Engineer

Fil

en namngiven mängd data som lagras som en enhet

Vad det är

En fil är det som ett program ber operativsystemet om när det vill ha tillbaka samma bytes i morgon: det öppnar filen med dess namn, läser eller skriver, stänger den och litar på att nästa program som öppnar det namnet ser det som lämnades där. Dokument, foton, program, inställningar och loggar är alla filer, och på Unix-liknande system gäller det också flera saker som inte lagras någonstans alls.

POSIX, standarden bakom Linux, macOS och de andra Unix-liknande systemen, definierar en fil som ett objekt som går att skriva till, läsa från eller både och, med attribut som bl.a. omfattar dess åtkomstbehörigheter och dess typ. Sedan räknar standarden upp typerna: vanliga filer, kataloger, symboliska länkar, FIFO:er, socketar samt tecken- och blockspecialfiler, som står för enheter som en terminal eller en disk. När folk säger fil menar de nästan alltid den första sorten, som standarden beskriver som en följd av bytes som kan läsas på valfri position och som systemet inte ger någon egen struktur. Systemet vet inte och bryr sig inte om de bytesen är ett kalkylark, ett foto eller en sida i en databas; strukturen tillhör programmet som läser dem. POSIX avstår till och med från att skilja textfiler från binära filer, men noterar att många verktyg bara ger meningsfull utdata när de får text.

Namnet hålls utanför filen. En katalog är en lista med poster, och varje post knyter ett filnamn till en fil. Ett filnamn är en enda komponent som får innehålla vilken byte som helst utom snedstreck och nollbyten, upp till en längd som systemet bestämmer; en sökväg som `/srv/site/index.html` är kedjan av namn som systemet går igenom för att nå filen. Eftersom namnen bor i kataloger kan en fil ha flera av dem, och systemet håller räkningen. När det sista namnet tas bort och ingen process längre har filen öppen frigörs utrymmet.

Allt annat som systemet vet om en fil är metadata, och en lång listning visar det mesta av den:

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

Det första talet är filens serienummer, unikt inom dess filsystem: det numret är filen, och namnen pekar bara på den. Sedan kommer behörighetsbitarna för ägaren, gruppen och alla andra, antalet namn som filen har, ägaren och gruppen, storleken i bytes och tidpunkten då innehållet senast ändrades. POSIX håller ytterligare två tider bredvid den: när innehållet senast lästes och när filens status senast ändrades. Ett program som har öppnat en fil håller en fildeskriptor, ett litet heltal som bara gäller inom processen. Deskriptorerna 0, 1 och 2 är standardinmatning, standardutmatning och standardfel, och det är därför ett skal kan skicka ett programs utdata till en fil utan att programmet vet om det.

Använd begreppet i ett beslut

Använd ”Vad är en fil?” i ett konkret beslut. Vi kan knyta begreppet till systemet ni bygger eller köper. Daglig skötsel av servrar, e-post, konton och uppdateringar, så att det som alla förlitar sig på förblir odramatiskt.

Se våra tjänster Hör av er →

Granska antagandena bakom ordet

Använd förklaringen och det länkade arbetet för att formulera frågor till projektet. Vi kan undersöka antagandena tillsammans innan ett verktyg väljs. Drev 10 PostgreSQL- och MS SQL-servrar på Ubuntu — 10 databasservrar som driftsattes och hölls i gott skick

Drev 10 PostgreSQL- och MS SQL-servrar på Ubuntu Hör av er →

Varför det spelar roll

Filer är det äldsta gränssnittet mellan program som aldrig har presenterats för varandra. Ett verktyg som skrevs för årtionden sedan kan lämna data till ett som skrevs i går genom en fil, utan någon annan överenskommelse än bytesen och namnet. Konfiguration, loggar, exporter, säkerhetskopior och innehållet på de flesta webbplatser färdas på det sättet, så att hålla filerna pålitliga är det mesta av att hålla ett system pålitligt.

Det är i metadatan som de lömska problemen bor. Ändringstiden säger när bytesen senast skrevs, vilket är ett annat faktum än när det som de säger senast ändrades, och de två går isär i samma ögonblick som ett verktyg skriver om en fil med identiskt innehåll. Ägarskap lagras som ett tal, så samma fil kan tillhöra en annan person på en annan maskin.

Sedan är det beständigheten. En skrivning som har returnerat har lämnat bytesen till operativsystemet, som kan behålla dem i en cache en stund. POSIX säger att cachade data når lagringen när ett program ber om det med ett anrop som `fsync`, och att de kan komma dit tidigare på egen hand. Standarden varnar också för att senare operationer efter en krasch kan finnas på disken medan tidigare saknas, och ger exemplet med ett namnbyte som lämnar det nya namnet skapat och det gamla fortfarande kvar.

Till sist är en fil en enda kopia på en enda enhet tills någon ordnar något annat. Lagring går sönder på sätt som beror på mediet, och det enda skydd som håller mot alla sätten är en andra kopia någon annanstans, en som har lästs tillbaka.

Så ser det ut i praktiken

Skillnaden mellan en fil och dess innehåll syntes en gång för varje sökmotor: webbplatskartan för den här webbplatsen uppgav samma ändringsögonblick för 172 av 176 sidor. Åtgärden spårade datumet tillbaka till själva filen: datumen lästes från filsystemets ändringstider, och eftersom innehållsexporten skapar alla filer på nytt varje gång den körs fick hela innehållet en ny stämpel vid en enda synkronisering. Nu letar generatorn efter datumet i tur och ordning i ett explicit fält, i commit-historiken och sist i filen själv, och sidan varnar för en detalj i den ordningen: fältet läses bara om dess bokstavliga namn står med i listan. Ett okänt datum lämnar fältet tomt i stället för att fyllas med en gissning, och i det dokumenterade implementeringsskedet berörde en ny synkronisering efter en månads ändringar åtta filer av 186 i stället för alla.

Företagets implementering av ett återställningstest slutar med att en fil läses tillbaka. Dess systemd-enheter stöder en databasdump där det är konfigurerat, en krypterad, deduplicerad restic-ögonblicksbild som skickas via SFTP till lagring hos en annan leverantör, gallring enligt bevarandepolicyn och en integritetskontroll. Ett separat månatligt arbetsflöde återställer en känd fil till ett tillfälligt arbetsutrymme och jämför dess byte, så det avsedda testet är en återställning snarare än en uppladdning. Implementeringsbeskrivningen daterad 2026-10-27 anger att timerenheterna var inaktiverade utan ett konfigurerat repo och redovisar ingen genomförd återställning i drift. Till och med lösenfrasen ligger avsiktligt i en fil: enheterna läser den ur en skyddad fil i stället för ett miljövärde som systemd tolkar, där hanteringen av escape-tecken skulle kunna ändra en lösenfras med ett omvänt snedstreck så att den skiljer sig från den som repot skapades med.

Flera år tidigare handlade jobbet om att rädda filer, ofta den enda kopian, ur hårdvara som hade gått sönder: dataräddning från SD-kort, snurrande hårddiskar, SSD:er, RAID-arrayer, externa diskar och Mac-system. Var och en av dem går sönder och lämnar ifrån sig sina data på sitt eget sätt, och ett Mac-filsystem har sina egna egenheter, så arbetet handlade lika mycket om rätt angreppssätt för mediet som om någon enskild teknik. Måttet var det enkla: filer som lämnades tillbaka i behåll.

Var det går fel

Varje sida ser ändrad ut på en gång. En kopiering, en utcheckning, en återställning eller en export skriver om filen och ger den en ny ändringstid, och varje system längre ned i kedjan som litar på den tiden ser en ändring som aldrig hände. En kopiering som bevarar tiderna gör det motsatta och döljer en ändring som faktiskt hände. Hämta innehållets datum från innehållet eller från versionshistoriken, och låt filens tid betyda det den betyder.

En läsare hittar en halv fil. Ett program som skriver om en fil på plats lämnar ett fönster då en annan process, eller nästa uppstart, ser en del av den gamla versionen och en del av den nya. Skriv i stället en temporär fil i samma katalog, anropa `fsync` på den och byt sedan namn på den så att den ersätter originalet. POSIX kräver att det ersatta namnet förblir synligt under hela namnbytet och pekar på antingen den gamla filen eller den nya.

Ett skript snubblar på ett tillåtet namn. Ett filnamn får innehålla mellanslag, radbrytningar och ett inledande bindestreck, och ett skript som delar upp namn vid blanktecken eller lämnar dem ociterade till ett kommando agerar på fel filer. Standardens egna råd är att undvika ett inledande bindestreck och att hålla sig till bokstäver, siffror, punkt, understreck och bindestreck i namn som ska kunna flyttas mellan system. Citera varje expansion, och avsluta tolkningen av flaggor med `--` före en lista med namn.

En hemlig fil kan läsas av alla. Behörighetsbitarna på en ny fil kommer från standardinställningarna i den process som skapade den, och de inställningarna låter oftast gruppen och övriga läsa. En nyckel eller lösenfras som skrivs på det sättet delas med varje konto på värddatorn. Skapa hemliga filer med behörighet för enbart ägaren från början i stället för att strama åt dem efteråt.

Läs mer

- NIST CSRC glossary: file — de definitioner som säkerhetspublikationer från NIST använder, var och en med det dokument den kommer från.
- POSIX.1-2024 Base Definitions: Definitions — den formella betydelsen av fil, vanlig fil, filnamn, sökväg, katalogpost och fildeskriptor.
- POSIX.1-2024 Base Definitions: General Concepts — behörigheter, filens tre tider, portabla filnamn och vad en krasch kan lämna efter sig i cachen.
- POSIX.1-2024: rename() — garantin att ett namn som ersätts hela tiden pekar på antingen den gamla filen eller den nya, som säker sparning bygger på.

Ordlista

Välj utifrån era faktiska förutsättningar

Ta med driftsättningsvägen, driftsriskerna och dem som har jour. Vi kan avgränsa en förändring med kontroller, en väg tillbaka och tydligt ansvar.

Se vad det kostar Hör av er →

Gör begreppen praktiska

Fråga hur ”Vad är en fil?” gäller i er organisation. Ett konkret exempel från ert nuvarande system räcker för att börja samtalet.

Se våra tjänster

welcome@engineer.company +45 50 10 61 84 Köpenhamn, Danmark

Ta med dig den här sidan

Visitkort

Omsorgsfullt skapad med kärlek av D.

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

Upphovsrätt © 2025 – nu · Engineer ApS · Alla rättigheter förbehållna.

Utskrivet från https://engineer.company/sv/glossary/file/