/0)	2672400
3	unassigned	wm	0	0	(0/0/0)	0
4	unassigned	wm	0	0	(0/0/0)	0
5	unassigned	wm	0	0	(0/0/0)	0
6	unassigned	wm	0	0	(0/0/0)	0
7	home	wm	694-1964	844.02MB	(1271/0/0)	1728560

2. Escribimos quit a solamente la letra q para salir nuevamente al comando format.

Uso del comando prtvtoc.

El comando prtvtoc permite ver la tabla de contenidos de un disco VTOC (volume table of content) desde la línea de comandos. A continuación veremos un ejemplo para ilustrar el uso de este comando:

```
# prtvtoc /dev/rdisk/c1t1d0s2
* /dev/rdisk/c1t1d0s2 partition map
*
* Dimensions:
* 512 bytes/sector
* 424 sectors/track
* 24 tracks/cylinder
* 10176 sectors/cylinder
* 14089 cylinders
* 14087 accessible cylinders
*
* Flags:
* 1: unmountable
* 10: read-only
*
*          First  Sector  Last
* Partition Tag  Flags  Sector  Count  Sector  Mount Directory
0      2  00          0 31464192 31464191
1      3  01 31464192 8395200 39859391
2      5  00          0 143349312 143349311
3      7  00 39859392 10491456 50350847
4      0  00 50350848 12587712 62938559
5      0  00 62938560 12587712 75526271
6      0  00 75526272 67823040 143349311
```

La información de la etiqueta del disco incluye los siguientes campos:

Dimensions	Describe las dimensiones lógicas del disco
Flags	Describe las banderas que son listadas en la tabla de partición
Partition	Es el numero de la partición
Tag	Un valor para indicar el uso que se le dará a la partición
Flags	00 significa lectura/escritura, montable; 01 significa lectura/escritura desmontable ; 10 es únicamente lectura.
First Sector	Define el primer sector de la partición
Sector Count	Define el número total de sectores en la partición
Last Sector	Define el último sector de la partición
Mount Directory	Indica si la partición está asociada a un punto de montaje

3.10 Reetiquetando un disco.

Al guardar la tabla de contenidos o la etiqueta de un disco utilizando el comando `prtvtoc` podemos reetiquetar el disco esto lo podemos hacer si usamos el comando `fmthard` en caso de que ocurra alguna de las siguientes situaciones:

- La tabla de contenidos VTOC ha sido eliminada.
- Accidentalmente se cambió la información de la partición en el disco y no se hizo un respaldo de la etiqueta en el archivo `/etc/format.dat`.

Para guardar o salvar la VTOC en un archivo ejecutamos el siguiente comando:

```
# prtvtoc /dev/rdisk/c1t0d0s2 > /vtoc/c1t0d0
```

El comando fmthard.

Para re-etiquetar un disco se puede guardar la salida del comando prtvtoc en un archivo en otro disco y usarlo como el argumento del comando fmthard. Ejemplo:

```
# fmthard -s <datafile> /dev/rdisk/c##t#d#s2
```

3.11 Manejo del sistema de archivos (*file systems*) en Solaris.

Un sistema de archivos (*que a partir de ahora lo llamaremos file systems por término técnico*) es un conjunto de archivos y directorios organizados. El sistema operativo Solaris maneja tres diferentes tipos de sistema de archivos:

- Basado en disco
- Distribuidos
- Pseudo

3.12 File Systems Basados en Disco.

Los file systems basados en disco están en los discos duros, CD-ROM's, diskettes, Vds., etc. Los siguientes son algunos ejemplos de este tipo:

- ufs - El sistema de archivos de UNIX que es también el de Solaris.
- hfs - El High Sierra que es un sistema de archivos con un propósito especial desarrollado para el uso del CD-ROM.
- pcfs - Una implementación de UNIX para el sistema operativo DOS el llamado FAT32.
- udfs - El Formato de Discos Universal para unidades de almacenamiento ópticos como el DVD y el CD-ROM. Permite el intercambio de datos universales y soporta operaciones de lectura y escritura.
- Vxfs - Sistema de archivos utilizado por Veritas File System.

3.13 File Systems Distribuidos

En este tipo de file systems el acceso a los datos se proporciona a través de la red.

- NFS- Este tipo de file systems permite a los usuarios compartir archivos entre varios tipos de sistemas sobre la una red. El sistema de archivos NFS permite acceder a los archivos remotos exactamente igual que si fueran locales.

3.14 Pseudo file systems.

Los pseudos file systems están basados en memoria (RAM). Estos file system aseguran un mejor rendimiento al sistema, además del acceso a la información del kernel . Los Pseudo file systems incluyen:

- tmpfs – Es un file system temporal, guarda archivos en memoria. lo cual elimina la sobre escritura en los file systems basados en disco. El file system tmpfs se crea y destruye cada vez que el sistema es reiniciado.
- swapfs – El file system swap es usado por el kernel para administrar el espacio de swap en los discos.
- fdfs – El archivo descriptor del file system proporciona nombres explícitos para archivos abiertos utilizando archivos descriptores (por ejemplo, /dev/fd/0, /dev/fd/1 , /dev/fd/2) in the /dev/fd directory.
- procfs – El file system de procesos contiene una lista de procesos activos en el directorio /proc. Los procesos están listados por número de proceso. La información en el directorio es usado por comandos, como el comando ps.
- mntfs – El file system de montaje proporciona información de lectura-escritura del kernel de los file system localmente montados localmente en el sistema.

3.15 Creación de un file system ufs.

Los usuarios y el sistema operativo Solaris ven al file system ufs de una forma distinta. Para los usuarios, un file system parece una colección de archivos y directorios usados para guardar y organizar los datos para ser accedidos por el sistema y sus usuarios. Para el sistema operativo, un file system es una colección de estructuras de control y bloques de datos que ocupa un espacio definido por una partición, esta colección guarda datos y los administra.

Solaris almacena datos en una jerarquía lógica de archivos que a menudo se compone de varios file systems. Esta jerarquía está referenciada como una jerarquía de directorios Solaris.

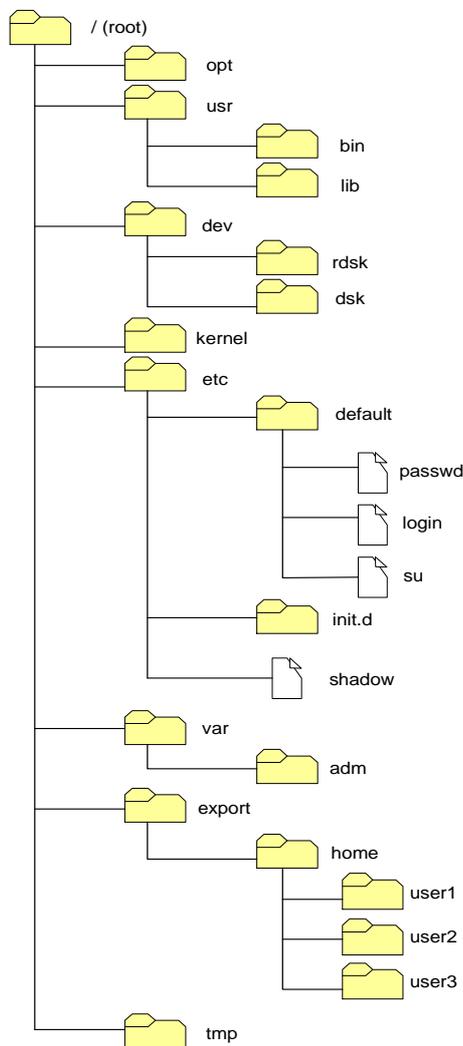


Figura 3.10 Jerarquía de directorios de Solaris.

La figura 3.10 muestra la jerarquía de Solaris comenzando con el directorio /root. (la figura no muestra la jerarquía completa)

Un file system ufs se crea en una partición de disco antes de ser usado por Solaris. La creación de un file system ufs en una partición permite a Solaris almacenar directorios y archivos UNIX.

La figura 3.11 muestra la forma en la que están localizados los file systems ufs sobre varias particiones.

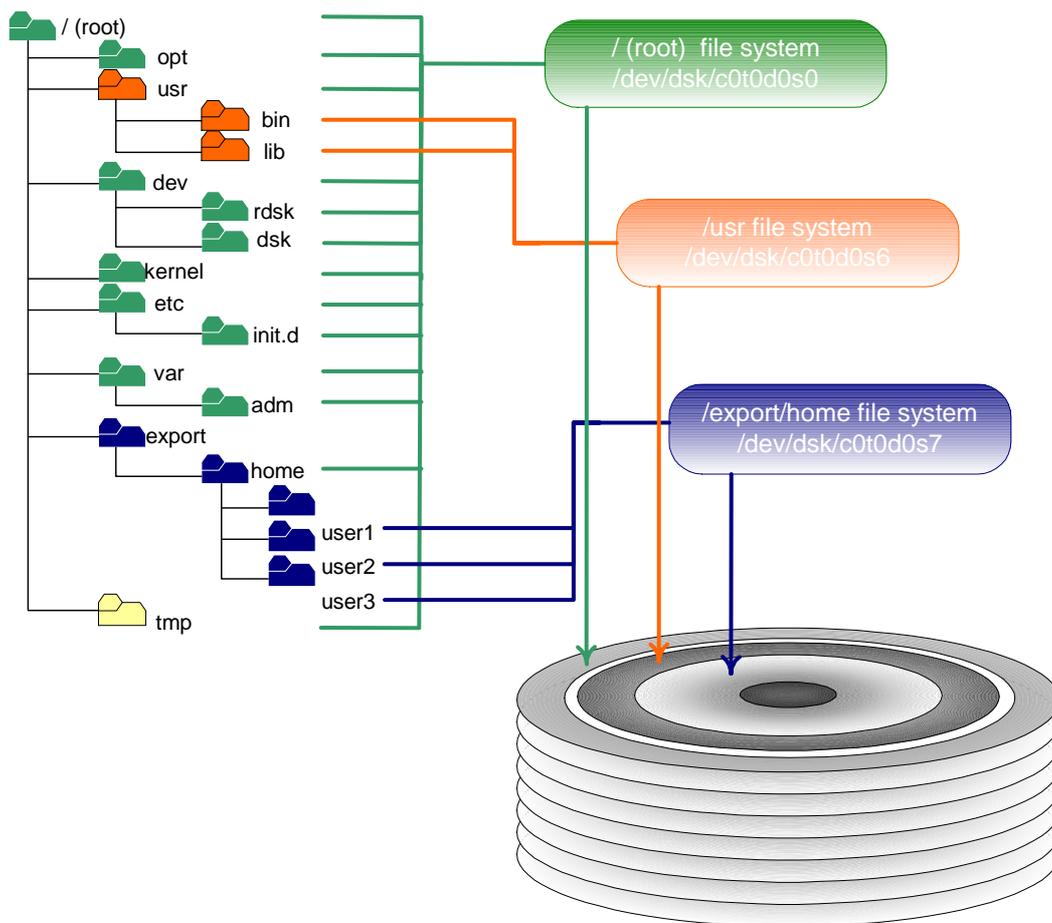


Figura 3.11 Los file systems ufs residen en particiones de discos.

El file system ufs de Solaris contiene las siguientes estructuras básicas de soporte.

3.16 Etiqueta del disco (VTOC)

La etiqueta del disco VTOC contiene la tabla de partición del disco. La VTOC se aloja en el primer sector del disco (512 Bytes). Solo la primer partición del disco contiene la VTOC, aunque el file system puede ser creado en cualquier partición descarta el primer sector ya que en el puede estar la VTOC.

3.17 Bloque de arranque.

El programa bootstrap (programa de arranque para levantar los sistemas de archivos o file system) se aloja o reside en los primeros 15 sectores del disco (sectores 1-15) que siguen después de la VTOC. Solo el file system /root tiene un bloque de arranque activo.

3.18 Super bloque primario.

El superbloque reside en los sectores del 16 al 31 después del bloque de arranque. Es una tabla de información que describe al file system, incluye:

- El número de bloques de datos
- El número de bloques de cilindros
- El tamaño de un bloque y fragmento de datos
- Una descripción de hardware, derivado de la etiqueta
- El nombre del punto de montaje
- La bandera de estado del file system: limpio, estable, activo, autenticado o desconocido.

3.19 Super bloque de respaldo.

Cuando un file system se crea, cada grupo de cilindros replica el super bloque comenzando en el sector 32. Esto protege a los datos críticos en el superbloque de una catastrófica pérdida.

3.20 Grupo de cilindros.

Cada file system está dividido en grupo de cilindros con un tamaño mínimo default de 16 cilindros por grupo. Los grupos de cilindros facilitan el acceso al disco.

El file system constantemente optimiza el desempeño del disco posiciona los datos de un archivo en un simple grupo de cilindros, lo cual reduce la distancia que tiene que recorrer la cabeza del disco para acceder a los datos del archivo. El file system almacena grandes archivos a través de varios grupos de cilindros, si es necesario.

3.21 Bloques de grupos de cilindros.

Es una tabla en cada grupo de cilindros que describe al mismo, incluye:

- El número de inodos
- El número de bloques de datos en el grupo de cilindros
- El número de directorios
- Bloques libre, inodos libres y fragmentos libres en el grupo de cilindros.
- El mapa del bloque libre
- El mapa del inodo libre

La figura 3.12 muestra una serie de grupos de cilindros en un file system ufs.

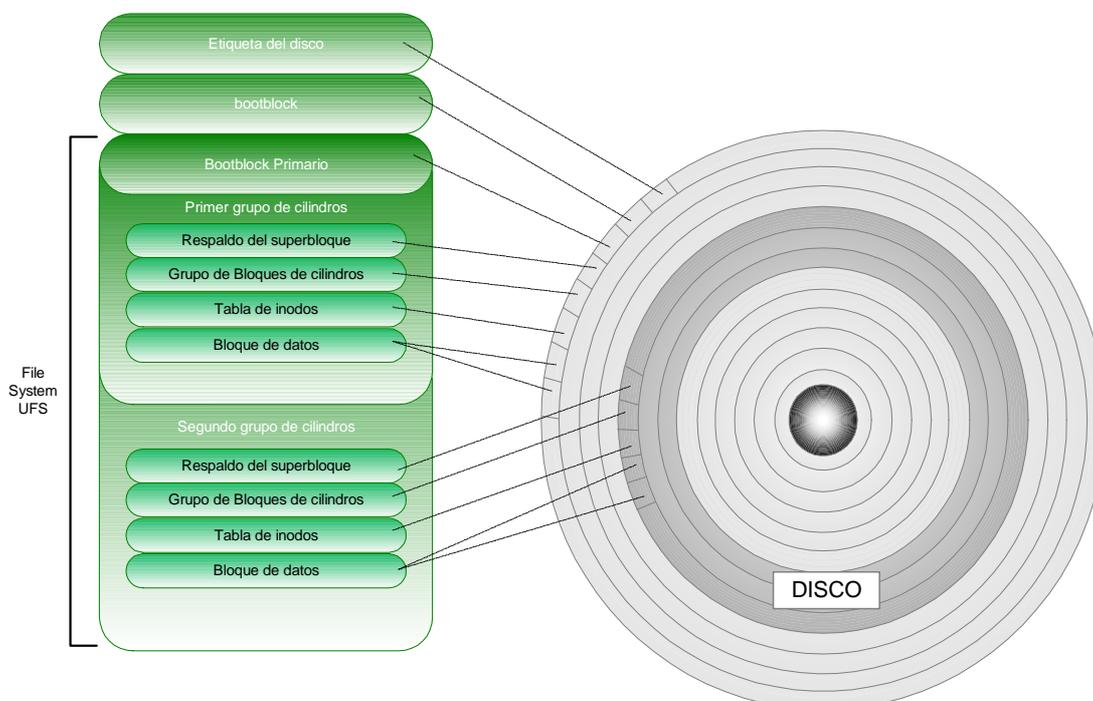


Figura 3.12 Estructura del file system ufs en Solaris

3.22 El Inodo ufs.

Un inodo contiene la siguiente información acerca de un archivo.

- El tipo de archivo y los modos de acceso
- La identificación de usuario (UID) y la identificación del grupo (GID) número de propietarios y grupos
- El tamaño del archivo
- El control de la liga
- La fecha del último acceso y modificación, y los cambios de inodo.
- El número total de bloques utilizados por el archivo.
- Dos tipos de apuntadores: apuntadores directos y apuntadores indirectos

Las figura 3.13 muestra alguna de la información contenida en un inodo.

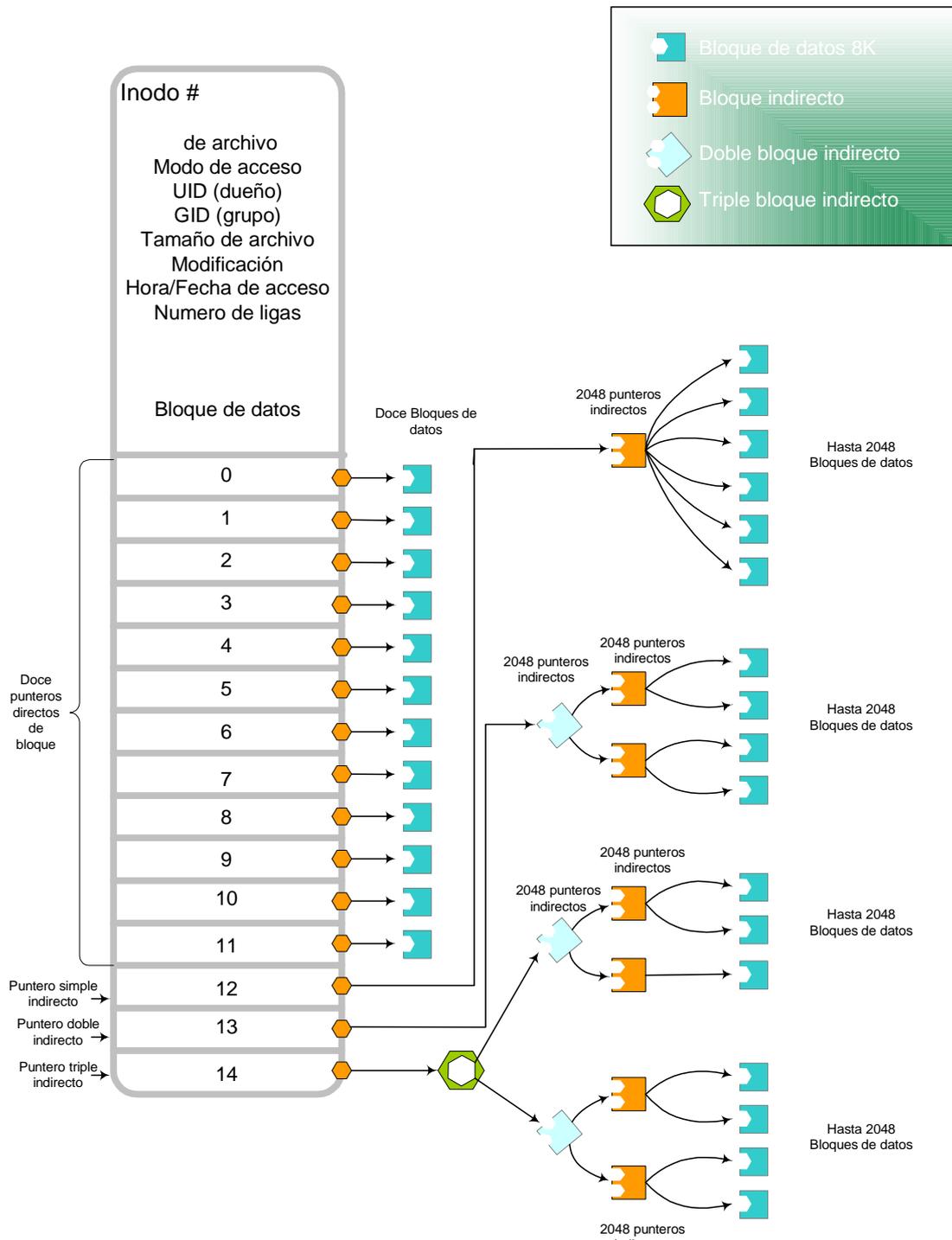


Figura 3.13 Estructura de un inodo ufs.

3.23 Apuntadores Directos.

Dentro del inodo hay 12 apuntadores o punteros directos, que contienen direcciones para los 12 primeros bloques de datos. Cada uno de estos apuntadores pueden referenciar bloques de datos de 8 kbytes de un archivo que sobrepasa de los 96 kbytes.

3.24 Apuntadores Indirectos.

Los tres tipos de apuntadores indirectos dentro de un inodo son:

- Apuntador indirecto individual. Refiere o apunta a un bloque del file system que contiene punteros de bloque de datos. Este bloque del file system contiene 2048 direcciones de bloques de datos de 8 Kbytes que pueden apuntar a otros 16 Mbytes de datos.
- Apuntador indirecto doble. Apunta a un bloque del file system que contiene apuntadores indirectos individuales. Cada apuntador apunta a un bloque del file system que contiene los punteros del bloque de datos. Apuntan a 32 Gbytes de datos.
- Apuntador indirecto triple. Puede referenciar mas de 64 Tbytes de datos. Sin embargo el tamaño máximo de un file system ufs esta limitado a 1 Tbyte debido al maximoel espacio de direcciones de 32 bits para los controladores del dispositivo.

3.25 Bloques de datos.

El resto del espacio asignado en el file system ufs tiene bloques de datos. Los bloques de datos están asignados, por default en bloques lógicos de 8 Kbytes. Los bloques mas alejados son divididos en fragmentos de 1 Kbyte. Para un archivo regular, los bloques de datos alojan al contenido del archivo. Para un directorio, los bloques de datos contienen entradas que asocian los números de inodo y los nombres de los archivos y directorios contenidos en ese directorio.

Dentro del file system, los bloques que no están siendo usados como archivos, bloques indirectos de direcciones o bloques de almacenamiento son marcados como libres en el mapa del grupo del cilindro. Este mapa también guarda la pista del fragmento como seguridad.

3.26 Fragmentación.

Es el método usado por el file system ufs para asignar espacio de disco eficientemente

Los archivos de menos de 96 Kbytes son almacenados usando fragmentación.

Por default los bloques de datos pueden ser divididos en ocho fragmentos de 1024 bytes cada uno. Los fragmentos guardan archivos y piezas de archivos menores a 8192 bytes. Para los archivos de mas de 96 Kbytes, los fragmentos nunca son asignados y solo se usan los bloques llenos.

La figura 3.14l muestra un fragmento en un bloque de datos.

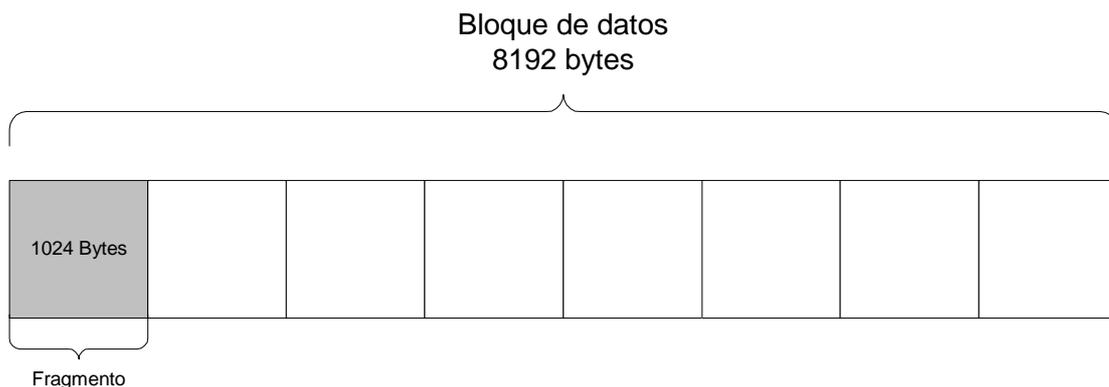


Figura 3.14 Bloque de datos dividido.

Si un archivo contenido en un fragmento crece, requiere más espacio, por lo tanto este se almacena en uno o más fragmentos adicionales dentro del mismo bloque de datos.

La figura 3.15 muestra el contenido de dos archivos diferentes guardados en fragmentos en el mismo bloque de datos.

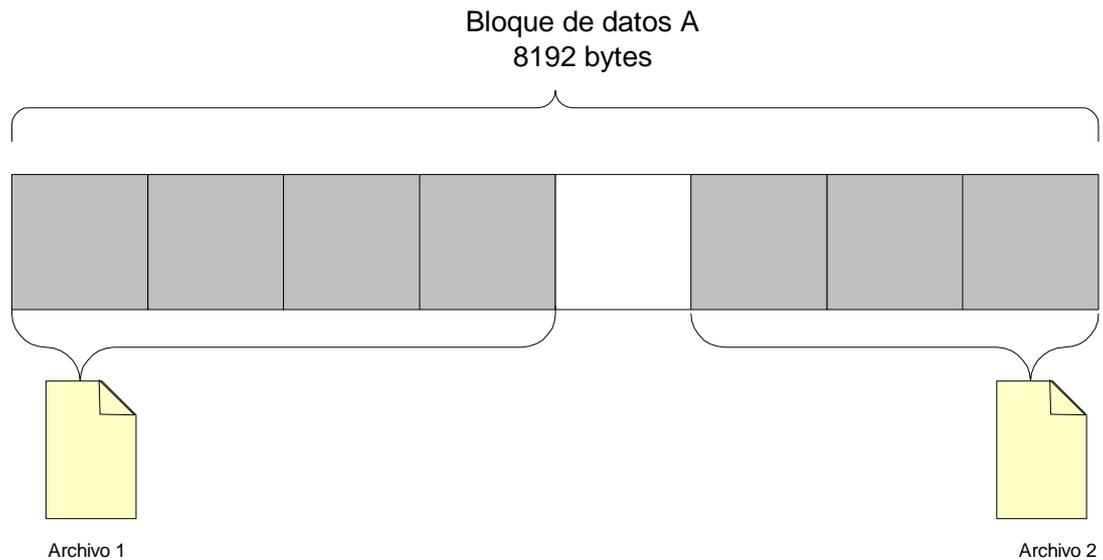


Figura 3.15 Dos archivos guardados en un bloque de datos.

Por ejemplo, si el Archivo 1 requiere más espacio que está disponible en el bloque de datos compartido, todo el contenido del archivo se mueve a un bloque de datos libre. Esto asegura que todos los fragmentos del archivo están dentro de un bloque de datos entero. El file system ufs no permite a los fragmentos de un mismo archivo estar en dos bloques diferentes.

En éste capítulo revisamos la forma en la cual se manejan y administran los discos duros utilizando el sistema operativo Solaris así como algunos de los comandos más utilizados en el manejo de discos, también se vió la creación de file systems, en el siguiente capítulo estudiaremos las formas de organizar la información dentro de los discos duros.

4

Tecnología RAID

El almacenamiento apropiado de la información es esencial para las operaciones de las empresas o negocios. Se requiere de una solución de almacenamiento que pueda administrarse fácilmente y que proporcione la más alta disponibilidad del sistema.

Desde hace más de 20 años los procesadores han incrementado su capacidad en un 40% cada año. En el mismo periodo de tiempo los discos duros han aumentado su capacidad de almacenamiento al doble, mientras que su costo se ha reducido. Desafortunadamente el rendimiento del sistema general de las computadoras a superado al rendimiento de los discos duros. Esto provoca una descompensación entre el trato que se le da a la información del sistema (rápido) y la lectura y escritura de los datos en los discos (lenta).

Fue en 1987 cuando David Patterson, Garth Gibson y Randy Katz presentaron un artículo en la Universidad de California titulado “A Case for Arrays of Inexpensive Disks (RAID)”, este artículo describía varios tipos de arreglos de discos, los cuales se referían al acrónimo RAID.

La idea básica de un RAID (arreglos redundantes de discos de bajo costo) es la de tener un sistema para guardar información en varios discos duros a la vez con ello hacer el acceso más rápido ya que la carga se distribuye entre los diferentes discos duros.

Como lo mencionamos, en un principio el término RAID se refería a Redundant Array of Inexpensive Disk (arreglos redundantes de discos de bajo costo) como una contraposición directa a los SLED Single Large Expensive Disk (discos independientes de alto costo). Sin embargo, las increíbles bajas en los precios de los discos duros han ocasionado que los SLED prácticamente desaparezcan, por tanto, el significado de las siglas ha cambiado y en la actualidad se le traduce como Redundant Array of Independent Disks (arreglos redundantes de discos independientes).

En éste capítulo se explicará que es un RAID, que tipos existen para cada necesidad, las aplicaciones, se trataran los casos de RAID por software y RAID por hardware así como sus diferencias.

4.1 ¿Qué es RAID?

Es combinar varios discos en un arreglo de discos el cual pueda desempeñarse de la misma forma en la que lo haría un disco de gran tamaño por lo regular de muy alto costo.

Adicionalmente a ésta ventaja del ahorro económico para el sistema operativo, aparenta ser un solo disco duro lógico.

Al colocar los datos en varios discos, las operaciones de lectura y escritura pueden superponerse de un modo equilibrado, mejorando el rendimiento del sistema.

Otra característica de los sistemas RAID es que el tiempo medio entre errores (mean time between failure, MTBF) de un arreglo es igual al MTBF de un disco único, dividido entre el número de los discos que conforman el arreglo. Por esto el MTBF de un arreglo de discos es mucho menor para muchos de los requerimientos de las aplicaciones. Por tanto podemos decir que el almacenamiento redundante de los datos en éste tipo de sistemas aumenta la tolerancia a fallos.

Desde el principio se definieron cinco niveles para los arreglos de discos cuyas funciones son ofrecer mejoras en el rendimiento, confiabilidad, tasa de transferencia y tasas de lectura escritura. Los fabricantes han introducido y/o desarrollado variaciones a los cinco niveles originales.

Cada nivel tiene sus propias ventajas y desventajas, las cuales iremos revisando en el transcurso del capítulo.

4.2 Técnica del sistema RAID.

La base del sistema RAID es la técnica conocida como “striping” (balanceado o creación de bandas), un método de concatenar varios discos en una única unidad de almacenamiento. Esta técnica incluye la partición de cada espacio de almacenamiento de los discos en unidades que pueden ser tan pequeñas como un sector (512 bytes) hasta varios megabytes. Estas unidades están interpaginadas (interleaved) y se accede a ellas en orden.

En los usuarios sistemas monousuarios en donde por lo general se almacenan grandes registros (como por ejemplo imágenes), las unidades en las que son divididos los discos son pequeñas, tal vez lo que

mide un sector de modo que un registro esté ubicado en todos los discos facilitando la lectura de todos los discos a la vez.

Al contrario en los sistemas multiusuario, un mejor rendimiento demanda que las unidades se establezcan en mayores tamaños. Esto permite acciones de lectura/escritura superpuestas en los distintos discos.

4.3 Sistemas RAID como una solución a la operación de los Discos Duros.

Como lo vimos en el capítulo 3 las operaciones de lectura/escritura de los discos son mecánicas e involucran, normalmente dos operaciones: la primera es el posicionamiento de la cabeza y la segunda es la transferencia desde o hacia el propio disco. El posicionamiento de la cabeza está limitado por dos factores: el tiempo de búsqueda y el retardo por el giro del disco hasta la posición de inicio de los datos. La transferencia de datos, por su parte, ocurre de un bit a la vez y se ve limitada por la velocidad de rotación y por la densidad de grabación del medio.

Una forma de mejorar el rendimiento de la transferencia es el uso de varios discos en paralelo; esto se basa en el hecho de que si un solo disco puede entregar una tasa de transferencia dada, entonces dos discos serían capaces, teóricamente, de ofrecer el doble de la tasa de transferencia anterior; lo mismo sucedería con cualquier operación.

Al principio se pensó en esto como una solución al problema del almacenamiento de grandes cantidades de información pero después de varios intentos se llegó a la conclusión de que no hay mejoras en la velocidades de archivos individuales, solo mejora la cantidad de archivos accedidos en forma concurrente.

Y el balance tampoco es posible de mantener en el tiempo debido a la naturaleza eminentemente dinámica de la información.

Un arreglo de discos es la forma más efectiva de conseguir el objetivo, el cual según la definición de RAID Consultory Board es “una colección de discos que integran uno o más subsistemas combinados con un software de control el cual se encarga de controlar la operación del mismo y de presentarlo al Sistema Operativo como un solo gran dispositivo de almacenamiento”.

El software de control de los arreglos de discos puede ser integrado directamente al Sistema Operativo o residir en el propio arreglo; así como éste último puede ser externo o interno.

En éste tiempo la mayoría de los sistemas operativos incluyen soporte para arreglos de discos.

Las soluciones de arreglos basadas en hardware son principalmente implementadas mediante el uso de controladoras SCSI especializadas, las cuales a menudo están dotadas de procesadores propios para liberar a la CPU del sistema de la tarea de control y de cachés para mejorar aun más el desempeño.

Así pues un arreglo de discos ofrece un mejor desempeño debido a que divide en forma automática los requerimientos de lectura/escritura entre los discos que lo conforman.

4.4 Ventajas del uso de RAID.

Si se manejan grandes cantidades de datos el uso de RAID ofrece las siguientes ventajas:

- Aumento de velocidad
- Aumento de la capacidad de almacenamiento mediante el uso de un disco virtual
- Eficacia en la recuperación de un fallo del sistema.

4.5 Niveles de arreglos.

4.5.1 RAID 0. Striping sin tolerancia a fallos.

Requiere de al menos dos discos para ser implementado.

4.5.1.1 Funcionamiento.

Llamado también “**striping**” (separación o fraccionamiento). Este nivel implementa un arreglo de discos en líneas en donde los datos son divididos en bloques y cada bloque es almacenado en un disco distinto.

La figura 4.1 nos muestra un ejemplo visual de cómo son almacenados los datos en éste nivel de RAID.

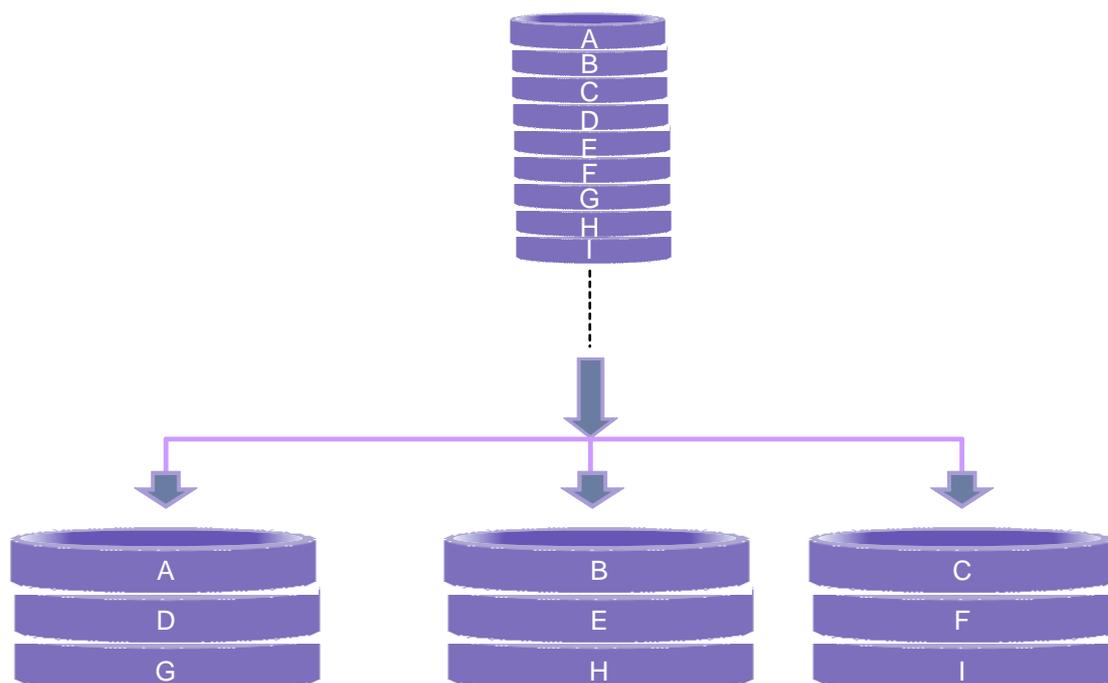


Figura 4.1 RAID 0

4.5.1.2 Ventajas.

Se mejora el desempeño de lectura/escritura al distribuir la carga de la escritura y lectura a través de varios canales y platos del disco. El mejor desempeño es alcanzado cuando los datos son distribuidos en varios controladores con solo un disco por controlador.

- No hay sobre carga por el cálculo de paridad.
- Son de un diseño simple.
- Tienen una sencilla implementación.

4.5.1.3 Desventajas.

- No se le puede llamar RAID ya que no es redundante ni tolerante a fallos.
- La falla de uno de los discos resulta la pérdida de los datos de todo el arreglo.

4.5.1.4 Aplicaciones

- No se usan en ambientes de misión crítica.
- Son útiles en producción y edición de video e imágenes, así como aplicaciones de preimpresión.
- En general para aplicaciones que requieran de un almacenamiento a gran velocidad pero que no requieran tolerancia a fallos.

4.5.2 RAID 1. Espejo (mirroring) y duplicación.

Requiere de al menos dos discos para ser implementado. La tarjeta controladora debe ser capaz de realizar dos lecturas concurrentes independientes por par de espejo o dos escrituras duplicadas por par de discos en espejo.

4.5.2.1 Funcionamiento.

Implementa un arreglo de discos en espejo (mirroring) en donde los datos son leídos y escritos de manera simultánea en dos discos distintos. Este nivel proporciona redundancia escribiendo datos idénticos en cada miembro del arreglo, dejando una copia "idéntica" en cada una de los discos. Como ya se mencionó opera con dos o más discos que pueden utilizar una modalidad de acceso paralelo para transferir de manera rápida datos de lectura, pero más comúnmente opera de manera independiente para ocuparse de valores altos de transacciones I/O.

La figura 4.2 nos muestra un ejemplo visual de cómo son almacenados los datos en éste nivel de RAID.

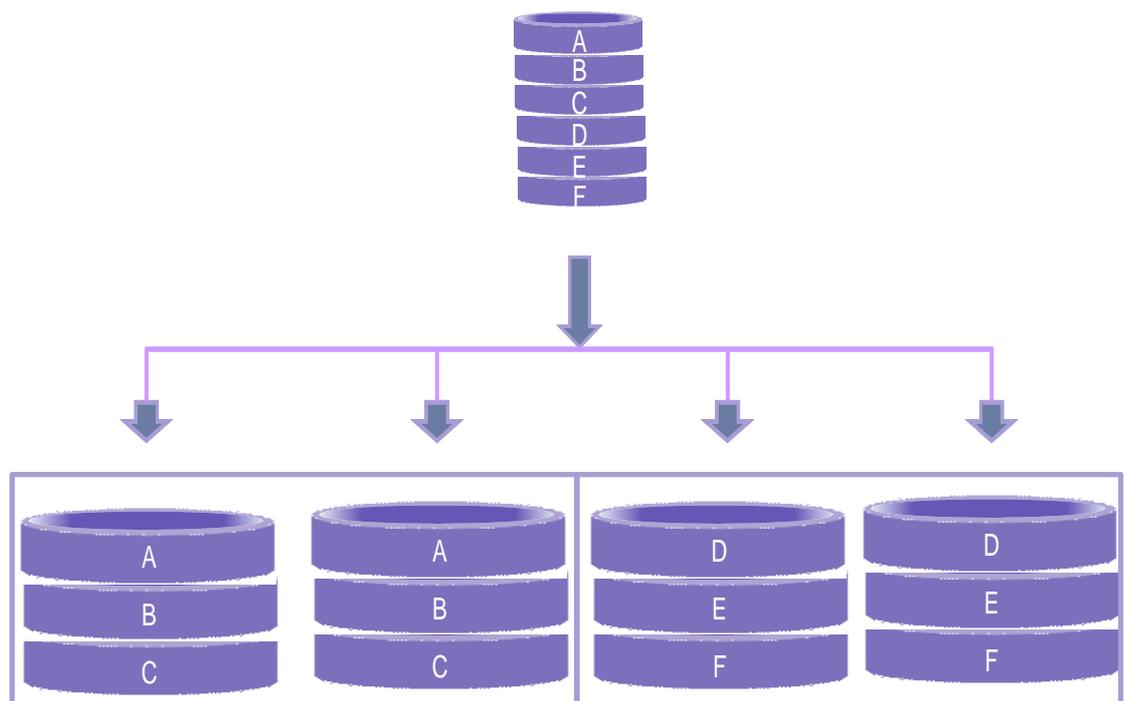


Figura 4.2 RAID 1

4.5.2.2 Ventajas.

En esta configuración un arreglo de discos puede realizar una escritura o dos lecturas por par de espejo, duplicando la tasa de lectura de discos simples con la misma tasa de escritura que los discos tradicionales. Una redundancia total de datos significa que no es necesaria la reconstrucción en caso de falla de algún disco, sino solo una copia.

- La tasa de transferencia por bloques es la misma que en los discos tradicionales.
- Bajo ciertas circunstancias RAID 1 puede soportar fallas simultáneas múltiples de discos.
- Es el diseño RAID más simple.
- Tienen una sencilla implementación.

4.5.2.3 Desventajas.

- Es el que tiene mayor desperdicio de disco.
- Típicamente la función RAID es llevada a cabo por el software del sistema degradando el desempeño del mismo.
- No soporta cambio en caliente de un disco dañado cuando se implementa por software.

4.5.2.4 Aplicaciones.

- Es mejor cuando se utiliza un software para manejar éste nivel.
- Sirve si la información que se almacenará es administrativa o financiera.
- En general se implementa para aplicaciones que requieren de una alta disponibilidad.

4.5.3 RAID 2 Código de Corrección de Errores con código de Hamming.

Requiere de al menos dos discos para poder ser implementado.

4.5.3.1 Funcionamiento.

Este nivel cuenta con discos para bloques de redundancia y corrección de errores (error checking and correcting ECC). La división es a nivel de bits, cada byte se graba con un byte de paridad en cada uno de los discos y un bit de paridad en el noveno. El acceso es simultáneo a todas las unidades tanto en operaciones de escritura como lectura. Algunos de estos discos son empleados para códigos de error, los cuales se emplean para

referencias de los datos en caso de que falle uno de los discos. Este nivel tiene un costo bastante elevado ya que si se desea mantener los códigos de error son necesarios muchos discos. Gracias a como están distribuidos los datos en los discos se consigue mejorar la velocidad de transferencia principalmente en la lectura ya que podemos emplear todos los discos en paralelo.

La figura 4.3 nos muestra un ejemplo visual de cómo son almacenados los datos en éste nivel de RAID.

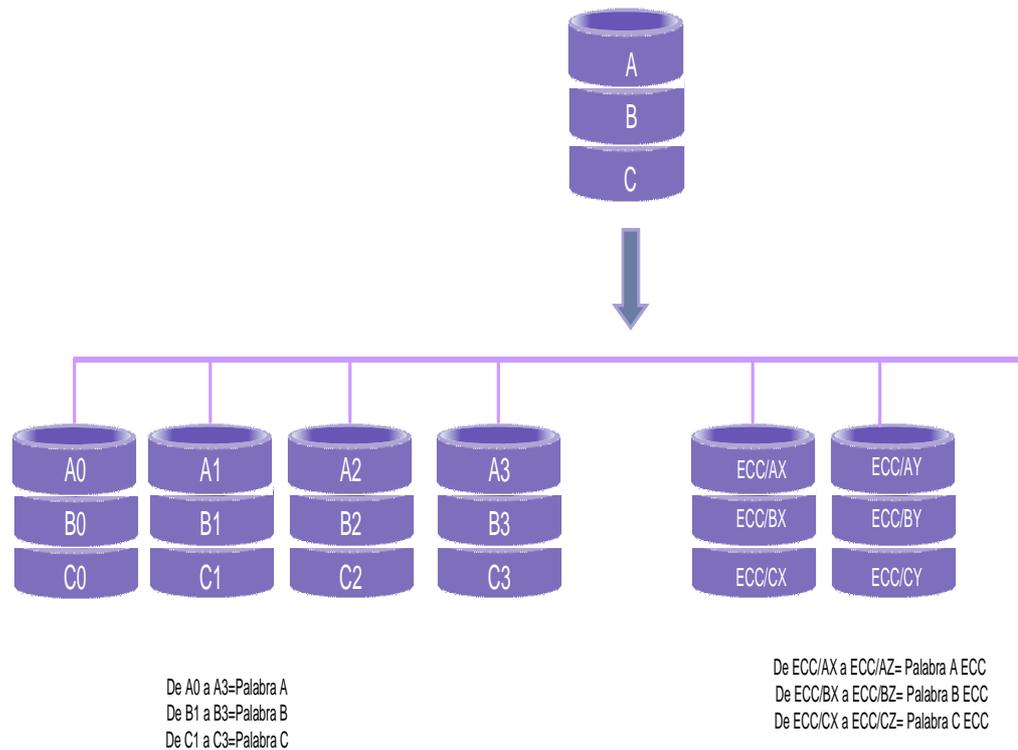


Figura 4.3 RAID 2

4.5.3.2 Ventajas.

- Se pueden recuperar los datos.
- Es posible alcanzar tasas de transferencias muy altas.
- El diseño es simple.

4.5.3.3 Desventajas.

- Costo elevado.
- No existen implementaciones comerciales.
- Tiempo de escritura lento. Lo que lo hace no viable.

4.5.3.4 Aplicaciones.

- Se emplea para mejorar la demanda y también la velocidad de transferencia.

4.5.4 RAID 3 Transferencia en paralelo con paridad.

Requiere de al menos tres discos para funcionar.

4.5.4.1 Funcionamiento.

Este nivel también es conocido como Striping con paridad dedicada. Utiliza también un disco de protección de información separado para almacenar información de control codificada con lo que se logra una forma más eficaz de proporcionar redundancia de datos. Este control de información codificada o paridad proviene de los datos almacenados en los discos y permite la reconstrucción de información en caso de fallos. Los datos se dividen, fragmentos que se transfieren a los discos que funcionan en paralelo, lo que permite enviar más datos de una sola vez, y aumentar en forma sustancial la velocidad general de transferencia de datos. Esta última característica convierte a este nivel en idóneo para aplicaciones que requieren de la transferencia de grandes archivos contiguos hacia y desde la computadora central o servidor.

La figura 4.4 nos muestra un ejemplo de cómo son almacenados los datos en éste nivel de RAID.

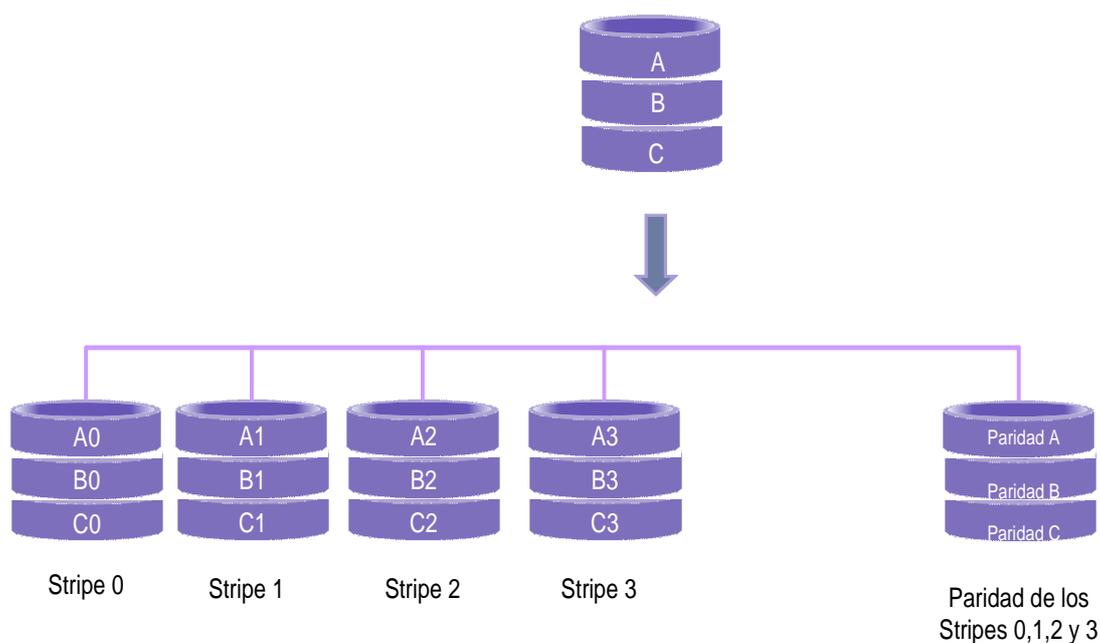


Figura 4.4 RAID 3

4.5.4.2 Ventajas.

- Se pueden recuperar los datos.
- Alto rendimiento para aplicaciones de velocidad de transferencia alta.

4.5.4.3 Desventajas.

- Si se pierde el disco de paridad se pierde toda la información redundante
- Tiempo de escritura lento.

4.5.4.4 Aplicaciones.

- Producción de video y transmisiones digitales en línea.
- Edición de imágenes.
- Edición de video.
- Aplicaciones de predicción.
- Cualquier aplicación que requiera alta densidad de transferencia de datos.

4.5.5 RAID 4 Arreglo de discos independientes.

Sistemas de discos independientes con disco de control de errores.

4.5.5.1 Funcionamiento.

En el nivel 4 de RAID los bloques de datos pueden ser distribuidos a través de un grupo de discos para reducir el tiempo de transferencia y explotar toda la capacidad de transferencia de datos de la matriz de disco. El nivel 4 de Raid es preferible al nivel 2 de Raid para pequeños bloques de datos, porque en este nivel, los datos son distribuidos por sectores y no por bits. Otra ventaja del nivel 4 de RAID frente a los niveles 2 y 3 es que al mismo tiempo puede estar activa más de una operación de lectura escritura sobre el conjunto de discos.

El nivel 4 de RAID tiene división a nivel de bloques y el acceso al arreglo de discos es paralelo, pero no simultaneo. Posee un disco dedicado a paridad y corrección de errores. La operación de escritura se realiza en forma secuencial y la lectura en paralelo.

La figura 4.5 nos muestra un ejemplo visual de cómo son almacenados los datos en éste nivel de RAID.

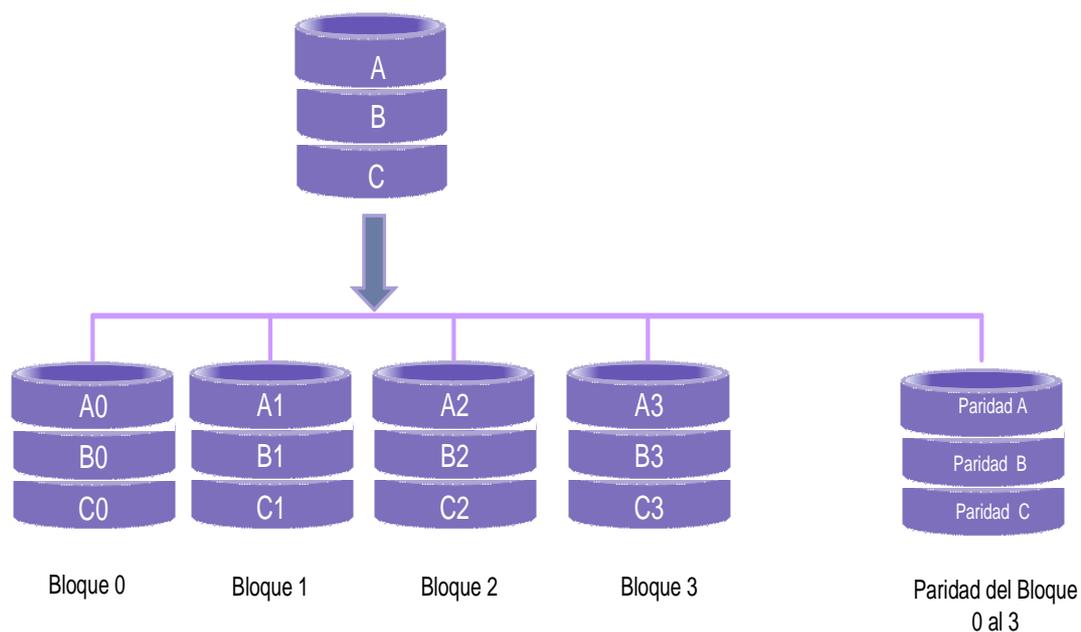


Figura 4.5 RAID 4

4.5.5.2 Ventajas.

- Buen rendimiento en las escrituras de datos.
- Tiene integridad de datos.

4.5.5.3 Desventajas.

- Si se pierde el disco de paridad se pierde toda la información redundante
- Menor rendimiento en la lectura de datos.

4.5.6 RAID 5 Arreglo de discos independientes.

Sistemas de discos independientes con integración de código de error mediante una paridad.

4.5.6.1 Funcionamiento.

Sistemas de discos independientes con integración de código de error mediante una paridad. En RAID 5 los datos y una paridad son guardados en los mismos discos por lo que conseguimos aumentar la velocidad de demanda, ya que cada disco puede satisfacer una demanda independiente de los demás. Con diferencia con el RAID 3, el RAID 5 guarda la paridad del dato dentro de

los discos y no hace falta un disco para guardar dichas paridades. En el nivel 5 de Raid las unidades de disco actúan independientemente, cada unidad es capaz de atender a sus propias operaciones de Lectura/Escritura, lo que aumenta el número de operaciones de entrada salida simultánea. Esta característica mejora considerablemente el tiempo de acceso, especialmente con múltiples peticiones de pequeñas operaciones de entrada salida.

El nivel 5 de Raid asegura un mejor rendimiento de operaciones de entrada salida para aplicaciones en la que el sistema realiza búsqueda aleatorias de muchos ficheros pequeños como sucede en las aplicaciones transaccionales, ofrece la posibilidad de soportar múltiples operaciones de escritura de forma que los datos pueden escribirse en un disco y su formación de paridad en otro. En este nivel no existe una unidad dedicada para paridad sino que el controlador intercala los datos y las paridades en todo los discos del subsistema. El inconveniente de este nivel es que presenta una operación adicional de escritura al almacenar los datos ya que tanto los datos como la información de actualidad se actualizan en operaciones distintas y en unidades de disco diferentes. Las aplicaciones implican numerosas operaciones de escritura y sufren descensos en el rendimiento.

La figura 4.6 nos muestra un ejemplo visual de cómo son almacenados los datos en este nivel de RAID.

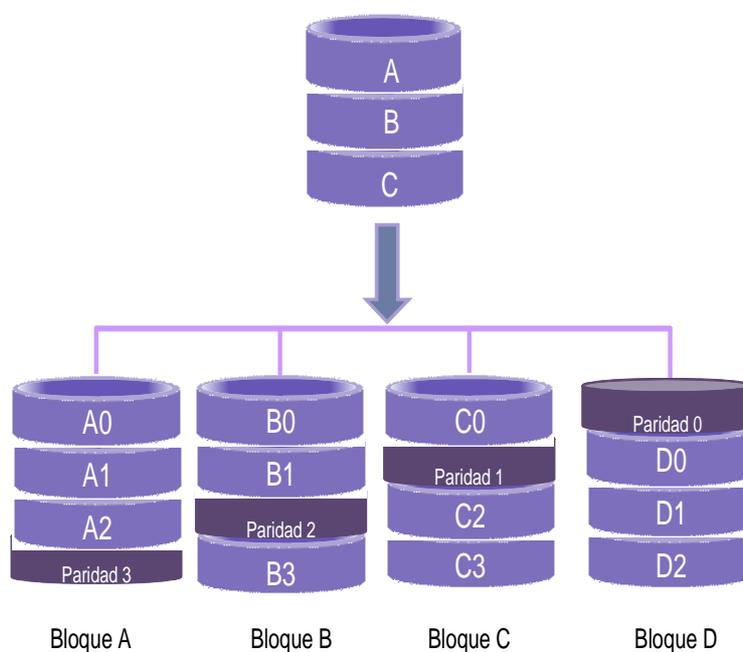


Figura 4.6 RAID 5

4.5.6.2 Ventajas.

- Alto rendimiento en aplicaciones de velocidad de demanda interactiva
- Costo efectivo .No desaprovecha un disco exclusivamente para paridad
- Se pueden recuperar datos Buen rendimiento en las escrituras de datos.
- Tiene integridad de datos.

4.5.6.3 Desventajas.

- El rendimiento en las escrituras de datos es bajo
- No aumenta el rendimiento en las aplicaciones , aunque la velocidad de transferencia de datos es alta.
- Si se pierde el disco de paridad se pierde toda la información redundante
- Menor rendimiento en la lectura de datos.

4.5.6.4 Aplicaciones.

Es recomendable para aplicaciones intensas de entrada salida y lectura escritura, tal como procesamiento de transacciones

4.5.7 RAID 6 Arreglo de discos independientes.

Sistemas de discos independientes con integración de código mediante una doble paridad.

4.5.7.1 Funcionamiento.

Raid 6 es esencialmente una extensión del RAID 5 , para ello guarda , una segunda paridad .Este nivel proporciona muy buena integridad de los datos y repara diversos errores en los discos .El nivel 6 de Raid añade un nivel mas de disco , resultando una organización con dos dimensiones de disco y una tercera que corresponde a los sectores de los discos la ventaja de este nivel consiste que no solamente se puede recuperar un error de entre dos discos , sino que es posible recuperar muchos errores de 3 discos .La operación de escritura es difícil debido a la necesidad de sincronizar todas las dimensiones .

La figura 4.7 nos muestra un ejemplo visual de cómo son almacenados los datos en éste nivel de RAID.

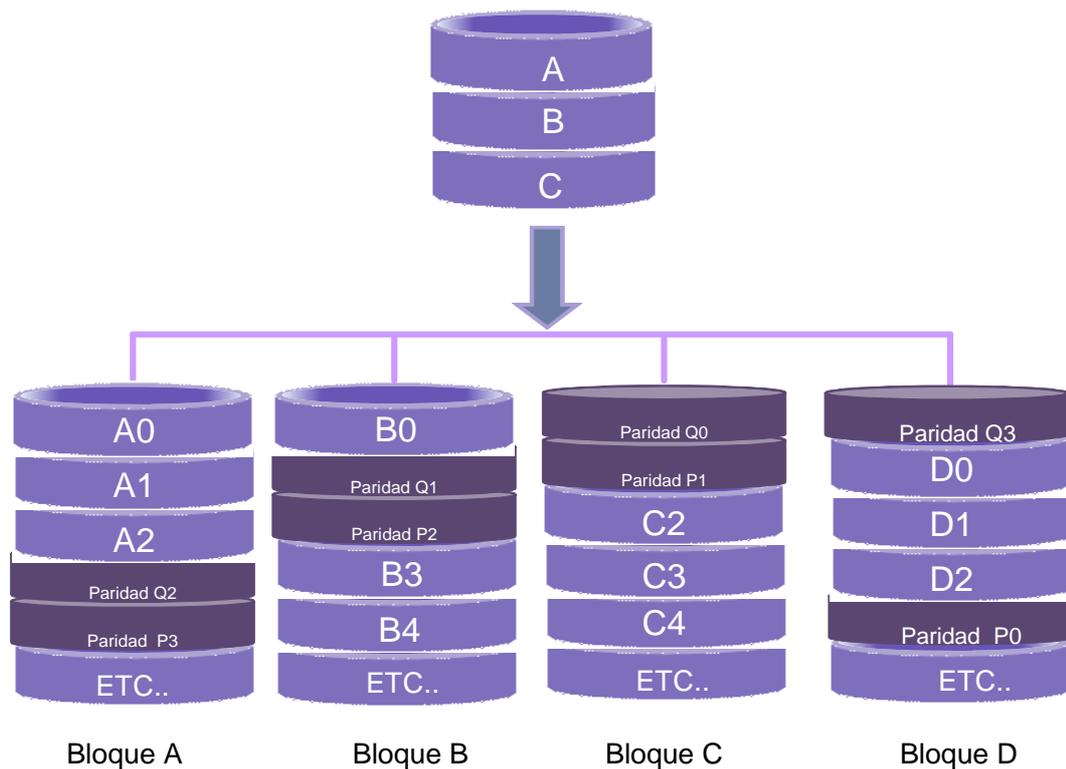


Figura 4.7 RAID 6

4.5.7.2 Ventajas.

- Podemos recuperar diversos errores simultáneamente.
- Nivel de integridad muy elevado solución perfecta para aplicaciones críticas

4.5.7.3 Desventajas.

- El rendimiento en escritura de datos es bastante lento
- No se dispone de muchas implementaciones comerciales en el nivel de RAID 6

4.5.7.4 Aplicaciones.

- Servidores de archivos y aplicaciones
- Servidores de bases de datos
- Servidores Web y de correos

4.5.8 RAID 10 o 1+0

4.5.8.1 Funcionamiento.

La información se distribuyen en bloques como el RAID 0 y adicionalmente, cada disco se duplica como raid 1 , creando un segundo nivel de arreglo se conoce como "Striping de arreglos duplicados " . Se requiere, dos canales, dos discos para cada canal y se utilizan el 50 % de la capacidad para información de control

La figura 4.8 nos muestra un ejemplo visual de cómo son almacenados los datos en éste nivel de RAID.

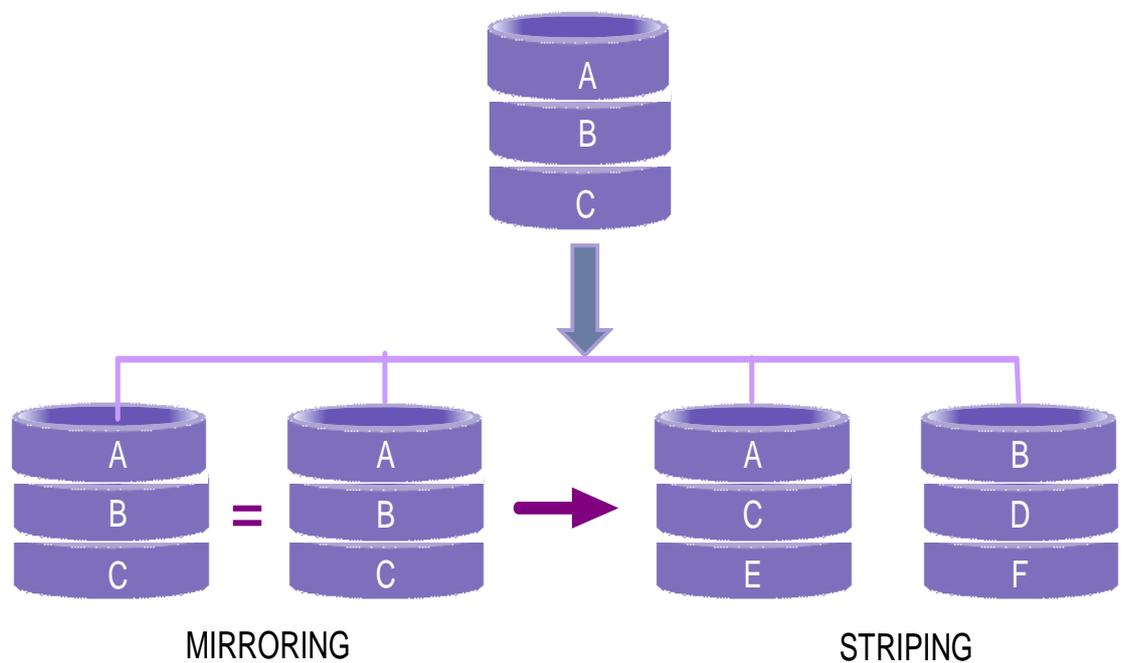


Figura 4.8 RAID 10

4.5.8.2 Ventajas.

- Este nivel ofrece un 100 % de redundancia de la información y un soporte para grandes volúmenes de datos, donde el precio no es un factor importante.

4.5.8.3 Desventajas.

- Alto costo de implementación
- Escalabilidad limitada a un alto costo

4.5.8.4 Aplicaciones.

Ideal para sistemas de misión crítica, donde se requiera mayor confiabilidad de la información, ya que pueden fallar dos discos inclusive (uno por cada canal) y los datos todavía se mantienen en línea. Es apropiado también en escritura aleatorias pequeñas.

Existen algunas combinaciones más de los arreglos anteriores, sin embargo los que hemos presentado son los más comúnmente usados en el campo y cada administrador de datos configura por software o hardware o ambos y depende de las necesidades que se elije el método conveniente.

4.6 Como seleccionar un disco RAID.

El seleccionar un disco raid puede ser una dura tarea, ya que no existen criterios establecidos, para saber, cual es la mejor solución. Una equivocación bastante común es juntar todos los discos Raid en el mismo nivel, por lo que debemos elegir una solución correcta para el grupo completo de disco, no metiendo todos en el mismo nivel sino estructurando los discos dependiendo de las necesidades en seguridad e integridad de los datos .

Para seleccionar un disco Raid en primer lugar tenemos que conocer nuestras necesidades, o saber los tipos de aplicaciones que se emplean. Podemos separa las aplicaciones en dos tipos:

- Aplicaciones de demanda:

Podríamos seleccionar el Raid 5 ya que ofrece una alta velocidad de demanda, tanto en escrituras como lecturas.

- Aplicaciones de Transferencias :

La mejor opción pasaría por Raid 3 ya que ofrece una alta velocidad de transferencia para gráficos e imágenes y aplicaciones en general en las que necesitamos gran transferencia de datos.

Además del tipo de aplicación, habremos de tener en cuenta los factores que nos guiaran a optar por la elección de un RAID u otro como

son :

-Integridad y costo :

Normalmente llegaremos a una solución Raid que asegure la integridad de los datos del disco, ya que de esta manera tendremos preservados, nuestros datos sobre posibles errores del disco .El Raid 0 es el único nivel que no proporciona integridad.

-Costo:

El costo vendrá marcado por nuestro poder adquisitivo será uno de los factores mas importantes a la hora de seleccionar nuestro disco .No solo tenemos que pensar en la integridad de los datos, sino en cuando se estropee un disco tenemos que apagar el sistema . A para ello existe varios dispositivos como Hot Swap , Hot Spare , Duplex Controler , Redundant Power Supplies .

La siguiente tabla muestra las opciones mas acertadas, a la hora de elegir el nivel Raid que mejor se adapte a nuestra necesidades teniendo en cuenta todos estos factores:

Costo	Rendimiento	Integridad	Nivel Raid
-	-	-	Raid 1
-	X	-	Raid 0
-	X	X	Raid 1
X	-	-	Raid 0
X	-	X	Raid 3/5
X	X	-	Raid 0
X	X	X	Raid 3/5

Máximas y mínimas cantidades de HD que se pueden ordenar para los diferentes niveles de RAID.

Nivel de Raid	Minimo	Maximo
5	3	16
4	3	N/A
3	3	N/A
2	N/A	N/A
1	2	2
0	2	16

En éste capítulo pudimos estudiar los diferentes tipos de arreglos existentes estos tipos pueden ser configurados tanto físicamente como por medio de programas especialmente diseñados para este propósito. En el siguiente capítulo describiremos de forma general un software el cual tiene entre muchas la capacidad de configurar arreglos de discos virtualmente.

5

Veritas Storage Foundation

5.1 Almacenamiento Virtualizado

El manejo del almacenamiento se hace cada vez más complejo, debido a

- Los distintos sistemas operativos
- El crecimiento desmedido de la información
- Varios y diferentes proveedores de Hardware
- Diferentes aplicaciones con distintas necesidades de recursos de disco

Para crear un ambiente realmente eficiente, los administradores deben tener las herramientas capaces de manejar ambientes grandes, complejos y heterogéneos. El almacenamiento virtualizado simplifica el complejo ambiente de TI y permite tener el control del capital así como la reducción de costos de operación ofreciendo una administración consistente y automatizada.

El almacenamiento virtualizado es el proceso de tomar varios dispositivos físicos y combinarlos formando dispositivos (virtuales) lógicos, los cuales son presentados como tales al sistema operativo, a las aplicaciones y a los usuarios.

Este proceso de almacenamiento crea una capa de abstracción por encima del almacenamiento físico. Los datos no están restringidos a dispositivos específicos, creando así un ambiente flexible. El almacenamiento virtualizado simplifica la administración y reduce potencialmente los costos mejorando el uso del hardware y la consolidación del mismo.

Con el almacenamiento virtual, los aspectos físicos de almacenamiento están enmascarados para los usuarios. Así los administradores pueden concentrarse menos en aspectos físicos y más en el otorgamiento de acceso necesario a los datos.

Los beneficios que se obtienen al utilizar el almacenamiento virtual son:

- Una mayor productividad de TI mediante la automatización de tareas manuales y la simplificación de la administración de entornos heterogéneos.
- Recuperación de la inversión por la mejora de rendimiento y el aumento de tiempo de actividad.
- Bajos costos en hardware por la utilización optimizada de recursos.

La manera en la que se use el almacenamiento virtual, y los beneficios derivados de esto, dependen de la naturaleza de la infraestructura de TI y las especificaciones requeridas por las aplicaciones. Los tres tipos del almacenamiento virtualizado son:

- Basado en el Almacenamiento
- Basado en Host
- Basado en la Red

La mayoría de las empresas utilizan una combinación de estos tres tipos de almacenamiento virtualizado para apoyar a su arquitectura elegida y las necesidades de sus aplicaciones.

El tipo de almacenamiento virtualizado que se usa depende de factores como:

- La heterogeneidad de los arreglos de almacenamiento
- La necesidad de las aplicaciones de acceder a datos contenidos en distintos dispositivos
- Importancia del tiempo de actividad cuando se reemplaza o se mejora un dispositivo
- La necesidad del acceso para múltiples hosts a datos dentro de un mismo dispositivo de almacenamiento.
- El valor de la madurez de la tecnología
- Inversiones en arquitecturas SAN
- El nivel de seguridad requerida
- El nivel de escalabilidad requerida

5.1.1 Almacenamiento Virtualizado basado en el Almacenamiento.

Se refiere a los discos dentro de un arreglo individual que se presenta virtualmente a múltiples servidores. El almacenamiento es virtualizado por el mismo arreglo.

Por ejemplo, los arreglos RAID virtualizan los discos individuales (que están contenidos dentro del arreglo) en LUNS lógicas las cuales son accedidas por sistemas operativos usando el mismo método de abordar un disco físico que está directamente conectado.

Este tipo de virtualización es útil bajo las siguientes condiciones:

- Si se necesita tener datos en un arreglo accesible a servidores de diferentes sistemas operativos
- Si las necesidades de almacenamiento para los datos de los servidores están cubiertas por la caja física.
- Si no se toma en cuenta la interrupción del acceso a los datos cuando se reemplaza o se mejora un dispositivo

La principal limitación de este tipo de almacenamiento es que los datos no pueden ser compartidos entre arreglos, creando almacenamiento aislado que debe ser gestionado.

5.1.2 Almacenamiento Virtualizado basado en Host.

Se refiere a discos dentro de múltiples arreglos y de diferentes proveedores que son presentados virtualmente a un único servidor.

El almacenamiento virtualizado basado en host es útil bajo las siguientes condiciones:

- Si los servidores necesitan acceder a datos almacenados en múltiples dispositivos.
- Si se requiere flexibilidad para acceder a datos almacenados en arreglos de diferentes marcas.
- Si se agregan servidores los cuales no necesitan acceder a datos asignados a un host particular.
- Si la madurez de la tecnología es un factor de alta importancia en la toma de decisiones.

5.1.3 Almacenamiento Virtualizado Basado en la Red.

Se refiere a discos de múltiples arreglos y diferentes marcas o proveedores que son presentados virtualmente a múltiples servidores.

El almacenamiento virtualizado basado en la red es útil bajo las siguientes condiciones:

- Si se necesita datos accesibles a través de servidores y

- dispositivos heterogéneos.
- Si es necesario centralizar la administración del almacenamiento a través de todo el sistema de red de almacenamiento (NAS) o dispositivos de red de área de almacenamiento (SAN)
- Si se quiere asegurar que un reemplazo o una mejora en los dispositivos de almacenamiento no interrumpen el acceso a los datos.
- Si se quiere virtualizar el almacenamiento proporcionando bloques de servicios para aplicaciones.

Veritas Volume Manager (VxVM) es parte de una suite de productos que proporciona las facilidades para poder lograr el almacenamiento virtual.

5.2 ¿Qué es Veritas Volume Manager (VxVM)?

Es el software desarrollado por Veritas Software (*ahora Symantec*) que permite administrar el almacenamiento de discos físicos como dispositivos lógicos denominados volúmenes.

VxVM es una solución del manejo de almacenamiento en línea para organizaciones que requieren de un acceso a datos de misión crítica.

Permite aplicar las políticas de negocio para configurar, compartir y administrar el almacenamiento sin importar las limitaciones físicas de los discos. Reduce el costo total de propiedad permitiendo a los administradores crear fácilmente configuraciones de almacenamiento que mejoran el rendimiento y aumentan la disponibilidad de datos.

A través del soporte de las técnicas de redundancia RAID, VxVM protege los datos ante la falla de hardware, mientras provee la flexibilidad para aprovechar las capacidades del hardware.

Trabajando en conjunto con VERITAS File System, VERITAS Volume Manager crea una fundación para otras tecnologías de valor agregado como ambientes SAN, clustering y failover, administración automatizada, respaldo y HSM (Hierarchical Storage Management) y administración remota basada en administrador.

5.3 ¿Qué es VERITAS File System?

Como ya lo vimos en el capítulo 3 un file system es una colección de directorios organizados en una estructura que permite localizar y almacenar archivos. Toda la información procesada es

almacenada eventualmente en un file system. Los principales propósitos de un file system son:

- Ofrecer acceso compartido a datos almacenados
- Proporcionar acceso estructurado a los datos
- Controlar de acceso a los datos
- Ofrecer una interface de aplicación común y portátil
- Posibilitar la manejabilidad de almacenamiento de datos

El valor de un file system depende de su integridad y desempeño.

- Integridad: La información enviada al file system debe ser exactamente la misma cuando se recupera del mismo.
- Desempeño: Un file system no debe imponer una excesiva sobrecarga cuando responde a peticiones de I/O desde las aplicaciones

En la práctica, los requisitos necesarios para proporcionar la integridad y el rendimiento se complican. Por tanto, un file system debe ofrecer un balance entre estos dos requerimientos.

VERITAS File System es una herramienta poderosa, de rápida recuperación programada, que proporciona un alto rendimiento y fácil manejabilidad en línea requerida por las aplicaciones de misión crítica.

VERITAS File System amplía la administración de archivos de UNIX con una disponibilidad continua y optimiza el funcionamiento. Proporciona escalabilidad, optimiza la capacidad para satisfacer las crecientes demandas de cargas de usuario en entornos cliente / servidor.

5.4 Beneficios de VERITAS Storage Foundation

Hoy en día los sistemas comerciales exigen que el tiempo de actividad sea continuo para muchas implementaciones. Los sistemas deben estar disponibles 24 horas al día, 7 días a la semana, y los 365 días del año. VERITAS Storage Foundation reduce el costo ofreciendo una administración escalable, la disponibilidad y mejoras en el desempeño del ambiente de la informática.

5.4.1 Manejabilidad

- El manejo de almacenamiento y del file systems se realiza en línea en tiempo real, eliminando la necesidad de inactividad programada.
- Volúmenes en línea y la administración de file system puede ser realizada a través de una interfaz de usuario gráfica, intuitiva y de fácil uso que está integrada en el software de VERITAS Volume Manager.

- VxVM proporciona una administración compatible con las plataformas Solaris, HP-UX, AIX, Linux, y Windows 2000.
- Los comandos de operación de VxFS son compatibles con las plataformas Solaris, HP-UX, AIX y Linux.

5.4.2 Disponibilidad

- La integridad del almacenamiento se mantiene por espejeo en todas las operaciones de escritura.
- A través de técnicas de RAID, el almacenamiento sigue estando disponible en caso de fallo del hardware.
- La redundancia de datos se mantiene por la reubicación en caliente, que protege contra múltiples y simultáneas fallas de los discos.
- El tiempo de recuperación se reduce con la resincronización logueada.
- El reconocimiento de los cambios de un file system ofrece una rápida recuperación del mismo.

5.4.3 Rendimiento

- El rendimiento en los dispositivos de I/O puede ser maximizado mediante la medición y la modificación de los volúmenes, mientras que el almacenamiento sigue siendo en línea.
- Los cuellos de botella pueden ser localizados y eliminados usando las herramientas de análisis de VxVM.
- La asignación de espacio para los archivos minimiza en gran medida el tiempo de acceso a los mismos.
- La lectura se hace de forma anticipada colocando la información en buffer dinámicamente para el acceso a los archivos.
- El uso de la lectura en caché reduce el número de accesos a disco.
- Los dispositivos de entrada y salida realizan E / S directamente a la entrada y la salida de buffers de usuario.

5.4.4 Escalabilidad

- VxVM corre sobre sistemas operativos de 32 bits y 64 bits.
- El almacenamiento puede ser exportado a un gran número de plataformas.
- Los dispositivos de almacenamiento pueden ser expandidos.
- VxVM está totalmente integrado con VxFS.
- Existen varios productos que se pueden agregar a VxFS para maximizar el ambiente de bases de datos.

5.5 Veritas Storage Foundation

Veritas Storage Foundation es una solución completa para la administración del almacenamiento en línea heterogéneo. Respaldo por los productos líderes Veritas Volume Manager y Veritas File System, ofrece un conjunto estándar de herramientas integradas que permiten centralizar la administración del crecimiento de los datos, maximizar las inversiones en hardware de almacenamiento, proteger los datos y adaptarse a las necesidades cambiantes de la empresa.

En este capítulo se explicará el manejo de los productos Veritas Volume Manager (VxVM) y Veritas File System (VxFS) así como sus principales características

Este capítulo describe los objetos del almacenamiento virtual que son utilizados por VERITAS Volume Manager para administrar discos físicos. Vamos a introducir en las disposiciones principales del almacenamiento virtual. Ilustrando como se relacionan los objetos del almacenamiento virtual con los del almacenamiento físico, y describiendo los beneficios del almacenamiento virtual.

5.6 Importancia.

Antes de instalar y ejecutar VERITAS Storage Foundation es necesario familiarizarse con los objetos virtuales que utiliza VxVM para administrar los discos físicos. Un conocimiento conceptual de los objetos virtuales pueden ser de ayuda para interpretarlos y manejarlos los cuales son representados por interfaces, herramientas y reportes de VxVM.

Temas

- Almacenamiento Físico de Datos
- Almacenamiento Virtual de Datos
- Objetos de Almacenamiento de Volume Manager
- Disposiciones de Almacenamiento de Volume Manager

5.6.1 Almacenamiento Físico de Datos.

Objetos Físicos de Almacenamiento.

El dispositivo físico y básico de almacenamiento que últimamente almacena los datos es el disco duro. Como parte del programa de instalación de Solaris los discos duros son formateados. El formato del disco es el método básico para organizarlo así se prepara para la escritura de archivos para y del disco. Un disco formateado tiene un patrón de

almacenamiento preorganizado que está designado para el almacenamiento y la recuperación de datos.

5.6.2 Estructura Física de Datos.

Recordando del capítulo 3 los discos físicos en Solares se componen de las siguientes partes:

- VTOC: Tienen una área llamada tabla de contenidos (VTOC) que guarda información acerca de la estructura y la organización de los discos. La VTOC también es llamada Etiqueta del Disco.

En un disco manejado por Solaris, la VTOC es menor a 200 bytes y reside en el primer sector de el disco. Recordemos que un sector es de 512 bytes en la mayoría de los sistemas.

En el disco de arranque, el bloque de arranque se encuentra dentro de los primeros 16 sectores (8K). El bloque de arranque contiene instrucciones que indican en donde se encuentra el segundo proceso de arranque.

- Particiones: Después de la VTOC, el resto del disco es dividido en unidades llamadas particiones. Una partición es simplemente un grupo de cilindros contiguos para un uso particular. La información del tamaño, la localización y el uso de las particiones se guarda en la VTOC en la tabla de particiones.

Por convención; la partición 2 refiere al disco entero, incluyendo la VTOC. Esta partición también se conoce como la partición de respaldo.

5.6.3 Nombres para los discos físicos.

Se localizan y se encuentran datos dentro de los discos especificando el controlador, el ID del target y el número del disco. Un nombre típico para un dispositivo utiliza el formato: C##t#d#.

- c# es el numero de controladora
- t# es el ID del Target
- d# es el número de la unidad lógica (LUN)

Si un disco está dividido en particiones, entonces también se especifica el número de la partición.

- s# es el número de la partición

Por ejemplo, el dispositivo con nombre c0t0d0s1 está conectado al controlador número 0 en el sistema, con el ID de target 0, el disco 0, y la

partición 1.

Recordando del capítulo 4

5.6.4 Arreglos de Discos

La lectura y la escritura sobre los discos físicos puede convertirse en un proceso lento, debido a que los discos son dispositivos físicos que requieren de tiempo para el movimiento de las cabezas a la posición correcta antes de leer o de escribir. Si todas las operaciones de lectura y escritura se realizan en discos individuales, una a la vez, el tiempo de lectura y escritura puede ser inmanejable.

Un arreglo de discos es una colección de discos físicos. Las operaciones de lectura y escritura en múltiples discos en un arreglo de discos puede mejorar la velocidad y el rendimiento.

Algunos arreglos de discos proporcionan múltiples puertos para acceder a los discos. Estos puertos se acoplan con el controlador del adaptador de bus del host (HBA) y cualquier bus de datos o I/O del procesador local en el arreglo, construyen varios caminos para acceder a los discos. Estos tipos de arreglos de discos son llamados arreglos de discos multicamino.

Se pueden conectar arreglos de discos multicamino a sistemas host en varias configuraciones diferentes, como:

- Conectando múltiples puertos a controladoras diferentes sobre un solo host.
- Encadenando puertos a través de una sola controladora sobre un host.
- Conectando puertos a diferentes host simultáneamente

5.7 Almacenamiento Virtual de Datos.

VERITAS Volume Manager crea un nivel virtual de administración de almacenamiento por encima del nivel del dispositivo físico creando objetos de almacenamiento virtual. El objeto de almacenamiento virtual que es visible para los usuarios y para las aplicaciones se llama *volumen*.

5.7.1 ¿Qué es un Volumen?

Un Volumen es un objeto virtual creado por Volume Manager, que almacena datos. Está formado por espacio de uno o más discos en los cuales los datos están almacenados físicamente.

5.7.2 ¿Cómo se accesa a un Volumen?

Para el sistema operativo los volúmenes creados por VxVM aparecen como discos físicos, y las aplicaciones que interactúan con volúmenes trabajan de la misma forma que si lo hicieran con discos físicos. Todos los usuarios y las aplicaciones accesan a los volúmenes como espacios de direcciones contiguas utilizando archivos de dispositivo especial de una manera similar a la del acceso a una partición de disco.

Los Volúmenes tienen bloques y nodos de dispositivo en el directorio /dev. Puedes dar el nombre de la ruta para un volumen con los comandos y programas, en el file system y archivos de base de datos de la configuración y en otro contexto donde de lo contrario se utilizaría la ruta de la partición física del disco.

5.8 ¿Por qué usar Volume Manager?

Los beneficios del uso de Volume Manager para el almacenamiento virtual incluyen:

- **Cupo en Disco:** Con el uso de volúmenes y otros objetos virtuales, Volume Manager permite colocar datos sobre múltiples discos físicos. El proceso de combinar lógicamente dispositivos lógicos para habilitar los datos para ser almacenados a través de múltiples dispositivos se llama *spanning*.
- **Balanceo de Carga:** Los datos pueden ser extendidos a través de varios discos dentro de un arreglo para distribuir o balancear operaciones I/O a través de los discos. Utilizando I/O paralelo a través de múltiples discos mejora el trabajo de I/O aumentando la velocidad de transferencia de datos y el rendimiento global del arreglo.
- **Configuración compleja de multidisco:** Los objetos virtuales de Volume Manager permite crear configuraciones complejas de disco en sistemas de multidiscos que mejoran el rendimiento y la fiabilidad. Las configuraciones multidisco, como *striping*, *espejo* y *RAID 5* (vistas en el capítulo 4), pueden brindar redundancia de los datos, mejoras en el rendimiento y alta disponibilidad.
- **Administración en línea:** Volume Manager usa objetos virtuales para desempeñar tareas administrativas en los discos sin interrupción del servicio para aplicaciones y usuarios.
- **Alta Disponibilidad:** Volume Manager incluye un *failover* automático y características de recuperación que aseguran el acceso continuo a los datos críticos. VxVM puede mover colecciones de discos entre hosts, traslada automáticamente los datos en caso de una falla de disco y detecta y usa también en forma automática arreglos de discos *multipathed*.

5.9 Controladora de discos de Volume Manager

Con Volume Manager se habilita el almacenamiento virtual al traer un disco bajo el control de Volume Manager, esto significa que Volume Manager crea objetos virtuales y establece conexiones virtuales entre estos objetos y los discos.

5.10 Cuando un disco es puesto bajo el control de Volume Manager:

Volume Manager quita todas las entradas a la tabla de particiones de la VTOC, excepto la entrada 2 de la tabla de particiones. Como ya lo hemos visto la entrada 2 de la tabla de particiones contiene el disco entero, incluyéndola VTOC, y se usa para determinar el tamaño del disco.

Después reescribe la VTOC y crea dos particiones en el disco físico. Una partición contiene una región privada y la otra contiene una región pública.

Región privada. Guarda información como la cabecera de disco, copias de la configuración y kernel logs, Esta información es usada por Volume Manager para manejar los objetos virtuales. La región privada representa una pequeña sobrecabecera de control.

El tamaño default para la región privada es 2048 bloques (sectores), y el tamaño máximo es de 524288 bloques (sectores). Con 512 bytes por bloque, el tamaño default es de 1048576 bytes (1M) y el tamaño máximo es de 268435456 bytes (256 MB).

Región pública. Consiste en el resto del espacio del disco. Representa el espacio disponible que puede ser utilizado por Volume Manager para asignar los volúmenes y es donde las aplicaciones almacenan datos. Volume manager nunca sobre escribe esta area a no ser que no se indique lo contrario.

Por convención la región pública de un disco controlado por Volume Manager es referida como un disco Volume Manager o disco VxVM. La verdadera definición de un disco VxVM es un disco controlado por Volume Manager.

La etiqueta de partición. VxVM configura las etiquetas de las particiones, los valores numéricos que describen el file system montado sobre una partición, para las regiones pública y privada.

- La etiqueta 14 es siempre utilizada para la región pública del disco.
- La etiqueta 15 es siempre utilizada para la región privada del disco.

Si un disco no tiene particiones bajo el control de Volume Manager, entonces VxVM crea primero la región privada y luego la región pública en el disco.

Una vez que el disco es controlado por Volume Manager este actualiza la VTOC con información acerca de la eliminación de particiones existentes y la adición de nuevas particiones, las cuales corresponden a las regiones públicas y privadas.

5.11 Objetos de Almacenamiento de Volumen Manager

Objetos Virtuales

Un Volumen es un Objeto Virtual que graba y recupera datos de uno y mas discos físicos. Los Volúmenes son uno de varios tipos de objetos virtuales usados por Volume Manager para administrar el almacenamiento.

Los objetos virtuales de Volume Manager incluyen:

- Grupos de discos
- Discos Volume Manager
- Subdiscos
- Plexes
- Volúmenes

Grupos de Discos.

Un grupo de discos es una colección de discos VxVM. Se agrupan discos con propósitos de administración, como para cargar los datos de una aplicación o de un conjunto de aplicaciones. Por ejemplo los datos de una aplicación de contabilidad pueden ser organizados en un grupo de discos llamado acctdg.

Una base de datos de la configuración es un set de registros con información detallada a cerca de todos los objetos de Volume Manager en un grupo de discos, incluyendo atributos de objetos y sus conexiones.

Los grupos de discos son configurados por el administrador del sistema y representan los límites de manejo y configuración. Los objetos de Volume Manager no pueden empalmar grupos de discos. Por ejemplo, un subdisco de volumen, plexes y discos deben ser derivados del mismo grupo de discos o disk group como el volumen.

Se pueden crear grupos de discos adicionales si es necesario. Un grupo de discos permite agrupar discos en una colección lógica.

Los grupos de discos ofrecen una alta disponibilidad, ya que un grupo de discos y sus componentes pueden ser movidos como una unidad de un servidor a otro. Las unidades de disco pueden ser compartidas por dos o más servidores, pero solo pueden ser accesadas por un servidor a la vez. Si un servidor tiene una falla, el otro servidor puede tomar las unidades de

disco así como los grupos de discos.

Discos Volume Manager.

Un disco Volume Manager (VxVM) se crea desde una región pública de un disco físico que está bajo el control de Volume Manager. Cada disco VxVM corresponde a un disco físico.

Cada disco VxVM tiene un único nombre de disco virtual llamado *disk media name*. El *disk media name* es un nombre lógico utilizado por Volume Manager con propósitos de administración. Volume Manager utiliza este nombre cuando asigna espacio a volúmenes. Un disco VxVM obtiene un *disk media name* cuando se agrega a un grupo de discos.

Se puede proporcionar el *disk media name* o se puede dejar que Volume Manager asigne un nombre por default que típicamente es de la forma *diskgroup#*, donde *diskgroup* es el nombre del grupo de discos. El *disk media name* es almacenado con un único ID para evitar colisiones de nombres.

Una vez que el disco VxVM tiene un nombre ya no es referido por su dirección física de *c# t# d#*. La dirección física de *c# t# d#* se conoce como el registro de acceso al disco.

Notas sobre el nombramiento de discos VxVM

- El grupo de discos o *diskgroup rootdg* es un grupo de discos especial que sigue convenciones diferentes en los nombres. Para los discos en el grupo *rootdg* los nombres de discos por default son: *disk01*, *disk02* y así continuamente.
- Si se utiliza la línea de comandos para administrar Volume Manager, entonces se utilizan los nodos del dispositivo para nombrar a los discos, a menos que se especifique el *disk media name*.

Subdiscos.

Un disco VxVM puede ser dividido en uno o más discos. Un subdisco es un conjunto de bloques contiguos que representan una porción específica del disco VxVM. Los subdiscos son también una subsección de la región pública de los discos. Son la unidad mas pequeña de almacenamiento en Volume Manager. Son la base de los objetos de Volume Manager.

Un subdisco de un disco VxVM está definido por un *offset* y una longitud en sectores. Los nombres por default para los subdiscos son de la forma *DMname-##*. El nombre se toma del nombre VxVM *disk media name*, un guión y un número de dos dígitos.

Un disco VxVM puede contener múltiples subdiscos, pero los

subdiscos no pueden traslapar o compartir la misma porción de disco VxVM. Cualquier espacio en el disco VxVM que no sea reservado o que no sea parte de un subdisco es espacio libre. Se puede utilizar el espacio libre para crear nuevos subdiscos.

Conceptualmente, un subdisco es como una partición: ambos los subdiscos y las particiones dividen a un disco en piezas definidas como una dirección y una longitud, cada una de estas piezas representa espacios contiguos en el disco físico.

Sin embargo existen diferencias:

- El número máximo de particiones en un disco son 8.
- No hay un límite teórico del número de subdiscos que pueden ser atados a un solo plex.

Plexes

Volume manager utiliza subdiscos para crear objetos virtuales llamados *plexes*. Un plex es una colección ordenada o estructurada de subdiscos, ésta colección representa una copia de los datos en el volumen. Un plex consiste en uno o mas subdiscos localizados en uno o más discos físicos.

Un plex es también llamado espejo. Los términos plex y espejo pueden ser utilizados intercambiamente, aunque un plex es solo una copia de los datos. Los términos “espejeado” o “espejeo” se aplican en dos o mas copias de los datos.

El tamaño de un plex está determinado por el último bloque que puede ser leído y escrito sobre el último subdisco en el plex. El tamaño del plex puede ser distinto al tamaño exacto de un sector, porque el plex está alineado a un límite de cilindros.

La convención por default para nombrar plexes es *volumename-##*. El nombre de plex por default consiste en el nombre del volumen, un guón y un número de dos dígitos.

Tipos de plexes.

Los plexes pueden ser categorizados en tres tipos.

- Plex completo o complete plex: contiene una copia completa de un volumen y por lo tanto mapas del espacio entero de direcciones del volumen. Un volumen debe tener al menos un plex completo. La mayoría de los plexes en VxVM son plexes completos. Por ejemplo si un volumen tiene el tamaño de 1 MB, entonces el plex completo debe tener al menos 1 MB de tamaño y el espacio de direcciones de 1 MB debe ser mapeado en uno o más subdiscos, cuyos tamaños combinados no sobrepasan de 1MB sin huecos en el espacio de direcciones.
- Plex escaso o sparse plex: Es un plex que tiene un tamaño

menor al tamaño del volumen o que mapea a sólo una parte del espacio de direcciones del volumen. Los plex escaso no son utilizados comúnmente en las versiones recientes de VxVM, con la finalidad de que sean usados para la mejora del desempeño. Por ejemplo la memoria RAM de un disco utiliza los plexes escasos para mapear a un punto específico dentro de un volumen para mejorar el desempeño de lectura. La memoria RAM de un disco tiene que ser de tamaño suficiente para cubrir el punto específico y que no necesite mapear el volumen completo.

- Plex registrado o log plex. Es un plex que está dedicado para registrar. Es usado para hacer más rápido el proceso de revisión de consistencia en los datos y la reparación después de un fallo en el sistema. Los volúmenes RAID-5 y el espejo típicamente usan un log plex.

Un volumen debe tener al menos un plex completo que contenga una copia de los datos en el volumen con al menos un subdisco asociado. Otros plexes en el volumen también pueden ser completos, escasos o registrados. Un volumen puede tener arriba de 32 plexes; sin embargo no se deben usar más de 31 plexes en un volumen. Volume Manager requiere un plex para las operaciones automáticas o temporalmente en línea.

Volúmenes

Los volúmenes son dispositivos de almacenamiento virtual que son utilizados por las aplicaciones en la misma forma que los discos físicos. Debido a su naturaleza virtual, un volumen no está restringido por las limitaciones del tamaño físico que aplican a los discos físicos. Un volumen está compuesto por uno o más plexes.

Un volumen puede expandirse a través de múltiples discos. Los datos en el volumen están almacenados en subdiscos de discos expandidos. Un volumen debe ser configurado de discos VxVM y subdiscos dentro del mismo grupo de discos.

Volume Manager utiliza la convención por default para los nombres *vol##* para los volúmenes. Donde *##* representa un número de dos dígitos. Se pueden asignar nombres significativos que reflejen la naturaleza o el uso de los datos en el volumen.

Escribiendo en un Volumen.

En el ejemplo, el plex llamado *vol01-01* consiste de dos subdiscos. Cada subdisco viene de un disco VxVM distinto.

- *datadg01-02*, tiene un tamaño de 20 MB viene del disco *datadg01*.

- datadg02-03, tiene un tamaño de 50 MB viene del disco datadg02.

El tamaño de un plex está definido como el último byte accesible en el plex, entonces en el ejemplo, el tamaño del plex es de 70 MB. El primer subdisco ocupa los primeros 20 MB del espacio de direcciones del plex y el segundo subdisco ocupa los siguientes 50 MB del espacio de direcciones, desde 20 MB hasta 70 MB.

Si una aplicación escribe en los bytes localizados en 10 MB en el plex, entonces los datos se escriben en el subdisco datadg01-02.

Si una aplicación escribe en los bytes localizados en 60MB en el plex, entonces los datos se escriben en el subdisco datadg02-03. Las aplicaciones empiezan escribiendo en 40MB desde el principio del subdisco. Los bytes localizados en 60MB dentro del plex son 40MB desde el comienzo del subdisco data02-03, porque 20MB mas 40 MB es igual a 60 MB.

RAID con Volume Manager.

La organización de los plexes en un volumen. Es la forma en la que los plexes están configurados para remapear el espacio de direcciones del volumen por lo cual los procesos de I/O están restringidos en el tiempo de ejecución. La organización de los plexes está basada en el concepto de discos con spanning, lo que es la habilidad de combinar lógicamente discos físicos para almacenar datos en múltiples discos.

Existen una variedad de formas disponibles para organizar los discos de forma lógica, cada una de éstas formas tienen sus ventajas y sus desventajas. La forma que se elija depende de los niveles de desempeño y confiabilidad requerida para el sistema.

Con Volume Manager se pueden cambiar las formas de organización de los discos sin interrumpir las aplicaciones o los file systems que está utilizando el volumen. Estas organizaciones pueden ser configuradas, reconfiguradas, se puede modificar el tamaño y sincronizar mientras que el volumen permanece accesible.

Las organizaciones soportadas por los volúmenes incluyen:

- Concatenado
- Striped
- Espejeado
- RAID-5
- Por capas

Concatenado.

En un volumen concatenado los subdiscos están arreglados secuencialmente y contiguos dentro de un plex. La concatenación permite que un volumen sea creado desde regiones múltiples de uno o mas discos si no existe suficiente espacio de un volumen entero en una sola región de un disco.

Striped.

En un volumen striped, los datos son extendidos uniformemente a través de múltiples discos. Stripes son fragmentos de igual tamaño que son asignados alternativamente en el subdisco de un solo plex. Debe haber al menos dos subdiscos en un plex striped, cada uno de ellos debe existir en un diferente disco. El rendimiento se incrementa con el número de discos en los que un plex es striped. Striping ayuda al balance de la carga de I/O en los casos en los que las areas de alto tráfico existen en ciertos discos.

Espejeado.

Un volumen espejeado utiliza varios plexes para duplicar la información contenida en un volumen. Aunque un volumen puede tener un solo plex, al menos dos son requeridos para un verdadero espejo (redundancia de datos). Cada uno de esos plexes pueden contener espacio en disco desde diferentes discos para que la redundancia sea útil.

RAID-5

Un volumen RAID-5 utiliza striping para extender los datos uniformemente a través de varios discos en un arreglo. Cada stripe contiene una unidad de paridad stripe y unidades de datos stripe. La paridad puede ser utilizada para reconstruir los datos si uno de los discos falla. En comparación con el trabajo de los volúmenes striped, el desempeño en la escritura de volúmenes RAID-5 decrece, porque la información de paridad necesita se actualizada cada vez que los datos son accedados.

Sin embargo, en comparación con el espejo, el uso de la paridad reduce la cantidad de espacio requerido.

Por capas.

Un volumen organizado por capas es un objeto virtual Volume Manager que anida otros objetos virtuales. Los volúmenes por capas ofrecen mejor redundancia por el espejeo en un nivel mas granular.

Hemos dado un vista general a la herramienta de software mostrando las principales características y capacidades del mismo.

6

Solución de Almacenamiento en una empresa del sector eléctrico

En este capítulo se plantea la solución ofrecida a una empresa mexicana del sector eléctrico.

Primeramente se mostrará el análisis de la situación inicial, posteriormente se presentara la propuesta ofrecida haciendo uso de herramientas mencionadas en el capítulos anteriores y finalmente se mostrara el resultado obtenido.

6.1 Descripción de la Compañía.

Es una empresa del gobierno mexicano que genera, transmite, distribuye y comercializa energía eléctrica para más de 26.2 millones de clientes, lo que representa a casi 80 millones de habitantes, e incorpora anualmente más de un millón de clientes nuevos.

En ella se produce la energía eléctrica utilizando diferentes tecnologías y diferentes fuentes de energético primario. Tiene centrales termoeléctricas, hidroeléctricas, carboeléctricas, geotermoeléctricas, eoloeléctricas y una nucleoeléctrica,

La empresa ofrece el servicio de energía eléctrica en la mayor parte del país, con excepción del Distrito Federal y algunas poblaciones cercanas a éste, donde el servicio está a cargo de otra compañía del mismo ramo.

Esta empresa es también la entidad del gobierno federal encargada de la planeación del sistema eléctrico nacional, la cual describe la evolución

del mercado eléctrico, así como la expansión de la capacidad de generación y transmisión para satisfacer la demanda en los próximos diez años, y se actualiza anualmente.

6.2 Situación Actual

La empresa cuenta con más de 80,000 empleados distribuidos en las diversas divisiones de la compañía a lo largo del país. Dada esta magnitud de trabajadores las áreas de recursos humanos como la mayoría de las empresas grandes se apoyan en sistemas de cómputo para la administración del su personal. La infraestructura a analizar será la del sistema que maneja la nómina de los empleados, funciones de recursos humanos así como de finanzas. Este está implementado utilizando el software peoplesoft

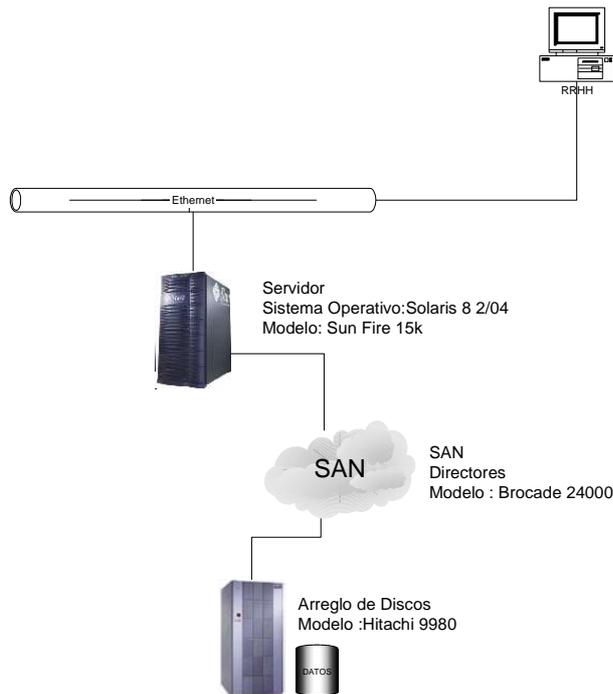
PeopleSoft es un software dentro de la categoría de un ERP (**sistemas de planificación de recursos** de la empresa inglés **ERP**, “*Enterprise Resource Planning*” son sistemas de gestión de información que integran y automatizan muchas de las prácticas de negocio asociadas con los aspectos operativos o productivos de una empresa) propietario de ORACLE.

Como se menciona en el párrafo anterior en este trabajo únicamente analizará la infraestructura que está utilizando el sistema de nómina, es decir, nos enfocaremos a la arquitectura a nivel hardware y manejador de volúmenes.

6.3 Análisis de la problemática.

El reporte inicial del usuario es que cuenta con una problemática que consiste en que sus file systems están a punto de llegar a su capacidad total y ya no cuenta con ninguna información que pueda ser depurada para liberar espacio. Comenta que cuando se ha presentado esta situación la acción que toma es bajar las aplicaciones, reestructurar los file systems para que posteriormente baje el respaldo en file systems de mayor capacidad.

El usuario nos dio acceso a su infraestructura y se recolecto toda información necesaria para dar un diagnostico sobre la problemática



A continuación se muestra una salida del comando `format` para indicar la cantidad de discos asignados:

```
Searching for disks...done
```

```
#format
```

```
AVAILABLE DISK SELECTIONS:
```

```
0. c0t0d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
   /pci@3c,700000/pci@1/pci@1/scsi@2/sd@0,0
1. c0t1d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
   /pci@3c,700000/pci@1/pci@1/scsi@2/sd@1,0
2. c0t2d0 <SEAGATE-ST373207LSUN72G-045A cyl 14087 alt 2 hd 24 sec 424>
   /pci@3c,700000/pci@1/pci@1/scsi@2/sd@2,0
3. c1t0d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
   /pci@5c,700000/pci@1/pci@1/scsi@2/sd@0,0
4. c1t1d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
   /pci@5c,700000/pci@1/pci@1/scsi@2/sd@1,0
5. c1t2d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
   /pci@5c,700000/pci@1/pci@1/scsi@2/sd@2,0
6. c4t50060E80042C9631d0 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
   /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,0
7. c4t50060E80042C9631d1 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
   /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,1
8. c4t50060E80042C9631d2 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
   /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,2
9. c4t50060E80042C9631d3 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
```

```

    /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,3
10. c4t50060E80042C9631d4 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
    /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,4
11. c4t50060E80042C9631d5 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
    /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,5
12. c4t50060E80042C9631d6 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
    /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,6
13. c4t50060E80042C9631d7 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
    /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,7
14. c4t50060E80042C9631d8 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
    /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,8
15. c4t50060E80042C9631d9 <HITACHI-OPEN-9-SUN-5005 cyl 10014 alt 2 hd 15 s
ec 96>
.
.
.
.
.

187. c4t50060E80042C9631d181 <HITACHI-OPEN-9-SUN-5009 cyl 10014 alt 2 hd 15
sec 96>
    /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,b5
188. c4t50060E80042C9631d182 <HITACHI-OPEN-9-SUN-5009 cyl 10014 alt 2 hd 15
sec 96>
    /pci@3c,600000/SUNW,qlc@1/fp@0,0/ssd@w50060e80042c9631,b6
189. c4t50060E80042C9631d183 <HITACHI-OPEN-9-SUN-5009 cyl 10014 alt 2 hd 15
sec 96>

```

184 discos Hitachi asignados únicamente por una sola tarjeta controladora como lo indica el siguiente comando:

```
# luxadm -e port
```

```
Found path to 4 HBA ports
```

```

/devices/pci@3c,600000/SUNW,qlc@1/fp@0,0:devctl          CONNECTED
/devices/pci@3c,600000/SUNW,qlc@1,1/fp@0,0:devctl      NOT CONNECTED
/devices/pci@5c,600000/SUNW,qlc@1/fp@0,0:devctl      NOT CONNECTED
/devices/pci@5c,600000/SUNW,qlc@1,1/fp@0,0:devctl      NOT CONNECTED

```

Salida del comando `df -k` donde muestra que todos los file system están montados sin ningún manejador de volúmenes.

```
#df -k
```

```

Filesystem      kbytes used  avail capacity  Mounted on
/dev/md/dsk/d0  8669357 317864 8264800 4%    /
/dev/md/dsk/d20 10081900   939217 9041864 10%   /usr
/proc           0    0    0%    /proc
mnttab         0    0    0%    /etc/mnttab
fd             0    0    0%    /dev/fd
/dev/md/dsk/d30 16136042   6706516 9268166 42%   /var
swap          59560432    32 59560400    1%   /var/run

```

```

swap 59833832 273432 59560400 1% /tmp
/dev/md/dsk/d40 15123627 5297983 9674408 36% /opt
/dev/md/dsk/d80 16536938 10 16371559 1% /var/crash
/dev/dsk/c4t50060E80042C9631d0 12032000 9130267 2720378 78% /db-psaurhum1
/dev/dsk/c4t50060E80042C9631d1 5781654 4362293 1330655 77% /db-rhum1
/dev/dsk/c4t50060E80042C9631d2 1740800 2157 1629985 1% /psft8
/dev/dsk/c4t50060E80042C9631d3 6502400 4905165 1497412 77% /db-psaurhum3
/dev/dsk/c4t50060E80042C9631d4 13541376 11248052 2221660 84% /db2-psfsprd
/dev/dsk/c4t50060E80042C9631d5 5781504 4393028 1301701 78% /db-rhum3
/dev/dsk/c4t50060E80042C9631d6 10874880 8234297 2475550 77% /db-rhum2
/dev/dsk/c4t50060E80042C9631d7 2764800 2076028 645730 77% /dbquest-psaurhum
/dev/dsk/c4t50060E80042C9631d8 13688832 10375118 3106611 77% /db-psaurhum2
/dev/dsk/c4t50060E80042C9631d9 481280 37206 416446 9% /Control-M
/dev/dsk/c4t50060E80042C9631d10 7219200 5533678 1580233 78% /db-psaurhum4
/dev/dsk/c4t50060E80042C9631d11 17405952 13099622 4171760 76% /db3-psfsprd
/dev/dsk/c4t50060E80042C9631d12 17264640 5180076 11330705 32% /quest
/dev/dsk/c4t50060E80042C9631d13 31141478 23622031 7049485 78% /db-psaurhum6
/dev/dsk/c4t50060E80042C9631d14 9211904 6690782 2363557 74% /db-psaurhum5
/dev/dsk/c4t50060E80042C9631d15 13977600 11482068 2417610 83% /db1-psfsprd
/dev/dsk/c4t50060E80042C9631d16 15845620 13196776 2566068 84% /db4-psfsprd
/dev/dsk/c4t50060E80042C9631d17 307200 235551 67311 78% /Tivoli
/dev/dsk/c4t50060E80042C9631d18 1198080 699592 470108 60% /prog
/dev/dsk/c4t50060E80042C9631d19 235520 194075 38878 84% /foglight
/dev/dsk/c4t50060E80042C9631d20 716800 389434 306941 56% /forte6
/dev/dsk/c4t50060E80042C9631d21 92740 18881 70160 22% /controlsa
/dev/dsk/c4t50060E80042C9631d22 10539008 8014407 2368743 78% /oracle
/dev/dsk/c4t50060E80042C9631d23 51200 1753 46360 4% /users_dba
/dev/dsk/c4t50060E80042C9631d24 20400128 14742844 5310279 74% /aplicaciones
/dev/dsk/c4t50060E80042C9631d25 4376576 2907107 1378242 68% /plansip2000
/dev/dsk/c4t50060E80042C9631d26 307200 171711 127042 58% /precise
/dev/dsk/c4t50060E80042C9631d27 266240 31722 219953 13% /tuxedo65j12
/dev/dsk/c4t50060E80042C9631d28 20561920 7331740 13024404 37% /resporacle
/dev/dsk/c4t50060E80042C9631d29 307200 70960 221537 25% /COBOL
/dev/dsk/c4t50060E80042C9631d30 8702689 4084139 4394331 49% /psft
/dev/dsk/c4t50060E80042C9631d31 1075200 598908 446535 58% /users
/dev/dsk/c4t50060E80042C9631d32 39602748 33074648 6324158 84% /db5-psfsprd
/dev/dsk/c4t50060E80042C9631d33 23324672 11801722 10897843 52% /psf8
/dev/dsk/c4t50060E80042C9631d34 13312000 10364896 2855012 79% /db6-psfsprd
/dev/dsk/c4t50060E80042C9631d35 11724800 8196822 3417736 71% /db11-psfsprd
/dev/dsk/c4t50060E80042C9631d36 15128576 10959688 4038616 74% /db12-psfsprd
/dev/dsk/c4t50060E80042C9631d37 204800 11486 181295 6% /changemanDS
/dev/dsk/c4t50060E80042C9631d38 13438976 10467228 2878886 79% /db8-psfsprd
/dev/dsk/c4t50060E80042C9631d39 13516800 11268072 2178462 84% /db10-psfsprd
/dev/dsk/c4t50060E80042C9631d40 204800 9045 183536 5% /chgmanPSOFT
/dev/dsk/c4t50060E80042C9631d41 204800 16071 176942 9% /monitor_space
/dev/dsk/c4t50060E80042C9631d42 204800 83718 113521 43% /dbreco2-sip2000
/dev/dsk/c4t50060E80042C9631d43 4929536 2562968 2218664 54% /dbundo1-sip2000
/dev/dsk/c4t50060E80042C9631d44 204800 83710 113529 43% /dbreco1-sip2000
/dev/dsk/c4t50060E80042C9631d45 16328704 13332349 2809088 83% /dbdata1-sip2000
/dev/dsk/c4t50060E80042C9631d46 5529600 411834 4797913 8% /dbtemp1-sip2000
/dev/dsk/c4t50060E80042C9631d47 16066560 13124404 2758276 83% /dbdata2-sip2000
/dev/dsk/c4t50060E80042C9631d48 1048576 8021 976358 1% /dbtrac-sip2000
/dev/dsk/c4t50060E80042C9631d49 78944152 61079880 17724776 78% /db14-psfsprd
/dev/dsk/c4t50060E80042C9631d50 409600 262060 138325 66% /dbsyst1-sip2000
/dev/dsk/c4t50060E80042C9631d51 491520 130751 340792 28% /soptec
/dev/dsk/c4t50060E80042C9631d52 11415552 9065984 2331224 80% /db13-psfsprd
/dev/dsk/c4t50060E80042C9631d53 307200 210560 95896 69% /dbsyst1-apweb
/dev/dsk/c4t50060E80042C9631d54 1355776 614168 735824 46% /dbtemp1-apweb
/dev/dsk/c4t50060E80042C9631d55 3428352 2668408 754016 78% /dbindx1-apweb
/dev/dsk/c4t50060E80042C9631d56 7315552 4954128 2342984 68% /dbdata1-apweb
/dev/dsk/c4t50060E80042C9631d57 102400 9480 92264 10% /dbtrac-apweb
/dev/dsk/c4t50060E80042C9631d58 102400 43856 58096 44% /dbreco1-apweb
/dev/dsk/c4t50060E80042C9631d59 102400 43856 58096 44% /dbreco2-apweb
/dev/dsk/c4t50060E80042C9631d60 307200 210560 95896 69% /dbundo1-apweb
/dev/dsk/c4t50060E80042C9631d61 1228800 90192 1068130 8% /planspsoft
/dev/dsk/c4t50060E80042C9631d62 102400 2432 99256 3% /brightstor
/dev/dsk/c4t50060E80042C9631d63 16433152 13211497 3020307 82% /psft_SO
/dev/dsk/c4t50060E80042C9631d64 8140800 4217579 3678026 54% /dbquest-psfsprd
/dev/dsk/c4t50060E80042C9631d65 17879040 12944026 4626594 74% /psrhr89
/dev/dsk/c4t50060E80042C9631d66 22183936 17491534 4399157 80% /dbdata1-fspr89sy
/dev/dsk/c4t50060E80042C9631d67 102400 43856 58096 44% /dbreco3-apweb
/dev/dsk/c4t50060E80042C9631d68 28262401 20569507 7212108 75% /psfspr89
/dev/dsk/c4t50060E80042C9631d69 512000 133252 355083 28% /dbreco1-fspr89sy
/dev/dsk/c4t50060E80042C9631d70 14014464 10660968 3143909 78% /dbindx1-fspr89sy

```

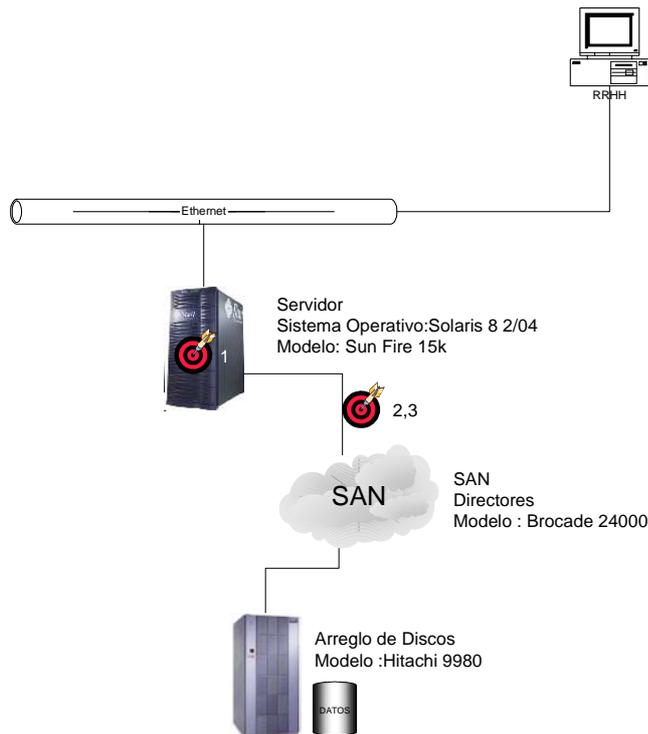
```

/dev/dsk/c4t50060E80042C9631d71 512000 338054 163080 68% /dbreco2-fspr89sy
/dev/dsk/c4t50060E80042C9631d72 3584000 2741821 789550 78% /dbsyst1-fspr89sy
/dev/dsk/c4t50060E80042C9631d73 6119424 2068847 3797422 36% /dbtemp1-fspr89sy
/dev/dsk/c4t50060E80042C9631d74 102400 66816 33364 67% /dbtrac-fspr89sy
/dev/dsk/c4t50060E80042C9631d75 16703488 12821226 3639627 78% /dbundo1-fspr89sy
/dev/dsk/c4t50060E80042C9631d76 33456128 26564105 6461316 81% /dbdata1-rhpr89sy
/dev/dsk/c4t50060E80042C9631d77 3072000 2065852 943271 69% /dbsyst1-rhpr89sy
/dev/dsk/c4t50060E80042C9631d78 512000 419415 86805 83% /dbreco1-rhpr89sy
/dev/dsk/c4t50060E80042C9631d79 15384576 10260873 4803479 69% /dbundo1-rhpr89sy
/dev/dsk/c4t50060E80042C9631d80 15630336 12871018 2586867 84% /dbindx1-rhpr89sy
/dev/dsk/c4t50060E80042C9631d81 512000 235850 258897 48% /dbreco2-rhpr89sy
/dev/dsk/c4t50060E80042C9631d82 16766976 13842970 2741262 84% /dbtemp1-rhpr89sy
/dev/dsk/c4t50060E80042C9631d83 7340032 852465 6082100 13% /dbdata2-apweb
/dev/dsk/c4t50060E80042C9631d84 2621440 2065740 520976 80% /dbquest-rhpr89sy
/dev/dsk/c4t50060E80042C9631d85 7340032 2347537 4680470 34% /dbdata3-apweb
/dev/dsk/c4t50060E80042C9631d86 3145728 503261 2477319 17% /dbindx2-apweb
/dev/dsk/c4t50060E80042C9631d87 3145728 768485 2228672 26% /dbindx3-apweb
/dev/dsk/c4t50060E80042C9631d88 399360 279738 112168 72% /productos
/dev/dsk/c4t50060E80042C9631d89 3401803 2227576 1100903 67% /planmaap
/dev/dsk/c4t50060E80042C9631d90 307200 87968 205535 30% /dbtrac_rhpr89sy
/dev/dsk/c4t50060E80042C9631d91 307200 97670 197799 34% /COBOL40
/dev/dsk/c4t50060E80042C9631d92 7739392 6266482 1380859 82% /dbquest-fspr89sy
/dev/dsk/c4t50060E80042C9631d93 8388608 984422 6941491 13% /arch_sip2000
/dev/dsk/c4t50060E80042C9631d94 23068672 1701216 21200592 8% /arch_rhpr89sy
/dev/dsk/c4t50060E80042C9631d95 70815744 4643784 65655120 7% /arch_fspr89sy
/dev/dsk/c4t50060E80042C9631d96 19922944 56031 18625301 1% /var/audit
/dev/dsk/c4t50060E80042C9631d97 10240000 6621255 3392627 67% /proarchivlog/WebSphere
/dev/dsk/c4t50060E80042C9631d98 5242880 1863039 3173236 37% /productos/WebSphere
/dev/dsk/c4t50060E80042C9631d99 61440 1738 55978 4% /var/log/sudo
/dev/dsk/c4t50060E80042C9631d100 1560576 462923 1029057 32% /dbsaux1-sip2k
/dev/dsk/c4t50060E80042C9631d101 409600 319282 84680 80% /dbsyst1-sip2k
/dev/dsk/c4t50060E80042C9631d102 204800 140029 60730 70% /dbreco1-sip2k
/dev/dsk/c4t50060E80042C9631d103 204800 140029 60730 70% /dbreco2-sip2k
/dev/dsk/c4t50060E80042C9631d104 5120000 4112623 944423 82% /dbtemp1-sip2k
/dev/dsk/c4t50060E80042C9631d105 6782976 5138776 1541444 77% /dbundo1-sip2k
/dev/dsk/c4t50060E80042C9631d106 18300928 14179556 3863790 79% /dbdata1-sip2k
/dev/dsk/c4t50060E80042C9631d107 1331200 834108 466091 65% /dbtrac-sip2k
/dev/dsk/c4t50060E80042C9631d108 1536000 17614 1423491 2% /tsm_psoft
/dev/dsk/c4t50060E80042C9631d109 614400 203857 384949 35% /var/opt/OV
/dev/dsk/c4t50060E80042C9631d110 1048576 381672 625269 38% /Tivoli62

```

Una vez realizado el análisis de su infraestructura de obtiene los siguientes puntos de mejora:

- 1) No se cuenta con ningún manejador de volúmenes.
- 2) Se tiene que distribuir los discos en distintas HBA para mejorar el tiempo de respuesta en el acceso a disco.
- 3) No se cuenta con canales redundantes en el acceso a disco.
- 4) Únicamente se cuenta con RAID por hardware RAID5.
- 5) Todos los File system son de tipo UFS
- 6) Se tienen cuellos de botella en el acceso a disco.



6.4 Propuesta.

Después del análisis en la infraestructura se dieron las siguientes recomendaciones :

- 1) Como primera acción se dio una presentación al cliente de las ventajas que obtendrá al adquirir el producto Veritas Storage Foundation y convertir sus datos en volúmenes controlados por Veritas Volume Manager y al convertir sus File Systems en tipo vxfs con el producto Veritas File System.
- 2) Se recomienda tener canales redundantes en el acceso a disco con esto se elimina el tener un punto único de falla en el acceso a disco previniendo cualquier falla física ya sea por fallas en las tarjetas HBA o falla en el cableado de fibra óptica, Veritas Volume Manager cuenta con un modulo integrado llamado Veritas DMP (Dynamic Multi-Pathing).
- 3) Distribuir los discos en dos canales distintos para distribuir la carga.

A continuación se comentará cada una de las recomendaciones mencionadas.

Para la primera recomendación se dio hincapié al explicarle al cliente que al utilizar el producto Veritas Storage Foundation obtiene las siguientes funciones y ventajas clave:

Funciones principales

- Administración del almacenamiento en línea compatible con el sistema operativo heterogéneo (Solaris, Linux, HP-UX y AIX) y un amplio conjunto de dispositivos y matrices de almacenamiento (Portabilidad de datos).
- Storage Foundation Manager permite la administración centralizada de diversas aplicaciones, servidores y almacenamiento.
- Las varias rutas dinámicas permiten que la entrada/salida se disemine en varias rutas, lo que brinda una protección contra fallas en las rutas y una rápida conmutación por error.
- El almacenamiento dinámico por niveles permite que los datos se trasladen dinámicamente a diferentes niveles de almacenamiento para responder rápidamente a las cambiantes necesidades de la empresa
-

Ventajas

- Aumenta la utilización del almacenamiento en todos los sistemas operativos heterogéneos y las matrices de almacenamiento.
- Aumenta la eficacia operativa gracias a la completa visibilidad y administración centralizada de las aplicaciones, los servidores y el almacenamiento en todos los hosts.
- Mejor disponibilidad y rendimiento de la ruta de E/S del almacenamiento y propagación más eficiente de la E/S en todas las rutas.
- Libera los datos de los cambios tecnológicos: migra datos dinámicamente a niveles diferentes de almacenamiento y realiza la migración de datos sin problemas en todas las diferentes arquitecturas del servidor.

Hablando de la segunda y tercera recomendación donde se menciona que existe una vulnerabilidad fuerte al contar únicamente con una sola ruta de acceso a los discos se resaltó al cliente que el contar con más de una ruta de acceso a sus discos es sumamente importante para la disponibilidad de su servicio.

Al momento de adquirir el producto Veritas Storage Foundation incluye el módulo de Veritas DMP (Veritas Dynamic Multipathing) .

Esta característica está disponible en arreglos de disco que permiten la conexión a múltiples controladores. DMP tiene dos grandes características. En primer lugar, ofrece múltiples "caminos" o conexiones físicas de cada uno de los discos de un arreglo, por lo que si un componente falla (ejemplo controladora, el cable, GBIC, I / O, etc.) el disco sigue siendo accesible. En segundo lugar, porque hay múltiples accesos al disco realizando un balanceo de carga en cada uno de los canales, obteniendo con ello un mejor desempeño en el tiempo de I/O.

6.5 Implementación.

El cliente finalmente optó por adquirir el producto de Veritas Storage Foundation una vez visto las bondades que el software le ofrece a su administración.

A continuación daremos la información completa sobre la instalación de Veritas Storage Foundation en el servidor de nómina de la empresa.

6.5.1 Instalación.

Antes de instalar VERITAS Foundation Suite, es necesario considerar el contenido de los discos y decidir cómo es que se desea que Volume Manager los manipule. Es necesario que la instalación sea de acuerdo a lo que el ambiente requiere.

Prerrequisitos:

Antes de los procedimientos de instalación se debe estar seguro de que la versión de VxVM y VxFS que se va a instalar sea compatible con la versión del sistema operativo Solaris que se está. Si es necesario se tendrá que actualizar el sistema operativo a la versión compatible.

El software de Veritas Volume Manager contiene los siguientes paquetes:

VRTSvxvm: Contiene los controladores del software VxVM, demonios y utilidades.

VRTSvlic: Contiene las utilidades de licenciamiento.

VRTSvmmdoc: Contiene copias en línea de documentación de VxVM.

VRTvmman: Contiene paginas del manual de VxVM.

VRTSob: Contiene el servicio VEA (Interfaz gráfica del usuario).

VRTSobgui: Contiene la interface grafica de usuario VEA.

VRTSvmpro: Contiene el proveedor del manejador virtual de discos

VRTSfspro: Contiene el proveedor de VERITAS File System.

Requerimientos de espacio para los paquetes de VxVM.

Es necesario confirmar si el sistema tiene el suficiente espacio libre en disco para la instalación. Los programas y archivos de VxVM se instalan en los file system /, /usr y /opt. La siguiente tabla muestra una aproximación del mínimo espacio requerido en cada file system para cada paquete.

Paquete	Contenido	Tamaño	File System
VRTSvxvm	Controlador y utilidades	59 MB	27 MB en / 32 MB en /usr
VRTSvlic	Utilidades de licenciamiento	2 MB	1 MB en /usr 1 MB en /opt
VRTSvmman	Manual	1 MB	/opt
VRTSvmmdoc	Documentación	30 MB	/opt
VRTSob VRTSobgui	GUI de Administrador	33 MB	/opt
VRTSvmpro	VxVM proveedor para VEA	5 MB	/opt
VRTSfspro	VxFS proveedor para VEA	8.5 MB	/opt

Tabla Espacio requerido para la instalación de VxVM.

Espacio total requerido:

27 MB en /

33 MB en /usr

78.5 MB en /opt

VERITAS File System consiste de los siguientes paquetes:

VRTSvxfs: Contiene el software de VERITAS File System y su manual.

VRTSfsdoc: Contiene la documentación en formato PDF de VERITAS File System.

Espacio requerido para los paquetes:

Los archivos y programas se instalan en los file systems /, /usr y /opt. Los espacios requeridos aproximados para cada directorio son:

Directorio	Tamaño	Contenido
/	1.5 MB	Binarios
/usr	2.25 MB	Librerías
/opt	3.5 MB	Comandos, Manuales

Tabla Espacio requerido para la instalación de VxFS

Licencias.

VxVM y VxFS son productos licenciados que requieren llaves de licencia válidas para uso. Es necesario contar con las licencias antes de la instalación ya que éstas son requeridas durante el proceso de instalación.

Se pueden agregar las licencias cuando se ejecuta el programa de instalación o utilizando el comando vxlicinst si el paquete VRTSvlic ya está instalado.

Para agregar una licencia nueva después de que el paquete VRTSvlic está instalado:

1. En la línea de comando se escribe vxlicinst.
2. Cuando se pida teclear el número de la licencia.
3. Después de que se introduce un número válido, el sistema verifica que la llave esté instalada. Se instala en /etc/vx/licenses/lic.

Para ver las licencias instaladas actualmente:

1. En la línea de comandos se teclea vxlicrep
2. La información acerca de las licencias instaladas es desplegada. Esta información incluye:
 - Número de la licencia.
 - Nombre del producto que habilita la licencia.
 - Tipo de licencia.
 - Características habilitadas por la llave.

Formas de agregar paquetes.

Existen dos métodos para agregar paquetes:

- a) Invocando la utilidad de instalador que se encuentra en el CD-ROM
- b) Manualmente agregando el software de los paquetes en la línea de comandos utilizando el comando pkgadd.

Se explicaran ámbra formas de instalación, sin embargo, en este proyecto se utilizo la forma manual.

a) Utilizando el instalador.

El instalador es una utilidad basada en menús que se puede utilizar para instalar cualquier producto contenido en el CD-ROM. Tiene mayor uso cuando se instalan varios productos VERITAS.

Si se necesita instalar un único paquete es preferible hacerlo desde la línea de comandos.

Para agregar paquetes usando el instalador:

1. Es necesario tener el superusuario.
2. Montar el CD-ROM VERITAS.

En los sistemas operativos Solaris el CD se monta automáticamente en /cdrom/CD_name

3. Ejecutar el instalador:

/cdrom/CD_name/intaller

4. La utilidad intenta localizar el paquete de licenciamiento, VRTSvlic. El paquete debe ser instalado para continuar con la instalación.

```
Looking for package VRTSvlic
```

```
Currently installed: 0
```

```
Minimum Version: 3.00.000
```

```
For Release Train installation to continue, VRTSvlic must be installed or upgraded.
```

```
Do you want to install it Now? [y,n]:
```

Se teclaea y para instalar VRTSvlic y continuar la instalación. Las utilidades de licenciamiento están instaladas en /etc/vx/licenses.

5. Después de que se instalan las utilidades de licenciamiento, Se despliega la pagina de status de productos VERITAS, muestra cuales productos están instalados así como su licenciamiento.
6. Para agregar la llave de licencia se teclaea 1 y Enter. Se requiere que se introduzca la llave y después se regresa a la página de status de producto.
7. En la página de status de producto se escribe 2 para instalar los productos que necesitamos después Enter.
8. Se despliega una lista de productos disponibles. Elegimos el número 3 que es el que requerimos.
9. Cuando la instalación se completa, regresamos al menú de instalación, se puede instalar productos adicionales o teclrear q para salir del menú.

b) Agregar paquetes manualmente con pkgadd.

Para instalar los paquetes manualmente utilizamos el comando `pkgadd`.

1. Firmarse como superusuario.
2. Montar el CD-ROM de VERITAS.
3. Agregar los paquetes VxVM usando el comando `pkgadd`:
`# pkgadd -d/cdrom/CD_name/product_name/pkgs packages`
4. Cuando se instalan los paquetes, el sistema despliega una serie de mensajes de status como progreso de instalación. Durante la instalación de VxFS, se reciben una serie de preguntas. Para continuar con la instalación escribimos `y`.
5. Cuando se termina la instalación se reinicia el sistema.

Modificaciones en `/etc/system`

El proceso de instalación de VxFS modifica el archivo `/etc/system` agregando las siguientes líneas.

```
Vxfs_START—do not remove the following lines:
VxFS requires a stack size greader than the default 8K.
The following values allow the kernel stack size
For all threads to be increased to 24K.
Set lwp_default_stksize=0x6000
Vxfs_END
```

El archivo original `/etc/system` se copia en `/etc/fs/vxfs/system.preinstall`.

Para verificar que los paquetes estén instalados se utiliza el comando `pkginfo` para ver la información acerca de la instalación.

Para listar los paquetes instalados en el sistema, se teclea el comando `pkginfo` en la línea de comandos.

```
#pkginfo
```

Para restringir la lista a los paquetes de VERITAS instalados tecleamos:

```
#pkginfo | grep VRTS
```

Primera configuración de Veritas Volume Manager

Como lo comentamos anteriormente antes de instalar y configurar Volume Manager es necesario conocer el contenido de los discos y decidir cómo es que se quiere que estos discos sean usados por Volume Manager. Durante la instalación se especificará como serán manejados los discos.

Cuando se corre el programa de instalación, se responden las siguientes preguntas:

- ¿Cuáles discos estarán bajo el control de Volume Manager?
- ¿Se quieren usar nombres basados en la residencia?
- ¿Se quiere excluir discos del control de Volume Manager?

- ¿Se quiere suprimir el Multipath para algún disco?
- Cuando se ponen discos bajo el control de Volume Manager,

¿Se quieren preservar o eliminar datos en el file system existente y las particiones?

- ¿Se quiere poner el disco de inicio del sistema bajo el control de Volume Manager?
- ¿Se quiere configurar cada disco en un arreglo diferente, o se quiere configurar los discos en un arreglo de la misma forma?

En la etapa de instalación, el único disk group que se crea es rootdg. Los demás se crean después de la instalación.

Consideraciones para responder las preguntas anteriores.

1. ¿Cuáles discos estarán bajo el control de Volume Manager?

El propósito de ejecutar el programa de instalación de VxVM es crear el grupo de discos raíz (root disk group rootdg), requerido para que el demonio (vxconfigd) pueda iniciar en modo habilitado. Antes de iniciar el programa de instalación se debe decidir cuales discos poner en el rootdg y cuales discos serán usados para otras funciones.

El rootdg se usa solo para el file system raíz y sus espejos.

- Si no se planea que el disco que contiene el sistema operativo esté bajo el control de VxVM, se debe poner otro disco en rootdg, de lo contrario el único disco que se deberá poner bajo el control de VxVM durante la instalación son el disco de inicio y los discos que se planeen de espejo para el mismo.
- Se pueden agregar otros discos y otros grupos de discos después de la instalación utilizando las utilidades.

Recordando del capítulo anterior:

Cualquier disco que se desee poner bajo el control de Volumen Manager deberá tener dos particiones libres (una para la región privada y otra para la región pública) y una pequeña cantidad de espacio en el principio o al final que no pertenezca a alguna partición. Este espacio es utilizado para guardar información de la configuración de disk group y Volume Manger. Este espacio asegura que Volume Manager puede identificar el disco aún si éste es movido a una dirección diferente o diferente controlador, y también ayuda a asegurar las recuperación en caso de falla en el disco.

La región privada es de 20487 sectores (1024 K) por default y es redondeado al límite más cercano del cilindro, esto es, un número entero de cilindros son siempre asignados a la región privada.

¿Se usarán nombres basados en enclosure?

Como una alternativa para los nombres estándar de discos (ejemplo c0t0d0), VxVM ofrece los nombres basados en enclosure. Al decir enclosure nos referimos a los nombres que controla el arreglo de discos.

Con VxVM los dispositivos pueden ser nombrados por su enclosure mas que por los controladores a través de los cuales son accedados.

En una SAN que utiliza hubs de canales de fibra o switches, la información acerca de la localización de los discos proporcionada por el sistema operativo puede no indicar la ubicación correcta de los discos.

Nombres basados en enclosure permite a VxVM acceder enclosures como entidades diferentes. Configurando copias redundantes de los datos en diferentes residencias podemos asegurar contra fallas en una o mas residencias.

Este tipo de nombres son de gran uso cuando se maneja la característica de dynamic multipathing (DMP) de VxVM.

Durante la instalación se requiere especificar el tipo de nombres que se van a utilizar, esto también se puede hacer después con las utilidades.

¿Se quiere excluir discos del control de Volume Manager?

Si hay discos que se desea excluir del control de Volume Manager se pueden especificar en el archivo excluido.

- Para excluir un disco específico, se crea el archivo `/etc/vx/disks.exclude` y se agregan los discos al archivo. El programa de instalación ignora cualquier disco que se encuentre dentro de éste archivo.
- Para excluir todos los discos que se encuentren en una controladora entera, se agrega el archivo `/etc/vx/cntrls.exclude` y se agrega el nombre de la controladora en el archivo. El programa de instalación ignora cualquier controladora que éste dentro de este archivo.
- Para excluir todos los archivos en una residencia específica se crea el archivo `/etc/vx/enclr.exclude` y se agrega el nombre de la residencia al archivo.

En este caso no se agregan discos al control de VXVM ya que contienen datos que queremos conservar.

¿Se quiere suprimir el multipath para algún disco?

Durante la instalación se puede suprimir el multipath (un disco físico se puede acceder con diferentes direcciones) para determinados dispositivos o todos si así se desea sin remover la capa DMP ya que es necesaria para el funcionamiento de VxVM.

Cuando se ponen discos bajo el control de Volume Manager, ¿Se quieren preservar o eliminar datos en el file system existente y las particiones?

Se llama encapsulación al proceso de preservar datos existentes en los discos o inicialización si se desea eliminar dichos datos.

- **Encapsulación.** Es guardar los datos en un disco que se va a poner bajo el control de Volume Manager. Los discos que serán encapsulados deben:
 - Contener el espacio libre mínimo requerido de 1024 sectores (512 K) (Por default, VxVM usa 2048 sectores (1024 K))
 - Contener una partición `s2` que represente al disco entero (La partición 2 no puede contener un file system)

Las particiones son convertidas en subdiscos que son usadas para crear volúmenes que reemplazan las particiones de solaris.

- **Inicialización.** Se le nombra a la eliminación de los datos existentes en los discos que se pondrán bajo el control de Volume Manager.

Cualquier disco que es encapsulado o inicializado durante la instalación es puesto en el disk group `rootdg`.

En este caso no se encapsulara ni se inicializará ningún disco ya que se requiere conservar los datos y se opto por crear file systems nuevos en diferentes discos.

¿Se quiere poner el disco de inicio del sistema bajo el control de Volume Manager?

Si se planea que el disco de arranque tenga un espejo, entonces se debe poner el disco de arranque bajo el control de Volume Manager, el cual requerirá de la encapsulación para mantener los datos originales en el disco.

Si no se planea espejear el disco de inicio o en caso de que el disco de sistema operativo este en RAID 1 con otro manejador de volúmenes, entonces no se debe poner el disco de inicio bajo el control de Volume Manager. La encapsulación y el espejeo del disco de arranque son recomendados para un ambiente de alta disponibilidad.

Cuando se encapsula un disco de arranque, se guarda la información existente y la información de arranque, y las particiones se convierten en volúmenes.

- Las particiones existentes `/usr` y `/var` son convertidas en volúmenes sin remover las particiones.
- Otras particiones son convertidas en volúmenes y las particiones son removidas.
- El área de swap existente es convertida en un volumen. Si hay suficiente espacio para la región privada en el disco de arranque, Volume Manager toma sectores del área de swap del disco, esto forma la región privada sobre poniéndose la región publica. La partición de swap mantiene el mismo tamaño, y el volumen swap es reasignado para ser mas pequeño que la partición swap.

Volume Manager preserva una copia original de la VTOC de cualquier disco que es encapsulado en `etc/vx/reconfig.d/disks.d/cxytdz/vtoc`, donde `cxytdz` es la dirección SCSI del disco.

¿Se quiere configurar cada disco en un arreglo diferente, o se quiere configurar los discos en un arreglo de la misma forma?

Cuando se corre el programa de instalación se puede elegir encapsular o inicializar los discos individualmente o todos los discos en cada

El programa vxinstall.

Después de que los paquetes están agregados, se puede configurar Volume Manager utilizando el programa de instalación interactivo llamado `vxinstall`.

El único propósito de ejecutar el programa `vxinstall` (se debe estar firmado como superusuario) es el de crear el `disk group rootdg`, Volume Manager requiere que existe este `disk group` y que contenga al menos un disco.

Se debe ejecutar `vxinstall` solo uno vez por sistema, excepto en situaciones de problemas (cuando no exista un `rootdg`).

El proceso vxinstall.

Vxintall es un programa interactivo que nos guía a través del proceso de instalación. Durante la instalación se presentan varias preguntas acerca de cómo se quiere que Volume Manager maneje los discos.

Los principales pasos en éste proceso son:

1. Introducir el comando vxinstall para comenzar el proceso.
2. Proporcionar una licencia válida cuando se requiera.
3. Seleccionar el formato de nombres para los dispositivos de host.
4. Especificar si se quiere prevenir algún dispositivo del multipath dinámico (DMP).
5. Seleccionar un método de instalación. Se puede seleccionar modo rápido o personalizado.
6. Especificar si se va a encapsular el disco de inicio.
7. Especificar la forma en la que Volume Manager manejará a los demás discos identificados en el sistema.
8. Se verifica las selecciones para los discos que se pondrán bajo el control de Volume Manager.
9. Se apaga y se reinicia el sistema.

Paso 1: Iniciar el programa vxinstall.

Se tecldea el comando vxinstall en la línea de comandos.

```
#vxinstall
```

Paso 2: Licenciar

El programa vxinstall ejecuta el comando vxlicense el cual despliega información de licenciamiento y solicita una llave válida para continuar con el proceso.

```
Some licenses are already installed. Do you wish to review them [y,n,q]
(default: y) y
```

```
Do you wish to enter another licence key [y,n,q] (default n)n
```

Paso 3: Seleccionar el método de nombres.

El programa examina y lista todos las controladoras conectadas en el sistema, y solicita se especifique cuales serán nombradas basadas en su residencia. Si se elije éste método se recibe el siguiente mensaje.

```
Generating list of attached enclosures....
```

```
VxVM will use the following format to name disks on the host:
```

```
<enclosurename>_<diskno>
```

In the above format, <enclosurename> is the logical name of the enclosure to which the disk belongs. VxVM assigns default enclosure names which can be changed according to the user requirements.

The Volume Manager has detected the following categories of storage connected to your system:

```
Enclosures: enc01 sena0 sena1 sena3 sena4 sena5
```

```
Others: others0
```

Paso 4: Suprimir el Multipath.

Después de una breve introducción en el proceso de instalación, se despliega un menú de opciones. Si se requiere prevenir algunos dispositivos del multipath es necesario especificarlo antes de continuar con la instalación.

1. Quick Installation
2. Custom Installation
3. Prevent multipathing/Suppress devices from VxVM's view
- ? Displays help about menu
- ?? Displays help about menuing system
- q Exit from menus

Cuando se selecciona la opción 3 del menú se despliega un menú más detallado que permite:

- Suprimir todas las rutas de una controladora, rutas específicas, discos específicos o todo, que esté a la vista de VxVM.
- Prevenir el multipath de todos los discos en una controladora o discos específicos.
- Listar los dispositivos actuales que no tienen multipath.

Si se elige hacer una exclusión en el multipath es necesario hacer un reinicio, después se continúa con la instalación.

Paso 5: Seleccionar un método de instalación.

Para continuar con la instalación se debe elegir un método se recomienda el método personalizado.

1. Quick Installation
2. Custom Installation
3. Prevent multipathing/Suppress devices from VxVM's view
- ? Displays help about menu
- ?? Displays help about menuing system
- q Exit from menus

Quick Installation. Esta opción permite inicializar todos los discos o encapsular todos los discos. No se recomienda

Custom Installation. Esta opción permite manejar cada disco y controlar cuáles discos quedarán bajo el control de VxVM. Se recomienda.

Paso 6: Encapsulamiento del disco de inicio o arranque.

El programa de instalación detecta el disco de arranque y proporciona información.

Se solicita la confirmación para encapsular el disco de inicio.

The c0t0d0 disk is your Boot Disk. You can not add it as a new disk. If you encapsulate it, you will make your root file system and other system areas on the Boot Disk into volumes. This is required if you wish to mirror your root file or system swap area. Encapsulate Boot Disk [y,n,q,?] (default:n)

Paso 7: Configuración de otros discos.

Después de que se especifica cómo se va a tratar al disco de arranque

hay que configurar los discos de datos, el programa identifica todos y cada uno de los discos en los arreglos y pide se especifique como se desea que sean tratados. Si los discos están listados como excluidos éstos se listan en forma separada.

Un menú muestra cuatro opciones para cada arreglo de discos:

Installation options for enclosure enc0

Menu: Volume Manger/Intall/Custom/enc0

1. Install all disks as pre-existing disks. (encapsulate)
 2. Install all disks as new disks. (discards data on disks!)
 3. Install one disk at a time.
 4. Leave these disks alone.
- ? Display help about menu
?? Displays help about the menuing system
q Exit from menus

Select an operation to perform:

Paso 8: Verificar las opciones configuradas.

Cuando se ha terminado con la instalación el programa despliega un resumen de los discos que se han inicializado o encapsulado por ejemplo:

The following is a summary of your choices.

c0t0d0 Encapsulate

c2t2d3 New Disk

Is this correct [y,n,q,?] (default: y)

Esta es la última oportunidad de hacer cambios si se confirma la información entonces los discos empiezan a ser inicializados o encapsulados según el caso.

Paso 9: Apagado y reinicio del sistema.

El ultimo paso de la instalación del producto es el reinicio del sistema.

Conversión de File System UFS a volúmenes con VXFS

Una vez instalado el producto Storage Foundation, la estrategia acordada con el usuario fue agregar nuevos discos al servidor. Posteriormente estos discos nuevos fueron inicializados bajo el control de Veritas Volume Manager:

```
# vxdisksetup -i <nombre del disco>
```

Se creo un disk group donde se crearan los volúmenes donde se migrara la aplicación people soft.

```
#vxdg init <nombre del dg> <nombre del disco=target>
```

```
#vxdg -g <nombre del dg> adddisk <nombre del disco=target>
```

Después se procede a la creación de los volúmenes

```
#vxassist -g <nombre del dg> make <nombre del volumen> <tamaño>
```

Una vez creados todos los volúmenes se crean los file system con Veritas File System

```
#mkfs -F vxfs /dev/vx/rdisk/<nombre del dg>/<nombre del volumen>
```

Se crean los puntos de montaje temporales con el comando mount

```
#mount <punto de montaje>
```

Y Finalmente se montan los nuevos file system

```
#mount -F vxfs /dev/vx/dsk/<nombre del dg>/<nombre del volumen>  
/<punto de montaje>
```

Una vez teniendo todos los file system nuevos , se solicita una ventana al usuario por 4 horas para copiar aproximadamente 900GB de bases de datos y aplicaciones hacia los nuevos file systems teniendo las bases de datos y aplicaciones abajo.

Se utilizo el siguiente script:

```
#cd /<file system origen ufs>  
#tar cvpf - .* |(cd /<file system destino>;tar xvpf - )
```

Esto se hizo por cada file system ufs con la intension de migrar la aplicación al producto Veritas Storage Foundation

Una vez terminada la migración de datos , se modifica el archivo /etc/vfstab donde se indica la tabla de file systems, haciendo referencia a los nuevos volúmenes. Se reinicia el servidor y con ello la información quedo migrada.

6.6 Resultados

Recordando la problemática inicial donde el cliente indicaba

“Los file systems están a punto de llegar a su capacidad total y ya no contamos con ninguna información que pueda ser depurada para liberar espacio”

Además del diagnostico inicial donde comentamos lo siguiente:

- 1) No se cuenta con ningún manejador de volúmenes.
- 2) Se tiene que distribuir los discos en distintas HBA para mejorar el tiempo de respuesta en el acceso a disco.
- 3) No se cuenta con canales redundantes en el acceso a disco.
- 4) Únicamente se cuenta con RAID por hardware RAID5.
- 5) Todos los File system son de tipo UFS
- 6) Se tienen cuellos de botella en el acceso a disco.

Y después de haber concluido la propuesta contamos con lo siguiente:

- El cliente cuenta con 25% de ahorro en espacio en disco ya que con el manejador de volúmenes ahora el espacio se administra de forma mas eficiente y no se tiene desperdicio alguno ya que ya no hacemos referencia a particiones físicas de los discos.
- El cliente quedo satisfecho al demostrarle que cuando un file system este por alcanzar su máxima capacidad se pueda crecer de forma dinámica con el comando `#vxresize -g <nombre del dg> <volumen> +<tamaño extra>` resolviendo con ello su principal problemática.
- El cliente cuenta con una infraestructura mas robusta ya que con la instalación de Veritas Storage Foundation se agrego la característica de Dynamic Multipathing teniendo con ello balanceo de carga en el acceso a disco quitando así el problema del cuello de botella en I/O. A su vez con esta característica habilitada se cuenta con canales redundantes en el acceso a disco previniendo cualquier falla de HW en la comunicación a los discos.

Por lo que se puede mencionar que la implementación fue un éxito.

Conclusiones

Del presente trabajo se desprende la conclusión de que tanto el almacenamiento como la administración de los datos que forman parte de información crítica para una empresa no es una labor que puede hacerse de forma tradicional, es necesario contar con conocimiento de las tecnologías de almacenamiento para poder tomar una decisión como administrador y persona responsable de dicha información y de esta forma implementar lo necesario para asegurar la integridad de los datos.

Con el desarrollo del trabajo pudimos darnos cuenta que el uso de la herramienta ofrecida por Symantec es una de las mejores soluciones para cuando se necesita manejar datos en gran cantidad y mantenerlos disponibles la mayor parte del tiempo, así como administrar de manera óptima los dispositivos. Con ello ofrecer también a la compañía una ventaja económica ya que el rendimiento de los discos es mayor.

Para la empresa del sector eléctrico descrita en el último capítulo ésta fue una solución vital ya que no importa el incremento en la cantidad de información se puede tener la seguridad de que los datos estarán íntegros

Glosario

ACTUADOR DE UN DISCO

Es un motor que mueve la estructura que contiene las cabezas de lectura entre el centro y el borde externo de los discos.

AIX

Advanced Interactive Executive. Sistema Operativo tipo UNIX propietario de IBM.

AMD64

Familia de procesadores diseñados por AMD compuesta por el procesador AMD Opteron, el procesador AMD Athlon 64 y la Tecnología Mobile AMD Turion 64. Esta familia ha sido diseñada para permitir la informática simultánea de 32 y 64 bits, sin pérdida de rendimiento.

ANSI

American National Standards Institute. Instituto Nacional Americano de Estándares. Organización no lucrativa que permite la estandarización de productos, servicios, procesos, sistemas y personal en Estados Unidos.

APUNTADORES

O punteros son variables que hacen referencia a una región de memoria.

ATA

Es una interfaz utilizada para conectar dispositivos como discos duros, unidades de CD-ROM y otros tipos de unidades para computadoras.

ATAPI

Advanced Technology Attachment Packet Interface. Interfaz de dispositivo conector de tecnología avanzada. Estándar que designa aquellos dispositivos que pueden conectarse a controladoras ATA. Por ejemplo las unidades de CD-ROM.

BACKUP

Es la copia total o parcial de información importante del disco duro, CDs, bases de datos u otro medio de almacenamiento. Esta copia de respaldo debe ser guardada en algún otro sistema de almacenamiento masivo.

BIOS

Basic Input/Output System. Sistema Básico de Entrada/Salida de datos. Programa que reside en la memoria EPROM. La BIOS es una parte esencial del hardware que es totalmente configurable y es donde se controlan los procesos de flujo de información en el bus de la computadora, entre el sistema operativo y los periféricos.

BITS

Binary Digit. Dígito Binario. Es la unidad digital más pequeña que puede manejar una computadora. Se maneja a través del sistema binario, es decir, puede tener dos estados 1 ó 0.

BLOQUES

O Clusters. Son unidades de almacenamiento en un disco con una determinada cantidad fija de bytes, Un disco está dividido en miles de bloques de igual tamaño y los archivos son repartidos y almacenados en distintos bloques. El tamaño se determina en el formateo de disco.

BOOT

Proceso inicial de una computadora en donde se carga la configuración (BIOS), los dispositivos de hardware y se busca el sistema operativo en la secuencia de boot.

BOOTSTRAP

Proceso donde un sistema simple active otro sistema más complejo para servir al mismo propósito. Hablando de sistema operativo Solaris es el programa de arranque para levantar los sistemas de archivos o file system) se aloja o reside en los primeros 15 sectores del disco (sectores 1-15) que siguen después de la VTOC.

BSD

Berkeley Software Distribution

BUFFER

Memoria intermedia. Memoria de almacenamiento temporal de información. Suele tratarse de una memoria intermedia entre un dispositivo y otro, por ejemplo, la computadora y la impresora, o la computadora y el disco duro.

BUS

Toda conexión entre los distintos dispositivos de una computadora por la cual se transmite información. Las velocidades de los buses influyen mucho en el rendimiento general de la máquina.

BYTE

Unidad de información formada por 8 bits. Según estén combinados los bits (ceros o unos), formarán un byte y, por tanto, un carácter cualquiera.

CABEZA DE UN DISCO

Son un conjunto de brazos alineados verticalmente que se mueven hacia dentro o fuera según convenga, todos a la vez. En la punta de dichos brazos están las cabezas de lectura/escritura, que gracias al movimiento pueden leer tanto zonas interiores como exteriores del disco.

CACHE

Conjunto de datos duplicados de otros originales. La duplicación se basa en que los datos originales son más costosos de acceder en tiempo con respecto a la copia en memoria caché.

CD

Compact disk o disco compacto. Disco óptico circular para el almacenamiento de información de forma binaria. Generalmente 12 cm. De diámetro y que pesa unos pocos gramos. La información se almacena de forma digital, o sea, unos y ceros. Almacena hasta 640 MB, aunque puede extenderse esa capacidad un poco más.

CD-ROM

Este tipo de CDs son de sólo lectura, por lo tanto es posible grabarles información si no son vírgenes.

CD-WR

Compact Disc Rewritable. Disco Compacto que permite múltiples escrituras.

CHS

Cylinder-Head-Sector. Cilindro-Cabeza-Sector es un modo de acceso al disco duro para discos menores de 8 GB.

CILINDROS DE UN DISCO

Grupo de pistas de igual diámetro en los discos internos de los discos duros. Se llama cilindro por la forma que tomaría si se conectara en el espacio a las pistas.

COMANDO

Command,orden,mandato. Es una orden que se le da a un programa de computadora que actúa como intérprete del mismo, para así realizar una tarea específica.

DAS

Direct Attached Storage. El método tradicional de almacenamiento y el más sencillo. Consiste en conectar el dispositivo de almacenamiento directamente al servidor o estación de trabajo, es decir, físicamente conectado al dispositivo que hace uso de él.

DDS

Digital Data Storage. Formato para el almacenamiento y el respaldo de datos de una computadora en una cinta magnética que evolucionó de la tecnología DAT (Digital Audio Tape)

DISCO DURO

Grupo de dispositivos de almacenamiento que permiten guardar grandes cantidades de datos. Poseen discos circulares magnéticos de uno o dos caras que rotan rápidamente cuando está encendido.

DISQUETES

Discos Flexibles. Cartucho plástico para almacenar información. Se tratan de una clase de discos magnéticos. Las dos versiones más conocidas para PC son: la más antigua 5 ¼ pulgadas y la de 3 ½, prácticamente sin uso en la actualidad.

DLT

Digital Linear Tape. Tecnología de almacenamiento de datos por cintas magnéticas. Es utilizado específicamente para copias de seguridad.

DRIVE

Es una palabra inglesa que se traduce como “unidad”. Un drive es un dispositivo que lee y/o escribe datos en un medio de almacenamiento.

DVD

Digital Versatile/Video Disc- Disco Versátil Digital. Formato de almacenamiento digital de datos. Tienen el mismo tamaño físico que un CD, aunque almacenan mucha más información. Los DVD guardan los datos utilizando un sistema de archivos denominado UDF, el cual es una extensión del estándar ISO 9660, usado para CD de datos.

ECC

(Error Checking and Correction) Chequeo y Corrección de errores

EEPROM

(Electrically Erasable Programmable Read Only Memory) Memoria de sólo lectura eléctricamente programable y borrrable. Tipo de tecnología de memoria moderna.

EIDE

Es una extensión del originalmente IDE, es la denominación que recibe la interfaz más empleada actualmente en los PC domésticos y cada vez más en aquellos ordenadores de altas prestaciones para la conexión de discos duros.

EJE DE UN DISCO

Es el soporte sobre el cual están montados y giran los platos del disco.

ENCLOSURE

Caja de Discos

ERP

Son sistemas de información gerenciales que integran y manejan muchos de los negocios asociados con las operaciones de producción y de los aspectos de distribución de una compañía comprometida en la producción de bienes o servicios.

ESDI

Enhanced Small Disk Interface. Es un estándar del equipamiento (hardware) desarrollado por un consorcio de 22 fabricantes de disqueteras. La ESDI brinda la posibilidad de contar con una velocidad máxima de transferencia de datos, desde y hacia el disco duro, de 10 megabits por segundo, y puede manejar (administrar) accionadores de disco de hasta 1 gigabyte en su capacidad de almacenamiento. Una tarjeta controladora de disco duro, ESDI (es decir, el circuito que controla las operaciones de un disco duro) solo puede ser usada con un disco duro que posee una interfaz ESDI.

FCAL

Es una tecnología de red utilizada principalmente para redes de almacenamiento, disponible primero a la velocidad de 1 Gb/s, y posteriormente a 2, 4 y 8 Gb/s.

FIBER CHANNEL SCSI

Se refiere a productos con capas físicas y protocolares Fiber Channel usando el set de comandos SCSI.

FIBRA DE VIDRIO

Es un material fibroso obtenido al hacer fluir vidrio fundido a través de una pieza de agujeros muy finos y al solidificarse tiene suficiente flexibilidad para ser usado como fibra. Sus principales propiedades son: buen aislamiento térmico, inerte ante ácidos, soporta altas temperaturas. Estas propiedades y el bajo precio de sus materias primas, le han dado popularidad en muchas aplicaciones industriales.

La fibra de vidrio, también es usada para realizar los cables de fibra óptica utilizados en el mundo de las telecomunicaciones para transmitir señales lumínicas, producidas por láser o LEDs.

FILE SYSTEM

Sistema de archivos. es un método para el almacenamiento y organización de archivos de computadora y los datos que estos contienen, para hacer más fácil la tarea encontrarlos y accederlos.

FLOPPY

Ver Disquete.

GB

Gigabyte. Unidad de almacenamiento. Existen dos visiones distintas de gigabyte (GB) dependiendo de la exactitud que se desee. Un gigabyte, en sentido amplio, son 1.000.000.000 bytes (mil millones de bytes), ó también, cambiando la unidad, 1.000 megas (MB o megabytes). Pero para más exactitud, 1 GB son 1.073.741.824 bytes ó 1.024 MB.

GRUPO DE DISCOS

Es una colección de discos. Se agrupan discos con propósitos de administración,

HBA

Controla la comunicación entre el sistema y la unidad de disco.

HD

Hard Disk. Ver Disco Duro

HOST

Máquina conectada a una red. Tiene un nombre que la identifica, el hostname. La máquina puede ser una computadora, un dispositivo de almacenamiento por red, una impresora, etc.

HP-UX

Es la versión de Unix desarrollada y mantenida por Hewlett-Packard desde 1983

HSFS

El High Sierra es un sistema de archivos con un propósito especial desarrollado para el uso del CD-ROM.

HVD

Differential o High Voltage Differential (HVD) Diferencial es un sistema de señales lógicas usado en algunos dispositivos SCSI.

IDE

El puerto IDE (Integrated device Electronics) o ATA (Advanced Technology Attachment) controla los dispositivos de almacenamiento masivo de datos como los discos duros.

INODOS

Es una estructura de datos propia de los sistemas de archivos tradicionalmente empleados en los sistemas operativos tipo UNIX como es el caso de Solaris. Un inodo contiene las características (permisos, fechas, ubicación, pero NO el nombre) de un archivo regular, directorio, o cualquier otro objeto que pueda contener el sistema de ficheros.

JAZ

Antiguo dispositivo de almacenamiento que utiliza cartuchos que internamente son muy parecidos a un disco duro (no poseen el cabezal lectograbador, pues se encuentra en la unidad Jaz). Permite un máximo de 1 o 2 GB de almacenamiento en sus cartuchos y una velocidad de transferencia de 5 MB/seg, con un tiempo de acceso de 15 ms.

KB

Unidad de medida de cantidad de datos. Con esta unidad se pueden medir velocidades de transferencia (agregándose una unidad de tiempo), capacidad de almacenamiento, tamaño de archivos, etc. Un byte corresponde a un caracter, 1 kb (kilobyte) corresponde a 1024

bytes.

De forma genérica y para redondear, 1 kb corresponde a 1000 bytes o sea, mil bytes.

KERNEL

Núcleo. Parte esencial de un sistema operativo que provee los servicios más básicos del sistema. Se encarga de gestionar los recursos como el acceso seguro al hardware de la computadora. Se encarga también del multiplexado, determinando qué programa accederá a un determinado hardware si dos o más quieren usarlo al mismo tiempo.

LATENCIA

Es el tiempo o lapso necesario para que un paquete de información se transfiera de un lugar a otro.

LBA

Logical Block Addressing Es el sistema dominante para direccionamiento de discos grandes, puesto que desde que alcanzaron el límite de 8.455 GB, se hizo imposible expresar su geometría en términos de Cilindro, Cabeza y Sector. Ver CHS.

LIGA

Es un elemento que hace referencia a otro recurso.

LINUX

Sistema operativo que posee un núcleo del mismo nombre. El código fuente es abierto, por lo tanto, está disponible para que cualquier persona pueda estudiarlo, usarlo, modificarlo y redistribuirlo. El término Linux se utiliza para describir al sistema operativo tipo Unix

LTO

(Linear Tape-Open o LTO). Tecnología de almacenamiento de datos en cinta magnética desarrollado como una alternativa abierta a las DLT. LTO fue desarrollada por Seagate, Hewlett-Packard e IBM. El factor forma de la tecnología LTO es llamado Ultrium.

LUN

En almacenamiento, una logical unit number o LUN es una dirección para una unidad de disco duro y por extensión, el disco en sí mismo. El término es originario del protocolo SCSI como una forma de diferenciar unidades de disco individuales dentro de un bus SCSI tal que un array de discos.

LVD

Low Voltage Differential- Es un esquema lógico diferencial pero usando niveles de Voltaje más bajos que HVD.

MB

Unidad que sirve para medir cantidad datos informáticos. Sirve para medir tamaño de archivos, capacidad de almacenamiento, velocidad de transferencia de datos (al agregarle una unidad de tiempo, generalmente segundos).Un megabyte equivale exactamente a 1024 KB (kilobytes) o a 1.048.576 bytes. 1024 (MB) megabytes equivalen a 1 GB. Para redondear se suele decir que un megabyte equivale a un millón de bytes.

MFM

Modified Frequency Modulation Es una línea de sistema de codificación utilizado para codificar la información en la mayoría de los formatos de disquetes.

MOTHERBOARD

(placa madre, placa base, tarjeta madre, mainboard, system board, logic board). Tarjeta central de circuitos integrados para la interconexión entre el microprocesador, las ranuras o slots para conectar tarjeta de expansión, memorias RAM, la ROM, dispositivos de almacenamiento y cableados.

MP3

Formato de audio que combina gran calidad de sonido y poco tamaño.Desarrollado en Alemania por Brandenburg, Popp y Grill, tres científicos del instituto tecnológico de Fraunhofer en Ilemenu en el año 1986. Luego en 1992 la Moving Picture Experts Group (MPEG) aprobó oficialmente la tecnología.El formato MP3 redujo el tamaño de los archivos de música conocidos hasta diez veces casi sin perder calidad por la compresión. Su nombre técnico es ISO MPEG Audio Layer 3.

MTBF

Acrónimo de mean time between failures, es la media aritmética (promedio) del tiempo entre fallos de un sistema. El MTBF es típicamente parte de un modelo que asume que el sistema fallido se repara inmediatamente (el tiempo transcurrido es cero), como parte de un proceso de renovación. En cambio, el MTTF (mean time to failure) mide el tiempo medio entre fallo con la suposición de un modelo en que el sistema fallido no se repara.

MULTIMEDIA

Cualquier sistema que utiliza múltiples medios de comunicación al mismo tiempo para presentar información. Generalmente combinan textos, imágenes, sonidos, videos y animaciones.

NAS

Network Attached Storage) es el nombre dado a una tecnología de almacenamiento dedicada a compartir la capacidad de almacenamiento de un computador (Servidor) con ordenadores personales o servidores clientes a través de una red (normalmente TCP/IP), haciendo uso de un Sistema Operativo optimizado

NFS

Network File System. Es un servicio que permite que una o varias máquinas en red (clientes) compartan un directorio o varios que se encuentran en una máquina principal llamada servidor. De esta manera, cualquiera de los clientes puede ingresar a directorios con archivos como si se tratase de su propia máquina.

OFFSET

Dentro de una estructura de datos es un entero que indica la distancia (desplazamiento) desde el inicio del objeto hasta un punto o elemento dado, presumiblemente dentro del mismo objeto. El concepto de distancia es solamente válido si todos los elementos del objeto son del mismo tamaño (típicamente dados en bytes o palabras).

PARTICION

Creación de divisiones lógicas (volúmenes) en un disco duro para aplicarles un formato lógico (sistema de archivos) del sistema operativo específico que se instalará.

PATA

(Parallel ATA). PATA es un estándar de interfaz para conexiones entre dispositivos de almacenamiento en computadoras, como discos duros, unidades de CD-ROM y unidades de estado sólido. El ATA paralelo o PATA es el ATA original, que comenzó a llamarse así en 2003 luego de que fuera lanzado el Serial ATA. El estándar PATA permite cables con una longitud de hasta 46 centímetros (18 pulgadas), por esta limitación esta tecnología sólo se utiliza para conectar dispositivos de almacenamiento dentro del gabinete de algunas las computadoras.

PCFS

Sistema de archivos, implementación de UNIX para el sistema operativo DOS el llamado FAT32.

PDA

(Personal Digital Assistant - Asistente Personal Digital). Son un tipo de handheld. Pequeñas computadoras que entran en la mano que tienen un sistema y programas con tecnologías de IA y que ayudan sus usuarios en ciertas actividades como la búsqueda de información, agenda electrónica, etc. Poseen reconocimiento de escritura.

PISTAS DE UN DISCO

Las pistas son círculos concéntricos de datos que se almacenan en los discos que revolucionan dentro de los discos magnéticos. En los discos duros hay varios discos y por lo tanto, varios círculos del mismo diámetro. Todas las pistas que tienen el mismo diámetro, constituyen un cilindro.

PLEXES

Es una colección ordenada o estructurada de subdiscos, ésta colección representa una copia de los datos en el volumen. Un plex consiste en uno o más subdiscos localizados en uno o más discos físicos.

Un plex es también llamado espejo. Los términos plex y espejo pueden ser utilizados intercambiamente, aunque un plex es solo una copia de los datos. Los términos “espejeado” o “espejo” se aplican en dos o más copias de los datos.

RAID

Originalmente proviene de Redundant Array of Inexpensive Disks - conjunto redundante de discos baratos. Actualmente de Redundant Array of Independent Disks - conjunto redundante de discos independientes. Sistema de almacenamiento de información que utiliza múltiples discos duros en donde se distribuyen y/o duplican datos. (Existen implementaciones RAID en un solo disco duro, pero las ventajas no son importantes y técnicamente no es RAID).

RAM

Random Access Memory - Memoria de acceso aleatorio. Tipo de memoria donde la computadora guarda información para que pueda ser procesada más rápidamente. En la memoria RAM se almacena toda información que está siendo usada en el momento. Su capacidad de almacenamiento se mide en megabytes y más recientemente en gigabytes. La información que contienen es renovada continuamente y cuando la computadora se reinicia o se apaga, toda la información contenida se pierde, por eso es llamada memoria volátil.

RLL

(Run Length Limited codes). Esquema de codificación empleado en el almacenamiento de datos en los discos duros, Hi-MD MiniDiscs y otros sistemas magnéticos. Sirve para codificar datos en pulsos magnéticos. RLL permite tiempos de accesos más veloces y aumenta la capacidad de almacenamiento de una unidad con respecto al sistema anterior, el MFM (Modified Frequency Modulation). Se calcula que RLL almacena un 50% más de datos por disco que el MFM, pero necesita más tiempo de procesamiento. También la circuitería lógica es más complicada que la MFM.

SAN

Definición de Red de área de almacenamiento (storage area network o SAN). Es una arquitectura para adjuntar dispositivos de almacenamiento de computadoras remotas como un conjunto de discos, librerías de tapes y conjunto de CDs, como si fuesen dispositivos locales.

SCSI

(Small Computer System Interface). Interfaz estándar para transferencia de datos entre periféricos en el bus de la computadora. Tanto la placa madre como el dispositivo deben soportar y disponer de un controlador SCSI. Es utilizado especialmente en dispositivos como los discos duros CD/DVD, etc., aunque, en principio, cualquier dispositivo podría soportar el SCSI, como impresoras o escáneres. El SCSI se utiliza en computadoras que necesitan de alto rendimiento. La compatibilidad para dispositivos SCSI puede agregarse a través de una tarjeta de expansión a la placa madre. Dos interfaces SCSI pueden ser incompatibles pues existen muchas variantes (incluso aunque esté estandarizado en ANSI).

SERVIDOR

En redes, computadora central en un sistema de red que provee servicios a otras computadoras.

SLED

Single Large Expensive Disk. Discos independientes de alto costo.

SOLARIS

Sistema operativo de la familia de Unix desarrollado por Sun Microsystems. Fue desarrollado originalmente para funcionar en computadoras SPARC de Sun, actualmente se ejecuta en otros sistemas.

SPARC

(Scalable Processor Architecture). Es una arquitectura RISC originalmente diseñada por Sun Microsystems en 1985. SPARC es una marca registrada de SPARC International, Inc., organización establecida en 1989 para promover la arquitectura SPARC.

SSD

Una unidad de estado sólido o SSD (solid state drive) es un dispositivo de almacenamiento de datos que usa memoria no volátil tales como flash, o memoria volátil como la SDRAM, para almacenar datos, en lugar de los platos giratorios encontrados en los discos duros convencionales. Aunque técnicamente no son discos a veces se traduce erróneamente en español la 'D' de SSD como disk cuando en realidad representa la palabra drive, que podría traducirse como unidad o dispositivo.

STRIPING

Reparto de bloques de un archivo a lo largo de múltiples unidades de disco. Esta técnica se emplea en RAID.

SUBDISCOS

Un subdisco es un conjunto de bloques contiguos que representan una porción específica del disco.

SUNOS

Fue la versión del sistema operativo Unix desarrollada por Sun Microsystems para sus estaciones de trabajo y servidores hasta el principio de los años 1990. Ésta estaba basada en los UNIX BSD con algunos añadidos de los System V UNIX en versiones posteriores.

SUPERBLOQUE

Reside en los sectores del 16 al 31 después del bloque de arranque. Es una tabla de información que describe al file system,

SYMANTEC

Symantec Corporation es una corporación internacional que produce software, especialmente relacionado a la seguridad. Fue fundada en 1982 y tiene su cuartel central en Cupertino, California (EE.UU.), y tiene operaciones en más de 40 países.

TARGET

Target (en español objetivo) es un anglicismo también conocido por público objetivo, grupo objetivo, mercado objetivo o mercado meta.

TB

Un terabyte es una unidad de medida de almacenamiento de datos cuyo símbolo es TB y puede equivaler a 1024 GB. Confirmado en 1960, viene del griego τέρας, que significa monstruo.

TI

Las tecnologías de la información y la comunicación (TIC) son un conjunto de servicios, redes, software y dispositivos que tienen como fin la mejora de la calidad de vida de las personas dentro de un entorno, y que se integran a un sistema de información interconectado y complementario.

TRASLAPE

Unión de dos elementos remontando una parte de un elemento sobre el otro.

UDFS

Universal Disk File System. El Formato de Discos Universal para unidades de almacenamiento ópticos como el DVD y el CD-ROM. Permite el intercambio de datos universales y soporta operaciones de lectura y escritura.

UFS

Unix File System (UFS) es un sistema de archivos utilizado por varios sistemas operativos UNIX y POSIX. Es un derivado del Berkeley Fast File System (FFS), el cual es desarrollado desde FS UNIX (este último desarrollado en los Laboratorios Bell).

UNIX

Sistema operativo multiplataforma, multitarea y multiusuario desarrollado originalmente por empleados de Bell de AT&T. Actualmente UNIX puede referirse a: Familia UNIX: grupo genérico de sistemas operativos que comparten determinados criterios en su diseño y por lo tanto son llamados de la familia (o tipo) UNIX. Son más de 100 sistemas operativos que se consideran de su familia.

USB

(Universal Serial Bus) Puerto de gran velocidad para comunicar computadoras y periféricos. Soporta plug & play y conexión en caliente (hot plugging). Soporta transferencias de 12 MBps. Un sólo puerto USB permite ser usado para conectar más de 127 dispositivos periféricos como ratones, módems, teclados, impresoras, etc.

VERITAS

Marca líder en soluciones de seguridad de datos. Propiedad de Symantec Inc.

VERITAS DMP

Veritas DMP (Dynamic Multi-Pathing). Característica que está disponible para arreglos de disco que permiten la conexión a múltiples controladores.

VERITAS FILE SYSTEM

Es una herramienta poderosa, de rápida recuperación programada, que proporciona un alto rendimiento y fácil manejabilidad en línea requerida por las aplicaciones de misión crítica.

VERITAS STORAGE FOUNDATION

Software que proporciona una solución completa para la administración del almacenamiento en línea heterogéneo. Respaldado por los productos Veritas Volume Manager y Veritas File System, ofrece un conjunto estándar de herramientas integradas que permiten centralizar la administración del crecimiento de los datos, maximizar las inversiones en hardware de almacenamiento, proteger los datos y adaptarse a las necesidades cambiantes de las empresa.

VERITAS VOLUME MANAGER

Veritas Volume Manager (VxVM) es parte de una suite de productos que proporciona las facilidades para poder lograr el almacenamiento virtual.

VOLUMEN

Es el nombre de cualquier división de un disco; siendo el disco duro el caso más común para esta técnica. La técnica consiste en dividir un disco en varias partes, las cuales actúan y son tratadas por el sistema como discos independientes; sin embargo, estas divisiones están presentes en un mismo disco físico. Cualquier unidad de almacenamiento completamente formateada es en realidad una partición primaria que ocupa todo el disco; en dicho caso, el término partición es poco usado, siendo sustituido por otro más familiar, como formato del disco. Esta técnica simple que puede considerarse como un precursor de la gestión de volúmenes lógicos, en inglés Logical Volume Management (LVM).

VTOC

Area llamada tabla de contenidos que guarda información acerca de la estructura y la organización de los discos. La VTOC también es llamada Etiqueta del Disco.

VXFS

Sistema de archivos utilizado por Veritas File System .

VXVM

Ver. Veritas Volume Manager.

WORM

(Write-Once-Read-Many), o sea, “Escriba una vez y lea muchas”.
Característica de algunos discos compactos.

ZIP

Dispositivo periférico que almacena datos en disquetes especiales llamados Zip con capacidades de más de 100 MB.

Bibliografía

Computer Architecture a Quantitative Approach, Patterson, David Hannessy, John,1996
Estructura de una computadora, Hillar, Gastón Carlos, 1997

Estudios de JISC/NPO en la preservación de materiales electrónicos: Un marco de los tipos y de los formatos de datos, y ediciones que afectan la preservación a largo plazo del material de Digital, Bennett, Juan C. 1997.

Basic Track I Solaris 9 Operating Environment , aut. Sun Microsystems Inc.. - San Antonio Road, Palo Alto, California . USA : Sun Educational, 2002.

Solaris Internals: Architecture and Techniques , James Mauro,1999

Veritas Foundation Suite 3.5 for Solaris: Administration and Troubleshooting, aut. Veritas Education. - [s.l.] : VERITAS Software Corporation, 2002. - Vol. 1.

[En línea]. - <http://docs.sun.com>.

[En línea]. - <http://support.veritas.com>.

<http://www.duiops.net/hardware/discos/discosd.htm> [En línea].

http://www.hitachigst.com/hdd/research/recording_head/pr/PerpendicularAnimation.html [En línea].

<http://www.recoverylabs.com/servicios/recuperacion/discos-duros.htm> [En línea].

<http://www.whatis.com> [En línea].

Wikipedia [En línea]. - <http://es.wikipedia.org>.

ÍNDICE

A

Almacenamiento Físico de Datos, 93

Almacenamiento Virtualizado, 86

Basado en Host, 88

Basado en las Red, 88

Apuntadores Directos, 66

Apuntadores Indirectos, 66

B

Bloque de arranque, 61

Bloques de datos, 66

Buffer y Caché. Véase Medidas que describen el desempeño de un Disco Duro

C

CD-ROM, 2

Command Overhead. Véase Medidas que describen el desempeño de un Disco Duro

D

DAS, 6

DDS, 4

Digital Data Storage. Véase DDS

Digital Linear Tape. Véase DLT

Direct Attached Storage. Véase DAS

Discos Duros, 6, 8

Actuador de un disco duro, 9

Arreglos de Discos Duros, 7

Cilindros de un disco duro, 10

Discos o Platters, 9

El eje de un disco, 9

Estructura física de un disco duro,
9

Fabricantes de, 28

Funcionamiento de la unidad de
disco duro, 11

Futuro de, 29

Las cabezas o heads, 9

Medidas que describen el
desempeño de un Disco Duro, 12

Particionamiento de, 32

Pistas de un disco duro, 10

Sectores de un disco duro, 10

Dispositivos de almacenamiento, 1

Distribución de un disco duro, 10

DLT, 4

DVD-ROM, 3

E

EMC2. Véase Fabricantes de Discos Duros

Enhanced IDE, 16

ESDI, 14

Estructura Física de Datos, 93

Etiquetas de los discos, 41

F

FCAL, 21

Fiber Channel, 21

File Systems, 57

Firewire, 25

floppy, 1

Fujitsu. *Véase* Fabricantes de Discos Duros

H

Hewlett Packard. *Véase* Fabricantes de Discos Duros

Hitachi. *Véase* Fabricantes de Discos Duros

HVD, 21

I

IBM. *Véase* Fabricantes de Discos Duros

IDE, 14

Inodo, 63

Interfaces y controladoras, 13

SCSI: Small Computer System Interface, 17

L

Latencia (latency). *Véase* Medidas que describen el desempeño de un Disco Duro

Librerías de cintas, 6

Linear Tape Open. *Véase* LTO

LTO, 4

LVD, 21

M

Memoria USB, 5

MP3, 6

N

NAS, 6

Network Attached Storage. *Véase* NAS

O

Objetos de Almacenamiento, 97

Grupos de Discos, 97

Subdiscos, 99

P

Plexes, 99

Pseudo File Systems, 58

R

RAID, 6

RAID 0, 71, 74

RAID 1, 73

RAID 10, 82

RAID 3, 76

RAID 4, 77

RAID 6, 80

Redundant Array of Inexpensive Disks.
Véase RAID

Reproductores de audio y video digital, 5

robots de respaldo. *Véase* Librerías de cintas

S

SATA Externo, 23

SCSI, 17

Seagate. Véase Fabricantes de Discos Duros

Serial ATA, 22

Sistema Operativo Solaris, 31

ST506, MFM y RLL, 14

Sun Microsystems. Véase Fabricantes de Discos Duros

Super bloque primario, 62

Super bloque de respaldo, 62

T

Tablas de partición de discos, 40

Tiempo de Búsqueda (seek time). Véase Medidas que describen el desempeño de un Disco Duro

Toshiba. Véase Fabricantes de Discos Duros

Transferencia. Véase Medidas que describen el desempeño de un Disco Duro

U

Unidades de cinta magnética, 4

Unidades de discos flexibles. Véase floppy

Unidades Jazz. Véase

Unidades Zip, 2

UNIX, 30

USB. Véase Interfaces y controladoras

V

VERITAS File System, 90

Veritas Storage Foundation, 86

Veritas Volume Manager 6, 89

Volumen, 95

VTOC, 61

W

Western Digital. Véase Fabricantes de Discos Duros

Write-Once-Read-Many. Véase WORM