

PostgreSQL のバックアップとリカバリー

技術を知る

データベースには、お客様の重要なデータが保管されています。そのデータを守るために、データベースの障害に備えてデータベースの内容を定期的にバックアップしておくことが重要です。

ここでは、PostgreSQL データベースと Fujitsu Enterprise Postgres（以降、Enterprise Postgres と略します）のバックアップとリカバリーについて解説します。バックアップの基礎から Enterprise Postgres を使用した場合のメリットまでを説明します。

1. PostgreSQL のバックアップ・リカバリーとは？

まずは、バックアップの基本から説明します。PostgreSQL データベースのバックアップには、「論理バックアップ」と「物理バックアップ」があります。それぞれの違いは以下のとおりです。

- 論理バックアップ

データベースに格納されているデータをファイルに取り出すバックアップ方法です。例えば SQL の形で取り出すと処理時間は長くなりますが、別バージョンや別システムへの移行などができるため柔軟性に優れています。

- 物理バックアップ

データベースのファイルそのものを物理的にコピーするバックアップ方法です。ファイルのコピーなので処理時間が短く、ハードウェア障害からのリカバリーなどに利用されています。

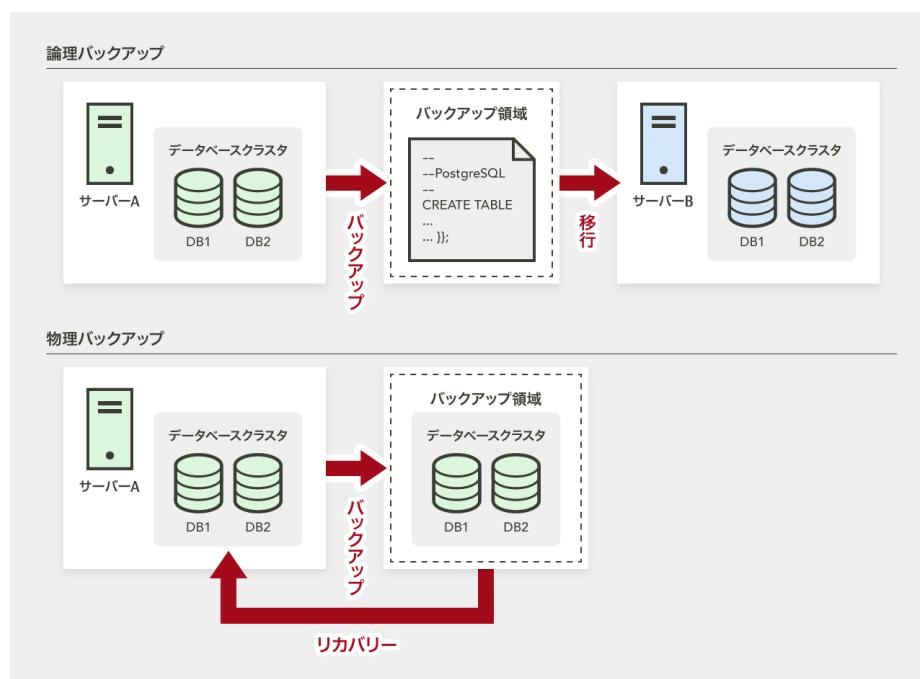


図 1：各バックアップ方法のイメージ

次に、それぞれのバックアップ方法について具体的に説明していきます。

1.1 論理バックアップ[†]

論理バックアップは、PostgreSQL データベースを起動した状態でデータベースのデータをファイルに取り出します。

論理バックアップを実施する方法として以下の 3 つがあります。

- **pg_dumpall コマンド**

データベースクラスタ全体の内容を SQL の形でバックアップします。取得したバックアップデータは、psql コマンドを使ってバックアップを取得した時点ヘリカバリーします。

- **pg_dump コマンド**

指定したデータベースの内容を SQL の形でバックアップします。スクリプト形式のバックアップデータは、psql コマンドを使ってバックアップを取得した時点ヘリカバリーします。アーカイブファイル形式のバックアップデータは、pg_restore コマンドを使ってバックアップを取得した時点ヘリカバリーします。

- **SQL コマンドの COPY**

指定したテーブルのデータを独自のテキスト形式、バイナリー形式、または CSV 形式でバックアップします。取得したバックアップデータは、COPY を使ってバックアップを取得した時点ヘリカバリーします。

1.2 物理バックアップ

物理バックアップは、データベースのファイルをそのまま物理的にコピーします。

PostgreSQL データベースを停止した状態でバックアップを取得する「オフラインバックアップ」と、起動した状態でバックアップを取得する「オンラインバックアップ」があります。

1.2.1 オフラインバックアップ

オフラインバックアップは、PostgreSQL データベースを停止した状態で、os 付属のコマンドなどを使ってデータベースのファイルをすべてコピーすることで、データベースクラスタ全体をバックアップします。取得したバックアップデータは、os 付属のコマンドを使って戻し、その後 PostgreSQL データベースを起動することで、バックアップを取得した時点ヘリカバリーできます。

1.2.2 オンラインバックアップ

オンラインバックアップは、PostgreSQL データベースを起動した状態でバックアップを取得します。

まず、起点となるデータのベースバックアップを取得します。ベースバックアップの取得中もデータは更新されていますので、更新処理はトランザクションログ（以降、WAL と略します）として保存されます。WAL は必要に応じてアーカイブされ、アーカイブログとして保存されます。データベース破壊など何らかのトラブルが発生した場合は、ベースバックアップ + アーカイブログ + WAL を使ってリカバリーします。

- トランザクションログとは？開く

以下に、オンラインバックアップを取得したあと、トラブルが発生して最新状態にリカバリーするまでの流れを示します。

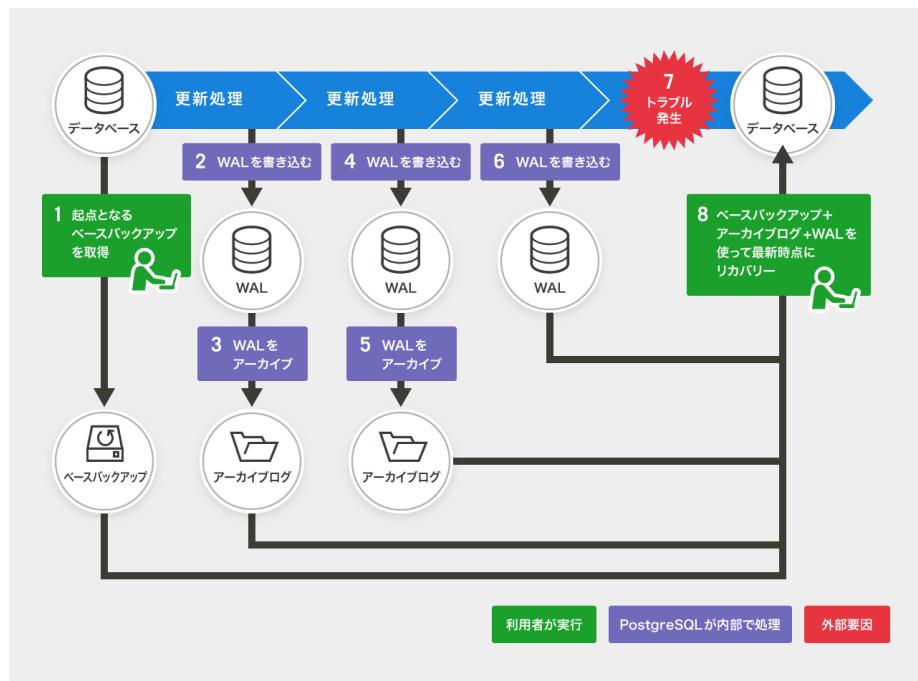


図2：オンラインバックアップの流れ

アーカイログと WAL を利用したリカバリーでは、バックアップ時点へのリカバリー、最新時点へのリカバリーのほか、任意の時点へのリカバリーも可能です。

PostgreSQL が持つオンラインバックアップ機能は以下です。

- 繼続的アーカイブによるオンラインバックアップ

データベースをバックアップモードにしたあと、OS の機能などでファイルをコピーし、コピー完了後にバックアップモードを終了する方法です。取得したバックアップデータは、ポイントインタイムリカバリーを使ってリカバリーします。復旧時点（リストアポイント）にはバックアップ時点、任意の時刻または任意の地点などが設定できます。

Enterprise Postgres では、PostgreSQL が持つ機能に加えて、さらに以下のオンラインバックアップ機能が使用できます。

- **pgx_dmpall コマンド**

Enterprise Postgres が提供するコマンドです。継続的アーカイブによるオンラインバックアップをワンタッチで行うコマンドです。取得したバックアップデータは、pgx_rcvall コマンド または WebAdmin でリカバリーします。復旧時点（リストアポイント）にはバックアップ時点、任意の時刻または任意の地点などが設定できます。

- **WebAdmin**

Enterprise Postgres が提供する GUI です。継続的アーカイブによるオンラインバックアップをワンタッチで行う GUI です。バックアップ取得時点と最新時点へリカバリーできます。

いずれの方法においてもデータベースクラスタ全体をバックアップします。

- ポイントインタイムリカバリーとは？開く
- pgx_rcvall コマンドとは？開く

Enterprise Postgres にはこんな使い方も！

大量データの高速バックアップ

オンラインバックアップにおけるベースバックアップのデータが大量だった場合はバックアップに時間がかかります。このような場合に、FUJITSU Storage ETERNUS などのストレージ製品にある高速コピー機能を活用すると、大量データを高速にバックアップすることができます。

Enterprise Postgres では、pgx_dmpall コマンドの実行時に、ユーザーが作成したコピー命令を呼び出すことで、任意のバックアップ先へバックアップできます。詳細については、「PostgreSQL のバックアップとリカバリー ~ 大量データの高速バックアップ ~」の記事や製品に添付されているマニュアル“Fujitsu Enterprise Postgres 運用ガイド”をご覧ください。

2. バックアップ方法の決め方は？

PostgreSQL データベースのバックアップには色々な方法があることがおわかりいただけたと思います。

では、業務に合ったバックアップ方法はどのように選ぶのが良いのでしょうか？選択基準となる観点には、「障害発生時にどの時点まで復旧する必要があるか、バックアップするために稼動している業務サービスを止めることができるか、どのような事象に備えるか」などがあります。お使いのシステムの運用要件に合った最適なバックアップ方法を選択してください。

例えば、以下のような運用要件があった場合における最適なバックアップ方法の選択例を示します。

- a. お客様の誤操作によりテーブルのデータ消失などが発生した際に、そのテーブルだけをバックアップ時点にリカバリーしたい場合は
 - ✓ テーブル単位のバックアップなので、論理バックアップの pg_dump コマンドを選択。
- b. 別サーバーに新システムを構築し、旧システムの全データを移行して同じ業務を継続したい場合は
 - ✓ システム全体の移行なので、論理バックアップの pg_dumpall コマンドを選択。
- c. ディスクのハードウェア障害が発生した際に、障害が発生する直前の最新時点にリカバリーしたい場合は
 - ✓ 物理的な復旧は物理バックアップで対応。最新時点への復旧は継続的アーカイブによるオンラインバックアップ、pgx_dmpall コマンドまたは WebAdmin を選択。
- d. データが論理的に破壊されてデータベースが正常に動作しない際に、破壊された時刻にリカバリーしたい場合は
 - ✓ データベース破壊の復旧は物理バックアップで対応。復旧時刻を指定したいので継続的アーカイブによるオンラインバックアップまたは pgx_dmpall コマンドを選択。

これらの要件に対応するバックアップ方法の選択例を以下に示します。図中の「独自機能」のバッジが付与された赤枠の機能が Enterprise Postgres の提供する独自機能です。

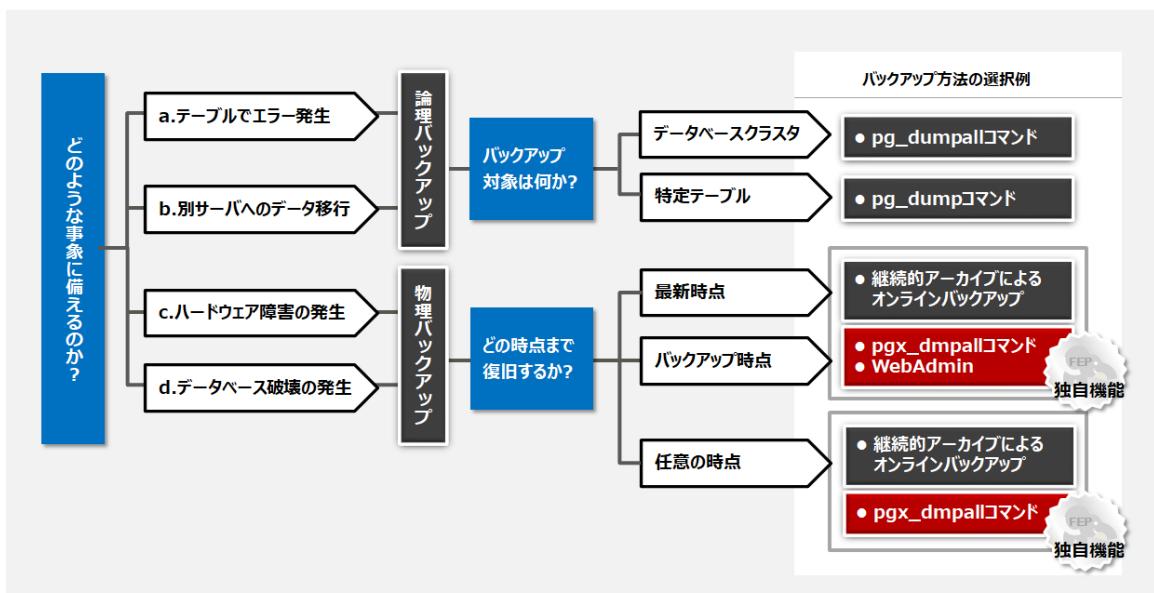


図 3: バックアップ方法の選択例

3. PostgreSQL のバックアップとリカバリーのまとめ

今まで説明してきたことをここでまとめます。バックアップとリカバリーには、PostgreSQL の機能と Enterprise Postgres の機能を合わせて 7 種類の方法があります。

以下に、各バックアップの方法と対象、リカバリー時の復旧時点について詳細を表にまとめます。

なお、表中記号の意味は以下のとおりです。

レ : 機能有り、N/A : 機能無し、赤塗りのセル : Enterprise Postgres の機能

分類		概要	バックアップ方法	バックアップ対象	リカバリーの復旧時点(リストアポイント)		
					バックアップ	最新	任意
論理 バック アップ		データベース起動状態で、データを取り出して保存。定義・データ単位のバックアップも可能。	pg_dumpall コマンド	データベースクラスタ	レ	N/A	N/A
			pg_dump コマンド	データベース	レ	N/A	N/A
			SQL コマンドの COPY	テーブル	レ	N/A	N/A
物理 バック	オ フ ラ イ ン	データベース停止状態で、データベースのファイルを物理的にコピー。	OS 付属 コマンド	データベースクラスタ	レ	N/A	N/A

分類		概要	バックアップ方法	バックアップ対象	リカバリーの復旧時点 (リストアポイント)		
					バックアップ	最新	任意
ア ッ プ	バ ッ ク ア ッ プ						
オ ン ラ イ ン バ ッ ク ア ッ プ	データベース起動状態で、データベースのファイルを物理的にコピー。WALも使用。	継続的アーカイブによるオンラインバックアップ	WebAdmin	データベースクラスタ	レ	レ	レ
				データベースクラスタ	レ	レ	N/A
		pgx_dmpall コマンド		データベースクラスタ	レ	レ	レ

データベースではシステムの運用管理要件にあった最適なバックアップ方法を選択して、業務データの保守運用を実施してください。

2025年1月27日