Engineer

Fil

et navngivet sæt data, der gemmes som én enhed

Hvad det er

En fil er det, et program beder styresystemet om, når det vil have de samme bytes tilbage i morgen: det åbner filen ved navn, læser eller skriver, lukker den og stoler på, at det næste program, der åbner det navn, ser det, der blev efterladt. Dokumenter, billeder, programmer, indstillinger og logge er alle filer, og på Unix-lignende systemer er flere ting, der slet ikke er gemt nogen steder, det også.

POSIX, standarden bag Linux, macOS og de andre Unix-lignende systemer, definerer en fil som et objekt, der kan skrives til, læses fra eller begge dele, med egenskaber, som blandt andet omfatter adgangsrettigheder og type. Derefter opregner den typerne: almindelige filer, mapper, symbolske links, FIFO'er, sockets samt tegn- og blokspecialfiler, som står for enheder som en terminal eller en disk. Når folk siger fil, mener de næsten altid den første slags, som standarden beskriver som en række bytes med vilkårlig adgang, som systemet ikke pålægger nogen yderligere struktur. Systemet ved ikke og er ligeglad med, om de bytes er et regneark, et foto eller en side i en database; strukturen hører til det program, der læser dem. POSIX afviser endda at skelne mellem tekstfiler og binære filer, men bemærker, at mange værktøjer kun giver meningsfuldt output på tekst.

Navnet ligger uden for filen. En mappe er en liste over poster, og hver post kobler et filnavn til en fil. Et filnavn er en enkelt komponent, der må indeholde enhver byte undtagen skråstreg og nulbyten, op til en grænse, som systemet fastsætter; et stinavn som `/srv/site/index.html` er den kæde af navne, systemet følger for at nå frem. Fordi navnene bor i mapper, kan én fil have flere af dem, og systemet fører en optælling. Når det sidste navn fjernes, og ingen proces længere har filen åben, frigives pladsen.

Alt andet, systemet ved om en fil, er metadata, og en lang oversigt viser det meste:

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

Det første tal er filens serienummer, unikt inden for dens filsystem: det nummer er filen, og navnene peger blot på den. Derefter kommer rettighedsbittene for ejeren, gruppen og alle andre, antallet af navne, filen har, ejer og gruppe, størrelsen i bytes og tidspunktet for den seneste ændring af data. POSIX gemmer to tidspunkter mere ved siden af: den seneste adgang til data og den seneste ændring af filens status. Et program, der har åbnet en fil, holder en fildeskriptor, et lille heltal, som kun gælder i processen. Deskriptorerne 0, 1 og 2 er standardinput, standardoutput og standardfejl, og derfor kan en shell sende et programs output til en fil, uden at programmet ved det.

Brug begrebet i en beslutning

Brug »Hvad er en fil?« i en konkret beslutning. Vi kan knytte begrebet til det system, I bygger eller køber. Daglig drift af servere, mail, konti og opdateringer, så det, alle er afhængige af, forbliver kedeligt.

Se vores ydelser Kontakt os →

Undersøg antagelserne bag ordet

Brug forklaringen og det tilknyttede arbejde til at formulere spørgsmål til jeres projekt. Vi kan undersøge antagelserne sammen, før et værktøj vælges. Har drevet 10 PostgreSQL- og MS SQL-servere — 10 databaseservere udrullet og holdt i god drift

Har drevet 10 PostgreSQL- og MS SQL-servere Kontakt os →

Hvorfor det betyder noget

Filer er den ældste grænseflade mellem programmer, der aldrig er blevet præsenteret for hinanden. Et værktøj skrevet for årtier siden kan give data videre til et, der blev skrevet i går, gennem en fil, uden nogen anden aftale end bytes og navn. Konfiguration, logge, eksporter, sikkerhedskopier og indholdet af de fleste websites rejser på den måde, så at holde filer troværdige er det meste af at holde et system troværdigt.

Metadataene er der, hvor de lumske problemer bor. Ændringstidspunktet fortæller, hvornår de bytes sidst blev skrevet, og det er en anden oplysning end, hvornår det, de siger, sidst blev ændret. De to skilles, i samme øjeblik et værktøj skriver en fil om med identisk indhold. Ejerskab gemmes som et tal, så den samme fil kan tilhøre en anden person på en anden maskine.

Så er der holdbarheden. En skrivning, der vender tilbage, har givet de bytes til styresystemet, som kan holde dem i en cache. POSIX siger, at data i cachen når frem til lageret, når et program beder om det med et kald som `fsync`, og at de kan komme derhen tidligere af sig selv, og standarden advarer om, at efter et nedbrud kan senere operationer ligge på disken, mens tidligere mangler.

Endelig er en fil én kopi på én enhed, indtil nogen sørger for andet. Den eneste beskyttelse, der holder, uanset hvordan mediet svigter, er en ekstra kopi et andet sted, som er blevet læst tilbage.

Sådan ser det ud i praksis

Sitets sitemap fortalte engang søgemaskinerne, at 172 af dets 176 sider var ændret i samme øjeblik. Rettelsen sporede datoen tilbage til selve filen: hver sides dato kom fra filens ændringstidspunkt, og indholdseksporten skriver hver fil om ved hver kørsel, så én synkronisering stemplede hele sitet. Nu kommer datoen fra versionshistorikken gennem en fallback-kæde, der prøver et eksplicit felt, så commit-historikken og til sidst filen. Siden noterer fælden i den kæde: selve feltnavnet skal stå på listen, ellers læser generatoren det aldrig. En manglende dato lader feltet stå tomt i stedet for at rumme et gæt, og på det dokumenterede implementeringstrin berørte en ny synkronisering efter en måneds redigeringer otte filer ud af 186.

Virksomhedens implementering af en gendannelsestest slutter med, at en fil bliver læst tilbage. Dens systemd-enheder understøtter et databasedump, hvor det er konfigureret, et krypteret, deduplikeret restic-snapshot sendt over SFTP til lager hos en anden udbyder, udtynding efter opbevaringspolitikken og et integritetstjek. Et separat månedligt workflow gendanner en kendt fil til et midlertidigt arbejdsområde og sammenligner dens bytes, så den tilsigtede test er en gendannelse og ikke en upload. Implementeringsbeskrivelsen dateret 2026-10-27 angiver, at timerne var deaktiveret uden et konfigureret backuplager, og dokumenterer ingen gennemført gendannelse i drift. Selv adgangssætningen ligger med vilje i en fil: enhederne læser den fra en beskyttet fil frem for en miljøværdi fortolket af systemd, hvor behandlingen af escape-tegn kunne ændre en adgangssætning med en omvendt skråstreg i forhold til den, backuplageret blev oprettet med.

År forinden gik arbejdet ud på at redde filer, ofte den eneste kopi, fra hardware, der var gået i stykker: datagendannelse fra SD-kort, roterende harddiske, SSD'er, RAID-arrays, eksterne drev og Mac-systemer. Hver af dem går i stykker og giver

sine data fra sig på sin egen måde, og et Mac-filsystem har sine egne særheder, så arbejdet handlede lige så meget om den rette tilgang til mediet som om nogen enkelt teknik. Målestokken var den enkle: filerne leveret tilbage i behold.

Hvor det går galt

Alle sider ser ændrede ud på én gang. En kopiering, en checkout, en gendannelse eller en eksport skriver filen om og nulstiller ændringstidspunktet, og alt længere nede i kæden, der stoler på det, ser en ændring, som aldrig skete. En kopiering, der bevarer tidspunkterne, gør det modsatte og skjuler en ændring, der faktisk skete. Tag indholdsdatoer fra indholdet eller fra versionshistorikken, og lad filens tidspunkt betyde det, det betyder.

En læser finder en halv fil. Et program, der skriver en fil om på stedet, efterlader et vindue, hvor en anden proces, eller den næste opstart, ser en del af den gamle version og en del af den nye. Skriv en midlertidig fil i samme mappe, kald `fsync` på den, og omdøb den derefter oven på originalen. POSIX kræver, at det navn, der erstattes, forbliver synligt under hele omdøbningen og peger på enten den gamle fil eller den nye.

Et script snubler over et lovligt navn. Et filnavn kan indeholde mellemrum, linjeskift og en indledende bindestreg, og et script, der deler navne ved mellemrum eller giver dem til en kommando uden anførselstegn, rammer de forkerte filer. POSIX råder selv til at undgå en indledende bindestreg og til at holde sig til bogstaver, cifre, punktum, understreg og bindestreg i navne, der skal kunne flyttes. Sæt anførselstegn om hver variabel, og afslut tilvalgslisten med `--` før en liste af navne.

En hemmelig fil kan læses af alle. Rettighedsbittene på en ny fil kommer fra standardværdierne i den proces, der oprettede den, og de lader som regel gruppen og andre læse. En nøgle eller en adgangssætning, der skrives på den måde, deles med hver eneste konto på værten. Opret hemmelige filer med rettigheder kun for ejeren fra begyndelsen i stedet for at stramme dem bagefter.

Læs mere

- NIST CSRC-ordliste: file — definitionerne fra NIST's sikkerhedspublikationer, hver med det dokument, den stammer fra.
- POSIX.1-2024 Base Definitions: Definitions — den formelle betydning af fil, almindelig fil, filnavn, stinavn, mappepost og fildeskriptor.

- POSIX.1-2024 Base Definitions: General Concepts — rettigheder, filens tre tidspunkter, flytbare filnavne og hvad et nedbrud kan efterlade i cachen.
- POSIX.1-2024: rename() — garantien for, at et navn, der erstattes, hele tiden peger på enten den gamle fil eller den nye, som sikker gemning bygger på.

Ordliste

Vælg ud fra jeres faktiske vilkår

Medbring udrulningsforløbet, driftsrisiciene og dem, der har vagten. Vi kan afgrænse en ændring med kontroller, en vej tilbage og tydeligt ansvar.

Se, hvad det koster Kontakt os →

Gør begreberne praktiske

Spørg, hvordan »Hvad er en fil?« gælder i jeres organisation. Et konkret eksempel fra jeres nuværende system er nok til at begynde samtalen.

Se vores ydelser

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

Tag denne side med dig

Visitkort

Lavet med kærlighed af D.

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

Copyright © 2025 – nu · Engineer ApS · Alle rettigheder forbeholdes.

Udskrevet fra https://engineer.company/da/glossary/file/