PostgreSQL 15 Cookbook

Efficient data management with 100+ recipes

Mohammad Samsad Hussain

www.bpbonline.com

First Edition 2024

ISBN: 978-93-55516-619

LIMITS OF LIABILITY AND DISCLAIMER OF WARRANTY

The information contained in this book is true to correct and the best of author's and publisher's knowledge. The author has made every effort to ensure the accuracy of these publications, but publisher cannot be held responsible for any loss or damage arising from any information in this book.

All trademarks referred to in the book are acknowledged as properties of their respective owners but BPB Publications cannot guarantee the accuracy of this information.

Dedicated to

The readers

About the Author

Mohammad Samsad Hussain is based in Malaysia. He is a Database Architect and consultant with more than 13 years of experience in different roles (Data Analyst, Database Solution Architect, Application Consultant) among software industries. He is an IBM DB2 DBA Certified Professional and PostgreSQL, MySQL & MariaDB database Consultant as he has worked on different technologies and architectures, such as SAP HANA, Oracle, MongoDB, MS SQL and IBMi (As400).

Samsad is currently employed as Database Infrastructure Solution Architect in a Utilities based Organization in Malaysia, and he is in charge of databases and systems management and infrastructure.

About the Reviewers

- **Sami Lekesiz**, born in Turkiye, graduated from Istanbul Technical University. He has been working exclusively with PostgreSQL for the last five years of his career. His expertise relates to various PostgreSQL projects where he has played a crucial role in numerous critical tasks such as migrations, building highly available cluster architectures, and system optimizations. With his extensive knowledge and hands-on experience, Sami has been instrumental in the successful delivery of multiple PostgreSQL projects. He has always focused on delivering excellence in tasks that are central to these projects. He has ensured seamless migrations, set up highly available clusters from scratch, and implemented critical system enhancements. Looking ahead in his career, Sami aims to continue his dedicated focus on PostgreSQL, aspiring to further specialize and contribute within this domain.

- **Regina Obe** is the president of Paragon Corporation, a PostgreSQL database consulting company located in Boston, Massachusetts, USA. She has over 25 years of experience in various programming languages and database systems, especially PostgreSQL and spatial databases. She is a member of the project steering committee and development teams of two popular PostgreSQL spatial extensions, PostGIS and pgRouting. She is also a charter member and System Administration support team member for Open Source Geospatial Foundation (OSGeo), which focuses on developing and evangelizing free and open-source geospatial software. She holds a BS degree in mechanical engineering from the Massachusetts Institute of Technology.

Acknowledgement

I want to thank everyone who helped bring this book to life. Special thanks to my friend and colleague, **Abhishek Chaturvedi** and **Mohd Zaidi b. Mohd Yusoff** for their valuable contributions and strong support. Thanks to my family and friends for always cheering me on. A big thank you to the readers.

I also want to give a special thanks to the publisher of this book, BPB Publications, for believing in this project and providing the platform to make it happen. Your expertise and dedication made this dream come true. Thanks to my colleagues too, for their support and insights, adding depth to this work. Together, we have created something special, and I am truly thankful for this collaborative journey.

Preface

Welcome to **PostgreSQL 15 Cookbook**, PostgreSQL 15 stands out as a robust and reliable SQL database server. This book is here to help you understand and master PostgreSQL 15 in a practical and straightforward way.

In this book, we cover various aspects of PostgreSQL 15, providing insights and hands-on guidance. Whether you are just starting with PostgreSQL 15 or delving into advanced topics like database hierarchy, cloud provisioning, migration, and more, each chapter is crafted to improve your understanding and skills.

PostgreSQL 15 is essential in managing data. Whether you are new to it or experienced, this recipe guide makes PostgreSQL 15 accessible and valuable. We focus on real-world applications, showing the importance of each topic in practical situations.

This book is divided into 13 chapters, each focused on a specific aspect of PostgreSQL 15. Whether you are exploring the basics, advancing your skills in managing PostgreSQL 15, or troubleshooting problems, our goal is to present information that is clear, concise, and practical. Our goal is to equip you with the knowledge to use PostgreSQL 15 effectively.

Chapter 1: Up and Running with PostgreSQL 15 - In this chapter, we will learn the basics of databases and their core architecture concepts, seeing how these elements work together. The focus is on starting with PostgreSQL, giving you simple instructions for installing and setting it up on Linux VMs. We will also cover database upgrades to both minor and major releases.

Chapter 2: Database Hierarchy - We will work on creating and using databases in PostgreSQL. We will cover best practices for user access and authorization, the hierarchy of database objects, tables inheritance, concurrent indexes, and workarounds with various database objects.

Chapter 3: Cloud Provisioning - In this chapter, we will work on setting up a PostgreSQL cloud instance and handling database connections using AWS EC2 and RDS instance. We will also explain how to perform native backup and restore procedures for both AWS EC2 and RDS instances. Additionally, we will explore managing PostgreSQL on the cloud, including aspects like connection management and replication.

Chapter 4: Migration - This chapter explores different migration options for moving SQL databases to PostgreSQL and transferring on-premises data to the AWS cloud. We will work on each migration method, including the use of tools like PgLoader. Additionally, we will explore the best practices to ensure a seamless and efficient migration process.

Chapter 5: Transaction Log - We will explore admin and application-specific logs, discover practical solutions, and gain a detailed understanding of PostgreSQL transaction logs. We will examine the crucial role of the Transaction Log in PostgreSQL replication, discussing topics such as Archive mode and WAL management.

Chapter 6: Partitioning and Sharding - We will discuss in detail the concepts of scaling up your PostgreSQL database based on your organization's limitations and operational needs. The goal is to provide practical concepts for enhancing the performance of your business-critical data, covering topics such as partitioning, sharding, and the role of Foreign Data Wrapper.

Chapter 7: Replication and High Availability - In this chapter, we will learn replication and High Availability solutions for PostgreSQL, understanding the best fit for different practical scenarios and business needs. We'll also dive into the concepts of Load balancing in PostgreSQL, offering a mix of theoretical understanding and practical solutions**.**

Chapter 8: Leveraging SQL - In this chapter, we will talk about how to get to know and work with data in the database. Once you understand the basics of PostgreSQL, this section will guide you in using SQL to interact with relational data stored in PostgreSQL databases. We will also cover how to organize query access using psql, explore Postgres JSON Query, and understand the Postgres CAST Operator.

Chapter 9: Server Controls and Auditing - We will explore the basic concepts of authentication and encryption in the PostgreSQL database server. We'll look into common methods to encrypt your client connections and provide practical solutions for user and group control to ensure flexible access control to your database. Additionally, we have a dedicated section on database authentication through SSL.

Chapter 10: Backup - We will explore the basics of database backup, offering practical examples and insights into various backup tools, such as pg_probackup and pgBackRest. We will also discuss planning for backups and strategies to improve backup performance.

Chapter 11: Recovery - This chapter covers the fundamentals of database recovery. We will explore the concept, discuss planning for recovery, understand crash recovery, and explore Point-in-Time Recovery (PITR) and table-level recovery.

Chapter 12: Monitoring and Diagnosis - In this chapter, we will discuss how to make your setup more reliable by adding monitoring for crucial actions in your database. We will look into different monitoring approaches for database components. Additionally, we will check out tools like Prometheus and Grafana, explore statistics, find monitoring scripts, and understand the significance of fsync.

Chapter 13: Troubleshooting - This chapter will explore various situations where a PostgreSQL database can be beneficial. We will provide practical solutions to help database administrators manage their databases effectively. The topics covered include transaction logs and checkpoints, benchmarking with pgbench, and utilizing data checksums with pg_checksums.

Code Bundle and Coloured Images

Please follow the link to download the
Code Bundle and the ***Coloured Images*** of the book:

https://rebrand.ly/j6vnarp

The code bundle for the book is also hosted on GitHub at
https://github.com/bpbpublications/PostgreSQL-15-Cookbook.
In case there's an update to the code, it will be updated on the existing GitHub repository.

We have code bundles from our rich catalogue of books and videos available at **https://github.com/bpbpublications**. Check them out!

Errata

We take immense pride in our work at BPB Publications and follow best practices to ensure the accuracy of our content to provide with an indulging reading experience to our subscribers. Our readers are our mirrors, and we use their inputs to reflect and improve upon human errors, if any, that may have occurred during the publishing processes involved. To let us maintain the quality and help us reach out to any readers who might be having difficulties due to any unforeseen errors, please write to us at :

errata@bpbonline.com

Your support, suggestions and feedbacks are highly appreciated by the BPB Publications' Family.

Piracy

If you come across any illegal copies of our works in any form on the internet, we would be grateful if you would provide us with the location address or website name. Please contact us at **business@bpbonline.com** with a link to the material.

If you are interested in becoming an author

If there is a topic that you have expertise in, and you are interested in either writing or contributing to a book, please visit **www.bpbonline.com**. We have worked with thousands of developers and tech professionals, just like you, to help them share their insights with the global tech community. You can make a general application, apply for a specific hot topic that we are recruiting an author for, or submit your own idea.

Reviews

Please leave a review. Once you have read and used this book, why not leave a review on the site that you purchased it from? Potential readers can then see and use your unbiased opinion to make purchase decisions. We at BPB can understand what you think about our products, and our authors can see your feedback on their book. Thank you!

For more information about BPB, please visit **www.bpbonline.com**.

Join our book's Discord space

Join the book's Discord Workspace for Latest updates, Offers, Tech happenings around the world, New Release and Sessions with the Authors:

https://discord.bpbonline.com

Table of Contents

Chapter 1
Up and Running with PostgreSQL 15

Introduction

PostgreSQL, aka Postgres, is an advanced open-source relational database and world's fastest-growing database management system. Being open-source, PostgreSQL is released under the PostgreSQL License, a liberal open source license, similar to the BSD or MIT licenses. As of the time of writing this book, the most recent major release of PostgreSQL is version 16.

This chapter provides an introduction to PostgreSQL 15 database systems. We begin by discussing the history of PostgreSQL. We then look at the installation, configuration and starting up PostgreSQL server. Finally, we will cover a couple of common configuration options and database client configuration.

Structure

In this chapter, we will cover the following topics:

- Introduction to PostgreSQL 15
- Exploring the history of PostgreSQL
- Learning about PostgreSQL architecture
- Learning about PostgreSQL database structure

- Installation methodology
- Database server and client environment
- Exploring backward compatibility

Objectives

This chapter covers a vast range of topic that provides an understanding of database overview and architecture of core concept and how these component work together. This chapter is all about getting up and running with PostgreSQL using basic recipes and step to cover the installation and setup of PostgreSQL on Linux VMs with step-by-step instructions. If you are already well versed with PostgreSQL basics, It is advisable to glance through the chapter quickly before proceeding to the next chapter.

Introduction to PostgreSQL 15

PostgreSQL is a powerful, flexible, and scalable general-purpose database. As it is highly stable, very low effort is required to maintain this DBMS. It runs on a wide variety of operating systems flavours, including Linux, Unix and windows.

PostgreSQL 15, the latest version released in October 2022, comes with nicer and consistent functionalities. With the recent release the new enhancement is essential reasons to upgrade. Enhancement with PostgreSQL 15 offers, few of them are:

- Addition of **`MERGE`** command that improves Developer experience.
- support for **`LZ4`** and **`zstd`** compression to **write-ahead log** (**WAL**) files
- Enhancement in logging format: **`jsonlog`**
- Improved collation support
- Enhancement with logical replication support for using **two-phase commit** (**2PC**)
- Support inspection on write-ahead log files with SQL interface
- Support for **`WAL`** compression with **`LZ4`** and **`Zstandard`**
- enhance performance of sorts that exceed **`work_mem`**

Beside of its new features, PostgreSQL is **ACID** (**atomicity, consistency, isolation, and durability**) complaint standard and transactional. The database programmed in C and employs a monolithic design, which means that the components are entirely united and work in a systematic manner.

Exploring the history of PostgreSQL

PostgreSQL version history dates back in mid-1980s, as a follow-up to INGRES. Postgres has undergone several significant changes since that time. Postgres used the POSTQUEL query language until 1994, In mid-1996s the project was renamed to PostgreSQL to reflect its support for SQL.

Following is the timeline of release history for PostgreSQL database with its enhancements or changes. Here, we cover the last three major versions that have not yet achieved the **end Of life** (**EOL**), please refer to *Table 1.1:*

Database Release	Enhancements
PostgreSQL 13	• Incremental sorting • Parallelized vacuuming of indexes • Improved performance for queries that use aggregates or partitioned tables
PostgreSQL 14	• Support in place tablespaces • Change in default password encryption to **scram-sha-256** • additions to remove expired index entries
PostgreSQL 15	• Support for **merge** command • Log output using JSON Format • Enhancement in in-memory • Support for **zstd** compression

***Table 1.1**: PostgreSQL database release history - enhancement*

Learning about PostgreSQL architecture

The PostgreSQL architecture defines the relationship between various PostgreSQL systems. The basic architecture follows the client-server model, where the key concept is that the client makes a request and initiates the connection. The server responds and provide the service.

This architecture enables a PostgreSQL system to service a wide variety of different clients that can connect locally or over the network. To each successful client, connection spawn a new child process via fork.

PostgreSQL contain many internal components, some are fundamental to PostgreSQL and some are optional. Let us take a look around a typical PostgreSQL internal and external components in the following figure:

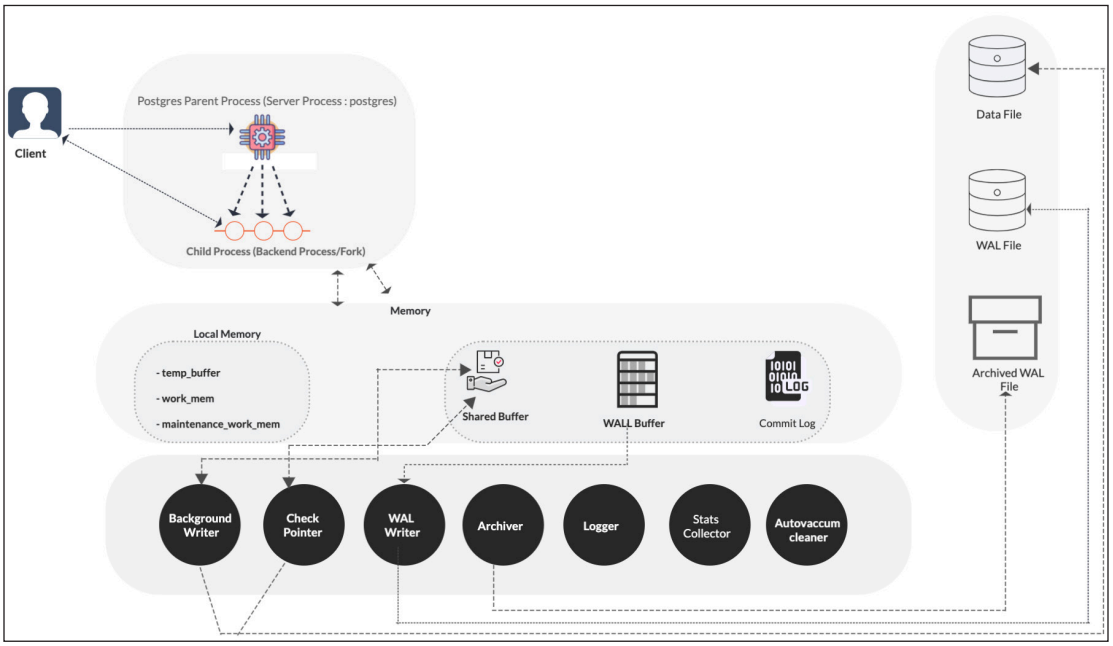

***Figure 1.1**: PostgreSQL database cluster architecture*

Learning about PostgreSQL database structure

PostgreSQL Cluster, refers to the collection of database all managing the shared data directory that contain one or more database, that are managed by single server instance. A cluster is generally created for you when you install PostgreSQL database. **`initdb`** initializes the database cluster with default locale and character set encoding, unless custom setting is specified with the required option.

After the initialization, a newly created instance has **`template0`**, **`template1`** and postgres as the built-in databases. The database is the **`template`** upon which the new database is built.

In PostgreSQL, the schema is the logical grouping of database objects or the collection of databases objects are maintained and organized in database. A SCHEMA is essentially a namespace, it makes it possible to have a table with the same name if it belongs to a different schema. Let get insight to PostgreSQL:

Every database objects has two parts:

```
schema_name.object_name
{prefix}.{suffix}
```

When you create a database object, it is always created within a schema. The prefix part logically distinguish your organizational data with other database object.

The syntax for creating a schema:

```
# create schema <schema_name>;
```

or

```
# create schema IF NOT EXISTS <schema_name>;
```

Note: IF NOT EXISTS clause is an optional parameter of the CREATE SCHEMA to check whether or not the SCHEMA of the same name already exist in the database before creating it. If it already exist then it will produce an error.

Installation methodology

Now that you have an overview of the PostgreSQL database, the next step is to install it. In this section we walk you through the installation methodologies with step-by-step instructions. To offer flexibility, in this book we will perform the installation through binary packages and from the source code.

In this recipe we will cover the installation of PostgreSQL in the following 2 aspects:

- working with installation from binaries
- working with installation from source code

Recipe 1: Working with installation from binaries

In this recipe, we will cover the installation of PostgreSQL from binary package. Most of the installs of PostgreSQL are on various flavours of Linux and Windows. For this recipe, we will use Red Hat version 8.4 x86_64 Linux server.

Determine the running Linux version. The following command will give us the information we need:

```
# cat /etc/os-release
NAME="Red Hat Enterprise Linux"
VERSION="8.4 (Ootpa)"
ID="rhel"
ID_LIKE="fedora"
VERSION_ID="8.4"
```

First, assuming that we have downloaded the PostgreSQL binaries from the download site. If not then you can download the latest version of PostgreSQL from their official website at **https://www.postgresql.org/download/** after select your host operating system and version and follow the installation instructions.