Engineer

コンピューターファイル

一つの単位として保存する、名前付きのデータのまとまり

どのようなものか

コンピューターファイルは、プログラムが、明日も同じバイトを取り出したいとき、OSへ求めるものです。名前で開き、読み書きし、閉じ、次に同じ名前を開くプログラムにも、残したものが見えると信頼します。文書、写真、プログラム、設定、ログは、すべてファイルです。Unix系では、どこにも保存しないものの一部も、ファイルです。

Linux、macOSなどのUnix系の基礎となる標準、POSIXは、ファイルを、読み取り、書き込み、または両方が可能で、アクセス権や種類などの属性を持つオブジェクトと定義します。種類には、通常ファイル、ディレクトリー、シンボリックリンク、FIFO、ソケット、文字型やブロック型の特殊ファイルを挙げます。最後のものは、端末やディスクなどのデバイスを表します。普通、ファイルと言うときは、最初の種類です。標準は、任意の位置へアクセスでき、システム自身は構造を課さない、バイトの列と説明します。表計算、写真、データベースのページのどれかを、システムは知りませんし、区別もしません。構造は、読むプログラムのものです。POSIXは、テキストとバイナリーすら分けません。ただし、多くのユーティリティーは、テキストだけで意味のある出力を返すと述べます。

名前は、ファイルの外にあります。ディレクトリーは、エントリーの一覧で、各項目が名前をファイルへ結びます。ファイル名は、スラッシュとヌルバイト以外のどんなバイトも含められる、一つの要素で、長さの上限はシステムが決めます。`/srv/site/index.html`のようなパス名は、到達するためにシステムがたどる、名前の連なりです。名前がディレクトリーにあるため、一つのファイルに複数を付けられ、システムが個数を保持します。最後の名前を消し、開いたままのプロセスもなくなると、領域を解放します。

システムが知る、それ以外の情報はメタデータで、詳細な一覧は、その大半を示します。

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

最初は、ファイルシステム内で一意の、ファイルシリアル番号です。その番号がファイルで、名前は、それを指すだけです。続くのは、所有者、グループ、その他に対する権限ビット、名前の数、所有者とグループ、バイト単位の大きさ、最後にデータを変更した時刻です。POSIXは、ほかに二つの時刻も持ちます。最後のデータアクセスと、最後のファイル状態の変更です。ファイルを開いたプログラムは、そのプロセスだけの小さい整数、ファイル記述子を持ちます。0、1、2は、標準入力、標準出力、標準エラーです。そのため、シェルは、プログラムが知らなくても、出力をファイルへ送れます。

概念を判断に生かす

「コンピューターファイルとは？」を実際の判断に生かしましょう。その概念を、開発または購入するシステムに結び付けて考えるお手伝いをします。 サーバー、メール、アカウント、更新を日々管理し、誰もが頼る仕組みを平穏に動かし続けます。

サービスを見る お問い合わせ →

用語の背後にある前提を検証する

解説とリンク先の実績を参考に、プロジェクトについての質問を整理してください。ツールを選ぶ前に、前提を一緒に検討できます。 UbuntuでPostgreSQL・MS SQLサーバー10台を運用 — 10 導入し正常に保ったDBサーバー

UbuntuでPostgreSQL・MS SQLサーバー10台を運用 お問い合わせ →

なぜ重要なのか

ファイルは、互いを知らないプログラムの間をつなぐ、最も古いインターフェースです。何十年も前の道具から、昨日書いた道具へ、バイトと名前以外の取り決めなしに、データを渡せます。設定、ログ、エクスポート、バックアップ、多くのサイトの内容は、この方法で移ります。ファイルの信頼性を保つことは、システムの信頼性を保つ仕事の大部分です。

微妙な問題は、メタデータにあります。更新時刻は、最後にバイトを書いたときを示し、最後にその意味が変わったときとは、別の事実です。同じ内容で書き直した瞬間に、両者は分かれます。所有者も番号で保存するため、別のマシンでは、同じファイルが別の人のものになる場合があります。

永続性も問題です。書き込みの呼び出しが戻っても、バイトをOSへ渡しただけで、キャッシュに残る場合があります。POSIXは、`fsync` などを呼んで求めたとき、キャッシュから記憶装置へ届き、それより早く、自動で届くこともあるとし

ます。クラッシュ後には、後の操作がディスクにあり、先の操作が欠ける場合があるとも警告します。例は、リネームで新しい名前を作ったのに、古い名前も残る場合です。

そして、誰かが手配しない限り、ファイルは、一台のデバイスの一つの複製です。記憶媒体の故障は、種類ごとに異なり、すべてに通用する唯一の保護は、別の場所へ二つ目を置き、読み戻しておくことです。

実際の活用例

ファイルの更新時刻は、最後にバイトを書いたときで、内容が最後に変わった瞬間とは違います。このサイトのサイトマップにも、かつて、その差が現れました。176ページのうち172ページが、同じ瞬間に変わったと示していました。修正は、日付の出所をファイル自体までたどりました。コンテンツのエクスポートが、毎回、全ファイルを書き直すため、内容にかかわらず、一回の同期でサイト全体が同じタイムスタンプになりました。今は、明示的なフィールド、コミット履歴、ファイルの順に試す、代替の連鎖を通じて、バージョン履歴から日付を取ります。事例は、そこでの落とし穴も記録します。フィールド名そのものを一覧へ入れないと、生成ツールは読みません。日付がなければ、推測を入れず空にします。記録された実装段階では、一か月の編集後の再同期で、186ファイル中、八つだけが変わりました。

会社の復元テストの実装は、ファイルを読み戻して終わります。systemdのユニットは、設定した場合のデータベースダンプ、暗号化し重複を除いたresticのスナップショットを、SFTPで別の事業者のストレージへ送る処理、保存方針に沿った削除、整合性の検査に対応します。別の月次ワークフローは、既知のファイルを一時的な作業場所へ復元し、バイトを比べます。意図したテストは、アップロードではなく復元です。2026-10-27の実装記録では、リポジトリーは未設定で、タイマーも無効で、完了した実環境の復元は記録されていません。パスフレーズまで、意図的にファイルにしています。systemdが解析する環境変数の値ではなく、保護したファイルから読みます。エスケープの処理で、バックスラッシュが変わり、リポジトリーを作ったときと違うパスフレーズになるのを避けるためです。

それより何年も前は、故障したハードウェアから、唯一の複製であることも多いファイルを救う仕事でした。データ復旧の対象は、SDカード、回転するハードディスク、SSD、RAIDアレイ、外付けドライブ、Macシステムです。壊れ方も、取り出す方法も異なり、Macのファイルシステムには、固有の癖もあります。単一の技法と同じくらい、媒体に合う方法が重要でした。尺度は、無傷で返したファイルという、明快なものでした。

問題が起きるところ

すべてのページが、一度に変わったように見えます。 コピー、チェックアウト、復元、エクスポートが書き直して更新時刻を変えると、それを信じる下流のシステムには、起きていない変更が見えます。時刻を保つコピーは、その逆に、実際の変更を隠します。内容の日付は、内容かバージョン履歴から取り、ファイル時刻には、本来の意味を保たせます。

読み手に、半分のファイルが見えます。 その場で書き直すと、別のプロセスや、次回の起動から、古い版と新しい版の一部ずつが見える時間ができます。同じディレクトリーへ一時ファイルを書き、そのファイルに対して `fsync` を呼び、元のファイルへリネームして置き換えます。POSIXは、置き換える名前が、リネームの全期間にわたり見え、古いファイルか新しいファイルの、どちらかを指すことを求めます。

正当な名前で、スクリプトが失敗します。 ファイル名は、空白、改行、先頭のダッシュを含められます。空白で分割したり、引用符なしでコマンドへ渡したりするスクリプトは、違うファイルを操作します。POSIX自身は、別の環境でも使う名前には、先頭のハイフンを避け、文字、数字、ピリオド、下線、ハイフンだけを使うよう勧めます。展開結果はすべて引用符で囲み、名前の一覧の前に `--` でオプションの解析を終えます。

秘密のファイルを、誰でも読めます。 新しいファイルの権限ビットは、作ったプロセスの既定値から決まり、通常、グループやその他も読めます。そのまま書いた鍵やパスフレーズは、ホスト上の全アカウントと共有されます。後から絞らず、最初から所有者だけが使える権限で作ります。

参考資料

- NIST CSRC用語集：ファイル — NISTのセキュリティ文書で使う定義と、それぞれの出典です。
- POSIX.1-2024基本定義：定義 — ファイル、通常ファイル、ファイル名、パス名、ディレクトリーエントリー、ファイル記述子の正式な意味です。
- POSIX.1-2024基本定義：一般概念 — 権限、三つの時刻、ファイル名の可搬性、クラッシュ後にキャッシュが残し得るものです。
- POSIX.1-2024：rename() — 安全な保存を支える、置き換える名前が、全期間を通じて古いファイルか新しいファイルを指すという保証です。

用語集

実際の制約に合わせて選ぶ

デプロイの流れ、運用上のリスク、緊急対応の担当者をお知らせください。検査、復旧手段、明確な責任の所在を伴う変更の範囲を定められます。

料金を見る お問い合わせ →

用語を実務に結び付ける

「コンピューターファイルとは？」がお客様の組織にどう当てはまるかをご相談ください。現在のシステムの具体例が一つあれば、話を始められます。

サービスを見る

welcome@engineer.company +45 50 10 61 84 デンマーク、コペンハーゲン

このページを持ち歩く

連絡先カード

心を込めて丁寧に制作： D.

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

著作権 © 2025 – 現在 · Engineer ApS · すべての権利を留保します。

印刷元： https://engineer.company/ja/glossary/file/