Engineer

バックアップ

消失に備えたデータのコピーと、その作成

どのようなものか

バックアップは、元のデータが削除、破損、またはマシンとともに消失したとき、復元するためのものです。そのため、その出来事が届かない場所に置く必要があります。NISTの用語集は、復旧を容易にするためのファイルとプログラムのコピーと定義します。英語の名詞backup、複数形backupsはコピーを、二語の動詞back upは作成を指します。夜間ジョブがデータベースをバックアップし、一つのバックアップを残す、という関係です。

近いものが混同されます。レプリカは数秒で全変更を受け、ディスク故障によるデータの消失を防げても、誤った `DELETE` も忠実にコピーします。スナップショットは、ある時点のボリュームを固定しますが、元の故障が届かない場所に保存して初めてバックアップになります。定義を決めるのは、履歴と分離です。元を失った出来事に耐えた場所から、過去の状態を返します。

データベースでは具体的になります。PostgreSQLのマニュアルは、根本的に異なる三つを示します。 `pg_dump` のSQLダンプは、開始時点で整合したスナップショットとして、再構築するSQLを書き出します。ファイルシステム単位では、データディレクトリをコピーします。継続的アーカイブは、そのベースバックアップと、以後の先行書き込みログWALを合わせます。そのログを再生すれば、ベースバックアップ以後の任意の時点にデータベースを戻せます。これをポイントインタイムリカバリーと呼びます。

概念を判断に生かす

「バックアップとは？」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。 あると信じるだけでなく、復元を試したバックアップと、サーバーが起動しない朝に使える手順書を用意します。

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。 Barmanで重要データのバックアップ100件を自動化 — 100 データベース全体で自動化した重要バックアップ

Barmanで重要データのバックアップ100件を自動化 お問い合わせ →

なぜ重要なのか

安全でありたいという希望を、二つの数値が設計へ変えます。NISTは、復旧時点目標RPOを、停止後にデータを戻すべき時点とします。実務では、事業として直近の作業結果をどれだけ失っても耐えられるかです。復旧時間目標RTOは、停止が組織の活動に支障を及ぼすまでに、システムの復旧へ費やせる時間です。

両者は異なる部分に影響します。RPOはコピー頻度で決まり、夜間ダンプなら一日分の書き込みを失う可能性がありますが、WALアーカイブなら、アーカイブへ転送中の直近のログ分まで損失を減らせます。RTOは復元時間で決まります。PostgreSQLは、最後のベースバックアップ以後の全WALセグメントを再生するため、そのバックアップが古いと時間がかかります。ベースバックアップを取る頻度は、保存容量だけでなくRTOにも関わる判断です。

どちらも復元に関する主張なので、復元して評価します。終了コードゼロのジョブは、どこかにバイトを書いたことを示すだけです。RTO内に動くシステムを戻せるかは、本番以外に定期復元し、結果を比べて分かります。

もう半分は分離です。同じアカウント、同じ認証情報のコピーは、閉鎖、鍵の盗難、誤った対象へのスクリプトという、耐えるべき故障を共有します。どんな一つの出来事が、元と一緒にコピーを消すかを問うのが有用です。

実際の活用例

最小規模では、自社のバックアップと復元テストの実装を、事業者のスナップショットを超えて復旧できるよう設計しました。Ansibleロールがresticのスクリプトとsystemdユニットを入れ、設定した場合のデータベースダンプ、暗号化・重複排除したスナップショット、保存期限による削除、整合性検査に対応します。別の毎月の処理で、既知のファイルを一時的な場所へ戻し、バイトを比較します。バックアップと復元の成功信号は分け、タイマーには設定済みのリポジトリが必要です。パスフレーズは、systemdが解析する環境変数でなく、保護したファイルから読みます。これらは実装済みの制御と有効化の前提です。2026-10-27付の記録は、リポジトリ未設定でタイマーが無効だったとし、実環境で完了した復元は記録していません。自動化の作業を裏付けるもので、復旧成功や、今すでにホストを安全に失えることの主張ではありません。

より大きい規模では、GISデータを基盤とする製品向けに、BarmanとGoogle Cloudで百の重要データのバックアップを自動化しました。BarmanがPostgreSQLを、Google Cloudが遠隔コピーを担い、BashとPythonで実行を調整・検証しました。単に存在するだけでなく、正常であることを整合性検査で確認しました。

三つ目の製品では、復旧可能な状態を仮定せず、定義しました。バックアップの保持設定を、データベースの隣のBicepテンプレートに置きます。特定時点への復元は、ポータルで誰かが一度選んだ設定ではなく、テンプレートの性質になり、周囲の環境もコードです。監査は不足も書きました。地理冗長性は無効で、RTOは実測でなく提案であり、測定する訓練は記述済みですが、まだ実行待ちです。

問題が起きるところ

復元が必要になるまで隠れがちなので、平穏な日に探す価値があります。

稼働中のファイルコピーから、壊れたクラスタが戻ります。 動いているPostgreSQLのデータディレクトリを `tar` でコピーすると、異なる時点のファイルが混ざります。マニュアルは、この方法で使えるバックアップを得るには、サーバーを停止する必要があるとしています。WALを含む、ある一時点の状態を不可分に取得したファイルシステムのスナップショットなら、起動時に障害後と同じく再生するため使えます。ただし、データとWALが別のファイルシステムなら、同時取得できない場合があります。 `pg_dump` や継続的アーカイブ用のベースバックアップなら、稼働中も整合したコピーを得られます。

WALの欠落で、特定時点への復旧がそこで止まります。 ベースバックアップの開始時点までさかのぼる、途切れないアーカイブ済みWALファイルの列が必要です。アーカイブコマンドの失敗が続くと再試行し、 `pg_wal/` に積もるため、最初の症状はディスク満杯になりがちです。最初のベースバックアップより前にアーカイブを設定・テストし、遅延を監視するよう、マニュアルは勧めます。

ジョブが止まり、静けさが成功に見えます。 タイマーの失敗が、読まれないログだけに書かれ、静かな日をよい日と考えます。監視は無音を警報にする必要があります。成功後だけ、遅れれば緊急通知する外部サービスへ報告すれば、派手な失敗だけでなく、停止したジョブも捕まえます。

鍵が失われるか、変わって届きます。 暗号化されたデータへのアクセスは鍵に依存します。resticでは一つのリポジトリに複数のパスワードを設定でき、使える鍵がそのリポジトリに一つも残っていない場合に、データを復元できないという警告が当てはまります。鍵は、マシンを失っても残る場所に置く必要があります。上記では、サービス管理機構自体が変更の原因になり得ました。環境変数のエスケープを解釈するため、バックスラッシュのあるパスフレーズを渡すと、作成時と異なるバイトが、警告なくresticへ届き得ました。ファイルから読むことで、この経路を閉じます。

保持ルールが必要なコピーを消します。 破損は数週間後に気付くことがあり、日次七世代だけなら、正常なものをすべて壊れたものに置き換えています。段階的なルールなら、低い費用で長く残せます。resticの `--dry-run` で、何も消す前に削除予定を見られます。

```
restic forget --keep-daily 7 --keep-weekly 5 --keep-monthly 12 \
  --prune --dry-run
```

`forget` がスナップショットを、 `prune` がそれらだけが参照するデータを消すため、prune後の古いコピーは完全になくなります。

整合性検査が通っても復元できません。 標準の `restic check` は、全ダウンロードを避け、全packファイルを読まずに構造を検証します。 `--read-data` は全体、 `--read-data-subset=1/5` は五分の一を読み、 `1/5` から `5/5` の五回で全体を覆えます。全データを読んでも、保存したバイトが破損していないことを示すだけです。正しいロールと設定を持つデータベースへ戻せるかは、復元で示します。

参考資料

- NIST用語集：backup：ここで使う定義と出典の文書です。
- NIST SP 800-34 Rev. 1：RPO・RTOを定義する継続計画のガイドです。
- PostgreSQL：バックアップと復元：三つの方法と、それぞれに必要な負担を説明します。
- PostgreSQL：継続的アーカイブと特定時点への復旧：ベースバックアップ、WALのアーカイブ、選んだ時点への復元を説明します。
- restic：新しいリポジトリの準備：パスワード、ファイルからの読み取り、紛失の警告です。
- restic：バックアップスナップショットの削除：保持ルール、 `forget` 、 `prune` 、 `--dry-run` を説明します。

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

「バックアップとは？」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/backup/