4 ORGANIZACIÓN Y GESTIÓN DE LA INFORMACIÓN

En este capítulo se estudiará todo lo concerniente a la persistencia de la información en un sistema informático. Se estudiarán los sistemas de archivo de un sistema operativo y cómo implementarlos ya sea con particionamiento clásico, protección RAID, gestión de volúmenes lógicos, etc. Además de estos conceptos se verán las políticas de salvaguarda y de seguridad para que el sistema además de eficiente sea seguro frente a cualquier circunstancia.

4.1 SISTEMAS DE ARCHIVO

El objetivo de un sistema de archivos es el gestionar los ficheros almacenados generalmente en un dispositivo físico como un disco duro, unidad SSD, *pendrive*, etc. El sistema de archivos mantiene unas estructuras internas para mantener organizados los ficheros de forma eficiente y para que el acceso a los mismos sea rápido.

El sistema de archivos es el responsable de gestionar el espacio en los diferentes dispositivos organizando ficheros y directorios en ellos y llevando un control de qué zonas del dispositivo están en uso y cuáles no. Si no se hace una gestión eficiente del espacio se generará fragmentación (es cuando el espacio sin utilizar no está contiguo). La fragmentación provoca una mala gestión del espacio en el dispositivo e incluso una pérdida de eficiencia del sistema cuando se reorganizan los ficheros para eliminarla.

Otra función importante de los sistemas de archivos es el control de acceso a los mismos. El sistema de archivos debe tener mecanismos para permitir o no el acceso a usuarios o grupos de usuarios dependiendo de los derechos que tengan sobre dichos archivos o directorios. Estos mecanismos utilizarán ACL o listas de control de accesos, permisos de ficheros en forma de bits como en Unix, u otro sistema. En ocasiones, los sistemas de archivo además de mecanismos de control de accesos, utilizan la encriptación para cifrar el contenido del fichero en un intento de mejorar la seguridad del sistema.

Los sistemas de archivo también son responsables de mantener la integridad de la información. Deben tener mecanismos para preservar la integridad del sistema de archivos para que en el caso en que pérdidas de conexión, reseteos del equipo, pérdidas de tensión, fallo del dispositivo, cuelgues en los programas, etc., no dejen los datos del archivo o del sistema de archivos corruptos o dañados y se pierda la información.

Existen muchos sistemas de archivos dependiendo de dónde reside la información de los mismos. Por ejemplo:

- **Sistemas de archivos de disco**. Son los más utilizados. Están pensados para que múltiples procesos accedan a una zona concreta del fichero sin importar la ubicación exacta en el disco. Estos sistemas de ficheros están diseñados para que trabajen de forma eficiente e incluso se puedan anticipar a los datos que los procesos puedan demandar. Los sistemas de archivos más utilizados son FAT, NTFS, EXT, ReiserFS o HFS+. A continuación se verán en detalle los más utilizados:
  - **FAT16**. También es conocido como FAT. Permite trabajar con particiones de hasta 2 GB, las unidades de asignación son de 32 KB, el tamaño máximo de archivo es de 2 GB y no distingue entre mayúsculas y minúsculas para nombres de archivo o directorio. Es reconocido por la práctica totalidad de los sistemas operativos.

- **FAT32**. Permite trabajar con archivos de hasta 4 GB y con particiones de más de 2 GB. La unidad de asignación es de 4 KB. El tamaño máximo de partición es de 2 TB. No distingue entre mayúsculas y minúsculas para nombres de archivo o directorio.
- **NTFS**. Sus volúmenes pueden llegar hasta los 16 EB al igual que el tamaño máximo del archivo, sin embargo la implementación actual es de 256 TB y 16 TB respectivamente. Distingue entre mayúsculas y minúsculas en los nombres de archivo y directorio. El tamaño de la unidad de asignación varía, dependiendo del tamaño de la partición desde los 4 KB a 64 KB.
- **EXT3**. También conocido como Sistema de Archivos Extendido 3. El tamaño máximo de la partición es de 32 TB y el tamaño máximo de archivo es de 2 TB. Distingue entre mayúsculas y minúsculas para nombres de archivo y directorio. Es propio de sistemas operativos Linux.
- **EXT4**. También conocido como Sistema de Archivos Extendido 4. Permite particiones de hasta 1 EB y archivos de hasta 16 TB. Distingue entre mayúsculas y minúsculas en los nombres de archivo y directorio. Es utilizado actualmente por muchos sistemas operativos Linux.
- **HFS+** (*Hierarchical File System Plus*). También conocido como Sistema Jerárquico de Archivos Extendido. El tamaño máximo de una partición es de 16 EB y el tamaño máximo de archivo es de 8 EB. Es el empleado por los sistemas operativos de Apple.

■ **Sistemas de archivos de discos ópticos**. Estos sistemas de archivos son utilizados en CD, DVD y discos Blu-ray. Los más utilizados son ISO 9660 y el UDF (*Universal Disk Format*).

■ **Sistemas de archivo de cintas**. Son sistemas de archivos muy particulares debido a la naturaleza de las cintas. Al ser un sistema de almacenamiento secuencial los accesos aleatorios llevan más tiempo que en los discos tradicionales puesto que implica el rebobinar la cinta para colocar la cabeza lectora sobre el punto exacto donde se va a leer o escribir.

■ **Sistemas de archivo de red**. Los archivos en los sistemas de ficheros en red generalmente están distribuidos en uno o varios puntos de una red local en los que un cliente accede a los ficheros proporcionados por un servidor. Para el acceso a dichos ficheros se necesitarán protocolos como NFS o SMB.

4.1.1 NOMENCLATURA Y CODIFICACIÓN

Generalmente los archivos suelen constar de un nombre y una extensión separadas por un punto (por ejemplo datos.txt). La extensión, que es algo que viene de los tiempos del MS-DOS, suele tener tres caracteres e identifica el contenido del archivo. Por ejemplo un archivo txt suele contener texto mientras que un fichero exe identifica un fichero binario ejecutable. En algunos sistemas de archivos como en Unix, la extensión no es necesaria.

En algunos sistemas la longitud del nombre del archivo no puede sobrepasar un cierto límite y generalmente se impide que formen parte del nombre ciertos caracteres como “*”, “/”, “\”, “;”, “&”, etc.

En cuanto a la **codificación** de los archivos existen distintos sistemas para codificar la información en un archivo. A continuación se describen los más utilizados:

■ **ASCII**. El código ASCII ha sido durante mucho tiempo el más empleado. Inicialmente era un código que utilizaba 7 bits para representar texto, lo que significaba que era capaz de codificar 127 caracteres. Por ejemplo el número 65 (1000001 en binario) se utiliza para representar la “A”. Poco después surgió un problema: este

código es suficiente para los caracteres de la lengua inglesa, pero no para otras lenguas. Entonces se añadió el octavo bit para representar otros 128 caracteres que son distintos según idiomas (Europa Occidental usa unos códigos que no utiliza Europa Oriental), llamándose así ASCII Extendido.

- **Unicode**. Unicode es una ampliación del código ASCII que puede utilizar hasta 4 *bytes* (32 bits) para representar cada carácter, con lo que es capaz de codificar cualquier símbolo en cualquier lengua del planeta utilizando el mismo conjunto de códigos.
- **UTF-8**. Es un formato que es capaz de representar cualquier carácter Unicode y que al incluir la especificación US-ASCII de 7 bits puede codificar cualquier información ASCII sin cambios. Se suele utilizar también mucho en la codificación de páginas web y *email*. Tiene ventajas al ahorrar espacio de almacenamiento en textos con caracteres latinos.
- **Binario**. Se utilizan los ficheros binarios para almacenar información de cualquier tipo. En binario se pueden almacenar textos, imágenes, sonido, vídeo, datos, etc. Generalmente los ficheros binarios suelen contener una cabecera en la que hay una serie de datos o información para poder interpretar el resto del fichero. Si el fichero carece de esta cabecera se denomina archivo binario plano.

4.1.2 JERARQUÍAS DE ALMACENAMIENTO

Según las características de las memorias y los propósitos para los cuales han sido desarrolladas, los medios de almacenamiento se pueden clasificar de forma jerárquica en distintos niveles. A esta clasificación se la ha denominado "jerarquía de memorias".

Figura 4.1. *Jerarquía de memorias*

- **Nivel 0: Registros**. Los registros son memorias muy veloces pero con poca capacidad, los cuales están integrados en el procesador. Estos registros almacenan datos transitorios utilizados por el procesador, generalmente resultados de operaciones matemáticas. Los valores muy utilizados también son objeto de almacenar en registros del procesador.
- **Nivel 1: Memoria caché**. La memoria caché es una memoria intermedia que se coloca entre un elemento rápido y otro más lento del equipo. Generalmente cuando se habla de memoria caché nos referimos a la memoria existente entre la memoria principal y el procesador. El objetivo de esta memoria es almacenar copia de datos situados en memoria principal los cuales o son muy utilizados o se prevé que van a ser utilizados en un futuro. Generalmente cuando se accede a un dato, este se coloca en la memoria caché. La segunda vez que se acceda a dicho dato no hará falta acceder a la memoria pricipal puesto que el dato ya lo tenemos en la caché.

El procesador, antes de escribir o leer de memoria siempre va a mirar en la memoria caché por si una copia del dato está en ella. Si el dato está en caché las operaciones se realizan mucho más rápido puesto que la memoria caché es mucho más rápida que la memoria principal.

La caché generalmente utiliza tecnología SRAM (*static* RAM) la cual es más rápida que la tecnología utilizada en la memoria principal.

En los discos duros también hay una memoria caché que funciona de modo análogo a la caché situada entre la RAM y el procesador. Esta memoria caché no utiliza SRAM, la cual es muy cara, utiliza la misma tecnología que en la memoria RAM tradicional.

- **Nivel 2: Memoria RAM o principal**. La memoria RAM es la insertada en el *slot* de la placa base. Está presente en muchos elementos internos del equipo y es mucho más rápida que la memoria secundaria o disco duro. Posteriormente se estudiará la memoria RAM en profundidad.
- **Nivel 3: Disco duro**. Este nivel se suele denominar almacenamiento secundario. En este nivel se incluye el mecanismo de memoria virtual el cual utiliza espacio en disco para gestionar la memoria.
- **Nivel 4: Almacenamiento externo en redes**. El almacenamiento en red se está popularizando mucho dado que el crecimiento y mejora de las redes en cuanto a velocidad y prestaciones lo hace posible. Actualmente las necesidades con respecto a la información son que sea accesible desde cualquier dispositivo y cualquier lugar, lo que implica que tengamos que almacenar los datos en la red.

Los sistemas más populares y generalmente más económicos que se han estado utilizando desde hace tiempo son los NAS (*Network Attached Storage*). En estos sistemas los datos residen en un servidor el cual los comparte con el resto de la red (casi siempre TCP/IP). Muchas veces se denomina NAS a un equipo servidor (ordenador convencional) con un sistema operativo Linux o Windows®. No obstante, los auténticos NAS son dispositivos dedicados específicamente diseñados a tal fin. Entre los protocolos más utilizados por los NAS están el NFS y el CIFS (de Microsoft®).

Últimamente se está popularizando mucho el almacenamiento en la nube. Con el almacenamiento en la nube lo que se hace es almacenar nuestros datos en un servidor externo de Internet con lo cual se puede recuperar dicha información con cualquier dispositivo o compartirla con otras personas.

Quizás la aplicación más utilizada de almacenamiento en la nube actualmente es Dropbox. La instalación y utilización de Dropbox es sumamente sencilla e intuitiva. La aplicación funciona para las plataformas más usuales (Linux, Windows®, Android, Mac OS X®, iPhone, iPad® o Blackberry®). Una vez instalado Dropbox se crea una carpeta en la cual se pueden almacenar documentos los cuales estarán en la nube sincronizándose también en local, pudiendo ser utilizados por otros usuarios o dispositivos. Además de Dropbox existen muchas otras aplicaciones similares como Box, Ubuntu One, iDrive, SOS Online *Backup*, Zumo drive, ADrive, etc.

4.1.2 MIGRACIONES Y ARCHIVADO DE DATOS

Las migraciones de datos se realizan cuando hay que traspasar datos de un sistema a otro. Cada migración es un mundo y por ello hay que planear de una forma concienzuda la misma. Generalmente las migraciones de datos se hacen entre una base de datos de un sistema y la base de datos de otro sistema.

Existen muchos tipos de migraciones de datos. Algunos tipos de migraciones son los siguientes:

- **Traslado del sistema a otra máquina más potente** con el mismo gestor de base de datos. En ese caso las mismas herramientas de *backup* y restauración o exportación e importación del propio gestor de bases de datos pueden servir para realizar la migración.
- **Migración por upgrade o mejora de la base de datos**. En ocasiones se migra de una base de datos en Access o ficheros ASCII a una base de datos de más entidad como Oracle, Informix, MySQL, SQLServer, etc. En ese caso la migración implica el tener que crear una nueva base de datos con los requisitos de la base de datos previa. Para ello, habrá que realizar un análisis a fondo del sistema previo antes de la migración.
- **Migración entre distintos gestores de bases de datos**. En este caso puede ser de bastante utilidad estudiar a fondo el esquema de la base de datos origen. Si se puede disponer de ambos gestores de bases de datos, se pueden realizar todo tipo de pruebas de verificación de una forma sencilla antes de poner el nuevo sistema en producción.

Algunas cuestiones a tener en cuenta al hacer una migración son las siguientes:

- **Antes**. Analizar a fondo los datos de origen y analizar el nuevo sistema, más si cabe si alguno de los sistemas que se va a migrar está en producción y los datos no son estáticos.
- **Durante**. Convertir o adaptar tipos de datos verificando que los valores que se vayan a grabar en la nueva base de datos son los que se necesitan. Cuidado con los espacios en blanco, tabuladores o caracteres extraños que puedan aparecer por culpa de la migración. En ocasiones hay que deshabilitar *triggers* o también restricciones que puedan impedir el correcto proceso de migración.
- **Después**. Realizar pruebas para verificar que la totalidad de registros en origen se ha copiado en el sistema destino. Estas pruebas tienen que estar previstas de antemano. Generalmente estas pruebas consisten en su mayor parte en conteos de registros en origen y destino. Hay que prever mecanismos de vuelta atrás o de emergencia por si la migración no concluye de forma exitosa.

Sobre el archivado de datos, decir que hay que tener mucho cuidado con los *backups* puesto que uno de los mayores errores de seguridad se cometen cuando se descuida su correcta custodia. En ocasiones acceder al fichero y enviarlo por Internet mediante correo electrónico u otro sistema es una práctica fácil y común.

Además, el concepto de archivado implica el almacenar datos que no sean actuales o no se van a utilizar de forma habitual (por ejemplo las facturas de dos años hacia atrás). Lo más normal es que para este tipo de prácticas se utilice un software de archivado de datos más que un software de *backup*. Aunque parecen diferentes no son lo mismo. El software de archivado de datos está específicamente diseñado tanto para archivar datos como para recuperar datos según unos criterios determinados.

Existen muchos software en el mercado de archivado de datos. Algunos son StorNext FX™, Autonomy Zantaz™, Iron Mountain™, Mimosa™, NearPoint™ o Enterprise Vault™ entre otros.

4.2 VOLÚMENES LÓGICOS Y FÍSICOS

En este apartado se estudia todo lo concerniente al almacenamiento físico de un sistema operativo. Se comenzará estudiando el particionamiento para luego ver conceptos interesantes como la gestión de volúmenes lógicos o los sistemas RAID.

4.2.1 PARTICIONAMIENTO

Prácticamente todos los discos incluso los dispositivos con memoria *flash* se pueden particionar.

Existen tres tipos de particiones principales:

- **Primaria**. Suelen utilizarse para instalar los sistemas operativos aunque también los sistemas operativos pueden instalarse en una partición lógica. Como mucho un disco puede tener hasta 4 particiones primarias.
- **Extendida**. Las particiones extendidas pueden albergar particiones lógicas.
- **Lógica**. Son particiones "secundarias" por así decirlo. Su finalidad en muchas ocasiones es almacenar información.

Las particiones extendidas son necesarias, porque si no un disco solamente podría tener 4 particiones.

En el particionamiento se siguen una serie de reglas y limitaciones que se van a ver a continuación:

- **Regla 1**: un disco solo puede tener hasta 4 particiones primarias.
- **Regla 2**: las particiones extendidas cuentan como si fueran particiones primarias.
- **Regla 3**: no puede existir más de una partición extendida.
- **Regla 4**: dentro de una partición extendida pueden existir una o varias particiones lógicas.

PRÁCTICA 4.1

Pregunta: ¿Puedo tener un disco con 2 particiones primarias y 2 extendidas?

Respuesta: No, según la regla 3 no puede existir más de una partición extendida.

Pregunta: ¿Puedo tener un disco con 2 particiones primarias y 5 lógicas?

Respuesta: No, puesto que no existe ninguna partición extendida.

Pregunta: ¿Puedo tener en un disco 3 particiones primarias, 1 partición extendida y 4 particiones lógicas?

Respuesta: Sí, siempre que las particiones lógicas estén dentro de la partición extendida.

Pregunta: ¿Puedo tener en un disco 7 sistemas de archivos diferentes o repetidos?

Respuesta: Sí. El disco anterior podría tener todos esos sistemas.

Para particionar o modificar las particiones de un disco es preciso utilizar una herramienta de particionado. En sistemas Linux, GParted es una de las herramientas más utilizadas.

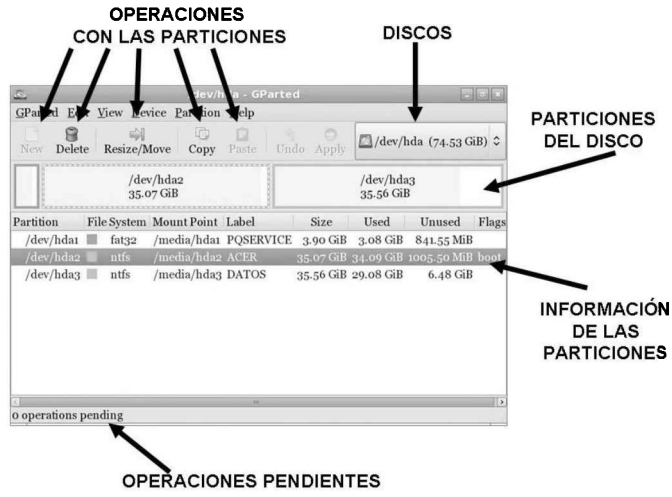

Figura 4.2. *Utilidad de particionamiento GParted*

4.2.1.1 Las particiones activas

Las particiones primarias son las utilizadas para instalar los sistemas operativos. Si un equipo no tiene ninguna partición activa, al arrancar dará un fallo. El sistema operativo de la partición activa será el que se cargue al arrancar desde el disco duro.

4.2.1.2 El sector de arranque

Un disco se compone de un sector de arranque y una serie de particiones y opcionalmente espacio sin particionar.

El sector de arranque es el primer sector del disco (cabeza 0, cilindro 0 y sector 1). Dentro de él está la tabla de particiones y el Master Boot o gestor de arranque. Este programa lee la tabla de particiones y cede el control al sector de arranque de la partición activa. Como se ha dicho antes si no hay partición activa, el equipo da un error al arrancar.

Figura 4.3. *Estructura del MBR*

El sector de arranque tiene 512 *bytes* (446 + 64 + 2 = 512) como se puede observar en la figura anterior.

4.1.2 SISTEMAS NAS Y SAN

Los servidores de almacenamiento masivo suelen emplearse en redes de área de almacenamiento concebidas para conectar servidores y arrays de discos. Existen distintas tecnologías para dar soporte a este tipo de redes como veremos a continuación.

Los Sistemas de Almacenamiento son fundamentales en los sistemas informáticos ya que serán los encargados de dar soporte a toda la información con la que trabajan. Existen diferentes tecnologías de almacenamiento entre las que destaca **NAS**.

4.2.2.1 Los sistemas NAS

El servidor NAS (*Network Attached Storage*) es un tipo de servidor de almacenamiento que da servicio a través de una red mediante el uso de protocolos estándar de comunicaciones como TCP/IP.

Es un modelo válido y necesario para sistemas que impliquen crecimiento y tamaño (escalabilidad).

Requieren altas prestaciones en accesos a disco y comunicaciones de red.

Emplean array de discos o *disk array*, conectados a la red de forma que todos los datos del sistema están asociados al dispositivo NAS y pasan por el mismo, constituyendo un punto sensible a fallos. Son también sistemas de alta disponibilidad.

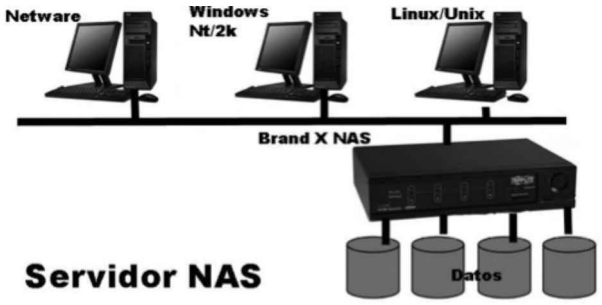

Figura 4.4. *Servidor NAS*

Muchos sistemas NAS cuentan con uno o más dispositivos de almacenamiento para incrementar su capacidad total. Normalmente, estos dispositivos están dispuestos en **RAID** (*Redundant Arrays of Independent Disks*).

NAS es muy útil para proporcionar el almacenamiento centralizado a ordenadores clientes en entornos con grandes cantidades de datos.

NAS puede habilitar sistemas fácilmente y con bajo costo con balance de carga, tolerancia a fallos y servidor web para proveer servicios de almacenamiento. El crecimiento del mercado potencial para NAS es el mercado de consumo donde existen grandes cantidades de datos multimedia.

El precio de las aplicaciones NAS ha bajado en los últimos años, ofreciendo redes de almacenamiento flexibles para el consumidor doméstico con costos menores de lo normal, con discos externos USB.

4.2.2.2 Los sistemas SAN

Los servidores SAN (*Storage Area Network*) son servidores de almacenamiento que dan servicio a través de una red y emplean acceso a través de fibra óptica no admitiendo enrutamientos.

Se emplean en sistemas de prestaciones muy elevadas, con alta velocidad de datos y crecimiento casi ilimitado (hasta 16 millones de dispositivos) y en acceso a discos y comunicaciones. Además permiten una escalabilidad y crecimiento más sencillos que en NAS pero más caros. Admiten grandes distancias de hasta 10 Km entre equipos.

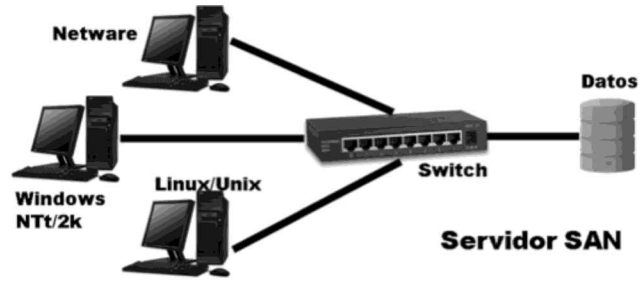

Figura 4.5. *Servidor SAN*

4.2.2.3 Gestión de volúmenes lógicos

La gestión de volúmenes lógicos es una solución para virtualizar el almacenamiento. Con la gestión de volúmenes lógicos se cambia el sistema tradicional de discos y particiones a una estructura más flexible.

Linux ofrece un gestor de volúmenes lógicos llamado Logical Volume Manager (LVM). LVM fue escrito originalmente por Heinz Mauelshagen en 1998 y ofrece las características típicas de cualquier gestor de volúmenes lógicos.

Antes de comenzar a ver la gestión de volúmenes lógicos hay que tener claros ciertos conceptos:

- **Número de unidad lógica o LUN (Logical Unit Number)**. Un LUN es un número que identifica una unidad lógica utilizado por el protocolo SCSI. Generalmente suele utilizarse para referirse a un disco lógico creado en un SAN aunque también se utiliza en unidades de cinta.
- **Volumen físico (VF) o Physical Volume (PV)**. Un volúmen físico puede ser un disco, una partición de un disco, un LUN, un dispositivo RAID (ya sea por software o por hardware). En grandes servidores no es raro ver volúmenes físicos como un grupo de discos conectados a una tarjeta controladora en RAID, ya sea para aumentar el rendimiento, para aumentar la seguridad o para ambas cosas. Generalmente en servidores no suelen utilizar RAID software sino RAID hardware.

- **Área física (AF) o Physical Extent (PE)**. Un volumen físico está dividido en áreas físicas. Generalmente todas estas áreas físicas tienen el mismo tamaño aunque en algunos sistemas el tamaño puede ser variable.
- **Área lógica (AF) o Logical Extent (LE)**. Un área lógica generalmente se corresponde con un área física. En el caso en que un área lógica se corresponda con varias áreas físicas de distintos volúmenes físicos se está creando una especie de redundancia.

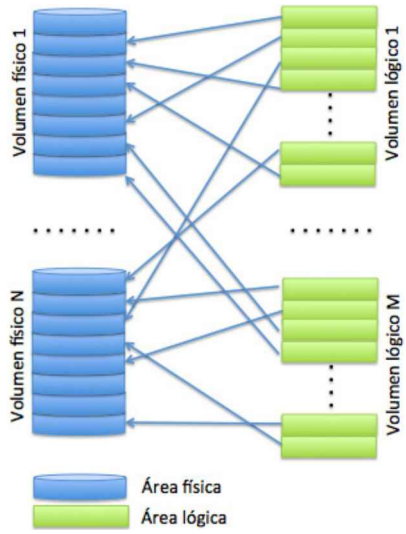

Figura 4.6. *Relación entre volúmenes físicos y lógicos*

Mapeo entre áreas físicas y lógicas
Generalmente el mapeo es uno a uno, esto quiere decir que un área física se corresponde a un área lógica. Con *mirroring* o RAID en espejo un área lógica se corresponderá con dos o más áreas físicas de discos separados. De esta forma, un grupo de áreas físicas puede corresponderse a un grupo de discos en RAID 1.

- **Volumen lógico (VL) o Logical volume (LV)**. Un volumen lógico es un grupo de áreas lógicas. Estos volúmenes lógicos funcionan como una partición de un disco duro. Se pueden montar sistemas de archivos en ellos o utilizarlos como partición de *swap*. Cada volumen lógico representa un espacio consecutivo de almacenamiento (aunque físicamente no esté contiguo).
- **Grupo de volúmenes (GV) o Volume Group (VG)**. Se trata de un grupo de varios volúmenes físicos que se corresponden con un conjunto de áreas físicas, áreas lógicas y volúmenes lógicos. Estos siempre pertenecen a un conjunto de volúmenes y no pueden intercambiarse entre distintos grupos de volúmenes si existieran. Los conjuntos de volúmenes se utilizan para realizar gestiones administrativas independientes sobre ellos como ponerlos *online* u *offline*.

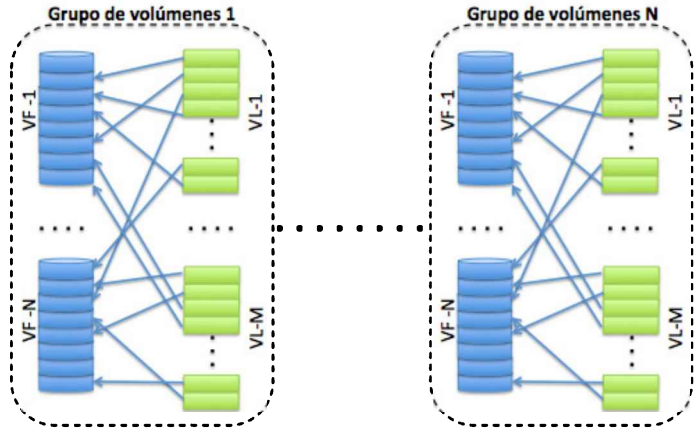

Figura 4.7. *Conjunto de grupos de volúmenes*

Crecimiento de volúmenes lógicos
Los volúmenes lógicos pueden agrandarse y encogerse cogiendo más áreas físicas o dejándolas en el grupo de áreas físicas no asignadas. Al no tener que estar las áreas físicas contiguas, un volumen lógico no tiene que moverse ni nada por el estilo.
Cuando se modifica el tamaño de un volumen lógico hay que cambiar también el tamaño del sistema de ficheros que reside sobre él. Existen sistemas de ficheros que pueden ajustarse on-the-fly como el ext3 y ext4 y de esa manera no se interrumpen los procesos que están almacenados en él.

- **Instantáneas o snapshots**. Las instantáneas lo que hacen es tomar como una fotografía de un volumen lógico en un momento dado. Haciendo un *snapshot* lo que se obtiene es una réplica del volumen lógico. Las instantáneas son muy útiles cuando se quiere hacer por ejemplo una actualización de un sistema operativo o un *backup* de una base de datos. Si se quiere hacer una actualización de un sistema operativo se toma una instantánea, se monta y se prueba la nueva actualización. Si no gusta el resultado se desmonta y se vuelve a montar el volumen lógico original. En versiones anteriores de LVM, las instantáneas eran solamente de lectura, actualmente las instantáneas son de lectura/escritura.

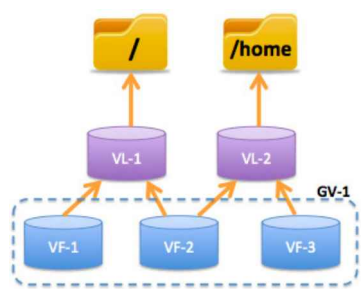

Figura 4.8. *Configuración de un gestor de volúmenes lógicos*

El montaje de un gestor de volúmenes lógicos no es un procedimiento muy complejo, simplemente basta con seguir los siguientes pasos:

- Primero se borran las tablas de particiones y se inicializan las particiones de los discos a utilizar. De esa manera se crean los volúmenes físicos (VF).
- Una vez creados los volúmenes físicos se crea el grupo de volúmenes (GV) que contendrá los volúmenes físicos inicializados anteriormente.
- El siguiente paso es crear los volúmenes lógicos, para ello se incluirán en los volúmenes lógicos el grupo de volúmenes o los volúmenes físicos que sean requeridos.
- Por último, habrá que crear un sistema de ficheros sobre el volumen lógico y montarlo para que esté disponible.

4.2.3 ACCESO PARALELO

En un sistema operativo multitarea generalmente existen muchos procesos ejecutándose al mismo tiempo. Eso implica que la probabilidad de que un fichero sea accedido por varios procesos es alta. Si todos los procesos se coordinasen a la hora de acceder a un mismo fichero no habría problemas pero como se sabe que esto no existe, el sistema operativo debe crear mecanismos para permitir un uso eficiente en el acceso a los ficheros.

Imaginemos por qué es importante la coordinación en el acceso concurrente a ficheros:

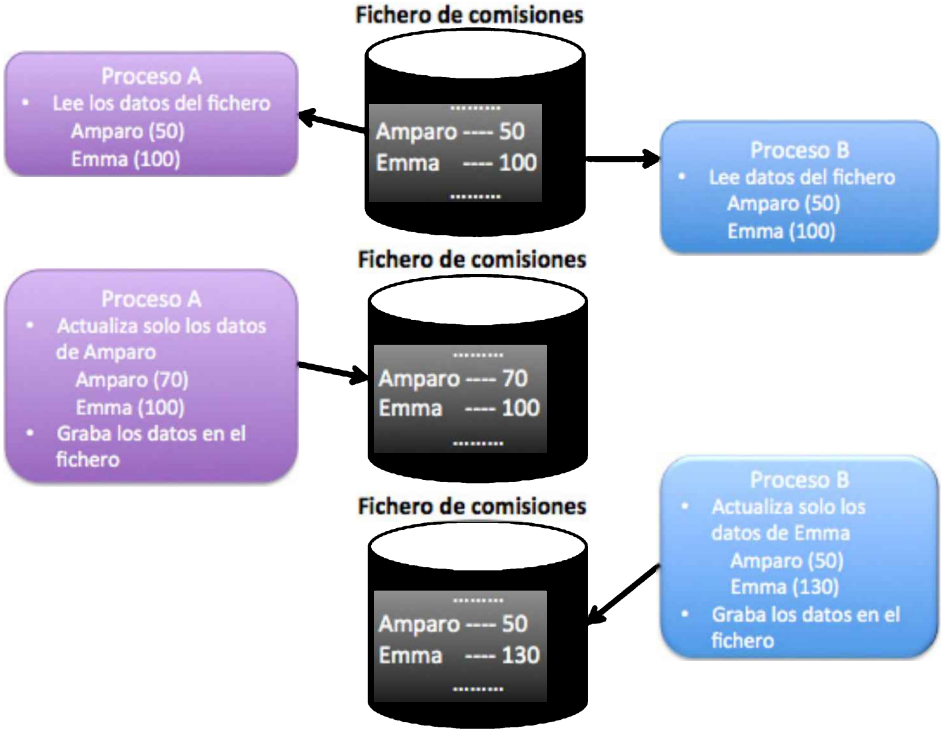

Figura 4.9. *Problema del acceso concurrente a ficheros*

Imaginemos que en una compañía existe un sistema de incentivos según el volumen de ventas que hagan sus comerciales. Al comienzo del día Amparo lleva ganados 50 euros en comisiones y Emma 100. Amparo realiza una venta y existe un proceso en el sistema para actualizar sus comisiones (ha ganado 20 euros más en comisiones) mientras que en ese mismo instante Emma realiza otra venta y automáticamente incrementa sus comisiones en 30 euros. Se lanzan dos procesos de actualización de comisiones de tal manera que el primero actualiza los datos de Amparo mientras que el segundo actualiza los datos de Emma.

Al terminar el segundo proceso, se observa que los datos de Amparo no quedan actualizados puesto que el proceso B trataba con unos datos del fichero no actualizados y por lo tanto una persona perdería su comisión.

Una solución a todo puede ser implementar bloqueos en el acceso a los ficheros. El proceso A como va a actualizar los datos del fichero de comisiones tendrá que bloquear el fichero en exclusiva a la vez que procede a hacer la actualización de las comisiones de Amparo. Mientras tanto, el proceso B espera a que el proceso A desbloquee el fichero y una vez ocurrido esto actualizar los datos de Emma.

Los sistemas operativos tienen distintos mecanismos para proteger los datos en el acceso simultáneo a ficheros como pueden ser:

- Utilizar bloqueos para todo el fichero o partes de un fichero. Estos bloqueos pueden ser meramente informativos (como sucede en Unix o Mac OS®) o ser forzosos (como ocurre en Windows®).
- Deshabilitar la escritura, borrado o actualización de ficheros.
- Utilizar mecanismos de control de accesos para que no todos los usuarios del sistema puedan hacer cualquier operación sobre el fichero.
- Utilización de ficheros de bloqueo. En ocasiones, aplicaciones y programas utilizan ficheros de bloqueo para el control de acceso a recursos (generalmente ficheros con extensión *lock*). Cuando un proceso va a acceder a algún recurso u archivo, crea un fichero de bloqueo para avisar al resto de procesos de que ese recurso está siendo utilizado y lo borra cuando deja de utilizarlo.

4.2.3.1 ¿Puede realizarse un backup mientras se están haciendo accesos concurrentes a ficheros?

Igual que en las bases de datos existen los *snapshots* o fotografías del estado de unos datos en un momento determinado, algunos sistemas operativos ofrecen un mecanismo de *snapshot* de volumen. Esto quiere decir que el sistema operativo realiza como una especie de fotografía de los ficheros del volumen en un momento determinado, para que en el *backup* no se vayan grabando las actualizaciones de ficheros realizadas mientras que dure el *backup*.

4.2.4 PROTECCIÓN RAID

RAID es una tecnología que generalmente consiste en duplicar la información de un sistema en varios discos con el objetivo de que en el caso de fallo de uno de ellos, el sistema pueda seguir funcionando sin pérdida de información.

4.2.4.1 ¿Qué es la tecnología RAID?

RAID significa *Redundant Array of Inexpensive Disks* (matriz redundante de discos económicos). Es una matriz de discos compuesta de dos o más discos, los cuales no tienen por qué ser caros (podemos utilizar la tecnología SATA), y organizados de una forma "redundante", lo cual indica que la información está repetida de alguna forma.

La repetición de esta información lo que implica es seguridad. Seguridad de que, en caso de que se estropee algún disco, la información no se pierda, pues está repetida. En el caso de fallo, el sistema puede seguir funcionando (en caso de que esté replicado por completo), cosa que no ocurre en un sistema en el que la información no esté repetida.

Los sistemas RAID se pueden implementar mediante:

- **Software**: es el sistema más económico, pero el menos eficiente. Normalmente se suele utilizar RAID para salvaguardar solamente los datos. El establecer RAID por software para salvaguardar el sistema operativo, por regla general, ralentiza mucho el sistema y los administradores no suelen optar por esta opción. No todos los sistemas operativos soportan RAID (Linux sí lo soporta y Windows Server® también).
- **Hardware**: es la forma más lógica de implementar RAID. De siempre se ha venido utilizando una tarjeta controladora conectada mediante un bus SCSI a discos SCSI con prestaciones mucho más altas que los discos normales (10.000-15.000 RPM frente a las 7.200 de los discos de un equipo de sobremesa), aunque actualmente existen controladoras integradas en las placas base que permiten realizar los métodos RAID más comunes con discos SATA, lo cual se convierte en algo más fácil de configurar y más económico.

Tipo de RAID	Ventajas	Limitaciones
RAID por software.	Bajo coste. Solo se requiere un sistema operativo correctamente configurado.	Mucho menos eficiente que el RAID por hardware. Bajo rendimiento del sistema.
RAID por hardware.	Más eficiente que el RAID por software. Menos carga de trabajo para el microprocesador.	Más costoso, aunque con una placa base que incorpore RAID los costes se reducen bastante (y también el rendimiento si es fake-RAID).

Tabla 4.1. *Ventajas y limitaciones de los tipos de RAID software y hardware*

4.2.4.2 RAID como complemento o no de las copias de seguridad

El administrador o responsable de sistemas que piense que RAID sustituye por completo las copias de seguridad está bastante equivocado.

RAID nos da la tranquilidad de que, ante cualquier fallo hardware, se puede recuperar el sistema de una manera ágil y rápida. De todas formas no previene sobre un borrado accidental de los datos, una corrupción de ficheros u otra desgracia parecida.

Por lo tanto, se deberán realizar copias de seguridad tanto si se utiliza RAID como si no se utiliza. RAID lo que da al administrador es una seguridad mayor ante cualquier posible fallo del hardware y la certeza de que el sistema puede seguir funcionando en menos tiempo.

4.2.4.3 Tipos de RAID

RAID 0 (*striping* o *duplexing*)

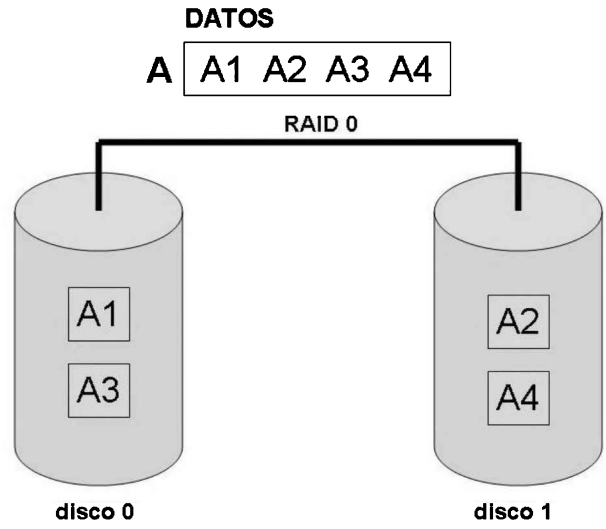

Figura 4.10. *Esquema del RAID 0*

Striping se puede traducir como entrelazado. En este caso no hay implementado ningún mecanismo de seguridad. La información se reparte en bloques por todos los discos que formen parte del *stripe*.

Se necesitan como mínimo dos discos.

El objetivo es aumentar el rendimiento, pues escribir o leer a la vez en varios discos hace que aumente la velocidad de escritura y lectura.

No es aconsejable utilizar este sistema cuando un fallo en el sistema o una parada prolongada del mismo pueden representar un problema grave.

RECUERDA

La pérdida o error en un disco de RAID 0 implican la pérdida de la información en el sistema. Esta pérdida será definitiva salvo que la información se encuentre en algún *backup*.

RAID 1 (*mirroring*)

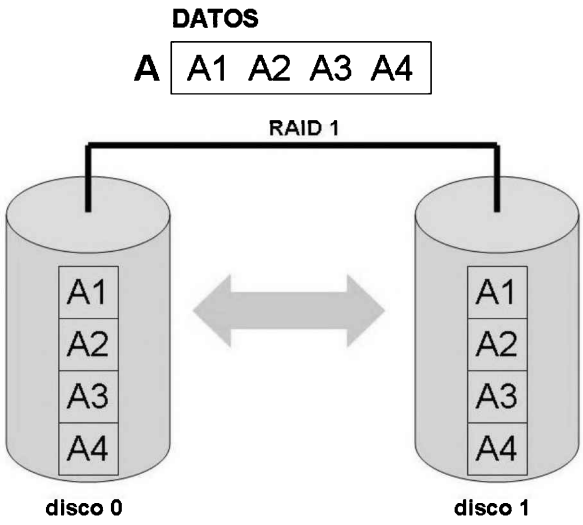

Figura 4.11. *Esquema RAID 1*

RAID 1 o *mirroring* (discos en espejo) consiste en la duplicidad de los datos. Por cada disco presente en el sistema se tiene otro con la misma información de tal manera que, cuando un disco falla, el sistema puede seguir funcionando dado que la información permanece duplicada.

RECUERDA

Se pueden utilizar en RAID 1 discos de diferentes capacidades o velocidades. Al final la pareja de discos tendrá la velocidad del menor y la capacidad del menor.
Aunque esto es posible, se aconseja utilizar discos exactamente iguales, de esa forma se evitarán problemas y se maximizará el rendimiento.

La sobrecarga u *overload* del sistema siempre será del 50%. Por cada disco se tendrá un disco extra, por lo tanto el sistema tendrá 2*n discos.

El rendimiento de lectura aumenta (hasta el doble como máximo) porque pueden leerse varios discos a la vez.

El rendimiento de escritura permanece constante.

RAID 2 (*bit striping + Hamming code*)

Utiliza una división de la información en bits por todos los discos del *stripe* y el código Hamming para la recuperación de errores.

RAID 2 no se utiliza debido a que existen sistemas más avanzados y eficientes como RAID 4 o 5.

RAID 3 (*byte striping* + *paridad*)

Utiliza una división de la información en *bytes* por todos los discos del *stripe* y la paridad para la recuperación de errores.

IMPORTANTE

La paridad
La paridad es una información adicional que se calcula antes de escribir los datos en disco. Si falla algún disco, la información se reconstruye con ayuda de la paridad. En el caso de que el disco que falle sea el de la paridad, la paridad se volverá a recalcular en el disco de repuesto.

La ventaja frente al RAID 2 es que la paridad ocupa menos que el código Hamming.

RAID 3 no se utiliza debido a que existen sistemas más avanzados y eficientes como RAID 4 o 5.

RAID 4 (*striping* + paridad)

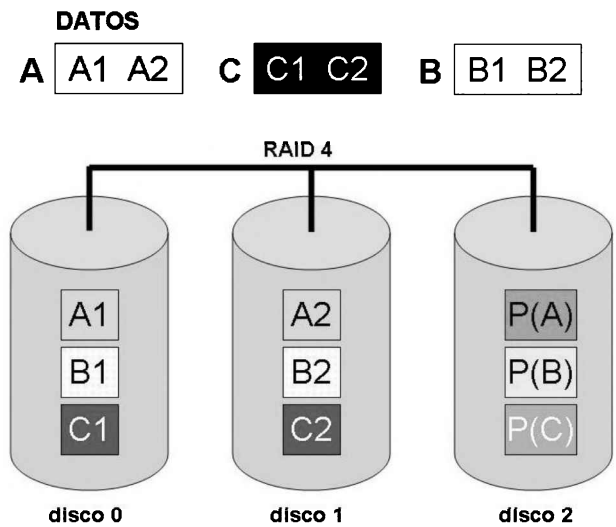

Figura 4.12. *Esquema del RAID 4*

Igual que RAID 3 pero en vez de distribuir la información por los discos por *bytes*, la distribuye por sectores. La paridad la almacena en un disco del *stripe* aparte de los datos.

Tiene un rendimiento mayor que RAID 2 y 3.

Necesita como mínimo 3 discos (2 para datos y 1 para paridad).

RAID 5 (*striping* + paridad distribuida)

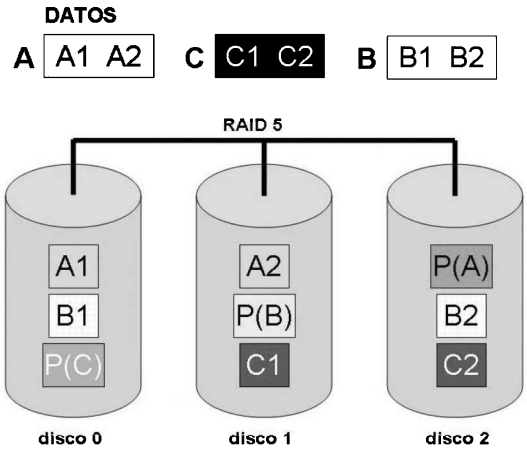

Figura 4.13. *Esquema del RAID 5*

RAID 5 funciona igual que RAID 4, pero distribuye la paridad por cada uno de los discos del *stripe* o matriz.

Tiene un rendimiento de escritura mucho mayor al escribir la paridad de forma distribuida.

Se necesitarán como mínimo 3 discos. Con 4 discos el *overload* o sobrecarga es de un 25% (por ejemplo, si tenemos 4 discos de 100 GB cada uno, el tamaño máximo disponible total de datos será de 300 GB).

Sobrevive al fallo de un disco (no al de dos).

Es el RAID más utilizado al ofrecer un mejor equilibrio coste-rendimiento-protección. Dado que RAID 5 es una solución mejorada a los sistemas RAID 2, 3 y 4, es fácil entender por qué estos últimos no se utilizan y prácticamente no se habla de ellos.

Tipo de RAID	N.º mínimo de discos	Ventajas	Limitaciones
RAID 0	2	Rendimiento.	No existe protección de datos.
RAID 1	2	Buena protección de datos y alto rendimiento.	Alto coste. Se necesita duplicar el número de discos (2*n). El *overload* es elevado, siempre un 50%.
RAID 5	3	Mejor relación rendimiento/precio. Necesita n+1 discos con un mínimo de 3 discos. Con 4 discos el *overload* es solo del 25%, mientras que con RAID 1 sería del 50%.	Escritura más lenta que con RAID 0 o 1, dado que el sistema tiene que calcular la paridad para cada dato que escribe.

Tabla 4.2. *Ventajas y limitaciones de los distintos sistemas RAID*

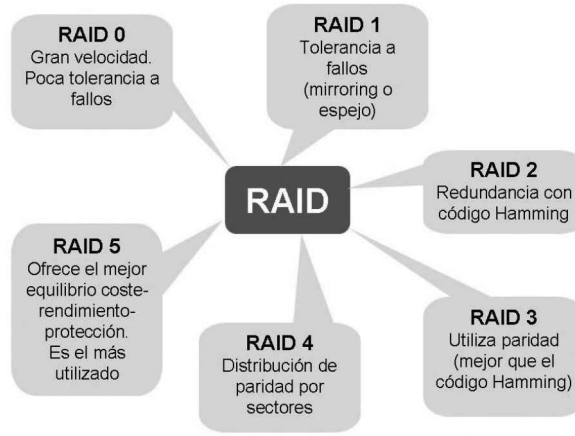

Figura 4.14. *Características principales de los distintos tipos de RAID*

4.2.4.4 Sistemas RAID anidados

En este caso hay que tener en cuenta el orden en que se enumeran los RAID. RAID 0+1 quiere decir que a un conjunto de discos en RAID 0 se le aplica RAID 1 (*mirroring*).

RAID 0+1

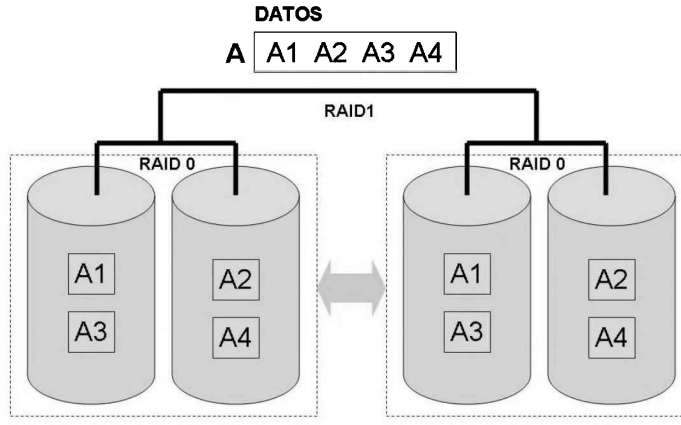

Figura 4.15. *Esquema del RAID 0+1*

RAID 0+1 corresponde a implementar un *stripe* y duplicarlo en espejo.

Se necesitarán como mínimo 4 discos.

La sobrecarga u *overload* es de un 50%.

RAID 10 (1+0)

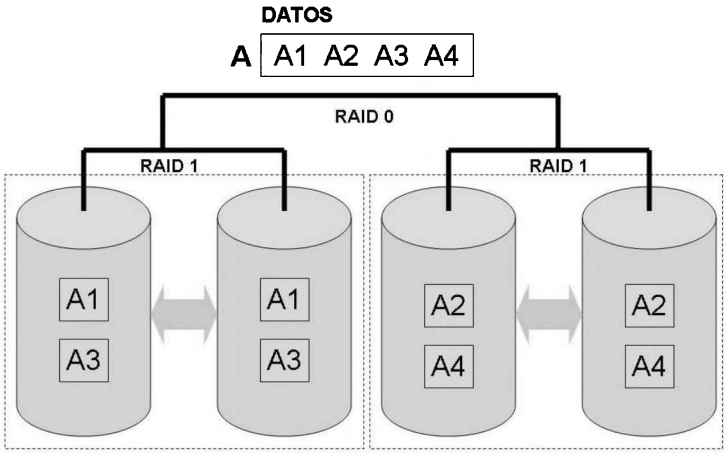

Figura 4.16. *Esquema del RAID 10*

RAID 10 corresponde a implementar varios espejos y luego realizar con ellos un *stripe*.

Se necesitarán como mínimo 4 discos.

La sobrecarga u *overload* es de un 50%.

Es una configuración más usada que RAID 0+1 puesto que es ligeramente más tolerante a errores. Permite múltiples fallos siempre que se produzcan en un espejo distinto.

Los niveles RAID más utilizados son los siguientes:

- RAID 0, 1 y 10 para equipos normales, *workstations* y servidores pequeños.
- RAID 5 y 50 para grandes servidores. También es común ver RAID 5 en equipos normales o servidores más pequeños.

4.3 POLÍTICAS DE SALVAGUARDA

La supervivencia de una empresa está sujeta a multitud de factores económicos, demográficos, de mercado y otros que son propios de su negocio pero también a otros muchos que son ajenos a esta. Es responsabilidad de la empresa disponer de las medidas que le permitan desarrollar su actividad en condiciones normales ante situaciones adversas.

En grandes empresas suele existir un documento llamado plan de continuidad de negocio, que refleja las capacidades, recursos y procedimientos de la empresa de cara a prevenir los efectos negativos sobre su negocio de riesgos o situaciones externas no controlables

Dicho plan se debe apoyar en tres estrategias principales: estrategias de prevención, de mitigación y de recuperación. Para garantizar esa continuidad de negocio hay que minimizar lo que se llama el tiempo de recuperación y el punto de recuperación. Esto se consigue mediante determinadas estrategias como la copia de datos en centros diferentes, el uso de conexiones de alta velocidad entre ellos o el empleo de una infraestructura paralela capaz de absorber la actividad del sistema ante incidencias.

En ese plan de continuidad existen diferentes elementos a contemplar en cuanto a la seguridad física aunque básicamente se pueden agrupar en:

- Control de acceso a la sala.
- Protección del sistema de cableado estructurado.
- Precauciones antiincendio.
- Sistemas de control ante inundaciones.

4.3.1 SALVAGUARDA FÍSICA Y LÓGICA

Los riesgos físicos pueden dividirse en riesgos naturales como aquellos procedentes del entorno natural (hundimientos, daños por viento, descargas eléctricas atmosféricas, nieve y hielo, deslizamiento del suelo, inundaciones o terremotos), y riesgos de vecindad, procedentes del entorno creado por el hombre (por proximidad de equipamientos o sistemas, transportes, uso de servicios públicos como luz, gas, agua o riesgos sociopolíticos como actos vandálicos o atentados). Todos estos riesgos que se han descrito pueden reducirse partiendo de la elección de un lugar o localidad apropiada.

Muchas de las medidas tomadas para garantizar la seguridad contra factores ambientales también pueden ser utilizadas para prevenir acciones humanas deliberadas o accidentales.

Sin embargo, la medida física más efectiva que se puede tomar para prevenir la intervención humana es la de ubicar la tecnología dentro de sitios seguros bajo llave, para restringir la entrada a edificios y lugares solo al personal autorizado mediante el control de acceso empleando los técnicamente llamados cerrojos.

Dentro de los cerrojos tenemos convencionales, operados por códigos de acceso, con bandas magnéticas, biométricos (reconocen rasgos físicos, como las huellas dactilares, de la mano o la retina) o que requieren una combinación de dos o más de estos dispositivos.

La vigilancia es otro mecanismo de seguridad. Se pueden utilizar guardias de seguridad para controlar el acceso a un recinto. Los guardias pueden utilizar cámaras de seguridad para monitorear distintas áreas de acceso. Se pueden emplear sensores (de movimiento o de temperatura, por ejemplo), para un control pasivo de la actividad y activar alarmas en caso de riesgo.

Si la seguridad fija es muy costosa, se pueden utilizar servicios de seguridad a petición expresa (llamada telefónica) a un coste más reducido que pueden patrullar periódicamente las instalaciones y atender llamadas de emergencia. Se pueden instalar sistemas que no solo activen alarmas a nivel local sino también en sitios remotos como una estación de policía o un puesto de seguridad.

4.3.1.1 Salvaguarda lógica

La salvaguarda lógica consiste en la aplicación de barreras y procedimientos que resguarden el acceso a los datos y solo se permita acceder a ellos a las personas autorizadas para hacerlo.

Los objetivos que se plantean serán:

- Restringir el acceso a los programas y archivos a usuarios sin los permisos correspondientes.
- Asegurar que los usuarios puedan trabajar sin una supervisión exhaustiva y no puedan modificar los programas ni los archivos que no les correspondan.
- Asegurar que se estén utilizando los datos, archivos y programas correctos en y por el procedimiento correcto.
- Controlar que la información transmitida sea recibida solo por el destinatario al cual ha sido enviada y no a otro.
- Comprobar que la información recibida sea la misma que ha sido transmitida.
- Que existan sistemas alternativos secundarios de transmisión entre diferentes puntos.
- Que se disponga de soluciones de emergencia para la transmisión de información.

Estos controles pueden implementarse en el sistema operativo, sobre los sistemas de aplicación, en bases de datos o en un paquete específico de seguridad.

Constituyen una importante ayuda para proteger al sistema operativo de la red, al sistema de aplicación y demás software de la utilización o modificaciones no autorizadas, para mantener la integridad de la información (restringiendo la cantidad de usuarios y procesos con acceso permitido) y para proteger la información confidencial de accesos no autorizados.

Los sistemas, además de implementar todo tipo de mecanismos para securizar el sistema deberán complementarlo con clonaciones o *backups* para evitar que los datos que se están protegiendo se pierdan por un error fortuito o un fallo de hardware o software.

4.3.1.2 Las copias de seguridad

Las políticas de respaldo o salvaguarda se basan fundamentalmente en las copias de seguridad o *backups*. No obstante, existen algunas otras alternativas como pueden ser la clonación o los puntos de restauración de Windows®.

Una copia de seguridad es hacer una duplicación de todo o parte del sistema.

En caso de un fallo en el sistema (borrado accidental, rotura de un disco, fallo del sistema operativo o alguna aplicación, etc.) se procederá a ejecutar el proceso de restauración del mismo. Este proceso permite dejar el sistema en el estado del mismo momento en el que se hizo el *backup*.

¿Cuándo se programan los *backups*?
Los *backups* generalmente se programan por la noche o en el momento que el servicio no tiene demanda. En ocasiones, los *backups* hay que hacerlos mientras que el servicio está funcionando. Estos *backups* llamados en caliente son más complejos que un *backup* en frío cuando el servicio se puede parar y los datos son estáticos.

Las copias de seguridad no deberían estar muy espaciadas en el tiempo. Cuanto más tiempo exista entre copia y copia mayor será el volumen de información que se pueda perder.

RECUERDA

Cuando hay un fallo en el sistema, se pierde la información existente entre el último *backup* y el momento del fallo. Esa información es posible que no se pueda recuperar.

Dependiendo de los datos y las necesidades de copia, la organización deberá establecer una política de copias, la cual marcará qué datos tienen que ser duplicados y con qué frecuencia.

Se establecerá la denominada **ventana de copia** o *backup window*, que es la franja de tiempo durante la cual se llevará a cabo el *backup*. En ocasiones estas ventanas no son muy grandes, luego se tiene un tiempo limitado para realizar la copia de seguridad. Es por tanto que hay que reducir la cantidad de datos a copiar. Una reducción de datos a copiar puede realizarse realizando copias incrementales o diferenciales o bien utilizando hardware mucho más potente y rápido.

En ciertas organizaciones lo que se hace es aumentar de forma artificial esta ventana de copia generando un *snapshot* de los datos o imagen congelada. De esa manera, las aplicaciones están trabajando con los datos reales mientras que el proceso de copia está trabajando con la copia ficticia o *snapshot*.

Además de la ventana de copia, la organización debe establecer el **período de retención** que es el período durante el cual se guardarán los *backups*. Pasado ese período no existirán *backups* normales salvo alguno que se haya almacenado por la razón que sea.

La realización de las copias de seguridad es un procedimiento muy importante. Los datos son el bien más preciado para la organización, con lo cual hay que poner un interés especial. Las copias deberán realizarse según el plan definido por la organización cumpliendo todas las políticas de seguridad de la empresa y la legalidad vigente. El nivel de seguridad para las copias de seguridad será igual o superior al de los datos en producción.

Además, los soportes deberán almacenarse de forma que se mantengan en **condiciones óptimas de conservación**. Cuando un soporte llega al fin de su vida útil (se ha grabado un número determinado de veces), este deberá destruirse de tal manera que sea imposible recuperar la información.

4.3.1.3 Tipos de copias de seguridad

Básicamente existen 3 tipos de copias de seguridad o *backup*:

- **Total o completa**: es aquella que copia toda la información almacenada en el sistema. Desactiva el atributo de modificado a todos los archivos.
- **Incremental**: copia solo los archivos que tienen el atributo de modificado activado. Una vez realizada la copia de seguridad ese atributo se desactiva.
- **Diferencial**: es igual que la incremental, lo único que el atributo modificado no se desactiva (este atributo se desactiva cuando se haga una copia de seguridad incremental o completa).

Cuando se trabaja con archivos y estos sufren alguna modificación el atributo de modificado se activa.

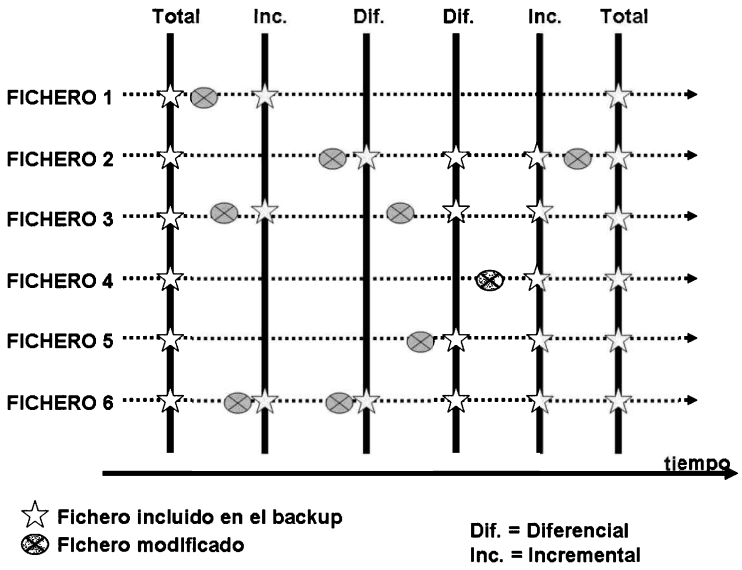

Figura 4.17. *Distintos backups en el tiempo*

En la figura anterior se puede ver cómo funcionan los *backups* totales, incrementales y diferenciales. El aspa indica que el fichero ha cambiado de contenido y la estrella marca los ficheros que se han incluido en el *backup*.

RECUERDA

El *backup* diferencial nunca desactiva el atributo modificado.

El proceso de **restauración** también llamado *restore* consistirá en:

1 Restablecer la última copia de seguridad total.

2 Posteriormente se restauran desde la más antigua hasta la más moderna aquellas copias de seguridad incrementales desde la última exportación total.

3 Por último restablecer la última copia de seguridad diferencial realizada siempre y cuando sea esta la última copia realizada (no existe ninguna copia incremental ni total posterior).

4.3.1.4 Las clonaciones

Una clonación es una copia exacta de algo. En el caso de un equipo informático será una copia exacta de un disco o una partición.

Las diferencias entre clonar un disco o una partición son las siguientes:

- **Clonación de un disco**. Una clonación de un disco permite clonar también el sector de arranque. De esa forma, si el disco tiene una avería, al recuperar la imagen en otro disco el equipo puede arrancar y funcionar sin problemas. Una clonación completa de un disco permite duplicar un equipo informático siempre que el hardware del nuevo equipo sea igual.
- **Clonación de una partición**. Solamente guarda los datos de una partición concreta pero no el sector de arranque. No permite duplicar equipos.

Cuando se tiene un disco que no contiene ningún sistema operativo desde el cual arrancar, basta con clonar las particiones y recuperarlas en otro disco y el disco contendrá los mismos datos.

Figura 4.18. *Parted Magic. Herramienta gratuita para la clonación*

4.3.2 ALTA DISPONIBILIDAD

Cuando se piensa en alta disponibilidad no necesariamente se debe pensar en hardware y software de alto coste. Existen soluciones de alta disponibilidad que se ajustan a la realidad de muchas empresas.

Una solución de alta disponibilidad va a ser aquella que permite que los sistemas de información operativos de nuestra empresa estén disponibles las 24 horas de los 7 días de la semana. Al implementar esta solución las empresas pueden contar con la seguridad de no perder negocios ni información debido a fallos en los sistemas.

La alta disponibilidad está asociada a dos términos: la fiabilidad y la disponibilidad, que en ocasiones se confunden.

4.3.2.1 Fiabilidad

Más allá del servicio que ofrezca un sistema informático, un sistema de alta disponibilidad debe ser fiable para que los usuarios puedan utilizarlo en condiciones óptimas.

El término fiabilidad se refiere a la probabilidad de que un sistema funcione normalmente durante un período de tiempo dado. A esto también se le llama continuidad del servicio.

Un fallo o problema surge cuando un servicio no funciona correctamente, es decir, que se genera un estado de funcionamiento anormal o que no se adecúa a las especificaciones.

Desde el punto de vista del usuario, un servicio tiene dos estados: apropiado (cuando satisface las expectativas) y no apropiado (cuando no lo hace).

Este tipo de problemas es atribuible a errores, es decir, a un funcionamiento incorrecto del sistema. Pero no todos los errores conducen a un fallo directo o interrupción en el servicio.

4.3.2.2 Disponibilidad

La alta disponibilidad consiste en una serie de medidas cuyo objetivo no es otro que garantizar la disponibilidad del servicio de una forma fiable, es decir, asegurar que el servicio funcione de forma correcta durante las 24 horas.

El término disponibilidad hace referencia a la probabilidad de que un servicio funcione adecuadamente en cualquier momento.

La disponibilidad se expresa con mayor frecuencia a través del índice de disponibilidad (un porcentaje), que se mide dividiendo el tiempo durante el cual el servicio está disponible por el tiempo total.

A continuación puedes comprobar ciertos valores del índice de disponibilidad evaluada anualmente:

Índice de disponibilidad	Duración del tiempo de inactividad
97%	11 días
98%	3 días
99%	3 días y 15 horas
99,9%	8 horas y 48 minutos
99,99%	53 minutos
99,999%	5 minutos
99,9999%	32 segundos

Tabla 4.3. *Índices de disponibilidad de sistemas*

¿SABÍAS QUE...?

Existe una normativa incluida en el estándar TIA-942 que determina que existe una serie de niveles de disponibilidad básicos llamados Tier I, Tier II, Tier III y Tier IV.
Es interesante profundizar en el estudio de este estándar para comprobar las diferentes tasas de disponibilidad de cada nivel, así como qué elementos son necesarios para implementar un CPD según el nivel de disponibilidad que presente.

4.3.2 *CLUSTERING* Y BALANCEO DE CARGA

Figura 4.19. *Sala con equipos en cluster. Cortesía Patrick Finnegan*

El balanceo de carga o equipos en *cluster* es una disposición de varios equipos los cuales realizan una tarea determinada compartiendo esfuerzos. En el caso de que uno no pueda hacerse cargo de la tarea, otro miembro del grupo o *cluster* se hará cargo de la misma. Generalmente el balanceo de carga se suele utilizar mucho en servidores webs, los cuales tienen que hacer frente a un número alto de peticiones de los usuarios en un momento muy concreto.

Los *cluster* tienen la característica de poder ampliar su capacidad añadiendo más equipos al grupo o cluster.

Otra característica que tienen los *cluster* es que, ante la caída de uno de los miembros del equipo, los demás equipos del *cluster* se hacen cargo de las tareas del mismo. De esa manera el servicio permanece inalterado, como mucho se notará una bajada en la eficiencia del mismo, pero nunca una paralización.

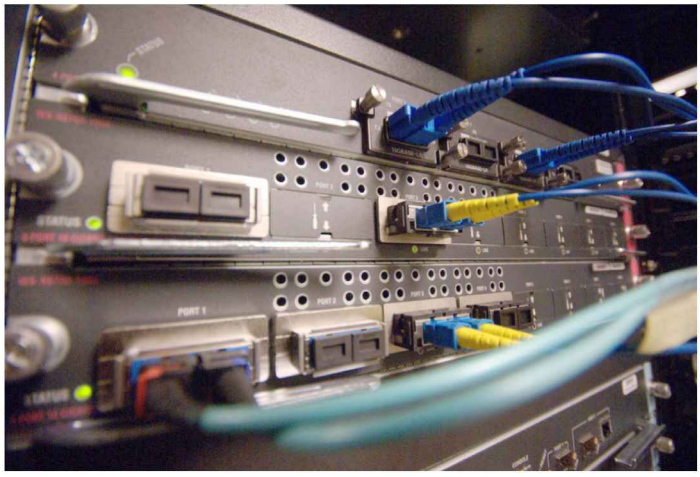

Figura 4.20. *Conexión de cluster. Cortesía Patrick Finnegan*

4.3.3.1 Diferencias entre *clustering* y *grid computing*

Aunque muchas veces se confunden estos dos términos, *clustering computing* y *grid computing* no son exactamente lo mismo. Ambos utilizan una serie de equipos para llevar a cabo una tarea o servicio determinado pero no de la misma forma. Mientras que un *clustering computing* es una misma aplicación que corre en muchos servidores con redundancia entre ellos, *grid computing* consiste en partir la aplicación en distintos módulos los cuales pueden funcionar en servidores distintos de forma independiente pero sin redundancia.

Mientras que los *grid* son conjuntos de servidores independientes unos de otros, los *cluster* suelen ser grupos de servidores con gran dependencia unos de otros. Los miembros del *cluster* hacen el mismo trabajo mientras que los miembros del *grid* no. El objetivo de un *cluster* es que un servicio o aplicación continúe a pesar de tener algún equipo un fallo de software o hardware, mientras que un grid se utiliza en aquellos casos con gran complejidad en los que un *cluster* no es capaz de solucionar el problema. Por ejemplo, en proyectos de investigación suelen emplearse *grids* para resolver problemas en los que hay que emplear una gran potencia de cálculo.

Muchas veces se utiliza *grid computing* para resolver problemas de forma distribuida. Una serie de servidores trabajan en paralelo sobre problemas independientes y luego ese trabajo es aunado por otros servidores. En realidad, el *grid computing* es como utilizar una supercomputadora o *mainframe* pero distribuyendo su capacidad de cálculo sobre una red de equipos menos potentes.

4.3.4 INTEGRIDAD DE DATOS Y RECUPERACIÓN DE SERVICIO

A continuación se estudiarán más a fondo los conceptos de integridad de datos y recuperación del servicio.

4.3.4.1 Integridad de datos

Siempre se dice que los datos en una empresa son el bien más preciado, por lo tanto el velar por la exactitud y la fiabilidad de los mismos es una de las funciones principales de toda empresa. Los datos pueden ver comprometida su integridad de muchas maneras, por ejemplo el factor humano puede ser una de ellas. Muchas veces los datos son accedidos por muchos miembros de la organización y pueden producirse de forma intencionada o no, cambios de los datos originales. Otras veces el origen de estas variaciones tiene su origen en fallos en los sistemas de transferencia o almacenamiento.

Cuando se diseña cualquier sistema, es necesario identificar las posibles amenazas que puedan ocurrir y partiendo de las mismas establecer todos los posibles mecanismos para que no ocurran y los datos queden preservados frente a las mismas. Una de las piezas fundamentales donde se implementa la integridad de datos es en las bases de datos. Al ser un sitio donde se centralizan todos los datos de la organización es el lugar idóneo donde implementar todos estos mecanismos de protección.

No solamente habrá que implementar mecanismos de protección, sino mecanismos que permitan reconocer los posibles problemas que puedan estar ocurriendo. En muchas ocasiones, el *malware* es el responsable de muchos problemas sobre la integridad de los datos, para ello habrá que instalar antivirus, cortafuegos u otras herramientas para contrarrestarlo.

4.3.4.2 Plan de recuperación de servicios

Un plan de recuperación del servicio forma parte del plan de contingencias de una empresa y determina las acciones y medidas a realizar una vez que la amenaza se ha materializado. El objetivo es restaurar el servicio a tal y como estaba antes de ocurrir el desastre o a un estado que se considere mínimamente aceptable.

Generalmente para crear el plan de recuperación del servicio se debe hacer un análisis de riesgos y sus posibles impactos así como evaluar los posibles daños que se puedan producir. Tras eso habrá que detallar las medidas a tomar para poder reanudar la actividad. Las políticas de *backup* o el prever un centro de respaldo son puntos a tener en cuenta en la elaboración de este plan.

En la mayoría de ocasiones, los planes de recuperación del servicio hacen referencia a los parámetros RTO y RPO.

4.3.4.3 RTO (Recovery Time Objective) y RPO (Recovery Point Objective)

Los desastres son impredecibles y muchas veces inevitables. No obstante, los administradores de un sistema tienen que velar por prevenirlos y en el caso de que no sea posible, minimizar su impacto. Una vez ocurrida la desgracia hay que tener en cuenta dos parámetros, el RPO (*Recovery Point Objective*) y el RTO (*Recovery Time Objective*) que forman parte casi siempre de los planes de recuperación frente a desastres o planes de protección de datos. El objetivo de esos planes es el recuperar el servicio y los datos de forma rápida y óptima.

- **RTO** (*Recovery Time Objective*). Es la duración de tiempo que puede sufrir un sistema tras un desastre sin estar operativo antes de ser restablecido. Sería el tiempo en recuperar el sistema tras la notificación de caída del servicio. Generalmente existe un tiempo máximo en el que el servicio puede estar caído, que si se supera puede tener graves o inaceptables consecuencias que hagan que el negocio no pueda continuar de forma normal. Generalmente el objetivo de los responsables de sistemas es el reducir el RTO al menor tiempo posible.
- **RPO** (*Recovery Point Objective*). Es el tiempo que puede pasar un sistema caído tras un desastre siempre que no exceda los máximos establecidos en el plan de continuidad de negocio. Cuando se exceden esos máximos, la caída del servicio suele tener consecuencias negativas para la empresa. Generalmente cuando hay una caída del sistema, habrá que recuperar la pérdida de datos que haya existido durante el tiempo de la caída. En ocasiones los diseñadores del sistema están dispuestos a aceptar una pérdida de datos de un tiempo determinado (ese será el RPO). Si no se está dispuesto a perder ningún dato el RPO será cero. Mientras menor sea el RPO más habrá que aumentar la frecuencia de copias de seguridad.

4.3.5 CUSTODIA DE FICHEROS DE SEGURIDAD

Los datos en una empresa son el bien más preciado, por lo tanto su custodia es una labor necesaria e indispensable. En muchos casos, las empresas hacen mucho hincapié en la seguridad en el acceso al sistema y a los datos pero relajan los controles a la hora de custodiar correctamente los *backups* y datos importantes. Esto constituye un grave error.

La custodia de los datos tiene que tener al menos el mismo o mayor nivel de seguridad que el acceso a los mismos. Además de impedir el acceso a personal no autorizado, las cintas de *backup*, discos duros, soportes ópticos, etc., deberán conservarse de la forma más óptima preservándolos de cualquier deterioro que pueda ocurrirles. Además, la custodia de los soportes debería hacerse en un emplazamiento distinto al de los datos para evitar posibles pérdidas a causa de algún incendio, inundación, etc.

IMPORTANTE

Hay que extremar las precauciones a la hora de destruir el material confidencial. Aenor tiene una norma, la UNE-EN 15713:2010, la cual es un código de buenas prácticas para la destrucción segura del material confidencial.
Sin una destrucción adecuada del material confidencial en una empresa, la seguridad de la misma puede verse afectada.

Muchas veces, las empresas no están preparadas para una custodia de datos y recurren a un servicio de *backup* externo. El *offsite backup*, *backup* remoto o *vaulting* es una solución necesaria en muchas ocasiones. Imagínese que se quema o se inunda una oficina donde residen los datos de una empresa y las copias de respaldo. En este desastre natural se perderían los datos de la empresa con el consiguiente perjuicio para la misma.

El realizar *backups* remotos permite asegurar la protección frente a infortunios locales. Sin *backups* adicionales en otros dispositivos o lugares, se corre el riesgo de pérdida de datos.

Existen dos posibilidades para hacer *backup* remoto:

- Realizar copias locales y custodiarlas en un lugar diferente.
- Realizar copias mediante un servicio de *backup* remoto. De esa manera los datos se envían electrónicamente a una ubicación distinta.

Algunas compañías grandes realizan sus *backups* remotos ellas mismas. No obstante, siempre va a ser más barato contratar algún servicio de *backup* remoto dado que estas compañías están especializadas y ofrecen un precio mucho más económico que el hacerlo uno mismo.

En muchas ocasiones estas compañías especializadas en custodia de datos almacenan la información en minas no utilizadas, infraestructuras militares de la guerra fría, etc.

Existen múltiples razones para realizar el almacenamiento externo:

- A veces, los datos almacenados por uno mismo en otra localización están sin encriptar, lo cual incurre en una vulnerabilidad dado que cualquiera puede acceder a ellos.
- El peligro de estar expuestos a desastres naturales, incendios, etc., es mucho menor.
- Se ahorran los siguientes costes:
  - Tiempo de los empleados en realizar y verificar los *backups*.
  - Costes de los soportes.
  - Costes del hardware.
  - Espacio de almacenamiento.
- No están expuestos a virus o *hackers*.

Aparte de este tipo de propuestas, existen compañías especializadas en el *backup offline* de empresas, las cuales suelen ofrecer los siguientes servicios:

- Planificación de *backups*.
- *Versioning* de *backups* (almacenan los x últimos *backups*).
- Realización de *backups* incrementales de forma automática.
- Posibilidad de recibir los *backups* en soportes como CD o DVD.
- Encriptación de datos.
- *Backup* especializado de bases de datos.
- Planes de recuperación de datos y del servicio.
- Otros servicios.

Según la ley orgánica de protección de datos (LOPD), existen tres niveles de seguridad dependiendo de la sensibilidad de los datos con los que se trata. Dependiendo del nivel de seguridad en los que se hayan englobado los datos, las medidas de protección serán más altas.
No obstante, para niveles básicos y medios, la ley dice que hay que conservar los ficheros de forma adecuada y crear mecanismos de seguridad para que el acceso sea restringido, registrando entrada y salida de soportes y autorizándose la salida de soportes en el caso que sea oportuno. Además, será necesario tomar medidas para impedir recuperar la información de aquellos soportes que vayan a ser desechados o reutilizados.

4.4 POLÍTICAS DE SEGURIDAD

Dentro de este apartado veremos ciertos aspectos de las políticas de seguridad como el acceso restringido por cuentas de usuario, los antivirus, las auditorías de seguridad o la utilización de cortafuegos y servidores *proxy*.

4.4.1 ACCESO RESTRINGIDO POR CUENTAS DE USUARIO. PROPIEDAD DE LA INFORMACIÓN

Para el desarrollo de este apartado nos vamos a centrar en los sistemas Linux/Unix.

En prácticamente cualquier sistema informático el acceso al mismo siempre es a través de un usuario y una contraseña. Los usuarios están registrados en el sistema por medio de cuentas (de usuario). De esa manera los usuarios podrán acceder a los recursos.

Los recursos de un sistema (ficheros, procesos...) están protegidos de tal manera que solamente pueden ser accedidos por aquellos usuarios que posean los privilegios necesarios para poder utilizarlos.

Como se ha dicho, los usuarios están registrados en el sistema a través de cuentas y generalmente estas cuentas pertenecen a uno o más grupos. Los grupos son utilizados para especificar derechos de acceso comunes a un conjunto de cuentas de usuario. En prácticamente todos los sistemas es necesario que la cuenta de usuario pertenezca al menos a un grupo.

Aunque las cuentas de usuario tienen un nombre especificado por el administrador, internamente los sistemas las identifican por un identificador único (ID). En los sistemas Linux la correspondencia entre usuario e identificador se encuentra en el fichero /etc/passwd. Igualmente ocurre para los grupos donde cada grupo tiene su GID o identificador de grupo correspondiente. Son los ID y GID los que realmente utiliza el sistema para determinar los permisos de acceso.

En los sistemas multiusuario, los ficheros y los procesos siempre tienen un dueño. El usuario de un fichero puede cambiar los permisos del mismo (lectura, escritura, ejecución) e incluso puede realizar otro tipo de acciones como traspasar la titularidad del mismo a otro usuario. También los usuarios de un proceso pueden realizar acciones sobre el mismo como cambiar la prioridad del mismo, cancelarlo, etc. Obviamente los superusuarios del sistema como Administrador en Windows® y *root* en sistemas Unix pueden saltarse todas estas restricciones.

Generalmente las cuentas de usuario tienen asociadas una serie de información como puede ser:

- El nombre del usuario en el sistema.
- Contraseña en el sistema.
- UID o ID de usuario.
- GID o ID de grupo al que está asociado.
- Directorio personal.
- Otra información como *email*, configuración específica de la cuenta, nombre real del usuario, *flags* que indican si el usuario está habilitado o no, si puede acceder al sistema de forma remota, etc.

En un sistema, generalmente existen las siguientes cuentas de usuario:

- **Usuarios**. Acceden a los recursos del sistema para los que se les han concedido acceso. Generalmente los usuarios se crean con el mismo tipo de permisos. Para establecer diferentes tipos de permisos, los administradores crean grupos. Una vez creado un grupo y asignados usuarios a él, al establecer algún permiso sobre el grupo automáticamente se asigna a los usuarios que pertenezcan a él.
- **Superusuarios o administradores del sistema**. Poseen todos los derechos sobre el sistema pudiendo modificarlos en caso de necesidad.
- **Invitados**. Suelen crearse este tipo de cuentas para permitir el acceso a usuarios eventuales para los cuales no tiene sentido crear una cuenta específica de usuario. Por motivos de seguridad, las cuentas de invitados suelen protegerse con contraseña para que no haya invitados sin identificar en el sistema. Este tipo de cuentas tiene un acceso reducido al sistema.
- **Usuarios específicos**. Usuarios especiales para realizar ciertas operaciones de mantenimiento concretas como copias de seguridad, etc.

4.4.1.1 Propiedad de la información

La información de una entidad pertenece a ella dado que es un activo de la misma y es función de ella el preservarla dictando **políticas de seguridad** y velando porque dichas políticas de seguridad se cumplan.

La información deberá estar disponible a las personas necesarias. Cada persona podrá acceder solamente a la información que le haga falta para el desarrollo o cumplimiento de sus funciones y de ese mismo modo dichas personas deberán hacerse responsables de la información a la que se les ha otorgado el acceso.

Para un mejor control de la información en muchas ocasiones es necesario tener un sistema centralizado. De esa manera centralizando la información se tiene un mayor control sobre ella. Los esfuerzos de seguridad se concentrarán en solo un punto mientras que si se utiliza un sistema descentralizado los esfuerzos de seguridad tienen que dividirse y por lo tanto el nivel de riesgo aumenta.

4.4.2 IDENTIFICADOR ÚNICO DE ACCESO

Cualquier usuario que se conecte a un sistema debe ser validado e identificado. En un sistema con un mínimo de seguridad, no deberían existir usuarios que no estén identificados de forma unívoca. Esto implica que para cada usuario en el sistema exista una persona real asociada al mismo. Al estar identificados los usuarios que acceden al sistema, cada operación realizada por un usuario deberá poder imputarse a la persona física a la que corresponde dicho usuario.

En el control de acceso a los sistemas deben seguirse una serie de reglas como son:

- Una persona <-> un usuario del sistema. Debería evitarse las cuentas grupales dado que no se identifica la persona que accede al sistema.
- Los usuarios tienen asignados un identificador único (UID) de tal manera que el sistema reconocerá al usuario por dicho identificador.
- Los identificadores no pueden reutilizarse.
- No pueden existir identificadores iguales. Por este motivo, el procedimiento para asignar un identificador de usuario no debe permitir duplicados.
- Los identificadores no pueden cambiarse.
- Los usuarios externos del sistema (invitados) deberán estar correctamente identificados.
- Se deberá adoptar una nomenclatura que identifique los usuarios internos, externos, administradores, etc.
- Cualquier operación realizada sobre los recursos del sistema se imputará al identificador único.
- Los administradores son los responsables de la gestión de los usuarios.

Los systemas Single Sign-On (SSO)
SSO es un procedimiento de autenticación que sirve para centralizar las autorizaciones y autenticaciones. La ventaja de estos sistemas es que solamente hay que identificarse una vez para poder acceder a varios sistemas. Estos sistemas también se denominan sistemas de autenticación reducida. Uno de los SSO más utilizados es Kerberos mediante el cual se puede externalizar la autenticación de usuarios. Una vez registrado en el servidor Kerberos se obtiene un ticket utilizado por las aplicaciones cliente para acceso a servicios o sistemas.

4.4.2.1 Identificación en el sistema mediante certificado

La posesión de un certificado puede identificar a los usuarios en un sistema. Los certificados de usuario contienen datos relativos a la identidad del mismo y por lo tanto permiten crear mecanismos de control y autenticación más avanzados.

Los certificados siempre deben estar protegidos con una contraseña y hay que instalarlos en el sistema del usuario y sirven para poder acceder a los datos ubicados en el servidor. También los certificados pueden almacenarse en una tarjeta inteligente haciendo el procedimiento de autenticación mucho más versátil.

Muchos sistemas operativos como Linux y Windows® Server incluyen servicios de certificación ofreciendo la posibilidad de que las propias organizaciones emitan sus propios certificados.

4.4.3 PROTECCIÓN ANTIVIRUS

Cuando el equipo se bloquea, se reinicia, no arranca o va mucho más lento que de forma habitual es posible que se haya contagiado con un virus u algún otro tipo de *malware*. No es fácil saber si el sistema tiene un virus o no, de ahí su peligrosidad. Normalmente el usuario se entera de que el equipo está infectado cuando ya es demasiado tarde.

Un antivirus es un software cuyas funciones son:

- Detectar los virus
- Prevenir las infecciones de los virus
- Analizar el sistema para comprobar la presencia de virus
- Eliminar los virus detectados en el sistema

Generalmente los antivirus funcionan en base a dos tipos diferentes de técnicas:

- **Técnica de scanning**. Los antivirus tienen una base de datos con los códigos de los virus conocidos. Cuando se escanea un archivo se comprueba el código del mismo con los códigos que existen en la base de datos (llamados firmas o vacunas) y si coinciden se conocerá el nombre del virus que ha infectado el archivo y el antivirus pasará a eliminarlo o si no es posible ponerlo en cuarentena. Esta técnica de *scanning* está en desuso precisamente porque no evita que el sistema sea infectado. Solo actúa a posteriori, cuando el sistema ya ha sido infectado.
- **Técnicas heurísticas**. Actualmente los antivirus además de hacer este tipo de comprobaciones monitorizan los programas en busca de comportamientos "sospechosos" propios de virus. El problema de esta técnica es que se puede sospechar de muchos programas que precisamente no son virus.

Lo más común es encontrarse en el mercado antivirus que combinen varias de estas técnicas para proteger el sistema. Además se analiza cualquier tipo de *malware,* no solo virus (*spam, adaware, spyware*, virus…).

RECUERDA

Hay que mantener el antivirus y el sistema operativo actualizado. De esa manera estaremos protegidos contra nuevos virus y agujeros de seguridad conocidos.
Un antivirus o sistema no actualizado es un sistema vulnerable y fácilmente atacable por un virus.

4.4.4 AUDITORÍAS DE SEGURIDAD

En las auditorías de seguridad de un sistema informático se realiza un trabajo riguroso, de forma independiente y sistemática, de tal manera que se puedan obtener evidencias de que el sistema cumple o no los requisitos de seguridad para los que fue creado. Al final de la auditoría se generará un informe en el cual se evaluarán de manera objetiva los resultados de la revisión del sistema.

Existen dos tipos de auditorías, las auditorías externas y las internas.

- **Auditorías externas**. En las auditorías externas, el auditor no tiene ningún tipo de responsabilidad en la actividad que está auditando y por lo tanto se puede demostrar de una manera clara su total independencia. Algunas auditorías externas pueden ser realizadas por los mismos clientes o por empresas externas contratadas a tal efecto.
- **Auditorías internas**. Son realizadas por personal de la propia empresa. El problema de este tipo de auditorías es que la independencia del equipo auditor puede estar en cuestión porque puede estar implicado en el sistema que está auditando. No obstante, al equipo de auditoría se le suele dar la máxima independencia posible para que el resultado de las auditorías sea el más objetivo posible.

Los trabajos de auditoría en un sistema informático de una empresa son llevados a cabo por personas especialistas. Estas personas se denominan auditores informáticos. Existen auditores informáticos especializados en temas de seguridad.

Generalmente los auditores informáticos son personas en las que es característico un amplio bagaje en el ámbito de la informática, son metódicos y se caracterizan por su capacidad de observación y sentido común. Algunas características de un auditor son las siguientes:

- **Independencia**. La imparcialidad y la objetividad en las actuaciones de un auditor al igual que en las conclusiones del informe final son fundamentales. El trabajo del auditor debe evitar el sesgo y las afirmaciones o conclusiones generadas estarán basadas única y exclusivamente en la evidencia. Se podrá verificar de una forma fehaciente las afirmaciones realizadas por el auditor y en el caso de que haya utilizado algún tipo de muestreo, este hecho deberá ser reflejado de forma clara y explícita.
- **Conducta ética**. La discreción en las actuaciones, al igual que la confianza e integridad de la persona es fundamental en el perfil del auditor.
- **Competencia, tacto y talante**. Dado que la tarea de un auditor puede ser sumamente importante, el trabajo se deberá hacer de forma impecable dada la confianza depositada en él tanto por parte del cliente auditado como por terceras personas.

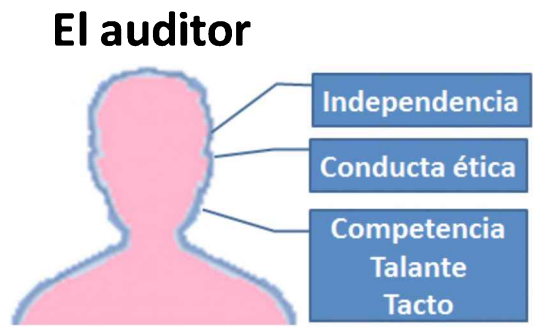

Figura 4.21. *Perfil del auditor*

- **Objetivos de las auditorías de un sistema informático**. Generalmente las auditorías de un sistema informático se centran en la seguridad del mismo y en determinar si el sistema cumple los requisitos necesarios evidenciando si existen deficiencias que subsanar o posibles mejoras a realizar para poder cumplir los objetivos propuestos. En el caso de encontrar alguna deficiencia o mejora, se deberá registrar de manera detallada en el informe de auditoría.
- **Ventajas de las auditorías**. La realización de auditorías de forma periódica en un sistema informático es fundamental. Las auditorías internas o externas ayudarán a mejorar el sistema y a corregir posibles deficiencias del mismo. La realización de auditorías hace un sistema eficaz y eficiente.

Entre otras cuestiones, las auditorías informáticas de seguridad revisarán algunos de los siguientes puntos:

- **Seguridad física**. La seguridad física tendrá como primer objetivo proteger la instalación. Se evitará que personas ajenas tengan acceso físico al sistema. Una persona ajena si accede al sistema podría por ejemplo robar los discos en los que se realiza el *backup* o robar un ordenador, etc.

Figura 4.22. *Sistema de control de accesos biométrico. Cortesía de Julián Correa*

- **Control de accesos al sistema**. Se comprobará de manera exhaustiva que el sistema de usuarios y claves es efectivo. De nada sirve tener el mejor sistema de claves posible si luego se apuntan en un cuaderno junto al ordenador.

- **Proteger la lógica del sistema**. La protección a nivel lógico de los sistemas implica una protección a:
  - **Nivel de aplicaciones**. Controlando los accesos no permitidos a las aplicaciones, verificando que se evita que desde las aplicaciones se hagan accesos indebidos a datos o a partes del sistema, controlando que las aplicaciones gestionan los datos de forma correcta, etc.
  - **Nivel de sistema operativo**. Verificando que el sistema operativo está protegido frente a todo tipo de *malware*, chequeando que el sistema tiene instaladas las últimas actualizaciones de seguridad, etc.
  - **Nivel de base de datos**. Verificando que el acceso a la base de datos es el establecido y no existe la posibilidad de acceder a los datos a no ser por las aplicaciones.
  - **Nivel de red**. Verificando que la seguridad a nivel de red está garantizada.

4.4.5 CORTAFUEGOS Y SERVIDORES PROXY

Un *firewall* o también llamado **cortafuegos** es un sistema para filtrar la transferencia de información que no cumpla unas reglas de seguridad determinadas. Los *firewalls* lo que hacen es examinar toda información entrante y saliente bloqueando aquella que no cumpla con los criterios de seguridad establecidos. Generalmente las reglas de seguridad más frecuentes son el bloqueo de ciertos puertos impidiendo la comunicación por los mismos. Además de este tipo de reglas sencillas puede crearse otras más complejas.

Actualmente, además de los cortafuegos se suelen utilizar otros mecanismos de seguridad para proteger redes y equipos de intrusos y software malintencionado. Existen dos tipos de *firewall* dependiendo del ámbito en el que se apliquen:

- **Firewall de equipo**. Su objetivo es filtrar el contenido entre el equipo y la red a la que está conectado.
- **Firewall de red**. Generalmente está asociado a un *router* y filtra el contenido entre dos redes, normalmente la red de área local e Internet.

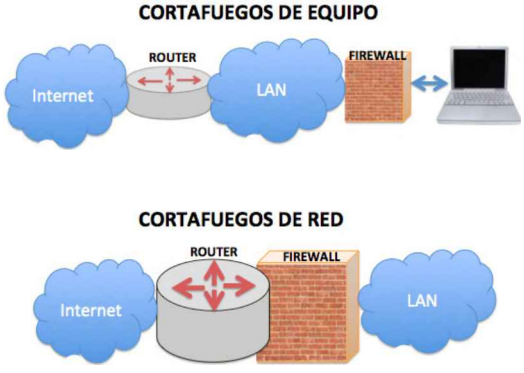

Figura 4.23. *Cortafuegos de equipo y cortafuegos de red*

Los *proxys* o **servidores proxy** son software, dispositivos o equipos cuya misión es interponerse entre la comunicación de dos máquinas. La finalidad de esta interposición puede ser garantizar el anonimato, mejorar la seguridad o rendimiento, mejorar el tráfico, filtrar contenidos, modificar contenidos, etc.

TEST DE CONOCIMIENTOS

1 Elige la respuesta falsa:

a) El objetivo de un sistema de archivos es el gestionar los ficheros almacenados generalmente en un dispositivo físico como un disco duro, unidad SSD, *pendrive*, etc.

b) Los sistemas de archivos son responsables de mantener la integridad de la información.

c) El RAID por hardware es más eficiente que el RAID por software pero genera más carga de trabajo para el microprocesador.

d) *Striping* en sistemas RAID se puede traducir como entrelazado.

2 Elige la respuesta falsa:

a) La caché generalmente utiliza tecnología SRAM (*static* RAM) la cual es más rápida que la tecnología utilizada en la memoria principal.

b) RAID 5 tiene una escritura más lenta que RAID 0 o 1, dado que el sistema tiene que calcular la paridad para cada dato que escribe.

c) FAT32 permite trabajar con archivos de hasta 4 GB.

d) Windows ofrece un gestor de volúmenes lógicos llamado Logical Volume Manager (LVM).

3 Elige la respuesta falsa:

a) El RAID por software es el sistema más económico, pero el menos eficiente.

b) Los *backups* generalmente se programan por la noche o en el momento que el servicio no tiene demanda.

c) En los *cluster*, ante la caída de uno de los miembros del equipo, los demás equipos del *cluster* se hacen cargo de las tareas del mismo.

d) Un plan de recuperación del servicio forma parte del plan de contingencias de una empresa y determina las acciones y medidas a realizar antes de que la amenaza se materialice.

4 Elige la respuesta falsa:

a) La salvaguarda lógica consiste en la aplicación de barreras y procedimientos que resguarden el acceso a los datos y solo se permita acceder a ellos a las personas autorizadas para hacerlo.

b) Unicode es una ampliación del código ASCII que puede utilizar hasta 4 *bytes* (32 bits) para representar cada carácter.

c) HFS+ (*Hierarchical File System Plus*) es el sistema de archivos empleado por los sistemas operativos de Apple.

d) En el *backup* incremental se copia solo los archivos que tienen el atributo de modificado activado. Una vez realizada la copia de seguridad ese atributo se desactiva.

5 Elige la respuesta falsa:

a) La alta disponibilidad es una disposición de varios equipos los cuales realizan una tarea determinada compartiendo esfuerzos.

b) La memoria caché es una memoria intermedia que se coloca entre un elemento rápido y otro más lento del equipo.

c) En RAID 1 por cada disco presente en el sistema se tiene otro con la misma información de tal manera que, cuando un disco falla, el sistema puede seguir funcionando dado que la información permanece duplicada.

d) NAS es muy útil para proporcionar el almacenamiento centralizado a ordenadores clientes en entornos con grandes cantidades de datos.

6 Elige la respuesta falsa:

a) La pérdida o error en un disco de RAID 0 no implica la pérdida de la información en el sistema.

b) Los registros del procesador son memorias muy veloces pero con poca capacidad.

c) La memoria caché de los discos duros utiliza la misma tecnología que en la memoria RAM tradicional.

d) Los servidores SAN (*Storage Area Network*) son servidores de almacenamiento que dan servicio a través de una red y emplean acceso a través de fibra óptica no admitiendo enrutamientos.

7 Elige la respuesta falsa:

a) Las particiones primarias suelen utilizarse para instalar los sistemas operativos.

b) El servidor NAS (*Network Attached Storage*) es un tipo de servidor de almacenamiento que da servicio a través de una red mediante el uso de protocolos estándar de comunicaciones como TCP/IP.

c) Una clonación de una partición es capaz también de clonar el sector de arranque permitiendo así duplicar equipos.

d) Un disco se compone de un sector de arranque, una serie de particiones y opcionalmente espacio sin particionar.

8 Elige la respuesta falsa:

a) LVM fue escrito originalmente por Heinz Mauelshagen en 1998.

b) El término fiabilidad se refiere a la probabilidad de que un sistema funcione normalmente durante un período de tiempo dado.

c) Las políticas de respaldo o salvaguarda se basan fundamentalmente en los sistemas RAID.

d) Cuando hay un fallo en el sistema, se pierde la información existente entre el último *backup* y el momento del fallo.

9 Elige la respuesta falsa:

a) Si un equipo no tiene ninguna partición activa, al arrancar dará un fallo.

b) Los sistemas SAN admiten grandes distancias de hasta 10 Km entre equipos.

c) Se pueden clonar tanto los discos como las particiones.

d) RAID previene sobre un borrado accidental de los datos, una corrupción de ficheros u otra desgracia parecida.

10 Elige la respuesta falsa:

a) RAID 5 es el más utilizado al ofrecer un mejor equilibrio coste-rendimiento-protección.

b) Mientras que los *grid* son conjuntos de servidores independientes unos de otros, los *cluster* suelen ser grupos de servidores con gran dependencia unos de otros.

c) *Recovery time objective* es la duración de tiempo que puede sufrir un sistema tras un desastre sin estar operativo antes de ser restablecido.

d) La paridad es una información adicional que se calcula después de escribir los datos en disco.