Engineer

Archivo informático

conjunto de datos con nombre guardado como una unidad

Qué es

Un archivo informático es lo que un programa pide al sistema operativo cuando quiere recuperar mañana los mismos bytes: abre el archivo por su nombre, lee o escribe, lo cierra y confía en que el siguiente programa que abra ese nombre vea lo que se dejó allí. Documentos, fotos, programas, ajustes y registros son archivos. En los sistemas de tipo Unix también lo son varias cosas que no están almacenadas en ninguna parte.

POSIX, el estándar que sirve de base a Linux, macOS y otros sistemas de tipo Unix, define un archivo como un objeto en el que se puede escribir, del que se puede leer o ambas cosas, con atributos que incluyen permisos de acceso y tipo. Después enumera los tipos: archivos normales, directorios, enlaces simbólicos, FIFO, sockets y archivos especiales de caracteres y bloques, que representan dispositivos como un terminal o un disco. Al decir archivo casi siempre se alude al primer tipo, que el estándar describe como una secuencia de bytes de acceso aleatorio sobre la que el sistema no impone estructura propia. Al sistema no le importa ni sabe si esos bytes son una hoja de cálculo, una foto o una página de una base de datos: la estructura pertenece al programa que los lee. POSIX incluso rehúsa separar archivos de texto y binarios, aunque advierte de que muchas utilidades solo producen resultados con sentido al trabajar con texto.

El nombre se guarda fuera del archivo. Un directorio es una lista de entradas y cada una vincula un nombre con un archivo. Un nombre de archivo es un único componente que puede contener cualquier byte salvo la barra y el byte nulo, hasta el límite que fije el sistema. Una ruta como `/srv/site/index.html` es la cadena de nombres que recorre el sistema para llegar a él. Como los nombres viven en directorios, un archivo puede tener varios y el sistema lleva la cuenta. Cuando se elimina el último y ningún proceso mantiene abierto el archivo, se libera el espacio.

Todo lo demás que el sistema sabe de un archivo son metadatos, y un listado detallado muestra la mayor parte:

```
$ ls -li notes.txt
1048577 -rw-r----- 2 ana staff 4096 Oct 16 09:14 notes.txt
```

El primer número es el número de serie del archivo, único en su sistema de archivos: ese número es el archivo y los nombres solo apuntan a él. Después aparecen los bits de permiso del propietario, del grupo y de los demás, el número de nombres del archivo, el propietario y grupo, el tamaño en bytes y la marca de tiempo de la última modificación de datos. POSIX conserva otras dos marcas de tiempo: el último acceso a los datos y el último cambio de estado del archivo. Un programa que ha abierto un archivo tiene un descriptor, un entero pequeño de ámbito local al proceso. Los descriptores 0, 1 y 2 corresponden, respectivamente, a la entrada estándar, la salida estándar y el error estándar, por lo que un shell puede dirigir la salida de un programa a un archivo sin que este lo sepa.

Utilice el concepto en una decisión

Aplique «¿Qué es un archivo informático?» a una decisión real. Podemos ayudarle a relacionar el concepto con el sistema que está construyendo o comprando. El cuidado diario de servidores, correo, cuentas y actualizaciones para que los servicios de los que depende todo el mundo sigan funcionando sin sorpresas.

Explore nuestros servicios Póngase en contacto →

Compruebe los supuestos detrás del término

Utilice la explicación y los trabajos enlazados para formular preguntas sobre su proyecto. Podemos examinar los supuestos con usted antes de elegir una herramienta. 10 servidores PostgreSQL y MS SQL en Ubuntu — 10 servidores de bases de datos desplegados y mantenidos

10 servidores PostgreSQL y MS SQL en Ubuntu Póngase en contacto →

Por qué importa

Los archivos son la interfaz más antigua entre programas que nunca se conocieron. Una herramienta escrita hace décadas puede pasar datos a otra de ayer mediante un archivo, sin más acuerdo que los bytes y el nombre. Configuración, registros, exportaciones, copias de seguridad y el contenido de la mayoría de los sitios viajan así. Mantener fiables los archivos es, por tanto, gran parte de mantener fiable un sistema.

Los problemas sutiles están en los metadatos. La marca de tiempo de modificación indica cuándo se escribieron por última vez los bytes, algo distinto de cuándo cambió por última vez lo que dicen. Ambas cosas se separan en cuanto

una herramienta reescribe un archivo con contenido idéntico. La propiedad se almacena como un número, de modo que el mismo archivo puede pertenecer a otra persona en otra máquina.

También está la durabilidad. Una escritura que termina ha entregado los bytes al sistema operativo, que puede conservarlos en caché. POSIX dice que los datos en caché llegan al almacenamiento cuando un programa lo pide con una llamada como `fsync` y que pueden llegar antes por su cuenta. También advierte de que, tras una caída, puede haber operaciones posteriores en disco mientras faltan operaciones anteriores. Pone el ejemplo de un cambio de nombre que deja creado el nuevo y mantiene también el antiguo.

Por último, un archivo es una copia en un dispositivo hasta que alguien dispone otra cosa. El almacenamiento falla de maneras distintas según el medio. La única protección que sirve frente a todas es otra copia en otro lugar que se haya vuelto a leer.

Cómo se aplica en la práctica

La marca de tiempo de modificación de un archivo registra cuándo se escribieron sus bytes por última vez, un momento distinto de cuándo cambió su contenido. El sitemap del sitio mostró una vez esa diferencia: sus fechas situaban 172 de las 176 páginas en un mismo instante de cambio. La corrección rastreó la fecha hasta el propio archivo: la exportación de contenido reescribe todos los archivos en cada ejecución, así que una sincronización daba al sitio entero una marca temporal única, dijera lo que dijera cada página. La fecha procede ahora del historial de versiones, con una cadena de alternativas que prueba primero un campo explícito, después el historial de commits y finalmente el archivo. La página documenta la trampa: el nombre literal del campo debe estar en la lista o el generador nunca lo lee. Si falta una fecha, el campo queda vacío en lugar de contener una suposición. En la etapa de implementación documentada, volver a sincronizar tras un mes de ediciones solo modificaba ocho archivos de 186.

La implementación de pruebas de restauración de la empresa termina leyendo de nuevo un archivo. Sus unidades systemd admiten un volcado de base de datos donde esté configurado, una instantánea restic cifrada y deduplicada enviada por SFTP a almacenamiento de otro proveedor, depuración según la política de retención y una comprobación de integridad. Un flujo mensual separado está diseñado para restaurar un archivo conocido en una ubicación temporal y comparar sus bytes, de modo que la prueba prevista es una restauración, no una

subida. El registro de implementación del 2026-10-27 dice que los temporizadores estaban desactivados sin un repositorio configurado y no documenta ninguna restauración real completada. Incluso la frase de paso se guarda deliberadamente en un archivo: las unidades la leen de un archivo protegido en lugar de un valor de entorno interpretado por systemd, cuyo tratamiento de secuencias de escape podría, si la frase contiene una barra inversa, producir una frase de paso distinta de la usada para crear el repositorio.

Años antes, el trabajo consistía en rescatar archivos, a menudo la única copia, de hardware averiado: recuperación de datos de tarjetas SD, discos duros mecánicos, SSD, matrices RAID, unidades externas y sistemas Mac. Cada medio falla y entrega sus datos de forma distinta, y un sistema de archivos de Mac tiene sus peculiaridades. El trabajo dependía tanto de elegir el enfoque adecuado al medio como de una técnica concreta. La medida era sencilla: archivos devueltos intactos.

Dónde surgen los problemas

Todas las páginas parecen haber cambiado a la vez. Una copia, un checkout, una restauración o una exportación reescribe el archivo y reinicia su marca de tiempo de modificación. Cada sistema posterior que confía en ella ve un cambio inexistente. Una copia que conserva las marcas de tiempo hace lo contrario y oculta un cambio real. Obtenga las fechas del contenido a partir del propio contenido o del historial de versiones y deje que la marca de tiempo del archivo signifique lo que significa.

Un lector encuentra medio archivo. Un programa que reescribe un archivo sobre sí mismo deja un intervalo en el que otro proceso, o el siguiente arranque, ve parte de la versión antigua y parte de la nueva. Escriba un archivo temporal en el mismo directorio, llame a `fsync` sobre él y después renómbrelo sustituyendo al original. POSIX exige que el nombre sustituido siga visible durante todo el cambio de nombre y apunte al archivo antiguo o al nuevo.

Un script tropieza con un nombre válido. Un nombre de archivo puede contener espacios, saltos de línea y un guion inicial. Un script que divide nombres por espacios en blanco o los pasa sin comillas a una orden actuará sobre archivos equivocados. POSIX aconseja evitar el guion inicial y limitarse a letras, dígitos, punto, guion bajo y guion para nombres que deban ser portables. Ponga entre comillas cada expansión y termine el análisis de opciones con `--` antes de una lista de nombres.

Un archivo secreto es legible por todos. Los bits de permiso de un archivo nuevo proceden de los valores predeterminados del proceso que lo creó, que suelen permitir leer al grupo y a los demás. Una clave o frase de paso escrita así se comparte con todas las cuentas del host. Cree desde el principio los archivos secretos con permisos solo para el propietario, en lugar de restringirlos después.

Lecturas adicionales

- Glosario NIST CSRC: archivo — Las definiciones usadas en publicaciones de seguridad del NIST, cada una con su documento de origen.
- Definiciones básicas de POSIX.1-2024: definiciones — El significado formal de archivo, archivo normal, nombre de archivo, ruta, entrada de directorio y descriptor de archivo.
- Definiciones básicas de POSIX.1-2024: conceptos generales — Permisos, las tres marcas de tiempo del archivo, portabilidad de nombres y lo que una caída puede dejar en la caché.
- POSIX.1-2024: rename() — La garantía de que el nombre sustituido apunta en todo momento al archivo antiguo o al nuevo, de la que depende guardar con seguridad.

Glosario

Elija según sus restricciones reales

Describa la ruta de despliegue, los riesgos operativos y quiénes se encargan de las guardias. Podemos definir el alcance de un cambio con comprobaciones, una vía de recuperación y responsabilidades claras.

Vea cuánto cuesta Póngase en contacto →

Haga práctica la terminología

Pregunte cómo se aplica «¿Qué es un archivo informático?» a su organización. Basta un ejemplo concreto de su sistema actual para iniciar la conversación.

Explore nuestros servicios

welcome@engineer.company +45 50 10 61 84 Copenhague, Dinamarca

Llévese esta página

Tarjeta de contacto

Creado con esmero y cariño por D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Derechos de autor © 2025 – actualidad · Engineer ApS · Todos los derechos reservados.

Impreso desde https://engineer.company/es/glossary/file/