

2020年4月15日
株式会社インプレスR&D
<https://nextpublishing.jp/>

ポストグレでデータベースを始めよう！
『データベース初心者のための PostgreSQL 教室』発行
技術の泉シリーズ、4月の新刊

インプレスグループで電子出版事業を手がける株式会社インプレス R&D は、『データベース初心者のための PostgreSQL 教室』(著者:目黒 聖)を発行いたします。

最新の知見を発信する『技術の泉シリーズ』は、「技術書典」や「技術書同人誌博覧会」をはじめとした各種即売会や、勉強会・LT 会などで頒布された技術同人誌を底本とした商業書籍を刊行し、技術同人誌の普及と発展に貢献することを目指します。

『データベース初心者のための PostgreSQL 教室』
<https://nextpublishing.jp/isbn/9784844378563>



著者:目黒 聖

小売希望価格:電子書籍版 1600 円(税別)／印刷書籍版 2500 円(税別)

電子書籍版フォーマット:EPUB3／Kindle Format8

印刷書籍版仕様:B5 判／カラー／本文 200 ページ

ISBN:978-4-8443-7856-3

発行:インプレス R&D

<<発行主旨・内容紹介>>

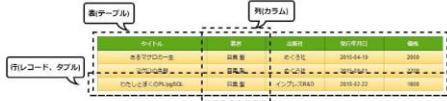
本書はデータベース初心者および PostgreSQL 初心者向けの入門書です。
データベースとは何か？から PostgreSQL のインストール、SQL の実行、トランザクションについて、レプリケーション、バックアップまでを解説しています。

(本書は、次世代出版メソッド「NextPublishing」を使用し、出版されています。)

データベースの考え方から丁寧に解説

です。一般的には「レコード」ということが多いですが、PostgreSQLでは「タプル」と呼ぶことも多いです。

図1.3: 関係モデルの用語



キー
テーブルで重要な意味を持つ列のことを「キー」や「キー列」といいます。特に、その列の値が決まればそのテーブルの中でレコードが1行に特定できることを、テーブルの「主キー」とか、「プライマリキー」などともいいます。1行に特定しなければならぬので、主キーはテーブル内で重複はありません。
たとえば世の中には同じタイトルのつけられた本が複数ありますが、ISBNという全世界の書籍一つ一つにつけられたコードがわかれば、そのなかから1冊の書籍を特定できます。同じISBNが全く異なる書籍に振られていたら、ちょっと世間が混乱するでしょう。

1.5 なぜDBMSを使うのか

データベースについて説明してきましたが、皆さんの中には「DBMSをいちいち使わなくても、表計算ソフトやCSVなどのテキストベースのファイル管理でも別にいいんじゃないの?」と思った方がいるかもしれません。
筆者自身データベースをよく知らなかったときはそう思いましたし、データベースにあまり詳しくない人にそう言われたこともあります。
DBMSを使うと何が便利なのか?とよは次のような点が挙げられます。

1.5.1 同時にデータにアクセスできる

たとえば、あなたの会社ではExcelで作成したプロジェクトの管理表がファイルサーバーに置いてあり、各自がそれぞれアクセスして更新していませんか?¹
そういう運用をしているところでは、「誰かが更新してるから自分が更新できない」とか、「自分がローカルで更新したファイルをサーバーにアップロードして、誰かが更新した内容を消してしまった」ということを体験した人も多いと思います²。

1. この例も、「一重で編集」とか「ロックで解決」といわれています。一部は目録ではまだ使われているようですが、RDBMSでは非同期で併発の恐れがあるのでよくない。
2. 実際はそういふ更新はしていません。ついでに、
3. 実際はよくやります。ついでに、

DBMSを使えばこのようなことを心配する必要はありません。同時にデータにアクセスする仕組みをきちんと備えているからです。

1.5.2 大容量データを扱える

あなたは何10メガも何100メガもあるExcelファイルを開かざるを得なかったことはありませんか? それなりの性能のPCでも、まづ開くだけで時間がかります。Excelは1シートにつき100万行以上のレコードを扱えますが、実際に100万行データを入れることはまず無理でしょうし、数千から10000行でもExcelにしては多い、と思います。
また、何100メガのテキストファイルをエディタで開き、必要な行を見つけデータの変更や更新をするのは一苦労というのは容易に想像できるでしょう。
RDBでは何10メガ何100メガは当たり前。千ガ単位はざらにあります。100万行のレコードを保持しているテーブルなんてゴロゴロあります。何億行だって存在します。
それでも高速に必要なデータを参照したり更新したりできる仕組みが存在しています。

1.5.3 データの矛盾・重複を防ぐ

Excelのような表計算ソフトやCSVのようなテキストファイルでは、データの分散と重複が防げません。そのため、あるデータを更新するために複数のシートやファイルを更新しなければならず、更新漏れを防ぐのは本人の努力以外ありません。
また、本来ならば数値を入れるべき箇所に誤って文字を入力した場合、エラーとして弾くのも難しいでしょう。
DBMSなら、このようなデータの矛盾や重複を防ぐ仕組みがあります。

1.5.4 別テーブルを関連付けて検索できる

この機能がRDBMSの真骨頂と言ってもいいでしょう。複数のテーブルを同時に検索し、検索結果として一つの表として表示できます。プログラムを書かずにExcelやCSVで同じ芸当をこなすのはかなり難しいでしょう。

どうでしょう、DBMS、特にRDBMSについて、なんとなく理解できたでしょうか。ではデータベースについて理解できたところで、次章では本書で扱うPostgreSQLというRDBMSについて、もう少し詳しく説明します。

PostgreSQLの歴史から概要を紹介

第2章 PostgreSQL概要

第1章「データベースを始めよう」ではRDBMSのごく基本的な知識を学習しました。世の中にはさまざまなRDBMSがありますが、本書ではPostgreSQLを扱います。本章ではPostgreSQLというRDBMSがどのようなものをご紹介します。

2.1 PostgreSQL前史

PostgreSQL。日本語では「はすとくれえすきゅーえーる」と読み、略して「はすくれ」とか、「はすとくれす」とかと言ったりもします。
PostgreSQLの大本をたどると、カリフォルニア大学バークレイ校の助教だったマイケル・ストーンブレイカー氏が、「1A 関係モデルとは」で登場したコッド氏の論文を読み、RDBMSの研究として1973年に開発を始めたIngres¹に突き当たります。
Ingresで採用した関係モデルに行き詰まりを感じたストーンブレイカー氏は、Ingresの後継となるプロジェクトを始めます。それがPOSTGRES(Post-INGRES)プロジェクトです。
POSTGRESは研究用や教材、商用として広まったのですが、ユーザー数が増えた結果ソースの保守開発に研究家のリソースの大半が取られてしまうということで、1993年にPOSTGRESプロジェクトはPostgres 4.2をリリースして正式に終了しました。
翌1994年、Postgres95がオリジナルのPOSTGRES/バークレイコードの後継としてWeb上でリリースされます。さらに1996年名前を「PostgreSQL」に変え、バージョンを6.0としてリリースします。
これが今日まで続くPostgreSQLの歴史の始まりです。

2.2 PostgreSQLの開発形態

ソフトウェアは、売り物として販売されている商用ソフトウェアと、ソースコードが公開され、ライセンスさえ守れば誰でも無償で自由に使用や修正、さらに再配布が可能なオープンソースソフトウェアのふたつに大別できます。
DBMSにもこのふたつが存在し、PostgreSQLは後者のオープンソースソフトウェアにあたります。同じオープンソースソフトウェアのRDBMSとして有名なMySQLは、開発はOracle社が中心となっていますが、PostgreSQLは、現在多くの企業にも属さずコミュニティが中心となって開発が行われています。メジャーバージョンアップは原則1年に1回行われ、今も機能の拡充と性能向上が

続けられています。

図2.1: 最近のバージョンアップと追加機能

バージョン	リリース日	主な改修点
9.0	2010/9/20	ストリーミングレプリケーション機能の追加。
9.1	2011/9/12	同期レプリケーション機能の追加。
9.2	2012/9/10	カスケードレプリケーション、Index Only Scan機能の追加。
9.3	2013/9/9	マテリアライズドビュー、更新可能ビュー機能の追加。
9.4	2014/12/18	JSONB型の追加、ALTER SYSTEM文の追加、レプリケーションロック機能の追加。
9.5	2016/1/7	UPSERT機能、行単位セキュリティ機能の追加。
9.6	2016/9/29	パラレルクエリ、レプリケーションで複数同時スキャンが可能。
10	2017/10/5	ロジカルレプリケーション、書き出しレプリケーション機能の追加。
11	2018/10/18	パラレルクエリの拡張、JITコンパイルの導入。
12	2019/10/3	JSON Pathに対応、パーティショニングの性能改善

一般、いや二倍くらい前までは、オープンソースソフトウェアという、無料で利用できる分、商用よりも劣る品質のものであると考える人も多かったのですが、PostgreSQLは常に機能の拡充と性能の向上を行っており、商用DBMSと遜色なくらいの機能と品質を持っています。今では実際のサービスでPostgreSQLや他のオープンソースソフトウェアのRDBMSが採用されていることは珍しくありません。

PostgreSQLのバージョン番号
PostgreSQLのバージョンが8から10になったときに、小数点以下の数字がなくなったことにお気づきでしょうか。これは、10がリリースされたときに、小数点第一位までメジャーバージョンを扱っていた従来のバージョン番号の形式を改め、整数部分がメジャーバージョン、小数点がセキュリティアップデートやバグフィックスでバージョンアップするマイナーバージョンとすることになったからです。
このPostgreSQLを使うようになったときは、とりあえず性能や機能が改善されている最新版を使うことをおすすめします。

2.3 PostgreSQLの特徴

2.3.1 ライセンス

オープンソースソフトウェアはソースコードを公開しているとはいえ、決して権利を放棄しているわけではありません。ちゃんとライセンスを公開しており、ライセンスに反しない場合は訴訟もあ

1. 1973年発表。計算機学、大学で研究。開発する方針が定まらず、いくつかの会社で結果はあり、開発を止まるまで未公開に閉じられていた。最終的には、商用で公開された。2. 開発者データベースのコミュニティ。多くのRDBMSの歴史になった。ストーンブレイカー氏の論文は関係データベース (Ingres) のことではない。
3. 当時開発者間で有名な Oracle Database を開発している Oracle 社が MySQL を開発したことが、オープンソースに傾いてはいたと見られる。

第5章 テーブルを作ろう

PostgreSQLのインストールし、起動を行い、psqlによる接続も行いました。本章では、データを格納するテーブルを作成していきます。

5.1 またデータベース？

「4.1 データベースの作成？」で、データベースという言葉には4つの意味があると書きましたが、本章ではその4つ目の意味「データを格納する単位」を説明します。

5.1.1 デフォルトデータベース

では一度、通常使用ユーザーでpsqlを使って接続し、次のコマンドを実行してください。

```
リスト 5.1: データベース一覧表示
\l
```

次のような結果が出力されたはずですよ。

```
postgres=# \l
```

List of databases						
Name	Owner	Encoding	Collate	Ctype	Access privileges	
postgres	postgres	UTF8	C	C		
template0	postgres	UTF8	C	C		rc/postgres
template1	postgres	UTF8	C	C		rc/postgres/postgres
						rc/postgres
						rc/postgres/postgres

(3 rows)

PostgreSQLでは、1つのデータベースクラスター、つまりは1つのインスタンスの中に、複数のデータベースを作成することができます。デフォルトで3つのデータベースが作成されます。ただし、データベースをまたがってデータをやりとりすることはできませんので注意してください。

PostgreSQLでinitdbした直後にデフォルトで作成される3つのデータベースについて説明します。

1. 「4.1 解説」でインスタンスを構築している場合は削除してから実行してください。

2. 複数のデータベースを構築している場合はデータベースクラスターという、名前に注意します。

postgres

デフォルトの接続先として用意されているデータベースです。削除してもPostgreSQLとしての動作には問題ありませんが、このデータベースの存在を前提としている外部のサービスや拡張機能などが多く存在しているため、安易に消すと危険です。

ただデータベースを試してみるだけならpostgresデータベースを使用すればいいのですが、本書運用などで使用する場合はpostgresという名称では何のためのデータベースなのかわかりにくいので、攻撃する側としてもpostgresというデータベースが存在することは想定しやすく大事なデータを入れるには攻撃しやすいですから、ちゃんとした目的を持った名称のデータベースを作成すべきでしょう。

template1

PostgreSQLでデータベースを作成する際にデフォルトのテンプレートとなるデータベースです。

新規にデータベースを作成するときには常に作成しておきたいテーブルや、有効にしておきたい設定などは、このtemplate1にあらかじめ作成しておくことで、データベース作成時に複製されます。

削除してもPostgreSQLの動作は可能ですが、データベース作成時に少し困るので、削除するのはおススメしません。

template0

initdbを実行した直後の、使用しているバージョンのPostgreSQLで定義済みの標準オブジェクトのみから構成されているデータベースです。このデータベースはこの初期状態で保存しておくべきものです。このデータベースを削除したり、手を加えたりしてはいけません。

```
リスト 5.2: template0をコピーしてデータベースを作成
```

```
CREATE DATABASE dbname TEMPLATE template0;
```

5.1.2 データベースの新規作成

<<目次>>

- 第1章 データベースを始めよう
- 第2章 PostgreSQL 概要
- 第3章 PostgreSQL をインストールしよう
- 第4章 データベースを作成しよう
- 第5章 テーブルを作ろう
- 第6章 SQL でデータを操作してみよう
- 第7章 トランザクションと同時実行制御
- 第8章 プログラムからPostgreSQLを操作しよう・準備編
- 第9章 プログラムからPostgreSQLを操作しよう・実践編
- 第10章 追記型アーキテクチャとVACUUM
- 第11章 止まらないデータベースシステムを構成しよう
- 第12章 データが壊れても元に戻せるようにしよう
- 第13章 性能問題の調査と改善のヒント
- 第14章 セキュリティと監査
- 第15章 他のDBMS の紹介
- 付録A PostgreSQL の Windows へのインストール
- 付録B GUI でPostgreSQLを操作してみよう

<<著者紹介>>

目黒 聖

某都内 SI 勤務のエンジニア。もともと開発で必要だったためにDBを学び始めたはずなのに、いつのまにか開発から離れDBAに。趣味で居合をやりながら、Postgre で面白いことができないか、日々考えています。

<<販売ストア>>

電子書籍:

Amazon Kindle ストア、楽天 kobo イーブックストア、Apple Books、紀伊國屋書店 Kinoppy、Google Play Store、honto 電子書籍ストア、Sony Reader Store、BookLive!、BOOK☆WALKER

印刷書籍:

Amazon.co.jp、三省堂書店オンデマンド、honto ネットストア、楽天ブックス

※ 各ストアでの販売は準備が整いしだい開始されます。

※ 全国の一般書店からもご注文いただけます。

【インプレス R&D】 <https://nextpublishing.jp/>

株式会社インプレス R&D(本社:東京都千代田区、代表取締役社長:井芹昌信)は、デジタルファーストの次世代型電子出版プラットフォーム「NextPublishing」を運営する企業です。また自らも、NextPublishing を使った「インターネット白書」の出版など IT 関連メディア事業を展開しています。

※NextPublishing は、インプレス R&D が開発した電子出版プラットフォーム(またはメソッド)の名称です。電子書籍と印刷書籍の同時制作、プリント・オンデマンド(POD)による品切れ解消などの伝統的出版の課題を解決しています。これにより、伝統的出版では経済的に困難な多品種少部数の出版を可能にし、優秀な個人や組織が持つ多様な知の流通を目指しています。

【インプレスグループ】 <https://www.impressholdings.com/>

株式会社インプレスホールディングス(本社:東京都千代田区、代表取締役:唐島夏生、証券コード:東証1部9479)を持株会社とするメディアグループ。「IT」「音楽」「デザイン」「山岳・自然」「旅・鉄道」「学術・理工学」を主要テーマに専門性の高いメディア&サービスおよびソリューション事業を展開しています。さらに、コンテンツビジネスのプラットフォーム開発・運営も手がけています。

【お問い合わせ先】

株式会社インプレス R&D NextPublishing センター

TEL 03-6837-4820

電子メール: np-info@impress.co.jp