

Administración de servidores

Jordi Serra Ruiz
Miquel Colobran Huguet
Josep Maria Arqués Soldevila
Eduard Marco Galindo

P08/81029/02266



Universitat Oberta
de Catalunya

www.uoc.edu

Índice

Introducción.....	7
Objetivos.....	8
1. Desmitificando el servidor.....	9
2. Funciones del servidor.....	10
2.1. Requisitos de los sistemas operativos en red	11
3. Elementos del servidor.....	13
3.1. Memoria RAM	13
3.2. Unidad de control de proceso	13
3.3. Placa base	14
3.4. Placa de comunicaciones	14
3.5. Disposición física del servidor	14
4. Configuraciones de servidores.....	16
4.1. <i>Host</i> o sistema centralizado	17
4.1.1. Servidores virtuales	18
4.1.2. Servidores de aplicaciones	19
4.2. Agregación de <i>hosts</i> o sistema distribuido	19
4.2.1. Balanceo de carga (<i>load balancers</i>)	20
4.2.2. Sistemas clúster	20
4.2.3. Computación en malla (<i>grid</i>)	23
5. Almacenamiento.....	26
5.1. Necesidades de la organización	26
5.2. <i>Direct Attached Storage</i> (DAS)	28
5.2.1. Discos <i>Intelligent Drive Electronics</i>	28
5.2.2. Discos <i>Serial ATA</i>	29
5.2.3. Discos <i>Small Computer System Interface</i>	29
5.2.4. Discos <i>Serial Attached SCSI</i>	30
5.2.5. Agrupaciones de discos en el servidor	31
5.2.6. Sistemas de ficheros	34
5.3. <i>Storage Area Network</i> y <i>Network Attached Storage</i>	36
5.3.1. <i>Storage Area Network</i>	36
5.3.2. <i>Network Attached Storage</i>	41
6. Copia de seguridad.....	43
6.1. Dispositivos de copia de seguridad	43
6.1.1. <i>Digital Audio Tape</i>	43

6.1.2.	<i>Digital Linear Tape</i>	43
6.1.3.	<i>Advanced Intelligent Tape</i>	43
6.1.4.	<i>Linear Tape Open</i>	44
6.1.5.	Librerías de copia	44
6.1.6.	Grabadora DVD	44
6.1.7.	Disco duro	45
6.1.8.	¿Dónde tienen que estar los dispositivos de copia?	45
6.2.	Políticas de copia de seguridad	46
6.2.1.	Tipos de copias de seguridad	47
6.2.2.	Políticas de copias de seguridad	48
6.2.3.	Información no variable	51
6.2.4.	Dónde se pueden guardar las copias de seguridad	52
6.2.5.	Recomendaciones	53
6.3.	Plan de contingencia	53
7.	Impresoras	54
7.1.	Impresoras láser	54
7.2.	Impresoras de inyección de tinta	55
7.3.	Impresoras remotas	56
7.4.	<i>Internet Printing Protocol</i>	57
8.	La corriente eléctrica	58
8.1.	La toma de tierra	59
8.2.	Sistema de Alimentación Ininterrumpida	60
9.	Seguridad de los servidores	63
9.1.	Seguridad física de los servidores	63
9.2.	Software	64
9.3.	Alta disponibilidad	64
9.3.1.	Sistemas tolerantes a fallos	65
9.3.2.	Clústers de alta disponibilidad	66
10.	Aspectos legales	67
10.1.	Colegios profesionales	67
11.	Tareas/responsabilidades	68
Resumen		70
Actividades		71
Ejercicios de autoevaluación		71
Solucionario		72
Glosario		73

Bibliografía..... 76

Introducción

Hoy en día, los servidores ya no son ordenadores “de película” que ocupan habitaciones enteras, sino que son ordenadores con características especiales de hardware y de software.

Si tenemos que mantenerlos, necesitamos saber en qué se diferencian de los ordenadores de sobremesa. Tenemos que saber qué podemos esperar de ellos y qué les podemos pedir que hagan. También es importante tener presente todo lo que tenemos que hacer para protegerlos, al menos físicamente. Finalmente, tendremos que escoger, configurar y mantener el sistema operativo.

Veremos qué hardware podemos conectar y cuál es la configuración más adecuada dependiendo de la función a la que lo queramos destinar.

Hay que tener presente que el servidor estará conectado a la red. Eso afecta a su configuración, y también se tiene que acordar que, como administrador de servidores, hay un conjunto de tareas y de responsabilidades que tenemos que conocer. Finalmente, tendremos que encargarnos de mantenerlo y vigilar que siempre funcione correctamente.

Objetivos

En los materiales didácticos de este módulo, presentamos los contenidos y las herramientas imprescindibles para alcanzar los objetivos siguientes:

- 1.** Conocer las características que han de tener los ordenadores que hacen de servidores, los cuales han de cumplir unos requisitos de funcionamiento bastante estrictos.
- 2.** Conocer las características que han de tener los sistemas operativos servidores, porque han de cumplir unas funciones diferentes y unos requisitos de seguridad bastante estrictos.
- 3.** Conocer las posibles configuraciones de servidores para obtener sistemas con mejor rendimiento. También, comprender las diferentes combinaciones y virtualizaciones de servidores con objetivos comunes o dispersos.
- 4.** Saber los diferentes tipos de almacenamiento, interno y externo, sus componentes, configuraciones y variedades para garantizar el rendimiento y la seguridad del servidor.
- 5.** Conocer los diferentes dispositivos y políticas de hacer copias de seguridad.
- 6.** Conocer los diversos componentes de hardware que se instalan en un servidor para poder obtener un buen rendimiento de ellos.
- 7.** Conocer las responsabilidades de un administrador de servidores.
- 8.** Saber cómo se ha de aplicar en los servidores el concepto de seguridad.

1. Desmitificando el servidor

Cuando se habla de servidores, hay una tendencia generalizada a creer que se trata de máquinas enormes que ocupan salas enteras y que se encuentran protegidas en ambientes especiales y con una seguridad de película. En un principio, los servidores sí que ocupaban grandes espacios y tenían ambientes especiales. Incluso ahora podemos encontrar algunos servidores centrales de tipo *host* o grupos de servidores dispuestos físicamente de manera que ofrecen este aspecto. Pero lo cierto es que la mayoría, individualmente, mantienen una apariencia muy parecida a una estación de trabajo cualquiera.

Así pues, aunque los servidores no son iguales en la imagen que tenemos predefinida de ellos, sí son sobradamente diferentes en funcionalidad y servicio en cualquier ordenador personal.

Hoy los servidores no son diferentes externamente. Lo que varía es el software y el hardware instalados dentro de la carcasa externa.

Un servidor es una máquina que funciona 24 × 7 (veinticuatro horas los siete días de la semana), y eso quiere decir que tiene que tener un hardware preparado para no parar nunca (problemas de calentamiento) y soportar reparaciones y la sustitución de discos averiados en caliente (sin apagar el ordenador). También tiene que poder aguantar centenares de peticiones de usuarios por medio de la red con tiempo de respuesta aceptable. Incluso tienen sistemas para que los usuarios accedan a la información de una manera selectiva, y gestionan colas de impresión, muestran páginas web, registran la actividad total que se hace, gestionan el correo de la organización y ya no ocupan habitaciones enteras.

2. Funciones del servidor

Un servidor es un sistema que pone recursos propios a disposición de otros ordenadores (los clientes). Por lo tanto, actualmente el concepto de servidor ya no está asociado necesariamente a un ordenador.

Podemos distinguir dos tipos de servidores:

- **Servidores físicos.** Muchas veces también se llaman servidores corporativos. Es la cantidad de ordenadores que hay en una organización dedicados exclusivamente a tareas de servicio.
- **Servidores funcionales.** La cantidad de tareas que hacen los servidores es muy grande. Conceptualmente, un servidor proporciona recursos y, por lo tanto, un ordenador físico puede servir muchas cosas. De la misma manera, un ordenador puede no estar dedicado a hacer de servidor, pero sí servir alguna cosa (**dar un servicio**).

Recursos de un servidor

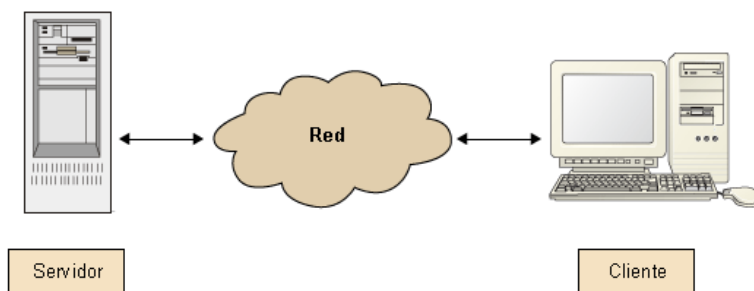
Los recursos que un servidor pone a disposición de otros ordenadores pueden ser datos, ficheros, aplicaciones, impresora, disco, correo...

Con las redes y la tecnología cliente/servidor, un servidor es una aplicación que da (sirve) información a un programa (cliente) que la pide mediante una conexión (normalmente la red) a partir de un protocolo.

Ved también

Sobre la arquitectura cliente/servidor, mirad más adelante el apartado 4 en este mismo módulo.

Esquema de la arquitectura cliente/servidor



Así, podemos encontrar servidores de muchas cosas: servidores de ficheros, servidores de impresión, servidores web, servidores de noticias, servidores FTP, servidores de correo, servidores DNS, servidores buscausuarios¹.

⁽¹⁾Servidores buscausuarios en inglés se expresa como *finger server*.

Como son aplicaciones, un ordenador puede ofrecer muchos servicios a la vez, es decir, puede hacer diversas funciones. Tendríamos, pues, un servidor físico que lleva a cabo funcionalmente el papel de diversos servidores.

Normalmente, un servidor físico da diferentes servicios; depende básicamente de qué destinación tiene y qué demanda tiene aquello que sirve.

Ejemplo de servicios de un servidor físico

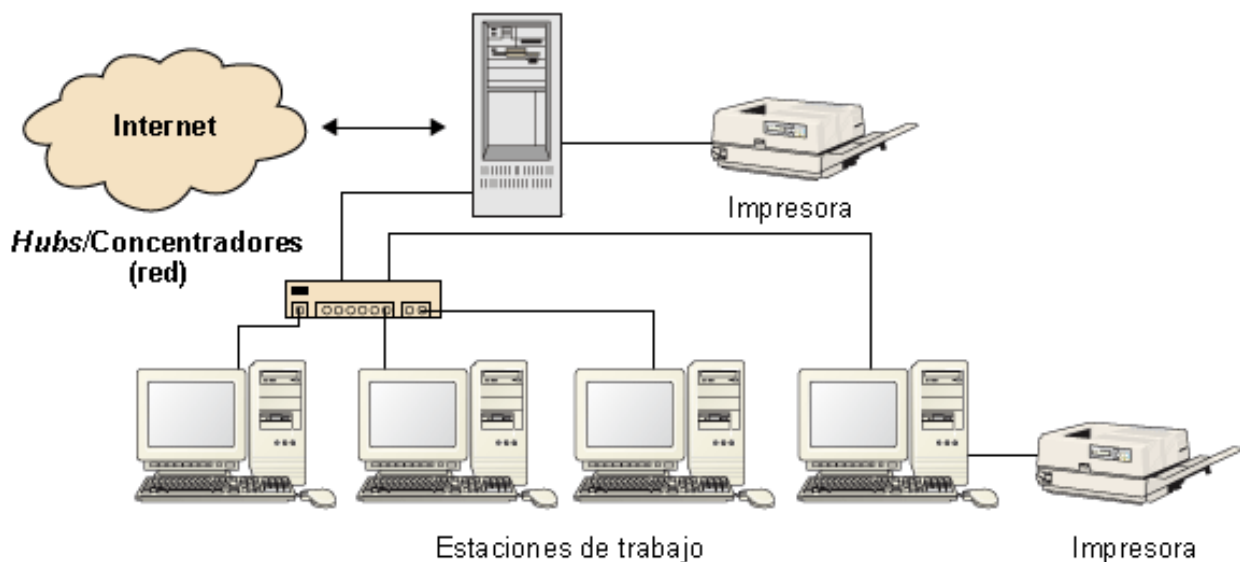
Una organización dedicada a la venta por Internet seguramente tendrá un servidor físico dedicado a hacer de servidor de web, mientras que en el caso de una empresa que sólo tenga el catálogo de sus productos, el servidor web estará en un servidor que también ofrezca otros servicios como, por ejemplo, la contabilidad (servidor de ficheros), los directorios de usuario (servidor de ficheros) y el correo electrónico (servidor de correo).

Por lo tanto, el número y la función de servidores físicos instalados depende de la actividad de la organización, del software que se utilice y del plan estratégico (las futuras ampliaciones de software y hardware).

2.1. Requisitos de los sistemas operativos en red

La instalación del sistema operativo (SO) en el servidor que escogemos nos tiene que permitir una gran variedad de funciones necesarias. La idea de servidor está basada en la tecnología cliente/servidor, pero si además permite una buena comunicación entre estaciones de trabajo, mucho mejor. Con la tecnología de base, tenemos la estructura siguiente:

Estructura de la tecnología cliente/servidor



Por lo tanto, el SO² de red tiene que poder proporcionar básicamente las funciones siguientes:

- Servidor de ficheros:
 - Poder definir grupos de usuarios.
 - Compartir ficheros entre todos los usuarios.
 - Compartir ficheros entre los grupos de usuarios.

⁽²⁾Recordad que "SO" es la abreviatura de sistema operativo.

- Que cada usuario tenga espacio personal para guardar la información. El hecho de que esté en el servidor facilita la movilidad y las copias de seguridad.
- Servidor de aplicaciones:
 - Compartir programas entre todos los usuarios.
 - Compartir programas entre los grupos de usuarios.
- Servidor de impresión:
 - Compartir las impresoras.
- Servidor de correo:
 - Enviar y recibir mensajes.

Todo eso con las restricciones de seguridad y permisos adecuados. Si además hace falta que en la red haya seguridad complementaria, como un cortafuegos³ u otros tipos de servidores –como un servidor web o un servidor de bases de datos–, se tienen que poder instalar o se tiene que poner otro servidor físico para instalarlos. En estos casos, la comunicación entre los servidores es una cuestión importante.

⁽³⁾Cortafuegos en inglés se expresa como *firewall*.

3. Elementos del servidor

Un servidor es un ordenador con una configuración de hardware y de software ajustada a la función que tiene que llevar a cabo.

De entrada, los componentes son los mismos que para un ordenador de sobremesa. Así, en un servidor podremos encontrar: monitor, teclado, ratón, lector óptico (DVD), memoria RAM, placa de comunicaciones, unidades de almacenamiento (discos duros), unidad de control de proceso (CPU⁴), fuente de alimentación, placa gráfica y placa base.

⁽⁴⁾La sigla “CPU” corresponde a la expresión inglesa *central processing unit*.

Algunos de los componentes no tienen que ser especiales (el monitor, el teclado, el ratón y la placa gráfica), mientras que con respecto a los otros componentes sí que hay prestaciones especiales y se tienen que mirar a fondo.

3.1. Memoria RAM

Todos los usuarios pedimos (hacemos peticiones) a un servidor. Por lo tanto, es importante que nos pueda responder cuanto antes mejor. Por este motivo, una buena cantidad de memoria RAM es muy importante, y cuanto más rápida sea la RAM que se instale, mejor. Si se trata de un servidor de bases de datos, entonces la cuestión es mucho más crítica y tenemos que instalar la cantidad de RAM que recomienda el vendedor del producto de bases de datos para asegurar un funcionamiento óptimo.

Es muy necesaria una gran cantidad de memoria RAM.

3.2. Unidad de control de proceso

En contra del pensamiento general, y exceptuando que sea un servidor de bases de datos con grandes transacciones y operaciones de bases de datos complejas, la CPU⁵ no es excesivamente crítica para el buen funcionamiento de un servidor. Basta con una buena CPU y no hacen falta sistemas multi CPU, en la mayoría de los casos. La CPU es necesaria en procesos que piden grandes cantidades de cálculo, pero no es el caso general de un servidor de ficheros, de un servidor de impresión o de un servidor web, por ejemplo. Podría ser el caso de un gran servidor de bases de datos al cual se hicieran muchas peticiones que implicaran consultas complejas y, por lo tanto, mucho movimiento en las tablas, pero seguramente entonces sería más un indicador de que algún

⁽⁵⁾Recordad que “CPU” es la sigla de unidad de control de proceso.

elemento de la base de datos está mal diseñado, porque este tipo de consultas no acostumbran a ser frecuentes sobre una base de datos (excepto si hay miles de usuarios).

Generalmente, la CPU no es crítica.

3.3. Placa base

Es esencial que la placa base sea de muy buena calidad para asegurar que hay una buena velocidad de transmisión entre todos los componentes del servidor. El bus del sistema forma parte de la placa base⁶ y es el componente que permite la comunicación entre todos los dispositivos dentro del ordenador. Entre una placa de buena calidad y una que no lo sea, el rendimiento puede bajar de una manera apreciable. El gran problema es que cuesta mucho detectarlo porque todo funciona, aunque ligeramente más lento.

⁽⁶⁾La placa base también se llama *placa madre*; su equivalente inglés es *motherboard*.

La placa base es vital para el servidor.

3.4. Placa de comunicaciones

La placa de comunicaciones es el punto de comunicación entre el servidor y “todo el mundo”. Por lo tanto, su calidad y velocidad determinan el comportamiento del servidor hacia la red. Es un componente crítico.

Una placa 10/100 de par trenzado en un concentrador o en un conmutador a 100 Mb es una buena solución para tener un servidor bien conectado (si se puede conectar a cualquier otra tecnología superior como, por ejemplo, fibra óptica, mejor). Cuanto más rápida sea la conexión del servidor con la red, antes podrá atender las demandas de las estaciones de trabajo e irá más descargado (o más carga podrá soportar sin colapsarse).

La placa de comunicaciones determina la capacidad de transmitir información a la red del servidor.

3.5. Disposición física del servidor

La disposición física de los servidores es variada. Desde cajas especiales para soportar el calentamiento (sobre todo si tienen muchas unidades de disco) hasta los sistemas rac, donde el teclado y la pantalla para controlar los ordenadores se implementan vía red.

Finalmente, encontramos el sistema *Blade*, en el que cada servidor se integra como una lámina dentro de una estructura (*blade center*) y se comparten recursos, como el acceso a la red, en una red Storage Area Network (SAN), fuentes de alimentación, ventiladores...



Servidores en *Blade* dentro de un *blade center*

4. Configuraciones de servidores

Las diferentes necesidades de una organización hacen que, a menudo, un equipo no sea suficiente. De manera que es habitual que las organizaciones tengan más de un servidor físico para alcanzar sus objetivos. Podemos encontrar, pues, un servidor que lleve a cabo una o muchas tareas o muchos servidores trabajando por un propósito común. También es posible encontrar servidores muy diferentes entre ellos agrupados en un mismo espacio llevando a cabo tareas diversas.

Estas combinaciones, muchas veces heterogéneas, de servidores se basan en la funcionalidad. Así, si por ejemplo queremos un servicio de correo que difícilmente falle, pondremos un cluster de correo en alta disponibilidad. Eso representa al menos dos servidores exclusivamente dedicados al correo. Si además nos hace falta un servicio de ficheros muy grande, entonces pondremos un servidor dedicado a *Network Attached Storage* (NAS) con una librería de copia de seguridad⁷.

⁽⁷⁾Copia de seguridad en inglés se expresa como *backup*.

Como podemos ver, es la necesidad de la organización lo que configura la estructura de los servidores. Debido al entorno dinámico de las organizaciones, se tendría que hacer una planificación inicial para prever, en la medida de lo posible, las ampliaciones que pueda haber para no hacer gastos y tareas de organización del sistema informático que sean insuficientes en poco tiempo.

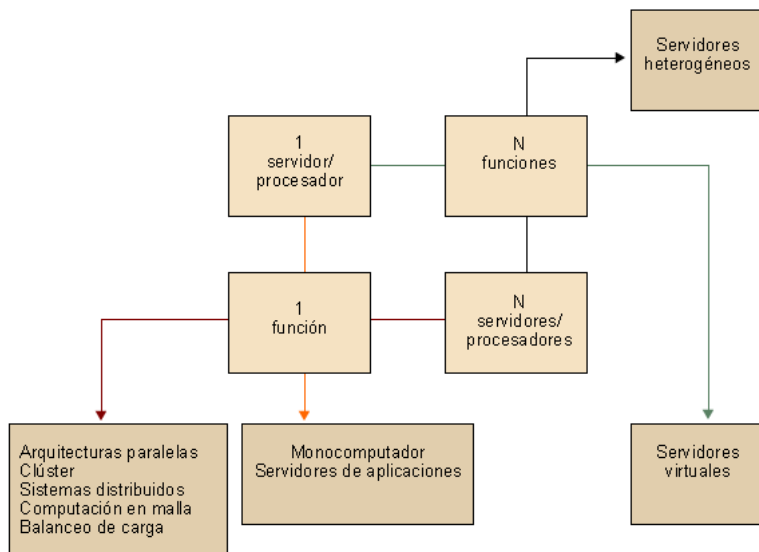
Ved también

Sobre cómo planificar el departamento, ved el módulo "El sistema informático dentro de la organización".

Esta gran variedad, atendiendo a la configuración y función, hace necesaria una clasificación de las configuraciones de los servidores. Esta clasificación no pretende ser exhaustiva, sino orientativa y didáctica, sabiendo que existen otras clasificaciones.

La configuración de los servidores tiene que cubrir las necesidades específicas de la organización.

Diagrama de cruces



Este diagrama de cruces define, conceptualmente, los diferentes tipos de servidores y los servicios que éstos pueden ofrecer a sus clientes conectados:

- **Un servidor/procesador, una función.** Es el nivel más sencillo de servidor, un sistema físico dedicado a una sola función. Por ejemplo, un ordenador realizando tareas de gestión de correo (servidor de aplicaciones).
- **Un servidor/procesador, N funciones.** Si disponemos de un ordenador poco utilizado en cuanto a recursos se refiere, podemos aprovechar este remanente con el fin de ofrecer otros servicios a los clientes. Así pues, tenemos un ordenador optimizando recursos y con diversas funciones de servicio.
- **N servidores/procesadores, una función.** En nuestra organización podemos tener servicios críticos, ya sea por necesidad de servicio, seguridad o rendimiento, que hacen necesarios un número de recursos muy importantes y escalables. Esta necesidad desarrolla las arquitecturas donde una sola tarea es tratada por más de un ordenador.
- **N servidores/procesadores, N funciones.** Cuando diversas funciones son tratadas por diferentes ordenadores, tenemos un sistema de servidores heterogéneo, en el que pueden aparecer un gran número de combinaciones posibles.

4.1. Host o sistema centralizado

Podemos distinguir dos tipos de sistemas centralizados: servidores virtuales y servidores de aplicaciones.

4.1.1. Servidores virtuales

Los servidores virtuales basan su funcionamiento en la tecnología de la virtualización. La virtualización, esencialmente, es dar a una computadora la posibilidad de realizar el trabajo de múltiples computadores, compartiendo los recursos por medio de diversos entornos. Típicamente, se ha referido a una sola computadora capaz de hacer trabajar, al mismo tiempo, a diferentes sistemas operativos y servicios de forma segura.

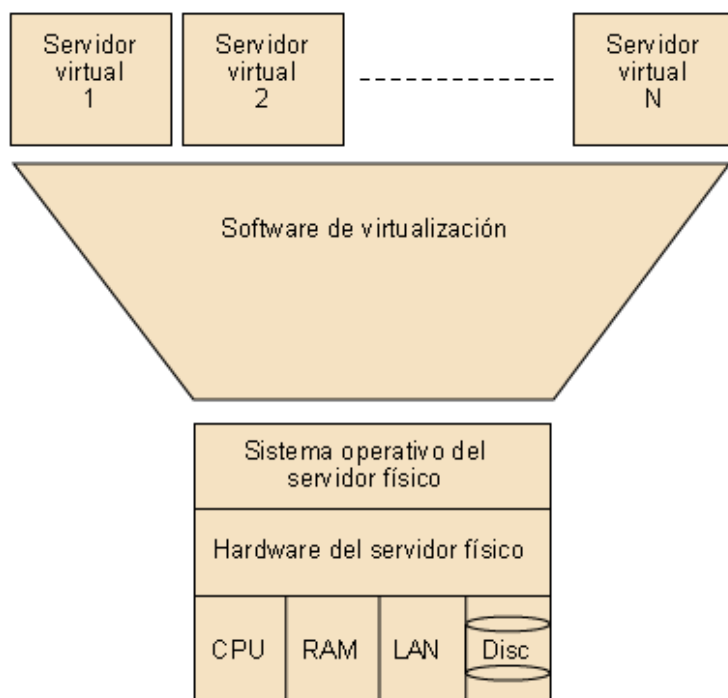
Podemos afirmar, pues, que un servidor virtual es aquel servidor capaz de realizar el trabajo de diversos servidores compartiendo los recursos del sistema, mediante uno o más sistemas operativos de forma segura.

Los servidores virtuales tienen bastantes ventajas que los hacen atractivos:

- Reducción del número de servidores físicos.
- Reducción del espacio dentro del centro de datos.
- Reducción del consumo de energía.
- Compartimentación de recursos y eficiencia de utilización.
- Centralización y simplificación de la gestión.

Hay diversos sistemas de virtualización, dependiendo de la plataforma tecnológica del servidor físico.

Esquema de servidor virtual



Virtualizadores comerciales

Algunos virtualizadores comerciales conocidos son VMWARE, Windows Virtual Server, Linux Virtual Server, etc.

4.1.2. Servidores de aplicaciones

Un servidor de aplicaciones es un servidor avanzado que permite gestionar aplicaciones y todos los recursos necesarios asociados, como el acceso a base de datos, seguridad, mantenimiento... Un servidor de aplicaciones se relaciona normalmente con un sistema de tres capas:

1) Primera capa⁸: capa de interacción con el usuario, basada en navegadores gráficos.

⁽⁸⁾Primera capa en inglés se expresa como *front-end*.

2) Capa intermedia⁹: servidor de aplicaciones en red local.

⁽⁹⁾Capa intermedia en inglés se expresa como *middle-tier*.

3) Tercera capa¹⁰: servidor de base de datos.

⁽¹⁰⁾Tercera capa en inglés se expresa como *back-end*.

Basado en la tecnología Java 2 Platform, Enterprise Edition (J2EE), el servidor de aplicaciones es una máquina virtual Java¹¹ (JVM) que ejecuta aplicaciones de usuario. El servidor de aplicaciones colabora con el servidor web para ofrecer una respuesta dinámica y personalizada a cada petición de cliente. Además, también da respuesta avanzada a código de aplicación, miniaplicaciones de servidor¹², JavaServer Pages (JSP), *enterprise beans* y las clases que les dan soporte.

⁽¹¹⁾Máquina virtual Java en inglés se expresa como *Java Virtual Machine*.

⁽¹²⁾Miniaplicaciones de servidor en inglés se expresa como *servlets*.

Ved también

Ved el módulo "Administración de la web" para repasar estos conceptos.

Un servidor de aplicaciones puede gestionar un gran número de aplicaciones, conexiones a base de datos y recursos, lo que hace que sea un sistema versátil, seguro y de futuro. Hay ventajas claras al utilizarlos:

- **Integridad.** Gracias a la centralización de las aplicaciones, se evitan los problemas de actualizaciones y migraciones en los sistemas cliente. Cualquier cambio se hace centralizado, disminuyendo al mínimo el riesgo.
- **Configuración centralizada.** Los cambios de configuración de la aplicación, como los accesos a la BD, se hacen de forma centralizada.
- **Seguridad.** El servidor de aplicaciones se convierte en el punto central de acceso a los datos, disminuyendo la diversidad y haciendo más fácil una buena defensa.
- **Rendimiento.** Se hace gestionando los accesos clientes al servidor de aplicaciones y de éste al servidor de base de datos.

Servidores de aplicaciones comerciales

Hay diversos servidores de aplicaciones en el mercado. Esta es una lista de los más utilizados: JBoss, IIS, WebSphere, Bea Weblogic, Tomcat...

4.2. Agregación de *hosts* o sistema distribuido

Podemos distinguir tres tipos de sistemas distribuidos: sistemas de balanceo de carga, clústers y *grids*.

4.2.1. Balanceo de carga (*load balancers*)

Balancear una carga significa dividir el total de trabajo que un sistema o computadora tiene que hacer entre dos o más sistemas o computadoras.

Así pues, esta división de carga permite realizar el mismo trabajo en una porción de tiempo más reducida, o lo que es lo mismo; permite realizar más carga de trabajo en el mismo tiempo total.

Los siguientes son algunos conceptos generales sobre el balanceo de carga:

- El balanceo de carga se puede implementar por hardware, software o una combinación de los dos.
- El balanceo de carga es especialmente indicado para entornos en los que es muy difícil prever el volumen de carga de trabajo.
- El factor de división de carga se puede definir dando más o menos carga a cada uno de los sistemas implicados. Esta característica es la **carga asimétrica**.

4.2.2. Sistemas clúster

Un clúster es un grupo de computadoras interconectadas que trabajan conjuntamente en la solución de un problema. Estos sistemas constituyen una solución flexible, de bajo coste y de gran escalabilidad para aplicaciones que requieren una elevada capacidad de computadora y memoria.

Un **clúster** es un grupo de equipos independientes que ejecutan una serie de aplicaciones, de forma conjunta, y aparecen ante los clientes y aplicaciones como un solo sistema.

Historia de los clústers

Si miramos la historia de los clústers, encontramos que si bien no se sabe la fecha exacta del primer clúster, se considera que la base científica del concepto del procesamiento en paralelo la estableció Gene Amdahl, que trabajaba en IBM, hacia 1967. El desarrollo de los clústers ha estado siempre unido al de las redes de computadores, ya que desde el comienzo se buscó la unión de los sistemas informáticos para obtener más rendimiento y capacidades. De todas formas, el primer clúster comercial fue ARCNet, desarrollado en 1977 por la corporación DataPoint. A partir de aquí, toda una serie de productos popularizaron el concepto, hasta la puesta en marcha del proyecto Beowulf, en 1994, que implicaba la interconexión en red local de computadores estándar, y gestionaba cómo éstos interactuaban entre sí. La idea tuvo tal éxito que incluso la NASA la adoptó.

Características de los clústers

Las siguientes son las principales características de los clústers:

- Un clúster consta de dos o más nodos conectados entre sí por un canal de comunicación.
- Cada nodo necesita únicamente un elemento de proceso, memoria y una interfaz para comunicarse con la red del clúster.
- Los clústers necesitan software especializado, ya sea en la aplicación o en el núcleo.
- Todos los elementos del clúster trabajan para cumplir una funcionalidad conjunta, sea ésta la que sea. Es la funcionalidad la que caracteriza el sistema.

Ventajas económicas de los clústers

Las **ventajas económicas** son una razón importante para la construcción de clústers. Reduce costes en el gasto inicial tanto de planificación, de instalación y también los costes asociados al mantenimiento (el coste total¹³) comparados con un “ordenador” de las prestaciones equivalentes.

(¹³) Coste total en inglés se expresa como *Total Coste of Ownership* (TCO).

Sencillez de los clústers

La tecnología que hace funcionar un clúster se basa en la unión de elementos sencillos (que pueden ser incluso ordenadores normales). Y esta sencillez es más beneficiosa cuando hablamos de disponibilidad de piezas de recambio (pueden ser piezas estándar) o de un tiempo de paro¹⁴ reducido (no hay tiempo de espera para un técnico enviado por la marca del equipo).

(¹⁴) Tiempo de paro en inglés se expresa como *downtime*.

Disponibilidad de los clústers

La interconexión de dos o más computadores trabajando, conjuntamente, en la solución de un problema, permite incrementar la disponibilidad de servicio, ya que se dividen aproximadamente los números de puntos críticos de servicio entre el número de nodos del clúster.

Escalabilidad de los clústers

Si el SO del clúster lo permite, sólo hay que conectar más equipos a la red del clúster, configurarlos correctamente, y ya tenemos un clúster ampliado y mejorado. Incluso mejorando alguno de los elementos que forman parte de cada nodo (memoria RAM o disco, por ejemplo), se obtiene una mejora del rendimiento o la disponibilidad.

Escalabilidad

La **escalabilidad** es la capacidad de un equipo para afrontar volúmenes de trabajo cada vez mayores, sin dejar de dar un rendimiento aceptable. Hay dos clases de escalabilidad:

- **Hardware** o escalamiento vertical: basado en el uso de un gran equipo con una capacidad que aumenta a medida que lo exige la carga de trabajo.
- **Software** o escalamiento horizontal: basado en el uso de un clúster hecho de diversos equipos de mediana potencia, que funcionan de manera muy similar a las unidades *Redundant Array of Inexpensive Disks* (RAID) de disco.

Rendimiento de los clústers

El incremento de recursos asignados con el fin de resolver la misma carga de trabajo permite aumentar el rendimiento del sistema como conjunto.

Balanceo de carga de los clústers

La tecnología de clúster de servidores por balanceo de carga mejora la respuesta a las peticiones, conmutándolas entre los diversos nodos del clúster.

Componentes de los clústers

Los componentes de un clúster son los siguientes:

- **Nodos.** Pueden ser simples ordenadores, sistemas multiprocesador o estaciones de trabajo.
- **Sistemas operativos.** Tienen que ser de fácil uso y acceso, y además permitir múltiples procesos y usuarios.
- **Conexiones de red.** Los nodos de un clúster pueden conectarse mediante una simple red *Ethernet*, o se pueden utilizar tecnologías especiales de alta velocidad, como *Fast Ethernet*, *Gigabit Ethernet*, *Myrinet*, *Infiniband*, *SCI*.
- **Software intermediario**⁽¹⁵⁾. El software intermediario es un software que generalmente actúa entre el sistema operativo y las aplicaciones con la finalidad de proveer una interfaz única de acceso al sistema, denominada *Single System Image* (SSI), la cual genera la sensación al usuario que utiliza un único ordenador muy potente.
- **Herramientas para la optimización y mantenimiento del sistema.** Migración de procesos, *checkpoint-restart* (parar uno o varios procesos, migrarlos a otro nodo y continuar su funcionamiento), balanceo de carga, tolerancia a fallos, etc.
- **Escalabilidad.** Tiene que poder detectar, automáticamente, nuevos nodos conectados al clúster para proceder a su utilización.

Ved también

Sobre el balanceo de carga, ved el subapartado 4.2.1 de este módulo.

⁽¹⁵⁾Software intermediario en inglés se expresa como *middleware*.

- **Ambientes de programación paralela.** Los ambientes de programación paralela permiten implementar algoritmos que hacen uso de recursos compartidos: CPU¹⁶, memoria, datos y servicios.

(16) Recordad que “CPU” es la abreviatura de Central Processing Unit, en castellano Unidad de Control de Proceso.

Tipos de clústers

Los clústers pueden clasificarse en base a sus características. Se pueden distinguir:

- **Clústers de alto rendimiento¹⁷.** Son clústers que ejecutan tareas que requieren gran capacidad computacional. Estas tareas pueden comprometer los recursos del clúster durante largos periodos de tiempo.
- **Clústers de alta disponibilidad¹⁸.** Son clústers diseñados para proporcionar disponibilidad y confiabilidad. La confiabilidad se provee mediante software que detecta fallos del sistema y permite recuperarse frente a éstos, mientras que en hardware se evita tener un único punto de fallo.
- **Clústers de alta eficiencia¹⁹.** Son clústers que están diseñados con el objetivo de ejecutar la mayor cantidad de tareas en el menor tiempo posible.

(17) Clústers de alto rendimiento en inglés se expresa como *High Performance clusters* (HPC).

(18) Clústers de alta disponibilidad en inglés se expresa como *High Availability* (HA).

(19) Clústers de alta eficiencia en inglés se expresa como *High Throughput* (HT).

4.2.3. Computación en malla (*grid*)

La computación en *grid* o malla es un nuevo paradigma de computación distribuida en el cual todos los recursos de un número indeterminado de computadores son englobados como un único superordenador de forma transparente.

Estos computadores englobados no están conectados o enlazados rígidamente, es decir, no tienen por qué estar en el mismo punto geográfico.

Los orígenes de la computación en *grid* se deben a la idea de la compartimentación de recursos. La práctica conocida como “computación distribuida” nos lleva a los inicios de la informática. A finales de los años cincuenta y principios de los sesenta, los investigadores se dieron cuenta de que necesitaban hacer más eficientes los sistemas que habían costado una fortuna; “los sistemas pierden mucho tiempo esperando que los usuarios introduzcan datos”. Los investigadores razonaron, entonces, que diversos usuarios podrían compartir el sistema aprovechando el tiempo de procesamiento no empleado.

En 1969 encontramos ya una primera aproximación a la definición de *grid* por parte de Len Kleinrock, que sugirió proféticamente:

“Nosotros probablemente veremos la extensión de las ‘utilidades de los ordenadores’, como las utilidades de la corriente eléctrica y telefónicas, que darán servicio a las casas y las oficinas por todo el país”.

En 1998, Carl Kesselman e Ian Foster intentaron otra definición en su libro *The Grid: Blueprint for a New Computing Infrastructure*.

Grid es la infraestructura de hardware y el software que proporciona un acceso serio, constante, penetrable y económico a capacidades computacionales de alta calidad.

En una revisión de la definición por los mismos autores junto con Steve Tuecke, se definió la computación *grid* como compartimentación de los recursos coordinados y de la solución de un problema en organizaciones virtuales dinámicas y multiinstitucionales.

El sistema de computación en malla es un sistema que tiene las siguientes características:

- 1) **Sus recursos coordinados no están sujetos a un control central.** Un *grid* integra y coordina recursos y usuarios que trabajan con diferentes dominios, por ejemplo, estaciones de trabajo de usuarios frente a computadoras centrales; unidades administrativas diferenciadas de la misma organización; o diferentes organizaciones.
- 2) **Utiliza un estándar, abierto, protocolos e interfaces genéricas.** Un *grid* está hecho de protocolos genéricos e interfaces que tienen, como principales inconvenientes, la autenticación, autorización, descubrimiento y acceso a los recursos. Es importante que estos protocolos sean estándares y abiertos.
- 3) **Entrega las cualidades no triviales de servicio.** Un *grid* permite a los recursos que lo constituyen ser utilizados de una forma coordinada entregando diferentes cualidades de servicio, relacionadas por ejemplo con el tiempo de respuesta, rendimiento, disponibilidad y seguridad, y/o la asignación de múltiples recursos para conocer las demandas de los usuarios; por lo tanto, esta utilización de los sistemas combinados es significativamente mayor que la suma de sus partes.

Grid ofrece nuevas y más potentes vías de trabajo, como los siguientes ejemplos:

- Portales científicos: permite aprovechar los métodos científicos de resolución de problemas.
- Computación distribuida: permite aprovechar la mayor capacidad que tienen las estaciones de trabajo para conseguir unos sustanciales recursos de computación.

- Computación en tiempo real de instrumentación: permite mejorar la utilización de aparatos en tiempo real.
- Trabajo en colaboración: permite trabajar en equipo compartiendo recursos, pero también los resultados de los diferentes estudios para su análisis.

5. Almacenamiento

El disco duro es el componente que almacena la información. Es crítico porque, además de contener toda la información de la organización, es el dispositivo que da más sentido a todo el concepto de las redes. Sin los discos duros, toda la expansión de las redes prácticamente no tendría sentido, dado que casi todas las peticiones que se hacen a servidores son directa o indirectamente peticiones al disco.

La capacidad y velocidad de los discos son los dos aspectos básicos y más importantes a tener en cuenta a la hora de escoger los discos que se quieren poner en los servidores.

¿Cuántos discos tiene que tener nuestro servidor? ¿Para qué los queremos?

Un disco es un espacio para guardar información que se divide en partes llamadas particiones.

Particiones

Normalmente, no se recomienda más de tres o cuatro particiones en un disco.

Si las particiones pueden ser de muchos Gb, ¿de qué sirve *particionar*?

Particionar un disco tiene dos utilidades básicas. La primera, y más importante, es que divide el disco en zonas independientes. Al estar formateada independientemente, cada partición del disco es un disco lógico (no físico) diferente para el SO. Por lo tanto, en caso de que por algún problema el sistema de ficheros quede corrompido y la información de dentro sea inaccesible, el contenido se pierde y la partición se tiene que reformatear. El resto de particiones son accesibles y la información se mantiene intacta. Incluso se puede recuperar toda la partición de la copia de seguridad.

La otra utilidad es que, al ser independientes, pueden estar formateadas en sistemas de ficheros diferentes. Por lo tanto, incluso podemos iniciar el ordenador desde diferentes particiones, a partir de sistemas operativos diferentes. Se utiliza mucho en la preparación de máquinas.

Tened presente que, si falla el disco físico, todas las particiones quedan inaccesibles y no se puede acceder a la información que contienen.

5.1. Necesidades de la organización

Las necesidades básicas de la organización, a grandes rasgos, son las siguientes:

- **Sistema.** La partición de sistema es necesaria para arrancar el servidor y para que funcione. Siempre se deja una partición sólo por el sistema operativo del servidor.
- **Usuarios.** La partición de usuarios contiene los directorios de los usuarios (las carpetas personales y si hay carpetas de grupo).

- **Datos.** En la partición de datos, normalmente hay directorios con datos de programas que tienen que estar instalados localmente en las estaciones de trabajo, datos compartidos por grupos de usuarios, y también puede haber un lugar para poner el “disco común”, que es una carpeta común a toda la organización para transferir cosas.
- **Aplicaciones básicas.** Las aplicaciones que utilizan todos los usuarios. El software base al cual necesitan acceder todos los usuarios y que tiene que estar en la red. El permiso tiene que ser de lectura y ejecución para todo el mundo.
- **Aplicaciones.** Esta partición contiene las aplicaciones que no son comunes a todo el mundo; por eso están separadas. Hay personas que las utilizan y otras que no. Se aplican permisos para grupos de usuarios. Además de las aplicaciones, muy posiblemente encontraremos datos asociados a las aplicaciones que funcionan.
- **Otros.** Teniendo en cuenta las necesidades reales de la organización, pueden hacer falta otras particiones: servidores de bases de datos, particiones por desarrollo, etc.

Gestión informática

Esta partición sólo tiene que ser vista por el departamento de informática. Contiene el software, las herramientas, las preinstalaciones, etc. necesarios para que el departamento pueda llevar a cabo su tarea y hacer funcionar todo el sistema.

Por ejemplo, se puede utilizar para ir a una estación de trabajo y reinstalar un software local (ofimática) sin llevar CD-ROM ni nada, sólo accediendo a esta partición del disco con los derechos adecuados. Los usuarios tienen que desconocer la existencia de esta partición.

La estructura final de todos los discos y las particiones está condicionada por la necesidad, y no tienen que ser forzosamente particiones. Siempre es la organización quien determina cómo se distribuye.

Las necesidades de la organización determinan las particiones que hacen falta.

5.2. Direct Attached Storage (DAS)

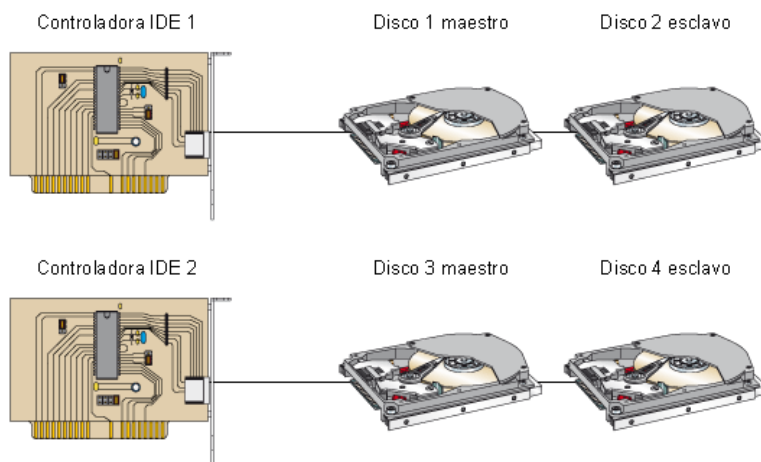
Las tecnologías “tradicionales” de almacenamiento se basan en la conexión directa (física) del dispositivo al servidor o estación de trabajo. Como consecuencia, las aplicaciones y los usuarios hacen las peticiones directamente al sistema de ficheros. Fundamentalmente, hay cuatro tipos de dispositivos:

- 1) *Intelligent Drive Electronics* (IDE).
- 2) *Serial ATA* (SATA).
- 3) *Small Computer System Interface* (SCSI).
- 4) *Serial Attached SCSI* (SAS).

5.2.1. Discos *Intelligent Drive Electronics*

Los discos *Intelligent Drive Electronics* (IDE) pueden ser de gran capacidad y son baratos, con el inconveniente de que no son muy rápidos. Esto hace que, para determinadas situaciones, no sea la mejor opción en el servidor. Además, las arquitecturas PC sólo permiten normalmente hasta cuatro discos IDE, porque tienen dos controladoras con capacidad de dos discos cada uno.

Discos soportados por las controladoras IDE



Si hay algún DVD, necesita uno de estos cuatro lugares, por lo cual nos quedamos con sólo tres espacios disponibles para discos.

Los IDE⁽²⁰⁾ se configuran de manera que uno de los dos discos de la controladora es el maestro, y el otro, el esclavo. Esto se hace modificando unos puentes⁽²¹⁾ del disco, antes de instalarlo físicamente dentro del ordenador. Normalmente, un ordenador con una arquitectura PC se pone en marcha a partir del maestro del IDE 1 (disco 1, que, por lo tanto, siempre tiene que estar), siempre que no haya controladores SCSI⁽²²⁾; entonces, se configura en la BIOS.

⁽²⁰⁾Recordad que “IDE” es la sigla de *Intelligent Drive Electronics*.

⁽²¹⁾Muchas veces, para referirnos a los puentes, se utiliza el término inglés *jumper*.

⁽²²⁾Recordad que “SCSI” es la abreviatura de *Small Computer System Interface*.

El estándar IDE surgió en 1981, con una velocidad de transferencia de 4 Mb/s aproximadamente. En la actualidad, con todos los cambios tecnológicos y modificaciones, se ha llegado a ATA/ATAPI 5 (1999), con una velocidad de transferencia de 66 Mb/s aproximadamente. A pesar de estos avances, tiene bastantes limitaciones, como por ejemplo, que mientras se utiliza el maestro del IDE 1 no se puede utilizar el esclavo del IDE 1 (no se pueden hacer lecturas paralelas en la misma controladora).



Vista posterior de un disco IDE

5.2.2. Discos *Serial ATA*

Los discos *Serial ATA* (SATA), con sus orígenes situados en torno al 2000, es la transición natural de los discos ATA o también llamados IDE.

El acceso a los discos se realiza en serie, sustituyendo al acceso en paralelo de los discos P-ATA (IDE). Este nuevo método de acceso proporciona mejoras con respecto a su antecesor:

- **Incrementa la velocidad de transmisión.** En una primera versión, esta velocidad se situó en 1,5 Gb/s (SATA 150),²³ pero actualmente se trabaja con una velocidad de transmisión de 3 Gb/s (SATA 300). Se está trabajando para conseguir 6 Gb/s en un futuro próximo.
- **Se incrementa la longitud del cable de transmisión.** La longitud soportada actualmente es de 2 metros.
- **Incrementa el número de dispositivos SATA conectados.** Para una estructura PC, se puede llegar hasta 16 dispositivos SATA conectados a cada controlador (situado normalmente en la placa madre).
- **Permite la conexión de discos “en caliente”.** Permite añadir discos a la configuración, mientras el sistema está funcionando con normalidad.

(23) Recordad que “SATA” es la sigla de *Serial ATA*.

External SATA

La tecnología SATA ha permitido la creación de una pequeña variante: *external SATA* (eSATA) que facilita la conexión de dispositivos externos.

5.2.3. Discos *Small Computer System Interface*

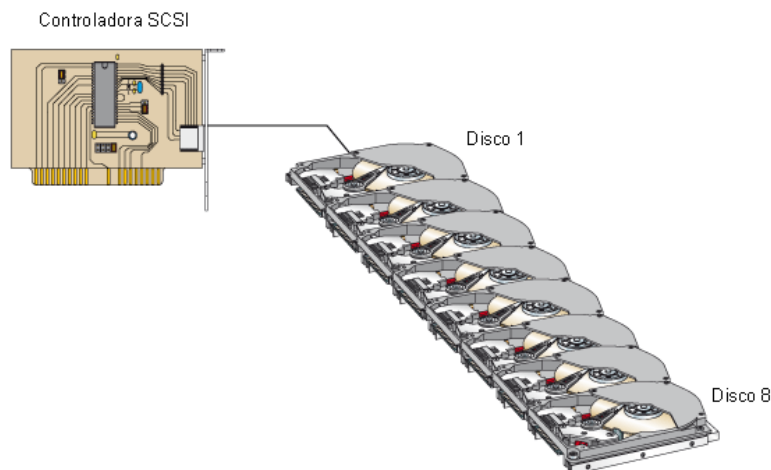
Small Computer System Interface (SCSI) es un tipo de disco que es mucho más rápido, pero también más caro. Es mucho más adecuado para servidores. Hay revisiones de SCSI, como por ejemplo la llamada Ultra Wide SCSI 2.

El SCSI soporta hasta 8 dispositivos por controladora (el SCSI II hasta 16).

Dispositivos en una cadena SCSI

Cuando se habla de dispositivos en una cadena SCSI, se hace referencia a discos duros, unidades de DVD, unidades de cinta, etc.

Cada controladora SCSI soporta 8 discos



La velocidad de transferencia de las cadenas SCSI ha variado bastante con las revisiones, desde el SCSI original, que iba a 5 Mb/s, hasta la última revisión, el Ultra 640 SCSI, que tiene una velocidad de transferencia de 640 Mb/s. Los buses SCSI permiten lecturas y/o escrituras simultáneas en la misma controladora y es, por lo tanto, el estándar que se utiliza en servidores corporativos.

En servidores de arquitectura Unix, también suele ser un estándar instalar discos duros SCSI, mientras que en arquitecturas PC se tiene que tener la precaución de hacerlo, ya que hay que añadir un componente de hardware (una placa controladora) para poder soportarlo.

5.2.4. Discos *Serial Attached SCSI*

Igualmente, como sucede con los discos SATA, los discos *Serial Attached SCSI* (SAS) son la evolución de los discos SCSI.

También, como sucedía con SATA, SAS⁽²⁴⁾ implementa la transmisión en serie entre el controlador y los dispositivos, lo que permite obtener significativas mejoras:

(24) Recordad que "SAS" es la sigla de *Serial Attached SCSI*.

- **Incrementa la velocidad de transmisión.** En su primera versión, esta velocidad se ha situado en 3 Gb/s (SAS 3.0). Se está trabajando para conseguir 6 Gb/s en un futuro próximo.
- **Se incrementa la longitud del cable de transmisión.** La longitud externa soportada actualmente es de 8 metros.
- **Incrementa el número de dispositivos SAS conectados.** Para una estructura PC, se puede llegar hasta 128 puertos de expansión de dispositivos y hasta 16.384 discos SAS conectados.

- **Conexión de discos “en caliente”.** Permite añadir discos a la configuración, mientras el sistema está funcionando con normalidad.

Una buena política y gestión de los discos puede determinar el rendimiento del servidor.

5.2.5. Agrupaciones de discos en el servidor

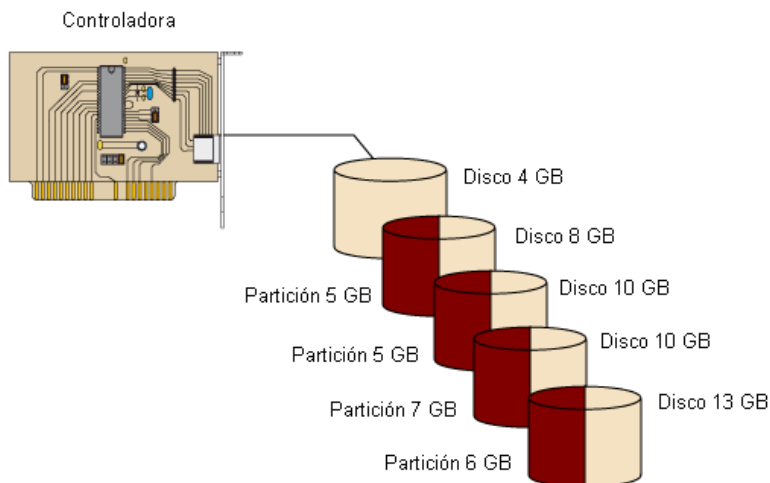
Los discos son siempre una de las piezas clave en los servidores. Esto ha provocado intentos tecnológicos para mejorar la capacidad y el rendimiento.

Multivolumen

¿Qué pasa si creemos que tendremos una base de datos que ocupará 18 Gb y sólo tenemos dos discos de 12 Gb? Hay una solución, por medio del SO, que consiste en convertir los dos discos de 12 Gb en uno de 24 Gb. Es la gestión multivolumen.

En general, la gestión multivolumen trata de juntar diversas particiones físicas en una sola partición lógica de un tamaño equivalente a la suma de los tamaños de las particiones.

Esquema de gestión multivolumen



En el esquema anterior la partición lógica total es $5 + 5 + 7 + 6 = 23$ Gb.

La principal ventaja es que se puede obtener una partición del tamaño que se quiera, juntando particiones de discos de otros tamaños. A menudo, las bases de datos necesitan discos muy grandes.

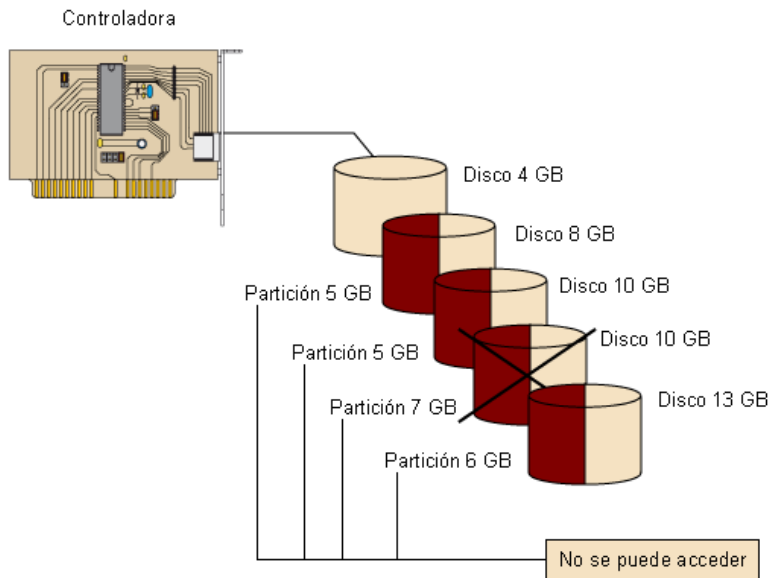
¿SATA o SAS?

Una pregunta natural que se nos plantea a continuación sobre los discos SAS/SATA es: ¿cuándo se tienen que utilizar discos SATA y cuándo se tienen que utilizar discos SAS?

La respuesta está basada en las inherentes diferencias entre los discos SCSI y los discos ATA hoy en día. Los discos SCSI han sido diseñados y fabricados para cumplir con los requisitos empresariales de alta disponibilidad y seguridad. Esta es la característica que nos tiene que llevar a una decisión.

El principal inconveniente es que si un disco falla físicamente (se estropea), no podremos acceder a ninguna de las particiones físicas que integran la partición multivolumen creada.

Si un disco falla, no podremos acceder a la partición multivolumen



Este problema, junto con el gran aumento de capacidad de los discos (y su notable reducción de precios), hace que la solución que se adopta sea comprar e instalar discos de la capacidad que se necesita, ya que si hay que hacer un multivolumen es síntoma de que la información que tiene que contener es crítica, y el gasto de un disco nuevo es pequeño en comparación con la seguridad que gana el servidor.

RAID

Ante los problemas que genera la gestión multivolumen, está la solución *Redundant Array of Inexpensive Disks* (RAID). El RAID permite, por ejemplo, tener cinco discos funcionando y sólo aprovechar cuatro, pero si falla cualquiera de estos cuatro el servidor continúa funcionando, porque el quinto contiene información redundante que lo permite.

Incluso permite cambiar el disco en caliente, es decir, sin cesar el servidor se puede sustituir el disco que ha dejado de funcionar por uno nuevo, que el propio RAID²⁵ vuelve a poner en funcionamiento.

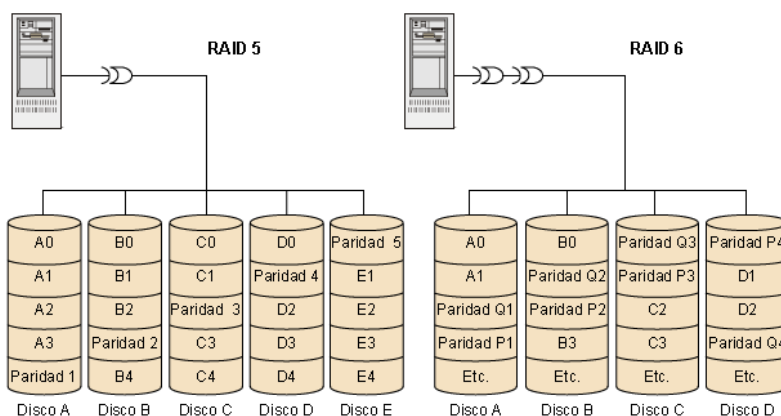
⁽²⁵⁾ "RAID" es la sigla de *Redundant Array of Inexpensive Disks*.

El RAID se puede hacer de diversas maneras, según el grado de velocidad y seguridad que se necesite. Se clasifican en niveles:

- **RAID 0.** La información se distribuye en diversas unidades, pero no hay redundancia. Por lo tanto, no hay protección en caso de fallo de disco.

- **RAID 1.** También llamado **espejo**. Cada unidad está duplicada con una unidad de apoyo. Por lo tanto, con seis unidades de disco, tres son de copia. La información se distribuye entre los discos.
- **RAID 2.** Hay distribución de datos con respecto a los bits sobre todas las unidades. No se utiliza porque el RAID de nivel 3 está mucho más extendido.
- **RAID 3.** Datos distribuidos, en el ámbito del bit (o del byte), en todas las unidades menos en una, que es la de paridad. Tiene muy buen rendimiento de lectura, pero en escritura cada vez se tiene que actualizar la unidad de paridad.
- **RAID 4.** Como el de nivel 3, pero todo se hace en el sector. Mejoran los tiempos de acceso.
- **RAID 5.** Se escriben en todos los sectores de todas las unidades, y se añaden códigos correctores a cada sector. Este nivel de RAID ofrece una escritura más rápida, porque la información de redundancia se distribuye en todas las unidades. Las lecturas en disco también tienen unos tiempos de acceso muy buenos.
- **RAID 6.** Este nivel de RAID es similar al 5, pero utilizando dos códigos correctores para cada sector y grupo de RAID. Las informaciones de paridad se distribuyen entre todos los discos del grupo.

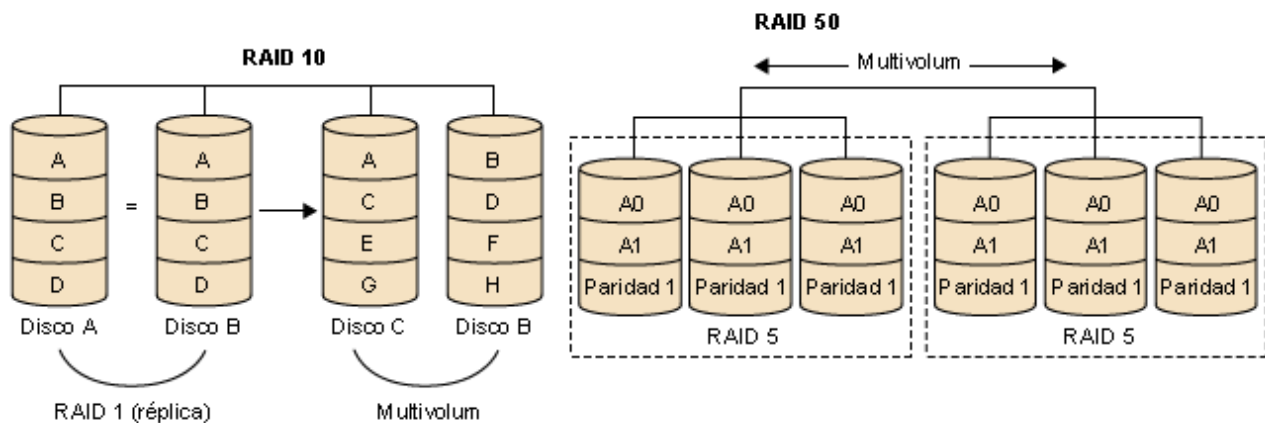
Esquema de los RAID 5 y 6



- **RAID 5e, 6e.** Estos dos niveles de paridad se basan en sus predecesores (5 y 6), pero añaden un elemento más de seguridad: **hot spare**. Éste es un nuevo disco en espera que entra a formar parte del grupo de RAID activamente si uno de los discos del mismo deja el grupo.

- **RAID 10.** Aparecen diversas combinaciones de niveles de seguridad, a partir de los niveles básicos comentados. Uno de ellos es el nivel 10 = 1 + 0. El cual replicaría un grupo de RAID 1 en un grupo de discos con RAID 0.
- **RAID 50.** Un grupo de nivel 50 = 5 + 0 distribuiría la información por multivolumen entre dos grupos de RAID 5.

Esquema de los RAID 10 y 50



Hay otras combinaciones posibles y otros niveles de RAID no estándares que proponen las compañías a sus clientes.

La técnica del RAID mejora el rendimiento, al distribuir la información entre diversas unidades, y puede ofrecer redundancia para aumentar la seguridad.

Una vez más, el RAID puede ser por software o por hardware. Si es por software es más lento, y si es por hardware es transparente en el SO.

5.2.6. Sistemas de ficheros

Una vez hemos hecho las particiones sobre los discos, necesitamos hacer una operación para que nuestro sistema pueda trabajar. Se tiene que formatear la partición. Este paso es imprescindible, porque informa al sistema operativo y al disco de cómo se reparte el espacio (tamaño del sector), de cómo se distribuirá lógicamente dentro del disco, y también se hacen operaciones para mejorar el rendimiento, aunque son dependientes del mismo sistema operativo. Por lo tanto, el resultado de este formateo genera el sistema de ficheros, ya vacío y preparado para poder poner información. El sistema de ficheros que se escoja para formatear la partición también es muy importante porque tiene diversas características asociadas, como por ejemplo cuestiones de seguridad.

RAID comerciales

Hay una gran cantidad de sistemas de RAID comerciales internos y externos, pero citamos algunos fabricantes que se pueden encontrar en la web: Dell (PowerVault), Compaq, StorageTek, Clarion, Hewlett Packard, IBM, Raid-Tec, etc.

File Allocation Table

El sistema de ficheros *File Allocation Table* (FAT) es bastante antiguo. Sólo soporta tamaños de hasta 2 Gb y no hay seguridad. Es el sistema de ficheros de los PC de sobremesa de antes.

El tamaño del clúster⁽²⁶⁾ (el número de sectores que guarda de golpe) puede ser muy grande, y los sistemas operativos tienen una gran cantidad de ficheros pequeños, por lo cual se pierde mucho espacio.

(26) Un clúster es un grupo de sectores.

FAT32

El sistema de ficheros FAT32 es el mismo que el FAT⁽²⁷⁾, pero soporta tamaños de disco superiores sin ningún problema. El tamaño del clúster es mucho más pequeño, por lo cual se aprovecha mucho mejor el espacio del disco. Tampoco tiene seguridad.

(27) Recordad que "FAT" quiere decir *File Allocation Table*.

Sistema de ficheros NT

El sistema de ficheros NT⁽²⁸⁾ (NTFS) se introdujo con la aparición de Windows NT. Tiene un tamaño de sector y de clúster muy pequeño, de manera que se aprovecha muy bien el espacio de disco. Lleva firma de la partición, por lo cual el disco no se puede leer en otro ordenador. Lleva seguridad en el sistema de ficheros, por lo cual los permisos están en el sistema de ficheros. Todo el conjunto, pues, hace que sea mucho más robusto.

(28) El sistema de ficheros NT se expresa en inglés como *NT File System*.

Sistemas de ficheros ufs, ext2 y ext3

Los sistemas de ficheros ufs, ext2 y ext3 son tres sistemas de ficheros que se utilizan mucho en sistemas Unix y GNU/Linux. El tamaño de sector es de 256 bytes (muy pequeño). Tiene una estructura de *inodes* para gestionar los ficheros y la seguridad de Unix Standard.

El sistema de ficheros **ext3** nos ofrece la posibilidad de trabajar con *journaling*, sistema mediante el cual se salvan periódicamente los archivos abiertos con el fin de evitar la pérdida de información o la corrupción de los datos si se produce una desconexión no planificada. Este sistema de ficheros aporta más seguridad, aunque por contra hace perder recursos de máquina, asignados precisamente a la tarea de *journaling*.

High Sierra File System

El sistema de ficheros *High Sierra File System* (HSFS) o ISO9660 es muy conocido. También es el formato de los CD-ROM/DVD. Todos siguen este formato, tanto los de datos como los de audio.

Sistemas de ficheros distribuidos

Un sistema de ficheros distribuidos permite almacenar ficheros en uno o más ordenadores (servidores), y permite que se hagan accesibles a otros, llamados clientes. Estos últimos pueden gestionar y manipular los ficheros como si fueran locales.

Normalmente, los sistemas de ficheros distribuidos incluyen herramientas automáticas de replicación y de tolerancia a errores. Totalmente transparente al usuario, el sistema es capaz de replicar los datos entre los servidores y dar servicio desde otro servidor en caso de fallo.

Hay diversas ventajas a tener en cuenta de los sistemas de ficheros distribuidos:

- **Ficheros fácilmente accesibles.** Actualmente, se puede acceder a un fichero desde cualquier punto del planeta. Sólo hace falta un ordenador conectado a la red y un navegador.
- **Compartimentación de ficheros.** Facilita claramente el trabajo en grupo y la interacción entre usuarios.
- **Simplificación de las copias de seguridad.** Simplifica la copia de seguridad de los ficheros, centrando la acción en los servidores y no en los clientes. Se evita en cierta manera la dispersión de la información.
- **Gran capacidad de almacenamiento.** Los servidores de ficheros proporcionan grandes volúmenes de espacio para almacenar ficheros, reduciendo el coste que supondría replicar el espacio en los ordenadores clientes.
- **Simplificación de la administración.** Desde el punto de vista del administrador, este sistema simplifica mucho la tarea, al tener centralizada toda la información.

Otros sistemas de ficheros

Existen muchos otros formatos y sistemas de ficheros. Así, haciendo un vistazo nostálgico al pasado, los disquetes, por ejemplo, también tienen un formato de sistemas de fichero. En contraposición, mirando al futuro, hay sistemas de ficheros distribuidos (The Google File System, WUALA, CODA, y muchos otros) que han hecho cambiar la visión de los sistemas de ficheros.

5.3. Storage Area Network y Network Attached Storage

Las redes SAN²⁹ y los servidores de ficheros NAS³⁰ son agrupaciones mayores de discos que la de los RAID.

⁽²⁹⁾Recordad que “SAN” es la sigla de Storage Area Network.

5.3.1. Storage Area Network

⁽³⁰⁾Recordad que “NAS” es la sigla de Network Attached Storage.

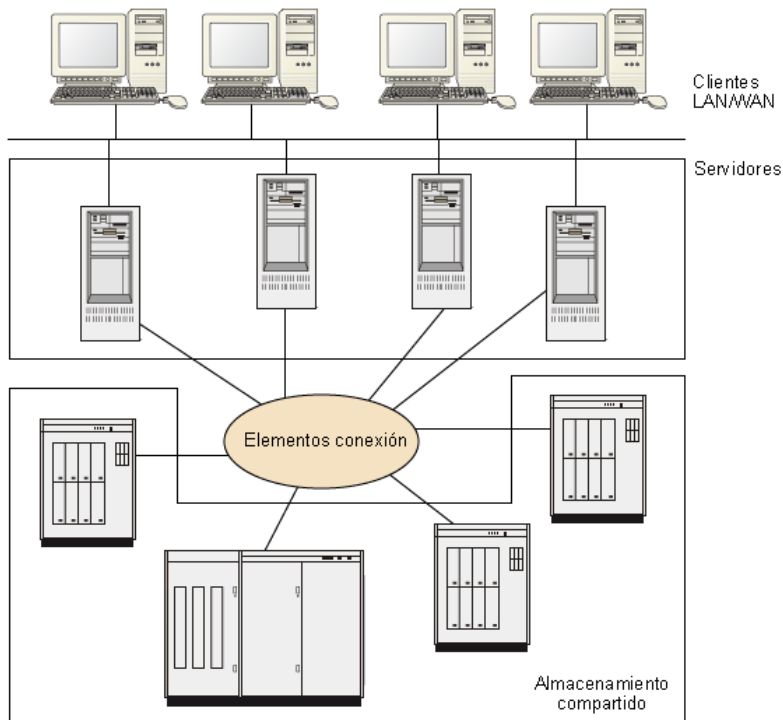
La asociación Storage Network Industry Association (SNIA) define *Storage Area Network* (SAN) como la red que tiene como objetivo principal la transferencia de datos entre sistemas o computadores y elementos de almacenamiento.

Otra definición más sencilla de SAN es: red especializada de alta velocidad que comunica servidores y dispositivos de almacenamiento. Una SAN también puede ser un sistema de almacenamiento formado por elementos y dispositivos de almacenamiento, computadores, aplicaciones, software de control, y todos estos elementos, comunicándose mediante una red.

Accesos a disco en una SAN

Los accesos a disco en una red SAN son normalmente (aunque no siempre) en el *Block I/O*.

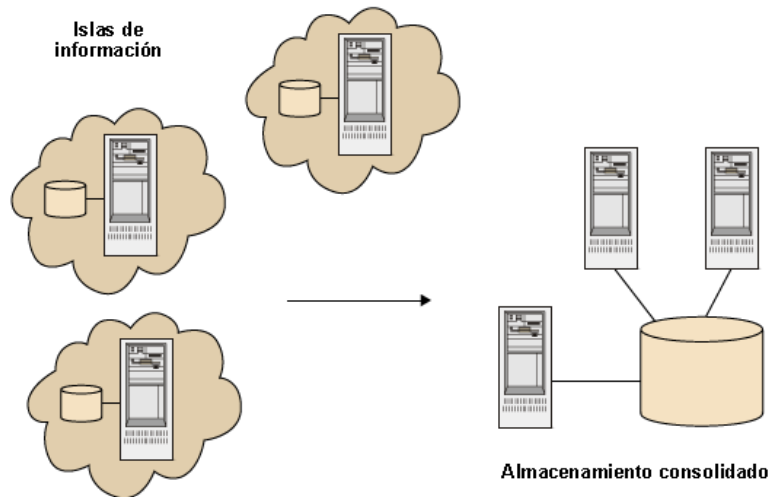
Esquema de SAN



Antes de hablar de los componentes, protocolos e implementaciones de las redes SAN, es bueno que definamos el porqué de su utilización. ¿Qué motivaciones tenemos para optar a esta tecnología?

La principal razón es que las organizaciones tienen cada vez un mayor número de servidores, y éstos a su vez necesitan constantes aumentos de capacidad de almacenamiento. Para evitar **islas de información**, donde cada servidor controla su almacenamiento independientemente del resto, hay que utilizar una nueva técnica que nos permita la compartimentación global de los recursos de almacenamiento.

Las redes SAN evitan las islas de información.



Si revisamos la historia, en los años setenta y principios de los ochenta, los sistemas *host* definían un modelo centralizado, en el que los datos residían internamente. La evolución nos lleva a los años ochenta y noventa a sistemas con datos distribuidos gracias al modelo cliente/servidor. El futuro nos lleva a un nuevo sistema de compartimentación global de recursos de almacenamiento por red. Es una nueva era, la SAN.

Elementos de una red SAN

Los elementos de una red SAN se pueden dividir en tres grandes grupos:

1) **Servidores.** Forman parte de una red SAN todos aquellos servidores que disponen de tarjetas específicas con el fin de establecer comunicación con los elementos de conexión.

2) **Elementos de conexión.** Forman parte de este grupo:

a) **Cableado:** específico para las redes SAN, suele ser cable de fibra óptica. Hay dos tipos, cableado multimodo de fibra de 50 micrones para distancias cortas, y monomodo para distancias largas (menos de 10 micrones).

b) **Conmutador:** conmutadores especializados en comunicación en redes SAN.

c) **Directores:** conmutador principal. Punto central de gobierno de las redes SAN.

d) **Concentradores:** hacen la misma función que los concentradores de redes, pero especializados para redes SAN.

e) **Encaminadores:** encaminadores especializados pueden convertir señales entre protocolos de SAN.

Ved también

Ved más información sobre los directores en el módulo "Administración de la red".

3) Almacenamiento:

- a) Sistemas de discos: dispositivos especializados en servir almacenamiento de disco.
- b) Sistemas de cintas, básicamente bibliotecas de cinta como elementos de gestión de gran volumen de datos para *backup* y dispositivos de cinta.

Conectividad SAN

Comprende todas las clases de hardware, software y componentes que permiten la interconexión de los dispositivos de almacenamiento y los servidores.

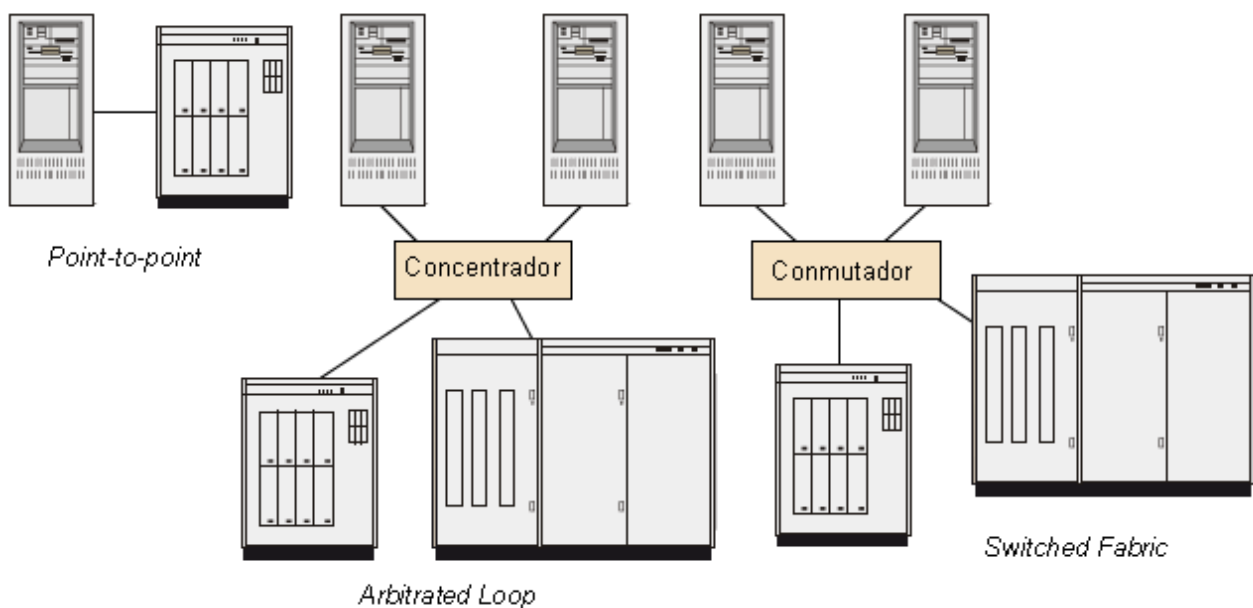
1) **Capa baja.** La capa baja comprende las capas de conexión física y redes. Tenemos diferentes posibilidades:

- *Ethernet*. Se construye una topología típicamente de bus que puede llegar a una velocidad de 10 Gbps de datos.
- *SCSI*. Es una interfaz en paralelo que puede trabajar hasta 25 metros a una velocidad de 160 Mbps.
- *Fibre Channel (FC)*. Es una interfaz en serie, normalmente implementada con cable de fibra óptica, que actualmente es la arquitectura más utilizada por las redes SAN. Mejora la velocidad y la distancia de SCSI. Sus posibles topologías son: *Point-to-point*, *Arbitrated Loop*, *Switched Fabric*.

Ved también

Revisad la estructura de capas de la torre OSI.

Esquema de topologías *Fibre Channel*



2) **Capa media.** La capa media comprende el protocolo de transporte y las capas de sesión.

- *Fibre Channel Protocol* (FCP). Es el protocolo de transporte para SCSI en *fibre channel*. Es una tecnología de Gigabit principalmente utilizada para redes de almacenamiento. Las señales de *fibre channel* pueden enviarse tanto en par trenzado de cobre como en cables de fibra óptica.
- iSCSI. Es un protocolo de transporte de datos que transporta los pedidos SCSI requeridos mediante la tecnología estándar de redes (TCP/IP).
- *Fibre channel* por IP (FCIP). También es conocida como *FC tunneling*. Método que permite la transmisión de FC³¹ mediante redes IP.
- Internet FCP (iFCP). Mecanismo que permite enviar datos en dispositivos de almacenamiento de una SAN mediante TCP/IP vía Internet.

⁽³¹⁾Recordad que “FC” es la sigla de *Fibre Channel*.

Protocolos propietarios

Aparte de los protocolos FCP³², iSCSI, FCIP³³, iFCP³⁴, hay otros protocolos propietarios como ESCON o FICON.

⁽³²⁾Recordad que “FCP” es la sigla de *Fibre Channel Protocol*.

⁽³³⁾Recordad que “FCIP” es la sigla de *Fibre Channel* por IP.

⁽³⁴⁾Recordad que “iFCP” es la sigla de Internet FCP.

3) **Capa alta.** La capa alta comprende las capas de presentación y aplicación:

- *Server Attached Storage*. Inicialmente, el almacenamiento era compartido directamente por el bus del servidor, utilizando una tarjeta de comunicaciones adecuada y el dispositivo de almacenamiento estaba dedicado a un solo servidor. El servidor controlaba las E/S hacia el dispositivo. Actualmente, los dispositivos de almacenamiento disponen de una inteligencia que les permite realizar acciones como la gestión de grupos de RAID, disponibilidad de memoria Cache E/S, control de unidades, etc.
- *Network Attached Storage*.

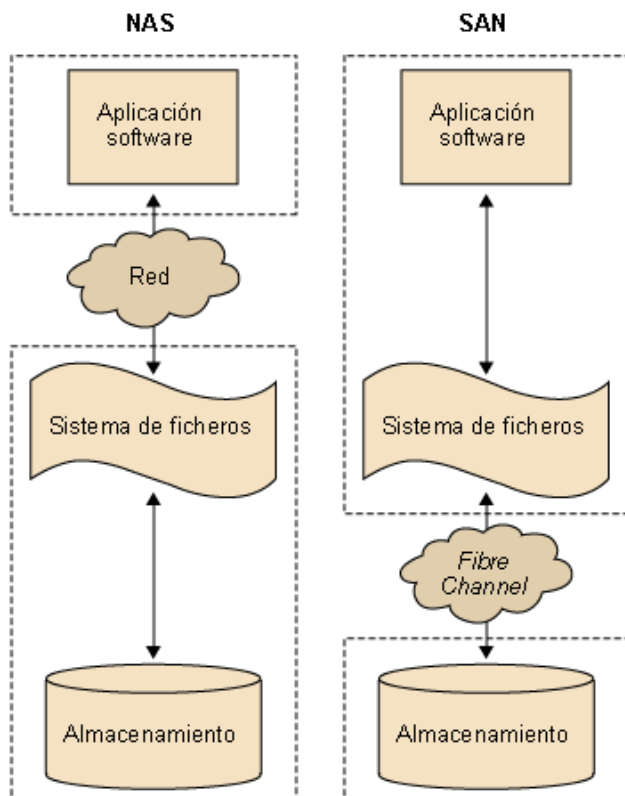
5.3.2. Network Attached Storage

Network Attached Storage (NAS) es básicamente un servidor de ficheros conectado a una red que sirve ficheros utilizando un protocolo. Un elemento NAS consiste en una máquina que implementa los servicios de ficheros (utilizando protocolos de acceso, como por ejemplo NFS o CIFS), y uno o más dispositivos, donde los datos son almacenados.

NAS proporciona capacidad de almacenamiento utilizando la misma red de comunicaciones o una adicional de bajo coste.

Así como los accesos que se realizan a una SAN son la mayoría de veces en el *Block I/O*, en una NAS se realizan en el sistema de ficheros. Es decir, las aplicaciones acceden al sistema de ficheros que proporciona el propio dispositivo NAS, mientras que, en una SAN, el sistema de ficheros pertenece al propio servidor.

Comparación entre una SAN y una NAS

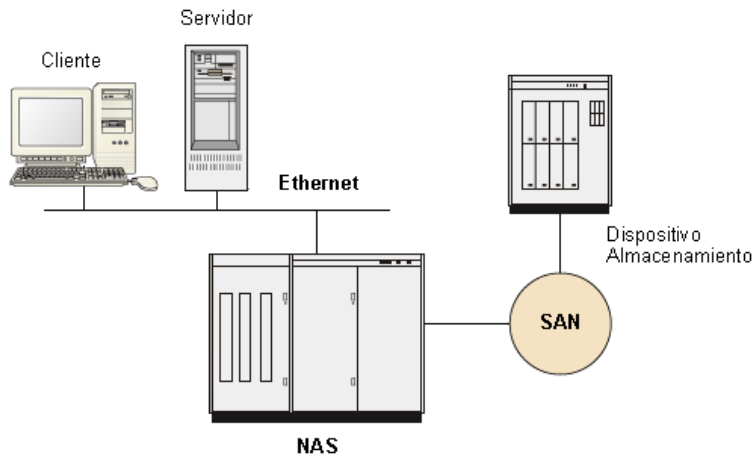


A menudo se tiene que tomar una decisión de diseño y escoger entre una SAN o una NAS. Esta decisión tiene que venir reforzada por dos premisas básicas: velocidad de comunicación (rendimiento de acceso a disco) y sistema de acceso (acceso al bloque o a ficheros).

Soluciones híbridas

Los sistemas NAS y SAN no son excluyentes; al contrario, pueden ser combinados dando todavía mayor flexibilidad y servicio de almacenamiento a los servidores. Un ejemplo sería tener un servidor conectado a una NAS, la cual tiene los discos en un sistema de almacenamiento conectado a una SAN.

Combinación de NAS y SAN



6. Copia de seguridad

Ante el problema de copiar la información de la organización para evitar pérdidas, hay muchos dispositivos (aparatos físicos) y técnicas. Tenemos que buscar los mejores para cada caso.

Los dispositivos de copia de seguridad⁽³⁵⁾ son los aparatos físicos que se utilizan para hacer copias de seguridad de la información de los servidores. Normalmente, las copias son procedimientos que tardan horas en terminarse, y también se tarda una hora o más en recuperar los ficheros del dispositivo en los que se ha almacenado la información.

⁽³⁵⁾Copia de seguridad en inglés se expresa como *backup*.

6.1. Dispositivos de copia de seguridad

Para hacer copias de seguridad hay disponibles diversos dispositivos: cintas, librerías de copia, grabadoras DVD y discos duros.

6.1.1. *Digital Audio Tape*

Las cintas son el dispositivo más habitual que hay en los servidores. Los *Digital Audio Tape* (DAT) son muy habituales, generalmente SCSI, y hay de varias capacidades, que pueden llegar hasta unos 20 GB por cinta.

Es normal que un servidor tenga una unidad de estas características y que diariamente se hagan copias de seguridad siguiendo alguna política de copia.

Capacidad de las unidades de cinta

Cuando se habla de capacidades de las unidades de cinta, hay con compresión o sin ella. Nosotros hablamos siempre de las que son sin compresión. Con compresión, la capacidad se puede duplicar o más.

6.1.2. *Digital Lineal Tape*

Otro tipo de cintas, las *Digital Lineal Tape* (DLT), también son generalmente SCSI, y las hay de diferentes capacidades, que pueden ir de los 20 GB hasta los 100 GB por cinta (sin compresión), utilizando Super DLT (SDLT).

6.1.3. *Advanced Intelligent Tape*

Un tercer tipo de cintas, las *Advanced Intelligent Tape* (AIT), también son generalmente SCSI y las hay de diversas capacidades, entre 25 GB y 100 GB por cinta (con AIT3). La variación de esta capacidad depende de la cinta, del tipo de AIT que se utilice y del nivel de compresión con que se hagan las copias.

6.1.4. Linear Tape Open

Un nuevo tipo de cintas, las *Linear Tape Open* (LTO), son una nueva tecnología desarrollada por Hewlett Packard, IBM y Seagate.

Estos tipos de cintas han ido evolucionando rápidamente. Mientras que en el año 2000 hablábamos de LTO 1, que permitía hasta 100 GB de copia por cinta, actualmente la capacidad de las cintas LTO4 llega hasta 800 GB sin compresión (1,6 TB con compresión).

Su velocidad de copia puede llegar a 120 Mb/s, y ya están planificadas las versiones LTO5 y LTO6 que permitirán almacenar hasta 3,2 TB a una velocidad de 270 Mb/s.

6.1.5. Librerías de copia

Se puede dar el caso de que nuestra organización manipule cantidades de datos que ocupen diversas cintas de copia al día. En este caso, una sola persona se pasaría el día haciendo copias de seguridad, y no acabaría nunca. ¿Cuál es la solución para estos volúmenes de información tan grandes? Hay unos dispositivos llamados librerías de copia. Son externos, con unos brazos articulados, y contienen desde 20 hasta 2.000 cintas de copia de seguridad (son como robots). Con el software adecuado, esto se ve, por ejemplo, como una unidad de 400 TB para guardar información. El software sabe en qué cinta está almacenada la copia, qué cintas están llenas, y maneja la política de sustitución de cintas. Las librerías de copia sólo tienen sentido para organizaciones de grandes dimensiones o bien que manejen cantidades de información muy grandes.

6.1.6. Grabadora DVD

Una unidad de cinta es un componente caro, y muchas veces resulta difícil acceder (pide un tiempo considerable) a los datos que contienen. Por eso, a veces se considera la grabadora de DVD como una opción de copias de seguridad. Permite hacer copias de los elementos siguientes:

- DVD grabables una vez de 4,7 GB.
- DVD grabables muchas veces de 4,7 GB.
- DVD grabables una vez de 9,4 GB.
- DVD grabables muchas veces de 9,4 GB.

Aunque estos son los volúmenes más estándar, podemos encontrar volúmenes de DVD grabables una vez de hasta 17 GB.

Actualmente, los DVD grabables una vez son una opción que hay que tener en cuenta al plantearse las copias de seguridad, ya que tienen las ventajas siguientes:

Otras cintas

Hay otras cintas, como por ejemplo el Hexabyte, pero las DAT (*digital audio tape*), DLT (*digital lineal tape*), AIT (*advanced intelligent tape*) y LTO (*linear tape open*) son las más extendidas.



Librerías de copia

Librerías de copia comerciales

Hay diferentes marcas que fabrican librerías en colaboración con marcas de software, para que puedan funcionar correctamente con los servidores en que se instalan. Algunas de estas marcas, con webs para poder ver los aparatos, son Qualstar, Adic, Hewlett Packard, StorageTek, Quantum (ATL), etc.

- El coste de compra del dispositivo es bajo.
- El coste de las unidades de copia es bajo.
- Son de gran capacidad (hasta 17 GB).
- Ofrecen una gran facilidad para acceder a la información guardada.

Que se utilice o no como mecanismo de copia depende siempre del tamaño de la organización, el volumen de datos, etc.

Tendencias

En poco tiempo, veremos cómo el Blue Ray sustituye al DVD como dispositivo de copia. El *Blu-Ray* es un disco del mismo tamaño que un DVD (12 cm) y con una capacidad de 25 GB por cara (50 GB en total a una velocidad de 36 Mbit/s). Existe ya el BD-R y el BD-RE (grabable y el regrabable) y hay prototipos para aumentar la velocidad (2X) y desarrollar un Blue Ray de 4 capas (100 GB por disco).

6.1.7. Disco duro

En sistemas críticos, y más teniendo en cuenta el coste y la capacidad actual de estos dispositivos, no se tiene que descartar nunca la posibilidad de hacer una copia de seguridad (o incluso de copiar toda la información) en otro disco duro sólo dedicado a esta función.

La estrategia es hacer una primera copia de seguridad en este disco duro (se puede hacer con un procedimiento automático y varias veces al día, si hace falta), y de este disco, posteriormente, se hará una copia de seguridad en otro dispositivo (que puede ser una cinta).

A veces, esta estrategia es necesaria si el procedimiento de copia necesita bloquear la información a la cual accede y es, por ejemplo, una gran base de datos de la cual depende toda la organización. La copia de disco en disco, al funcionar internamente, por los buses del sistema y con velocidades de transferencia muy elevadas, necesita bloquear muy poco tiempo la información para hacer la copia. Así pues, la interrupción para hacer esta tarea es prácticamente imperceptible.

6.1.8. ¿Dónde tienen que estar los dispositivos de copia?

El disco duro de que hablábamos, o los dispositivos de copia de seguridad (las unidades de cinta), ¿por qué han de estar en el mismo servidor?

Esta cuestión, que desde el punto de vista lógico (frente al físico) es perfectamente plausible, presenta en estos momentos graves inconvenientes tecnológicos. Desde el punto de vista de la red, podríamos tener la unidad de copia en cualquier sitio y transferir la información por la red. Sin duda funcionará, pero estos son algunos de los problemas que representa hacerlo:

Tendencias

Gracias a la proliferación de redes SAN o dispositivos NAS, que permiten una gran cantidad de espacio por almacenar, se utilizan cada vez más los discos como dispositivo de copia. Los propios proveedores ofrecen herramientas específicas que permiten hacer estas copias transparentes al propio sistema.

- Podemos colapsar la red, ya que si hacemos una copia de todo el servidor (la opción habitual) transferiremos por la red una cantidad de información del orden de diversas decenas de Gb.
- Toda la información del servidor (supuestamente segura) atraviesa la red, de manera que potencialmente está el peligro de que la vean personas ajenas al proceso (de dentro o de fuera de la organización). Un riesgo de seguridad.
- Para poder hacerlo, se tiene que hacer accesible por red (con permisos, contraseñas, etc.) todo el disco del servidor. Eso implica otro riesgo de seguridad, porque ni en el caso de que después de hacer la copia el disco deje de estar accesible, este disco estará allí unas cuantas horas al día (esperamos no olvidarnos ningún día de eliminar el acceso).
- El ancho de banda de la red y la de los buses de sistema (IDE, SCSI, bus interno, etc.) no son comparables, por lo que la velocidad de transferencia haría el tiempo de copia extraordinariamente superior.

Por lo tanto, lo más aconsejable es tener el dispositivo de copia en el servidor donde esté la información que se tiene que copiar.

Copias por red

¿Tenemos que descartar las copias por red? Rotundamente, no. Se utilizan para hacer copias de seguridad de información que hay en las estaciones de trabajo de los usuarios. Aunque, a lo largo de los materiales, se dice que se tiene que procurar que haya cuanto menos mejor, muy a menudo hay información en las estaciones de trabajo. En este caso, sin embargo, son del orden de algunas decenas de MB como mucho. No tardan en hacerse y no penalizan mucho la red. En cambio, sí que se tiene que ir con cuidado con las cuestiones de protección y seguridad.

6.2. Políticas de copia de seguridad

Una buena política de copias de seguridad es la clave para tener segura la información de la organización.

Algunos de los motivos para hacer copias de seguridad son los siguientes:

- Proteger la información contra un fallo del sistema o algún desastre natural.
- Proteger la información de los usuarios (los ficheros) contra borrados accidentales.
- Proteger la información de los usuarios y de la organización contra ataques por parte de terceros.

- Duplicar la información de los usuarios por seguridad, ya que se pueden dar casos de usos incorrectos que la dejen inconsistente o la modifiquen incorrectamente.
- Posibilitar un traspaso de la información cuando se actualiza o se reinstala el sistema.

6.2.1. Tipos de copias de seguridad

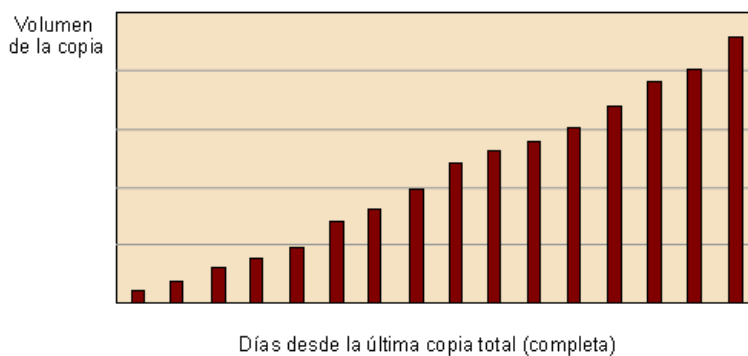
Podemos distinguir los siguientes tipos de copia de seguridad:

- **Copia de seguridad completa.** También se conoce con el nombre de *copia de seguridad total* o *copia hoja dump*. Se hace una copia de toda la partición del disco en cinta. A menudo, la copia se hace considerando el formato del disco y sin tener en cuenta el sistema de ficheros, ya que sólo hay que conocer la tabla de particiones del disco y en qué parte está la partición para duplicarla en un dispositivo de cinta. En estos casos, la restauración no puede ser selectiva; se tiene que restaurar toda la partición y no se puede seleccionar sólo un fichero. Está también la copia de seguridad completa del sistema de ficheros, en la que sí que es posible una restauración selectiva.
- **Copia de seguridad incremental.** En este caso, se guardan sólo los ficheros que se han modificado desde la última copia de seguridad que se ha hecho. Las copias de seguridad incrementales se utilizan conjuntamente con las copias de seguridad completas en lo que se llaman políticas de copias de seguridad.
- **Copia de seguridad selectiva.** También es posible hacer una copia de sólo unos ficheros determinados. Normalmente se lleva a cabo con ficheros de pedidos.
- **Copia de seguridad diferencial.** Este nuevo tipo de copia realiza una copia de todos los ficheros que se han modificado desde la última copia total. Así pues, si realizamos una copia total cada sábado y diferencial el resto de los días, la copia del viernes contendrá todos los ficheros modificados desde el sábado.

Copias de seguridad en cinta

Generalmente, las copias de seguridad se hacen sobre cinta, aunque no es el único dispositivo posible.

Esquema de una copia de seguridad diferencial



La copia diferencial presenta varias ventajas con respecto a la copia total. La primera, como es natural, es que requiere menos espacio, y la segunda, asociada a la primera, es que reduce el tiempo o ventana de copia.

Con respecto a la copia incremental, aporta la ventaja de que en el proceso de recuperación sólo necesitaremos la última copia total y la última copia diferencial. Sin embargo, la copia diferencial, a partir del segundo día, requerirá más espacio y más tiempo o ventana de copia.

6.2.2. Políticas de copias de seguridad

La estrategia de cómo hacer las copias de seguridad es crítica para asegurar que se haga todo correctamente y que se pueda restaurar la información cuando haga falta.

La necesidad de crear estrategias de copias de seguridad proviene del hecho de que, actualmente, en los servidores los discos son de mucha capacidad y, por lo tanto, hay mucha información (tanto de usuarios como de sistema), y toda esta información no cabe en un solo dispositivo de salida (en una sola cinta, por ejemplo). Finalmente, la transferencia dura horas y, por lo tanto, se tienen que buscar soluciones para optimizar su uso.

Analicemos la variabilidad de la información. Con un simple vistazo nos podemos dar cuenta de lo siguiente:

- Hay información que varía diariamente.
- Hay información que se modifica muy poco a lo largo del tiempo.
- Hay información que no hay que guardar en copias de seguridad (los ficheros temporales, por ejemplo).

Así pues, una estrategia de copia que lo copie todo diariamente no parece muy acertada.

Ved también

Ved el módulo “Administración de los datos”, que habla sobre cómo y dónde pueden estar los datos.

Sí que parece claro que tenemos que hacer una copia diaria de la información que varía cada día (acostumbra a ser la información de la organización). Se puede encontrar en los servidores o distribuida por toda la organización. En cualquier caso, hace falta que hagamos una copia diaria de estos datos.

Con la información sobre la cantidad de datos que hay que copiar (el volumen), y sabiendo el dispositivo en el que queremos hacer la copia, tenemos una idea aproximada de las cintas que necesitamos. Una posible política de copias es la siguiente:

Política de copias de seguridad

Lunes	Martes	Miércoles	Jueves	Viernes
 Tot.	 Inc.	 Inc.	 Inc.	 Inc.
 Tot.	 Inc.	 Inc.	 Inc.	 Inc.
 Tot.	 Inc.	 Inc.	 Inc.	 Inc.
 Tot.	 Inc.	 Inc.	 Inc.	 Inc.
 Tot.	 Inc.	 Inc.	 Inc.	 Inc.

Guardaremos la primera copia total de cada mes. De esta manera, siempre es posible recuperar los datos de meses anteriores.

Las ventajas son:

- La copia es rápida porque no hay copias totales diariamente, y las incrementales sólo copian los ficheros modificados durante el día, que son pocos.
- Se ahorran cintas, ya que las copias incrementales ocupan poco en relación con las totales.






Los problemas son:

- Recuperar un fichero requiere tiempo, porque se tiene que pasar por el juego de cintas desde la última copia total y todas las incrementales hasta llegar al fichero de la fecha que se quiere.

- Si falla una cinta incremental, no se puede recuperar nada de los juegos de cintas incrementales posteriores.

Ejemplo sobre los problemas de una cinta incremental

Es viernes, y hay que recuperar un fichero del jueves. Una cinta del juego del miércoles ha fallado, y, por lo tanto, no es posible recuperar la copia incremental de este día. Como consecuencia de ello, no es posible recuperar la copia incremental del jueves, aunque el juego de cintas esté en perfecto estado.

Lunes	Martes	Miércoles	Jueves	Viernes
 Tot.	 Inc.	 Inc.	 Inc.	 Inc.
...









































Este problema de las cintas incrementales hace que muchas organizaciones no utilicen la opción de las copias incrementales por el riesgo que comporta y, por lo tanto, se decanten por opciones de copia **diferencial** o **completa** de los datos.

En el mismo ejemplo anterior, si cambiamos las copias incrementales por copias diferenciales, la copia del jueves contiene los datos de las copias anteriores hasta la total; por lo tanto, la pérdida de la copia del miércoles, en este caso, no sería un problema.

Ahora bien, ¿cuál podría ser una manera de ahorrar cintas? Una posible política de copias de seguridad es la siguiente:

- Llamamos a los juegos de cintas A, B, C, D, E, etc.
- Cada mes, guardaremos un juego de cintas, de manera que tendremos una copia de la información mensual.

Política de copias de seguridad

Lunes	Martes	Miércoles	Jueves	Viernes
A  Tot. 	B  Tot. 	C  Tot. 	D  Tot. 	E  Tot. 
A  Tot. 	B  Tot. 	C  Tot. 	D  Tot. 	E  Tot. 
A  Tot. 	B  Tot. 	C  Tot. 	D  Tot. 	E  Tot. 
A  Tot. 	B  Tot. 	C  Tot. 	D  Tot. 	E  Tot. 

- El juego de cintas E del último viernes del mes lo guardaremos para no utilizarlo. Al final del año, tendremos doce juegos de cintas con la información de la organización. Tendremos que decidir si algunos de estos juegos los volvemos a utilizar o los continuamos guardando.

6.2.3. Información no variable

La información no variable necesita otros criterios de valoración para decidir la manera de cómo hacer una copia de seguridad. La estrategia que se acostumbra a seguir es la siguiente:

1) **Información de sistema.** La información de sistema de los servidores (las particiones con los operativos) se considera crítica. Perderla implica un fallo crítico de la estructura informática. Por lo tanto, como los ficheros de *log*, de registro, etc. también varían bastante, se acostumbran a considerar como en el caso anterior y se hace una copia diaria.

2) **Aplicaciones.** Las particiones con las aplicaciones de los usuarios, teniendo en cuenta que ya se ha gestionado correctamente la información que varía diariamente, no es una información que varíe con mucha frecuencia, por lo que hacer una copia diaria quiere decir cargar mucho el sistema. Por lo tanto, normalmente sólo se hace una copia manual (controlada por los administradores) cuando hay modificaciones sobre la partición.

Observación

La estrategia que presentamos aquí es una de las posibles, pero en ningún caso la única ni la mejor.

3) Estaciones de trabajo. También tienen información de sistema, datos y aplicaciones. Normalmente, la información de sistema (el sistema operativo) y de aplicaciones es prácticamente igual en todas las estaciones. Hacer una copia diaria desbordaría el sistema de copias y colapsaría la red para guardar prácticamente la misma información. En caso de desastre, está el mecanismo de restauración a partir de imágenes de las estaciones de trabajo. Además, en principio, en las estaciones no tendría que haber información, pero si la hubiese, ya se ha dicho en el apartado anterior que hay que hacer una copia diaria exclusivamente de esta información de la estación de trabajo.

Ved también

Ved el apartado 3.3.1 de imágenes de disco en el módulo "Administración de usuarios".

6.2.4. Dónde se pueden guardar las copias de seguridad

Las copias de seguridad tienen dos finalidades:

- Protegernos de fallos de los servidores.
- Proteger la información de la organización.

Con todo lo que hemos hecho hasta ahora, sólo hemos alcanzado el primer punto. El segundo, no, porque si se consigue llegar hasta donde están las copias de seguridad, la información está comprometida.

Normalmente, los administradores tienen las copias con los servidores para poder recuperarlos rápidamente en caso de fallo. Pero tiene que haber una política en un segundo nivel para proteger la información en caso de intrusión física en la zona de los servidores o de desastre de la organización que puede destruir esta zona (por ejemplo, un incendio). En estos casos, y si la información está en el mismo sitio, las copias se convierten en inútiles, por lo cual desmitificaremos previamente algunas recomendaciones:

- "Como la zona de servidores está cerrada bajo llave, no hay peligro". Si la intrusión física consigue llegar a esta zona, tendrá acceso a estropear los servidores y destruir (o hacer desaparecer) las copias de seguridad.
- "Pongámoslas dentro de la caja fuerte, que es ignífuga y no les puede pasar nada". Seguramente a los papeles y monedas, no, pero a las cintas y al material magnético (informático en general), sí: cuando están cerrados en un receptáculo metálico, y si tenemos la mala suerte de que hay un incendio, este receptáculo puede llegar a diversos centenares de grados de temperatura. Con un desastre de esta magnitud los servidores perderán la información, pero las copias de seguridad habrán estado dentro de un "horno" que las podría hacer completamente inútiles y, por lo tanto, también se habría perdido la información.
- "En el peor de los casos se vuelve a entrar todo, ya que está en papel". Hace un tiempo todavía era cierto. Ahora, cada vez menos. En cualquier caso, ciertas precauciones físicas sobre las cintas pueden evitar un problema de un coste muy elevado. Además, visto el uso creciente de las tecno-

logías intranet, mucha información ya no está en papel, sino que se hace directamente sobre sistemas informáticos. Muchos sensores y máquinas de producción recogen directamente los datos en el sistema informático. El uso de las tecnologías de Internet hace que haya información que llega por esta vía sin soporte en papel. El papel ya no lo refleja todo.

6.2.5. Recomendaciones

Siempre tendría que haber, aunque no estuviera actualizada, una copia de seguridad físicamente fuera de la organización. Podría estar en una caja fuerte de un banco o en manos de alguna persona de la dirección, por ejemplo. En caso de desastre prácticamente tendríamos toda la información, no se habrá perdido casi nada.

En caso de que la organización cierre en algunos periodos como, por ejemplo, durante las vacaciones de verano, o en general en periodos en que la seguridad global del edificio se relaja, es muy importante que haya una copia fuera de la organización para prevenir un desastre.

Actualmente hay empresas que se dedican a almacenar copias de seguridad siguiendo protocolos de seguridad pactados conjuntamente además de pactos de confidencialidad. Así pues, esta última opción puede ser la más adecuada para conservar las copias de nuestra organización.

6.3. Plan de contingencia

Planificar todos los pasos necesarios para permitir una recuperación ante un desastre o una situación de crisis es lo que llamamos plan de contingencia³⁶.

Los desastres son inevitables, en la mayoría de los casos impredecibles y varían de tipo y de magnitud. La mejor estrategia para las organizaciones es la confección de un plan de contingencia. Para una organización, un desastre significa una disfunción brusca de parte o de todas sus operaciones comerciales, que pueden provocar una pérdida irreparable o incluso el cierre.

Con el fin de minimizar las pérdidas provocadas por los desastres, es muy importante tener un buen plan de recuperación para cada organización, departamento y operación.

¿Obsesión?

La alta seguridad recomienda, especialmente si hay datos críticos, que la copia de seguridad esté guardada en otra placa tectónica diferente de donde se ubica la organización. El objetivo es que, si hay un terremoto, no pueda llegar a destruir la copia. Se trata de aproximadamente un centenar de kilómetros del sitio original.

⁽³⁶⁾Plan de contingencia en inglés se expresa como *disaster recovery plan*.

Ved también

Ved el módulo “El sistema informático dentro de la organización”, sobre la organización y el sistema informático para la creación de un plan de contingencias.

7. Impresoras

Las impresoras son otros dispositivos que se conectan al sistema informático y son controladas por los servidores de la organización. Las estaciones de trabajo no tienen impresoras conectadas físicamente, y la organización tiene muy pocas en relación con el número de estaciones de trabajo, por lo que es un recurso compartido, gestionado por el servidor mediante una cola de impresión.

La cola de impresión es un recurso de software para conseguir que una impresora (inherentemente no compartible) pueda ser compartida.

Por lo tanto, en el servidor se tienen que crear tantas colas de impresión como impresoras haga falta gestionar. La manera de hacerlo varía según el sistema operativo, pero se acostumbra a seguir dos pasos:

1) Informar al sistema operativo de qué impresora física está conectada. En algunos SO, como Windows, se dice que es necesario instalar el controlador³⁷.

⁽³⁷⁾Controlador en inglés se expresa como *driver*.

2) Crear la cola de impresión y asociarla a la impresora.

Actualmente, hay muchos tipos de impresoras, pero básicamente son dos las que se utilizan más: las impresoras láser y las de chorro de tinta.

7.1. Impresoras láser

Las impresoras láser son las más extendidas y funcionan según el principio de dibujar la página en un tambor especial con un rayo láser y, después, transferirlo al papel con un polvo que se fija con calor. A grandes rasgos, sus características son las siguientes:

Características de una impresora láser	
Coste	Alto
Calidad de impresión	Alta
Velocidad	Alta
Duración cartucho	Alta
Resolución	1.200 × 1.200 ppp
Páginas por mes (aprox.)	15.000

Actualmente, ya hay impresoras láser de doble cara y también impresoras láser de color, por lo que los resultados, por un coste muy aceptable, son totalmente profesionales. Incluso hay papeles especiales, como filmes plásticos transparentes, para hacer transparencias para presentaciones.

7.2. Impresoras de inyección de tinta

Tienen otra utilidad. Su coste es bajo, y muchas veces están instaladas en mesas de despacho. Todas son de color (es inherente a estas impresoras). Funcionan según el principio de lanzar una gota de tinta de forma electrostática sobre el papel.

A grandes rasgos, tienen estas características:

Características de una impresora de inyección de tinta	
Coste	Bajo
Calidad de impresión	Media
Velocidad	Baja
Duración cartucho	Baja
Resolución	1.200 × 1.200 ppp (pero entonces es muy lenta)
Páginas por mes (aprox.)	1.500

Coste del cartucho

Además de tener poca duración, el coste del cartucho de tinta es elevado.

Actualmente, ya hay impresoras de inyección de tinta a doble cara por un coste muy ajustado. El ahorro de papel es notable, y el pequeño incremento en la compra se puede amortizar en poco tiempo.

Si se quieren imprimir transparencias, no hace falta ningún film especial. También está el papel fotográfico, un papel especial y de coste elevado, pero si se imprime con calidad fotográfica (alta resolución), la calidad es extraordinaria.

Otras impresoras

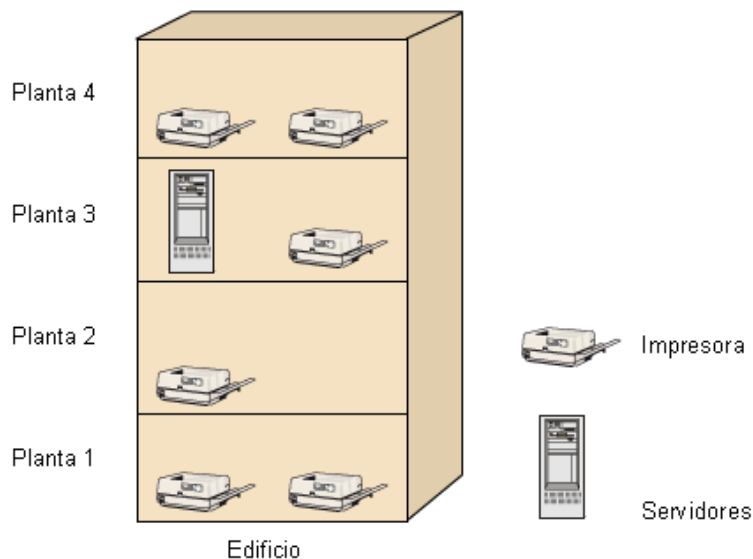
No cabe olvidar que hay otras impresoras, como son las matriciales, que si bien han caído en desuso son imprescindibles para determinadas aplicaciones. Hay otras especiales, como las de inyección de tinta con papel fotográfico A0 (como la Hewlett Packard DesignJet 10000 S), que sirven para hacer pósteres para ferias y congresos. Estas últimas son consideradas por HP como trazadores³⁸ y no como impresoras.

⁽³⁸⁾ Trazadores se expresa en inglés como *plotters*.

7.3. Impresoras remotas

Es muy habitual que la organización necesite impresoras que estén controladas por el servidor, pero que tengan que estar alejadas físicamente de él. Esto hace que no puedan estar conectadas directamente (en local).

Supongamos que nuestra necesidad para imprimir sea la siguiente:



Impresoras locales

Podemos conectar dos impresoras a un ordenador, y con una placa añadida hasta cuatro, pero entonces se necesita un cable físico del ordenador hasta la impresora.

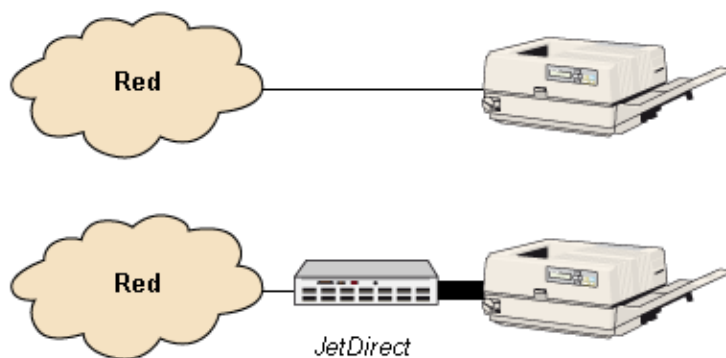
Tenemos seis impresoras y, por lo tanto, las queremos compartir entre todo el personal del edificio. La solución más sencilla es conectarlas al ordenador más próximo y compartirlas. En este caso, son locales para el ordenador al cual están conectadas y remotas para el resto. El principal inconveniente es que sólo se pueden utilizar cuando el ordenador a que se conectan funcione, cosa que hace que en la práctica esta solución sea poco útil, especialmente para las impresoras corporativas.

La solución que se utiliza para las impresoras remotas es la de conectarlas directamente a la red. Hay un cable de red que conecta directamente la impresora a la red de la organización.

En este caso, se tiene que configurar la impresora para que se comporte como un dispositivo de red, y después se tiene que configurar correctamente en el servidor. Básicamente, los pasos son los siguientes:

1) Conectar la impresora a la red. La conexión física se reduce a conectar la impresora a la red. Los modelos de gama alta ya llevan una conexión de red, mientras que en los otros se puede hacer con un dispositivo que enlaza la red con la impresora. En este último caso, se hace por medio del puerto paralelo.

Esquema de conexión de una impresora en la red



2) **Configurar el dispositivo de red de la impresora.** Se tiene que configurar el dispositivo de red. Siempre funcionan sobre protocolo TCP/IP, por lo que tienen una dirección IP. La configuración del dispositivo, una vez conectado a la red y puesto en marcha, se hace vía web, mediante una conexión *telnet* al dispositivo o con un programa específico. A partir de aquí, se configuran todos los parámetros.

JetDirect

A menudo, al dispositivo de red de la impresora se le llama *JetDirect*, aunque estos dispositivos de red son los de la marca Hewlett Packard.

3) **Declarar en el servidor la impresora física** (modelo, etc.). Se tiene que informar en el servidor de que hay una impresora remota y de la dirección IP que tiene, el tipo y el modelo de impresora y sus características relevantes.

4) **Asociar una cola de impresión a esta impresora declarada.** Finalmente, hay que asociar una cola de impresión a la impresora remota que se ha creado y ponerla en marcha.

Con todo esto, los usuarios ya podrán enviar trabajos –que el servidor gestionará sin problemas– por la red a la impresora. Los administradores gestionan esta impresora como si fuera local, dado que la cola está en el servidor y, por lo tanto, la podremos detener, arrancar, eliminar trabajos, etc.

7.4. Internet Printing Protocol

El *Internet Printing Protocol* (IPP) define un método estándar de envío de trabajos de impresión utilizando Internet. Fue desarrollado por el consorcio de compañías del sector: Printer Working Group.

IPP³⁹ provee un único y simple estándar para gestionar los procesos de impresión. Al trabajar con TCP/IP se pueden dirigir a una red local, a una intranet o bien a Internet.

⁽³⁹⁾Recordad que “IPP” es la sigla de *Internet Printing Protocol*.

8. La corriente eléctrica

La corriente eléctrica es uno de los grandes olvidados en el momento de diseñar la disposición de los equipamientos. A pesar de ello, resulta que los servidores, las estaciones de trabajo, la electrónica de red, las impresoras, los monitores, todos los dispositivos y toda la electrónica asociada a la informática van conectados.

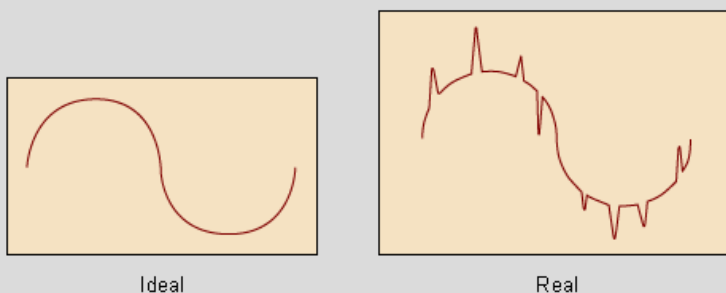
Nos limitamos a enchufar los equipos y suponemos que tendremos una corriente perfecta de 220 V, 60 Hz, 24 horas al día, 7 días la semana, 365 días al año. Éste es un planteamiento completamente irreal. Tenemos que mirar la corriente eléctrica desde una perspectiva mucho más realista.

La corriente eléctrica alimenta todos nuestros dispositivos y ordenadores; dependemos completamente de ella, y puede funcionar mal y ocasionar errores en los sistemas, incluso estropear aparatos.

Empecemos a estudiar la corriente eléctrica que pasa por la organización. ¿Cuáles son los problemas más habituales que nos puede dar?

- Picos de tensión.
- Caídas de corriente o microcortes.
- Proximidad con otras líneas. Las señales de otras líneas próximas (de tensión o de datos) influyen en la calidad global de la tensión.

Ruido es la suma de picos de tensión y caídas de corriente.



¿Qué genera este ruido en la línea? En general, puede venir de todas partes (cualquier aparato eléctrico), pero especialmente los motores eléctricos son muy propensos a generar ruido (por ejemplo, los ascensores, y también determinados elementos de iluminación, como los fluorescentes).

He aquí algunas posibles consecuencias:

- **Pérdida o corrupción de datos.** Si afecta al equipo, puede ocasionar cambios al azar en la electrónica, por ejemplo, cambiar el valor de alguna posición de memoria, por lo que algún programa (o el sistema entero) puede fallar.
- **Daños en el equipamiento.** Si hay grandes sobretensiones, pueden destruir los chips de las placas del ordenador y también estropear controladoras de disco (con la consiguiente pérdida de información), memorias, placas base, etc., de manera que el equipo ya no funcionará.
- **Desgaste prematuro.** Si un equipo está alimentado con corriente eléctrica de mala calidad (ruido), los circuitos electrónicos se desgastan antes de lo que es normal y el equipo falla sin motivo y de una manera aleatoria. Los chips degeneran de una manera desconocida y los resultados son imprevisibles. Entonces pueden pasar cosas como que haya errores de paridad al cabo de pocos minutos de haber arrancado el ordenador, cuando en principio ha pasado correctamente los diagnósticos.

8.1. La toma de tierra

Según el informe *Power and Ground for Distributed Computing*, de David Fench y Larry Fish, de ONEACH Corporation:

Los edificios tienen una toma de tierra de baja resistencia para proteger a la gente de choques eléctricos. La finalidad de la toma de tierra es que la corriente la siga porque hay menos resistencia y, por lo tanto, en caso de tocar algún aparato electrificado, la descarga no pase a través de la persona”.

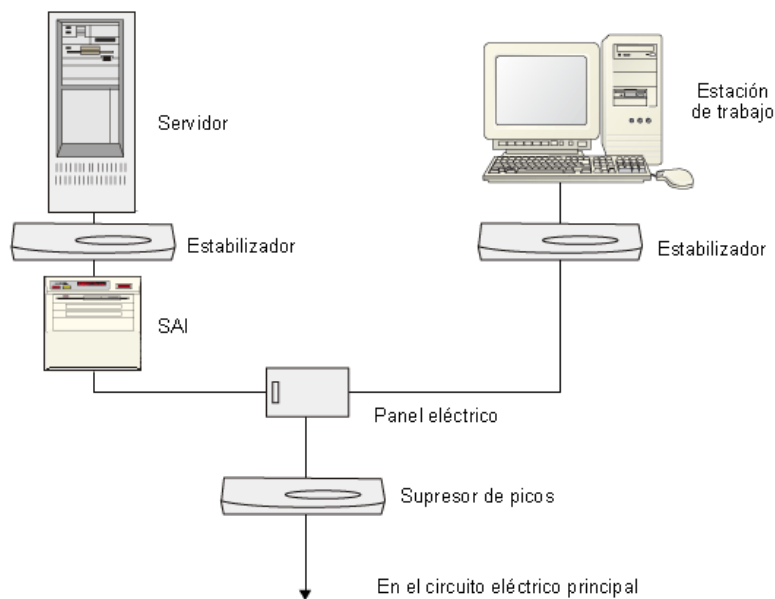
Derivación a tierra

La toma de tierra es una derivación a tierra de la corriente eléctrica.

Ahora bien, las tomas actuales de tierra no funcionan bien con los requisitos de los dispositivos electrónicos de la informática. El motivo es que muchas veces el ruido de la línea atraviesa mejor el equipo para llegar a la toma de tierra que el hilo de la alimentación. Esto pasa porque en grandes redes no hay un único punto de toma de tierra, de manera que el mejor camino es atravesar un ordenador.

La solución se propone con el esquema de Fench y Fish.

Esquema de Fench y Fish



Los estabilizadores aíslan el transformador de la corriente y ofrecen una electricidad de calidad y una buena toma de tierra en el equipo. Si se puede, también se tendrían que instalar estabilizadores en las estaciones de trabajo.

El sistema de suprimir los picos en la entrada está motivado porque, en cualquier otro lugar, desviaría los picos que llegaran al equipo hacia tierra, por lo cual podrían volver a entrar en el circuito por la misma toma de tierra.

Finalmente, se tiene que ir con cuidado cuando se haga el cableado para no instalarlo paralelo a otros circuitos de potencia. A veces, el neutro se deriva a tierra con la intención de solucionar problemas de ruido. Eso puede provocar una onda de baja frecuencia repetitiva en el cable de red que puede afectar a los datos.

8.2. Sistema de Alimentación Ininterrumpida

El Sistema de Alimentación Ininterrumpida (SAI) protege los servidores de cortes de corriente y otros problemas con la tensión.

La importancia de una buena corriente para los servidores se debe al hecho de que una falta de corriente repentina (corte) no les permitirá detenerse correctamente. Eso hará que las memorias intermedias⁴⁰ se pierdan y no se hayan actualizado en el disco, hayan quedado ficheros abiertos y las transacciones no se hayan completado. Es posible que al volver a poner en marcha el sistema, no se pueda arrancar completamente y se pierda información y/o ficheros. Si algún fichero es una base de datos, las consecuencias pueden ser desastrosas:

⁽⁴⁰⁾La memoria intermedia en inglés se expresa como *cache*.

Memorias intermedias con baterías

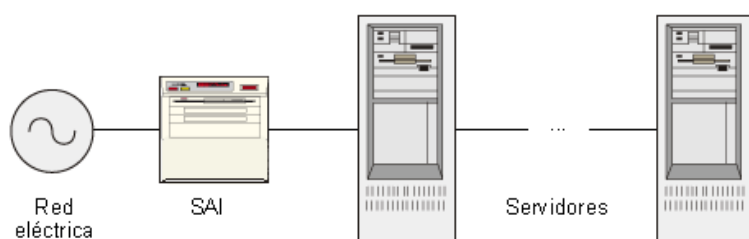
Hay sistemas con memorias intermedias alimentadas por baterías que permiten guardar hasta 72 horas las transacciones pendientes.

se tiene que recuperar de la copia de seguridad, pero desde que se ha hecho hasta que ha sucedido el corte, se ha perdido la información y el tiempo invertido en entrarla.

Un SAI⁽⁴¹⁾ suministra corriente cuando la red eléctrica no la da, de manera que el ordenador continúa funcionando correctamente, sin verse afectado por el hecho de que no hay suministro eléctrico general. Esto permite apagar los sistemas con total normalidad.

(41) Recordad que “SAI” es la sigla de Sistema de Alimentación Ininterrumpida.

Esquema de red con SAI



Las características más relevantes de un SAI son las siguientes:

- **Potencia que hay que suministrar.** Están los watts de potencia que puede dar el SAI cuando no hay corriente de entrada. Determina el número de servidores que podremos conectar.
- **Tiempo de duración de las baterías.** Los SAI llevan baterías que se cargan con la corriente eléctrica y son las que después dan electricidad cuando falla la corriente general. El número de baterías determina el tiempo que podrán suministrar corriente antes de agotarse.
- **Estabilizador.** Esta característica significa que el SAI, además, es capaz de suprimir el ruido. A pesar de ello, necesita una toma de tierra para desviar este exceso de corriente.
- **Tiempo de vida de las baterías.** Un SAI sirve de poco si falla cuando tiene que funcionar. Las baterías tienen una vida útil determinada. Traspasado este tiempo, no hay garantías de que funcionen y respondan correctamente cuando sea necesario. Es el fabricante del SAI quien dice cada cuántos años se tienen que cambiar estas baterías.
- **Aviso al servidor.** Actualmente los SAI tienen una línea (USB o serie) que llega al ordenador. De esta manera, cuando entra en funcionamiento es capaz de enviar una señal al servidor, que con el software adecuado (suministrado con el SAI) sabe que se mantiene con la alimentación eléctrica del SAI. Se mantiene un diálogo que informa del estado de las baterías del SAI y de su duración. Cuando falta poco para agotar la carga de las baterías, el SAI informa al servidor y puede proceder a enviar mensajes a los usuarios y a hacer una parada correcta, ordenada y automática del ordenador. Los

Consumo de un PC

Un ordenador tipo PC suele consumir entre 200 W y 300 W.

servidores acostumbran a estar preparados para arrancar solos, sin intervención del administrador, por lo que, cuando se restablezca el suministro eléctrico normal, el servidor se pondrá en marcha y todo volverá a funcionar correctamente.

9. Seguridad de los servidores

La seguridad es un tema muy amplio. Aquí sólo comentamos dos aspectos genéricos referidos a la seguridad de los servidores. Esta seguridad la tiene que conocer, aplicar y tener en cuenta el administrador de servidores, y afecta básicamente al buen funcionamiento de los servidores corporativos.

Ved también

El tema de la seguridad está perfectamente cubierto en el módulo "Administración de la seguridad".

9.1. Seguridad física de los servidores

Todo sistema de seguridad, aunque parezca muy evidente, empieza por la seguridad física. No sirve de nada proteger todo el sistema informático contra todo tipo de ataques por red si es muy sencillo llegar a los servidores físicamente.

Si podemos acceder físicamente a un ordenador, podremos acceder a la información que contiene.

Esta premisa indica que la información está segura en la medida en que el servidor está físicamente seguro. Éstas son algunas de las precauciones que se pueden tomar:

- Cerrar el recinto donde está el servidor cuando no trabaja nadie.
- Si se tiene que dejar el servidor sin nadie a su cargo, hay que bloquear el sistema con protectores de pantalla con contraseñas.
- Establecer algún procedimiento en caso de avería para que no desaparezcan componentes con información (discos, cintas, etc.).
- Fijar físicamente el equipo para evitar robos (cadenas, entre otros).
- Si está conectado a la red, puede ser interesante tener el equipo conectado a un conmutador⁴² configurado para que controle que sólo determinadas direcciones de placa se puedan conectar al servidor.
- Evitar que el ordenador arranque desde el disquete. Configurarlos para que sólo pueda arrancar desde el disco duro. Si el servidor no la necesita, incluso se puede sacar la unidad de disquete.
- Si el servidor tiene alguna clave para abrir la carcasa, tiene que estar guardada y lejos del servidor.

⁽⁴²⁾ Conmutador en inglés se expresa como *switch*.

- Poner contraseñas para encender el ordenador y para entrar en la configuración de la BIOS.

Actividad

¿Se os ocurren otras recomendaciones de seguridad? Comentadlas en el foro de la asignatura.

9.2. Software

También hay unas precauciones genéricas que se pueden aplicar a todos los sistemas operativos. Esta seguridad de software se orienta a dar unas indicaciones sobre las medidas generales que hay que tomar para tener una máquina físicamente más segura.

Ved también

La seguridad ante ataques en un sistema la hemos explicado en el módulo “Administración de la seguridad”.

- Las cuentas de administrador o superusuario que tengan contraseñas bien hechas, con una política de cambio periódica. De hecho, se tiene que seguir este criterio para todas las cuentas con privilegios especiales.
- Las cuentas con privilegios especiales que no tengan los nombres esperados. Eso quiere decir que, a ser posible, en un ordenador Unix la cuenta de superusuario no tendría que ser *root*, y en una máquina NT, la cuenta de máximos privilegios no tendría que ser *administrator* o administrador, porque de alguna manera significa dar pistas a los posibles atacantes. Por ejemplo, en el caso de NT es posible (y recomendable) cambiar el nombre de la cuenta de administrador por otro.
- No ejecutar en el servidor ni instalar software en el servidor, porque hay peligro de instalar un virus o programas maliciosos.
- Tener una política de grupos y usuarios para evitar agujeros de seguridad en este nivel.

9.3. Alta disponibilidad

Alta disponibilidad es la capacidad de mantener operativas las aplicaciones de la organización, eliminando las paradas de los sistemas de información. Los sistemas informáticos se tienen que haber configurado con el fin de reducir al mínimo porcentaje el tiempo de inactividad o de falta de disponibilidad, con el fin de conseguir la máxima cota de utilidad. La alta disponibilidad de un sistema se consigue al reducir al mínimo la posibilidad de que un error de hardware o un defecto de software comporte la interrupción de uso del sistema o la pérdida de datos del sistema. Por lo tanto, la disponibilidad de un sistema y de sus datos se puede mejorar gracias a la utilización ventajosa de los componentes de hardware o software que sirven para amortiguar el impacto de los errores.

Mito de los 9

El mito del 9 es el tiempo que un sistema está activo al año. Se buscan los 5 nueves, un 99,999% que el sistema tiene que estar disponible. Eso quiere decir que en un año puede no estar activo durante 5 minutos, no necesariamente consecutivos.

99%	3 días y 15 horas
99,9%	8 horas y 15 minutos
99,99%	53 minutos
99,999%	5 minutos
99,9999%	32 segundos

Cada 9 que se añade representa un incremento de costes muy considerable.

Para conseguirlo, se utilizan componentes redundantes y aislados como, por ejemplo, buses dobles, dispositivos de E/S y copias dobles de los datos.

El objetivo es eliminar los periodos de falta de servicio al usuario. Estas paradas pueden ser de dos tipos:

- **Paradas planificadas.** Aquellas debidas a actualizaciones de software o hardware.
- **Paradas no planificadas.** Son las causadas por un mal funcionamiento del hardware o bien por un desastre, ya sea de cariz natural (como inundaciones o incendios) o de carácter no natural (sabotaje, error humano...).

Hay organizaciones en que no es imprescindible un servicio ininterrumpido del sistema informático. En éstas, es necesario un plan de recuperación de datos con el fin de garantizar que el tiempo y el coste de la interrupción serán mínimos. En caso contrario, hace falta que dispongamos de una solución de alta disponibilidad, teniendo en cuenta las necesidades reales de la compañía.

Podemos conseguir alta disponibilidad a través de sistemas tolerantes a fallos, o bien mediante técnicas de *clustering*. Los sistemas tolerantes a fallos son sistemas muy costosos porque hay que asegurar la redundancia de los componentes de su hardware, y eso implica un alto coste. Los sistemas que usan técnicas de *clustering* son más económicos, ya que no hay que utilizar hardware específico. Además, estos sistemas ofrecen balanceo de carga, por lo que extraemos doble provecho con un coste menor.

9.3.1. Sistemas tolerantes a fallos

Éstas son algunas de las cuestiones a tener en cuenta en un sistema tolerante a fallos:

Ved también

Sobre el plan de recuperación de datos, ved el subapartado 6.3 de este mismo módulo.

Servicios de alta disponibilidad

La alta disponibilidad se puede aplicar a cualquier servicio. Los más comunes son:

- Servidor DNS.
- Servidor web.
- Servidor de bases de datos.
- Servidor de ficheros.
- Servidor de correo.

- **Redundancia en el suministro eléctrico.** Un corte en el suministro eléctrico, aunque sea de pocos segundos, provocará que durante un tiempo nuestra máquina esté fuera de servicio. Por lo tanto, es vital conseguir que nunca falte el suministro eléctrico. Aparte de garantizar el suministro, hay que tener en cuenta también las fluctuaciones de tensión, que también pueden afectar negativamente a nuestros equipos. Por lo tanto, hay que valorar la instalación de SAI⁴³, grupos electrógenos, fuentes de alimentación redundantes en el propio equipo (intercambiables en caliente) o incluso contratos con dos compañías eléctricas.
- **Discos duros redundantes.** Para conseguir un sistema tolerante a fallos, los discos tienen que ser redundantes, ya que están sometidos a errores electrónicos (subidas de tensión, por ejemplo) y a errores mecánicos (averías de cabezales, por ejemplo).
- **Conexiones de red.** La red se ha convertido en un elemento indispensable para las aplicaciones actuales. Es indispensable, y por eso hay que garantizar que la red estará disponible en todo momento. Para conseguir una red tolerante a fallos, hay que utilizar dispositivos de red tolerantes a fallos.

⁽⁴³⁾Recordad que “SAI” es la abreviatura de sistema de alimentación ininterrumpida.

Ved también

Sobre la corriente eléctrica, ved el apartado 8 de este mismo módulo.



Fuente de alimentación redundante

9.3.2. Clústers de alta disponibilidad

Los clústers de alta disponibilidad y tolerancia a fallos están destinados a proporcionar disponibilidad ininterrumpida de recursos y servicios mediante la redundancia. Si un nodo del clúster falla, las aplicaciones y servicios que se ejecutan pasarán a ejecutarse en uno de los nodos disponibles.

Algunas de las ventajas de este tipo de configuraciones son:

- **Escalabilidad.** Puede aumentar la capacidad de cálculo del clúster si se añaden más procesadores o equipos.
- **Alta disponibilidad.** El clúster está diseñado para evitar un único punto de error. Las aplicaciones pueden distribuirse en más de un equipo, consiguiendo un grado de paralelismo y una recuperación de errores y proporcionando más disponibilidad.

Ved también

Sobre los discos duros redundantes, ved el apartado 5, y en especial los subapartados 5.2.5 y 5.3 de este mismo módulo.

Ved también

Sobre las conexiones de red, ved el módulo “Administración de la red”.

Utilidad de los clústers de alta disponibilidad

Los clústers de alta disponibilidad se suelen utilizar para sistemas de bases de datos de aplicaciones críticas, servidores de mail, ficheros o aplicaciones.

10. Aspectos legales

El administrador de servidores es una figura que tiene a su cargo, de una manera directa e/o indirecta, una gran cantidad de información de la organización. Toda esta información es sensible, por lo que, además de velar para que esté disponible y al alcance de las personas que la tienen que utilizar, es información que el administrador tiene para manipular. ¿Dónde están las fronteras legales de todo esto? ¿Qué tiene que hacer si le piden que extraiga información de cierto lugar? ¿O que la mire? ¿Y si le dicen que instale un programa que controle la actividad de los usuarios sobre cierta información? ¿Qué puede hacer y qué no un administrador de servidores con toda esta responsabilidad?

A pesar de que, actualmente, la cuestión va variando bastante, y que la legislación se mueve en un panorama muy cambiante, intentaremos hacer un repaso a estas cuestiones en el módulo “Administración de la seguridad”.

Somos conscientes de que, en el momento en que aparece el problema, uno mismo tiene que buscar asesoramiento legal para resolverlo, pero consideramos que una de las cuestiones más importantes es saber reconocer, en materia legal, cuándo hay un problema real y cuándo no.

Ved también

En el subapartado 1.3 del módulo “Administración de la seguridad”, encontraréis el protocolo técnico que hay que seguir en caso de ataque a los servidores.

10.1. Colegios profesionales

Actualmente, hay colegios técnicos de informáticos. Algunos de sus objetivos son los siguientes:

- Peritar trabajos.
- Dar apoyo legal a los informáticos ante problemas.

Eso permite saber, en cualquier momento, cuándo una acción que han llevado a cabo los administradores se ha hecho dentro o fuera de la legislación (si es legal o no) y las consecuencias que puede tener. Muchas acciones, aparentemente inocuas, esconden situaciones potencialmente problemáticas. Una cosa tan sencilla como copiar una imagen de Internet para utilizarla o abrir un fichero del directorio personal (carpeta personal en terminología Windows) de un usuario, puede violar la legislación vigente. Tener claras las cuestiones, los límites y las consecuencias que se pueden derivar en caso de transgredirlos es una de las muchas funciones de estos colegios.

11. Tareas/responsabilidades

Una posible relación de las tareas/responsabilidades del administrador de servidores podría ser la siguiente:

- Velar por el funcionamiento correcto de los servidores.
- Cuidar de la protección física de los servidores.
- Cuidar de la copia de seguridad de los servidores.
- Procurar el buen funcionamiento de los subsistemas asociados a los servidores (colas de impresión, correo electrónico, etc.).
- Asegurar la disponibilidad de espacio para el trabajo de las aplicaciones y los usuarios.
- Velar por unos tiempos de respuesta de los sistemas correctos.
- Asignar los grupos de usuarios y permisos en relación con lo que se ha acordado con el responsable de informática.
- Velar por la seguridad del sistema.
- Mantener el sistema operativo actualizado.
- Mantener actualizadas las aplicaciones bajo su responsabilidad.
- Garantizar que la información del sistema esté protegida contra fallos, desastres naturales y eliminaciones accidentales. Normalmente eso se hace mediante la copia de seguridad.
- Proteger los datos/el contenido de los servidores.
- Asegurar la disponibilidad de la información que contiene.
- Asegurar el acceso al correo electrónico (desde el punto de vista de los servidores).
- Configurar los servidores corporativos.

Actividad

¿Consideráis que falta alguna tarea del administrador de servidores interesante? Lo podéis comentar en el foro de la asignatura.

Resumen

Hemos visto cómo tiene que ser físicamente un servidor y las características de hardware que hay que tener en cuenta. Hemos profundizado en las diferentes configuraciones de los servidores que nos permiten obtener funciones y rendimientos mucho mejores que un servidor aislado. Nos hemos dado cuenta de la importancia de los discos y de cómo se pueden configurar y ajustar a las necesidades de la organización, ya que es una de las cuestiones clave.

Hemos hablado mucho de los dispositivos de copia de seguridad, sus posibilidades y políticas posibles, dependiendo del tamaño y las necesidades de la organización. Hemos remarcado la importancia de la corriente eléctrica para asegurar el funcionamiento y la vida de los servidores.

Finalmente, hemos comentado aspectos de los sistemas operativos y las responsabilidades del administrador de servidores.

Tampoco hemos olvidado la seguridad física de nuestros servidores, porque contienen toda la información de la organización. Toda precaución es poca para nuestra información.

Actividades

1. ¿Conocéis algún sistema de ficheros distribuidos de los que actualmente hay en la red? (Para los que no conozcáis ninguno, escoged uno y conectaros a él). Haced pruebas con este sistema y comparadlo con los sistemas de ficheros tradicionales que conozcáis, por ejemplo, el de vuestra estación de trabajo.
2. Buscad en la red algún software de virtualización de pruebas y cread una máquina virtual. Intentad instalar en esta máquina virtual un sistema GNU/Linux y verificad cómo se pueden aprovechar los recursos físicos del sistema, como por ejemplo la tarjeta de red, el disco, la memoria, etc.

Ejercicios de autoevaluación

1. Suponiendo que en una organización hay un servidor con protección de discos RAID-6e y una capacidad de almacenamiento muy grande, aparte de una gran cantidad de memoria RAM, ¿qué tipo de servidor crees que puede ser y qué tipo de datos sería lógico que almacenara?
2. ¿Qué ventaja nos proporciona la copia de seguridad diferencial frente a una copia de seguridad incremental?
 - a) La copia se realiza en menos tiempo, pero ocupa más espacio.
 - b) La copia se realiza en más tiempo a partir del segundo día de copia, pero ocupa menos espacio.
 - c) Salva todos los objetos modificados desde la última copia diferencial, ocupa más espacio desde el segundo día, pero tarda menos que una copia incremental.
 - d) Salva todo los objetos modificados desde la última copia total, ocupa más espacio desde el segundo día, y tarda más tiempo en hacer la copia.
3. ¿Por qué crees que un administrador de un servidor escogería discos SAS para su servidor en vez de discos SATA?
4. Os proponen implementar un sistema dedicado al servicio de correos de usuario, que tiene que ser lo más seguro posible y escalable teniendo en cuenta que la organización crece en número de empleados constantemente. ¿Qué opción de las siguientes escogeríais?
 - a) Virtualización de un servidor físico en servidores de correo.
 - b) Clúster de servidores físicos, todos ellos dedicados a dar el servicio de correo.
 - c) *Load balancer* de servidores de correo.
 - d) Las tres respuestas anteriores son correctas.
 - e) Las respuestas b) y c) son correctas.
5. Escribid un pequeño texto en el que se relacionen los siguientes elementos: NAS, estación de trabajo, acceso a nivel de *block*, LAN, SAN, datos de usuario y elemento de almacenamiento.

Solucionario

Ejercicios de autoevaluación

1. Según los datos que nos han suministrado, podemos deducir diferentes aspectos funcionales del servidor.

Una gran cantidad de almacenamiento y una gran cantidad de RAM nos conducen a un servidor destinado a servir datos.

La utilización de un sistema de seguridad de discos RAID-6e utiliza dos códigos correctores para cada sector y grupo de RAID, además de un disco *hot spare* (en espera) por si uno de los discos falla. Esto significa que los datos almacenados son muy importantes y hay que protegerlos.

Respondiendo, pues, a las preguntas, estamos hablando de un servidor de almacenamiento, posiblemente de un servidor de base de datos con datos muy importantes para la organización.

2.d. La ventaja mayor que tenemos es que salvamos todos los objetos modificados desde la última copia total, y a la hora de restaurar el sistema, sólo tendremos que restaurar la copia total y la última diferencial.

3. Tal y como se indica en este módulo, los discos SCSI han sido diseñados y fabricados para cumplir con los requisitos empresariales de alta disponibilidad y seguridad. Así pues, los discos SAS, que son la evolución, también cumplen con este objetivo.

4. e. Teniendo en cuenta que tendremos que dar un solo servicio (correo) y que tiene que ser escalable, es decir, tiene que poder servir un número indeterminado de usuarios, la mejor opción es un clúster, que puede ser, cómo no, un *load balancer* para aprovechar la potencia de trabajo.

5. Un usuario que trabaja con su estación de trabajo tendría que tener sus datos más importantes remotamente en un servidor de almacenamiento, como por ejemplo una NAS. Para acceder a este servidor, utilizará la red LAN de comunicaciones.

Aunque el usuario acceda a la NAS para buscar sus datos, puede ser que éstos realmente estén en un elemento de almacenamiento al cual accede el servidor NAS, en el bloque, mediante una red SAN.

Glosario

advanced intelligent tape *f* Ved **cinta avanzada inteligente**.

AIT *f* Ved **cinta avanzada inteligente**.

arsenal redundante de discos económicos *m* Un doble intento de tener unidades de discos muy grandes y redundancia en el sistema por si falla un disco. Se trata de distribuir la información entre diversas unidades de disco.

en redundant array of inexpensive disks.

sigla: **RAID**.

alta disponibilidad *f* Instalación que intenta conseguir el máximo de disponibilidad de un sistema (24 × 7).

backup *m* Ved **copia de seguridad**.

blade *f* Hoja o lámina. Se aplica a servidores en una tarjeta o lámina.

blade center *m* Cabina específica para gestionar *blades*.

cinta audio digital *f* Uno de los dispositivos para hacer copias de seguridad en cinta.

en digital audio tape.

sigla: **DAT**.

cinta digital inteligente *f* Uno de los dispositivos para hacer copias de seguridad en cinta.

en digital intelligent tape.

sigla: **DIT**.

cinta avanzada inteligente *f* Uno de los dispositivos para hacer copias de seguridad en cinta.

en advanced intelligent tape.

sigla: **AIT**.

clúster *m* Agrupación de servidores que dan servicio a una tarea única.

copia de seguridad *f* Método para duplicar la información de la organización sobre otro soporte que sea más seguro.

CPU *f* Ved **unidad de control de proceso**.

DAT *f* Ved **cinta audio digital**.

descarga completa *f* Copia de seguridad completa de una partición de disco.

en full dump.

digital audio tape *f* Ved **cinta audio digital**.

digital intelligent tape *m* Ved **cinta digital inteligente**.

directorio *m* Espacio lógico dentro de un disco en el que se guardan ficheros y directorios.

sin. **carpeta**.

disco duro *m* Dispositivo físico que sirve para guardar información.

DIT *f* Ved **cinta digital inteligente**.

full dump *f* Ved **descarga completa**.

IDE *m* Ved **lector electrónico inteligente**.

grid *m* Computación en malla, permite interconectar ordenadores dispersos por la red para aprovechar su potencia de cálculo.

impresora remota *f* Impresora que está conectada directamente a la red informática en lugar de estarlo a un ordenador. El servidor la gestiona a través de la red, no localmente, porque no hay cable.

intelligent drive electronics *m* Ved **lector electrónico inteligente**.

interfaz pequeña del sistema informático *f* Tipo de controladora de dispositivos de altas prestaciones. Se pueden conectar a ella muchos dispositivos diferentes, y las diversas revisiones permiten conectar hasta dieciséis dispositivos a la misma controladora.

en small computer system interface.

sigla: **SCSI**.

J2EE *f* *Java to enterprise edition*. Estándar java orientado a arquitecturas de empresa.

JVM *f* *Java virtual machine*. Máquina virtual de java. Interpreta los pedidos java en un sistema.

lector electrónico inteligente *m* Tipo de controladora de dispositivos de bajo coste. Normalmente se conectan a ella discos duros o CD-ROM. Cada controladora puede soportar sólo dos dispositivos.

en intelligent drive electronics.

sigla: **IDE**.

memoria de acceso aleatorio *f* Memoria volátil que utilizan todos los ordenadores.

en random access memory.

sigla: **RAM**.

motherboard *f* Ved **placa base**.

NAS *f* Servidor de ficheros. Acceso al fichero.

partición *f* División del espacio interno del disco duro.

placa base *f* Componente del ordenador que tiene los buses de sistema y el árbitro del bus. Controla toda la comunicación entre los diferentes componentes. Contiene la BIOS, el espacio para montar la CPU, la RAM y las ranuras de expansión (*slots*) para la placa gráfica, la placa de red, etc.

en motherboard.

plan de contingencia *m* Estudio del impacto de posibles contingencias y su tratamiento con el fin de recuperar la normalidad funcional.

placa de red *f* Componente del ordenador que permite la comunicación entre la red y los buses internos. El software hace toda la lógica de sobre.

RAID *m* Ved **arsenal redundante de discos económicos**.

RAM *f* Ved **memoria de acceso aleatorio**.

random acces memory *f* Ved **memoria de acceso aleatorio**.

redundant array of inexpensive disks *m* Ved **arsenal redundante de discos económicos**.

SAI *m* Ved **sistema de alimentación ininterrumpida**.

SAN *m* *Store Area Network*. Red especializada en comunicación entre servidores y elementos de almacenaje.

SAS *m* *Serial Attached Scsi*. Protocolo de acceso en serie a discos SCSI.

Ved **SCSI**.

SATA *m* *Serial-ATA*. Protocolo de acceso en serie a discos ATA.

Ved **IDE** o **P-ATA**.

SCSI *f* Ved **interfaz pequeña del sistema informático**.

servidor institucional *m* Ordenador que usa aplicaciones con tecnología cliente/servidor y sirve peticiones por la red, bajo demanda de los clientes (estaciones de trabajo).
sin. **servidor corporativo**.

servidor virtual *m* Ordenador servidor físico que puede dar servicio a diferentes servicios virtuales, cada uno con su propio sistema operativo.

sistema de alimentación ininterrumpida *m* Componente que evita la caída de los servidores por falta de corriente eléctrica, porque se encarga de suministrarla cuando no hay.
sigla: **SAI**.

sistema de ficheros *m* Configuración consistente en una partición para poner los ficheros.

small computer system interface *f* Ved **interfaz pequeña del sistema informático**.

toma de tierra *f* Conductor que se pone en contacto íntimo con el suelo.

unidad de control de proceso *f* Cerebro del ordenador.
sigla: **CPU**.

velocidad de transferencia *f* Velocidad en Mb/segundo en la que viaja la información entre dos dispositivos o componentes.

Bibliografía

Cockcroft, A. (1995). *Sun Performance and Tuning SPARC & Solaris*. Estados Unidos: Sun Microsystems.

Halliday, C. (1996). *Los secretos del PC*. Madrid: Anaya Multimedia.

IBM (2004). *SAN Survival Guide*. [Disponible en línea].

Microsoft Corporation (1997). *Windows NT 4.0 Workstation Kit de Recursos*. Madrid: McGraw Hill.

Muellers, S. (1999). *Upgrading and Repairing PCs*. Estados Unidos: Que Corporation.

Sheldon, T. (1994). *Novell NetWare 4 Manual de Referencia*. Madrid: McGraw-Hill.

Sun Microsystems Ibérica (1994). *Administración de Sistemas Solaris 2.x*. Madrid.

National Institute of Standards and Technology (2002). *Risk Management Guide for Information Technology Systems*. NIST Special Publication 800-30.