UF1466: Sistemas de Almacenamiento

Contenido

1. Sistemas de Archivo

Un sistema de archivos es la estructura y los métodos que utiliza un sistema operativo para organizar, almacenar y gestionar los datos en dispositivos de almacenamiento, como discos duros o memorias USB.

El **sistema de archivos** organiza y gestiona los datos en dispositivos de almacenamiento (discos duros, SSD, pendrives, etc.). Se encarga de:

- **Administrar el espacio** y reducir la fragmentación.
- **Controlar accesos** con permisos, ACL y encriptación.
- **Garantizar integridad** frente a cortes eléctricos, errores de programas o fallos del sistema.

Tipos principales

1. **FAT (File Allocation Table)**
   - FAT16, FAT32, exFAT:
     - Son sistemas de archivos desarrollados por Microsoft.
     - Muy utilizados en dispositivos pequeños como memorias USB y tarjetas SD.

    - Compatibles con la mayoría de sistemas operativos, pero tienen limitaciones.

2. NTFS (New Technology File System)

- Desarrollado por Microsoft y utilizado principalmente en sistemas Windows.
- Soporta archivos grandes, permisos avanzados y encriptación.
- Ideal para discos duros modernos y sistemas operativos Windows.

3. exFAT (Extended FAT)

- Similar a FAT32 pero sin sus limitaciones de tamaño.
- Optimizado para memorias flash como pendrives y tarjetas SD.
- Compatible con Windows y macOS.

4. HFS+ (Hierarchical File System Plus)

- Sistema de archivos usado en macOS hasta macOS High Sierra.
- Optimizado para discos duros y soporta características como la compresión.

5. APFS (Apple File System)

- Sustituyó a HFS+ en macOS.
- Diseñado para discos SSD, con funciones como snapshots y cifrado avanzado.

6. EXT (Extended File System)

- Utilizado en sistemas Linux.
- Versiones: **EXT2, EXT3, EXT4**.
- **EXT4** es la versión más moderna, con soporte para archivos grandes y un rendimiento mejorado.

¿Qué es FAT32?

FAT32 es un sistema de archivos de la familia FAT desarrollado por Microsoft y se introdujo con Windows 95 OSR2. Algunas características clave:

- Compatibilidad:
    - Funciona en la mayoría de los sistemas operativos (Windows, macOS, Linux).
    - Es ampliamente usado en dispositivos como memorias USB y tarjetas SD.
- Limitaciones:
    - Tamaño máximo de archivo: **4 GB.**
    - Tamaño máximo de partición: **32 GB** (en algunos sistemas operativos, aunque técnicamente puede manejar hasta 2 TB).
- Ventajas:
    - Alta compatibilidad con diferentes sistemas operativos.
    - Simple y eficiente para dispositivos de almacenamiento pequeños.
- Desventajas:
    - No admite permisos avanzados ni cifrado.
    - Limitaciones de tamaño de archivo y partición, lo que lo hace menos ideal para discos duros grandes.

Nomenclatura y Codificación

Los ficheros se nombran con nombre y extensión. Se utilizan distintos sistemas de codificación como ASCII, Unicode, UTF-8 o binario.

1. Codificación ASCII (American Standard Code for Information Interchange)

- Representa caracteres básicos (letras en inglés, dígitos y símbolos).
- Usa 7 bits por carácter (128 caracteres posibles).
- Limitado a caracteres del inglés y algunos símbolos comunes.

2. Codificación Unicode

Unicode es un estándar universal para representar texto en casi todos los idiomas y sistemas de escritura.

Variantes comunes de Unicode:

- **UTF-8 (Unicode Transformation Format - 8 bits):**
    - Usa entre 1 y 4 bytes por carácter.
    - Compatible con ASCII (los caracteres ASCII se representan en 1 byte).
    - Es eficiente para idiomas basados en caracteres latinos y ampliamente usado en la web.

- **UTF-16: ○ Usa 2 o 4 bytes por carácter.**
    - Eficiente para idiomas con muchos caracteres, como chino, japonés o coreano.
    - Es más pesado que UTF-8 para idiomas latinos.

- **UTF-32: ○ Usa 4 bytes por carácter.**
    - Representa todos los caracteres Unicode de forma directa, pero ocupa más espacio.

3. Codificación ISO/IEC 8859

- Serie de estándares diseñados para diferentes lenguajes y regiones (por ejemplo, Europa occidental, Europa oriental, árabe, etc.).
- ISO-8859-1 (Latin-1): ○ Compatible con ASCII y añade caracteres especiales para idiomas europeos (acentos, diéresis, etc.).

4. Codificación EBCDIC (Extended Binary Coded Decimal Interchange Code)

- Usado principalmente en sistemas IBM Mainframe y servidores antiguos.
- Codificación menos común hoy en día.

2. 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.

Un RAID (Redundant Array of Independent Disks) es una tecnología que combina múltiples discos duros para mejorar el rendimiento, la tolerancia a fallos, o ambos, dependiendo del nivel de RAID que se implemente. RAID puede ser gestionado por hardware (controladores RAID dedicados) o software (implementado por el sistema operativo).

Tipos de RAID más comunes:

1. RAID 0 (Stripe):

- Propósito: Mejorar el rendimiento.
- Funcionamiento: Divide los datos en bloques y los distribuye entre varios discos.
- Ventaja: Muy rápido y utiliza el 100% del almacenamiento disponible.
- Desventaja: No hay redundancia; si un disco falla, se pierde toda la informacion.

2. RAID 1 (Mirror):

- Propósito: Tolerancia a fallos.
- Funcionamiento: Duplica los datos en dos discos.
- Ventaja: Si un disco falla, los datos están a salvo en el otro.
- Desventaja: Solo se utiliza el 50% del almacenamiento total.

3. RAID 5:

- Propósito: Tolerancia a fallos y eficiencia de almacenamiento.
- Funcionamiento: Distribuye los datos y la paridad (información para recuperar datos) entre al menos tres discos.
- Ventaja: Puede soportar la falla de un disco sin pérdida de datos.
- Desventaja: Requiere al menos 3 discos y más potencia de cálculo.

4. RAID 6:

- Propósito: Mayor tolerancia a fallos.
- Funcionamiento: Similar a RAID 5, pero almacena dos bloques de paridad.
- Ventaja: Soporta la falla de hasta dos discos.
- Desventaja: Necesita al menos 4 discos.

5. RAID 10 (1+0):

- Propósito: Combina rendimiento y tolerancia a fallos.
- Funcionamiento: Combina RAID 1 (mirroring) y RAID 0 (striping).
- Ventaja: Excelente rendimiento y redundancia.
- Desventaja: Muy costoso porque requiere al menos 4

3. Volúmenes Lógicos y Físicos

El almacenamiento de un sistema se puede analizar en dos niveles:

- **Volúmenes físicos**: son las particiones reales que dividen un disco o dispositivo de almacenamiento. Se crean sobre el hardware (disco duro, SSD, memoria flash). Pueden ser:
    - **Primarias** (hasta 4 por disco, suelen albergar el sistema operativo).

    - **Extendidas** (contenedor que puede incluir varias lógicas).
    - **Lógicas** (subdivisiones dentro de la extendida, usadas para almacenar datos).
- **Volúmenes lógicos**: son estructuras virtuales que permiten gestionar el espacio de los volúmenes físicos de forma más flexible. Con herramientas como **LVM (Logical Volume Manager)** en Linux, se pueden agrupar varios discos físicos y crear un único volumen lógico que el sistema ve como una unidad. Esto permite:
    - Ampliar o reducir volúmenes sin depender de las limitaciones físicas.
    - Crear volúmenes que abarquen varios discos.
    - Gestionar el almacenamiento de forma más eficiente y escalable.

En resumen: **los volúmenes físicos representan las divisiones reales del disco**, mientras que **los lógicos son "capas virtuales" que permiten mayor flexibilidad de gestión**.

4. Sistemas NAS y SAN

- **NAS**: almacenamiento conectado a red TCP/IP. Centralizado, económico, usa RAID, y está muy extendido en empresas y usuarios domésticos.
- **SAN**: redes de almacenamiento de alto rendimiento con fibra óptica. Más costosas, permiten gran escalabilidad y distancias de hasta 10 km.

Diferencias entre SAN y NAS

- La principal diferencia radica en cómo gestionan los datos y su propósito. Un SAN opera a nivel de bloques, lo que significa que se presenta como un disco físico directamente conectado a los servidores, ofreciendo mayor rendimiento y flexibilidad para aplicaciones
- críticas. Por otro lado, un NAS funciona a nivel de archivos, siendo más simple de configurar y enfocado en el intercambio de archivos en red. En términos de costo, el SAN tiende a ser más caro y complejo debido a su infraestructura dedicada, mientras que el NAS es una solución más económica y fácil de administrar, adecuada para tareas menos exigentes.

5. Políticas de Salvaguarda

Las **políticas de salvaguarda** son el conjunto de medidas, procedimientos y estrategias que una organización define para **proteger sus datos, sistemas y servicios** frente a fallos, desastres o incidentes, y asegurar que la actividad pueda continuar con la menor interrupción posible.

Aparecen ligadas al **plan de continuidad de negocio**, y abarcan tres grandes ejes:

1. **Prevención**: evitar que ocurran problemas (ej. copias de seguridad frecuentes, control de accesos, redundancia de sistemas).
2. **Mitigación**: reducir el impacto cuando el problema ocurre (ej. alta disponibilidad, balanceo de carga, sistemas RAID).
3. **Recuperación**: restablecer el servicio lo antes posible y con la menor pérdida de datos (ej. planes de recuperación, parámetros RTO y RPO).

Incluyen distintos aspectos:

- **Salvaguarda física**: proteger equipos e instalaciones frente a riesgos naturales (incendios, inundaciones, terremotos) o humanos (robos, sabotajes).
    - Cerraduras, controles biométricos, videovigilancia, sensores
- **Salvaguarda lógica**: proteger los datos mediante barreras de acceso, autenticación de usuarios, cifrado y restricciones en programas y archivos.
    - Control de accesos a programas y datos, transmisión segura, planes alternativos de comunicación.
- **Copias de seguridad y clonaciones**: asegurar que los datos puedan restaurarse en caso de pérdida.
- **Alta disponibilidad**: garantizar que los sistemas estén accesibles 24/7.
- **Integridad de los datos**: mantener la exactitud y fiabilidad de la información.
- **Custodia de ficheros de seguridad**: almacenar y proteger los backups, incluso en lugares externos (offsite backup).

Copias de seguridad y clonaciones

- **Backup completo**: copia todo.

- **Backup incremental**: copia solo lo modificado desde el último backup.
- **Backup diferencial**: copia lo modificado desde el último completo.
- **Clonación**: copia exacta de un disco o partición (incluye sector de arranque en caso de disco).

6. Políticas de Seguridad

Las **políticas de seguridad** son el conjunto de normas, procedimientos y medidas que una organización establece para **proteger su información, sus sistemas y sus usuarios** frente a accesos no autorizados, ataques, errores humanos o fallos técnicos.

Las políticas de seguridad incluyen varios aspectos clave:

1. **Acceso restringido por cuentas de usuario**
    - Cada persona debe tener su propia cuenta e identificador único (UID).
    - Los usuarios se organizan en grupos con permisos definidos.
    - Existen distintos tipos de usuarios: normales, superusuarios (root/Administrador), invitados y cuentas especiales para tareas concretas.
2. **Propiedad de la información**
    - La información pertenece a la organización.
    - Cada persona accede solo a lo que necesita para su función.
    - Se recomienda centralizar la información para mejorar el control.
3. **Identificación y autenticación**
    - Validar de forma unívoca a cada usuario (usuario/contraseña, certificados digitales, smartcards, biometría).
    - Evitar cuentas compartidas y duplicadas.
4. **Protección antivirus y contra malware**
    - Detección y eliminación de virus mediante técnicas de firmas y análisis heurístico.
    - Prevención y monitorización de comportamientos sospechosos.

5. **Auditorías de seguridad**
    - Revisiones internas o externas para comprobar el cumplimiento de las normas de seguridad.
    - Evalúan seguridad física, lógica, control de accesos, aplicaciones, bases de datos y redes.
6. **Firewalls y servidores proxy**
    - **Firewalls**: filtran tráfico entrante/saliente según reglas (pueden ser de equipo o de red).
    - **Proxys**: intermediarios en las comunicaciones que aportan anonimato, seguridad y control del tráfico.