Create a Backup

by Sophia

WHAT'S COVERED

This lesson explores using the command line to back up a PostgreSQL database, in two parts. Specifically, this lesson will cover:

1. Introduction

Depending on the choice of database, backing up a database can take various forms. In most cases, there is a command line option and a graphical user interface option that functions by using the command line behind the scenes. These steps are important to save the data and state of the database so that it can be recovered if there are any issues. There are different types of database backup options, such as full or partial data, or being able to choose to back up structures.

The tools you will learn about in this lesson perform an operation known as a dump. **Dumping** is the process of extracting data or information from a system or application and storing it in a file or memory space for backup, analysis, or debugging. Dumping is roughly analogous to backing up, but dumps can have uses other than backups (for example, debugging or analysis) and may not always be comprehensive or suitable for directly restoring a system or dataset without some type of manipulation.

TERM TO KNOW

Dumping

The process of extracting data or information from a system or application and storing it in a file or memory space for backup, analysis, or debugging.

2. PostgreSQL Backup Tools

© 2024 SOPHIA Learning, LLC. SOPHIA is a registered trademark of SOPHIA Learning, LLC.

In PostgreSQL databases, we have the **pg_dump** and **pg_dumpall** tools. These will not work within our web interface tool, as the tool is logged into PostgreSQL already. However, you can test the commands if you have PostgreSQL installed locally.

KEY CONCEPT

The pg_dump and pg_dumpall tools do not actually perform the backup by copying the database files. Instead, they extract the structure and data from the database and represent them in a text format that can be executed as SQL commands to rebuild the database. This method provides a flexible and efficient way to back up and restore PostgreSQL databases.

There are a lot of different parameters that can be used. We will explore some of the more common options.

Let us first look at a complete command that will back up the mydb database to a mydb.sql file in the c:\backups\ folder on a Windows system, using the user adminrole:

```
pg_dump -U adminrole -W -F t mydb > c:\backup\mydb.sql
```

Let's break this down:

- pg_dump is the command line tool. The operation that pg_dump performs—while we might say it is a type of backup—is more commonly referred to as a dump.
- U adminrole specifies the user role that will connect to the database. In this case, we are using the adminrole to login to perform the backup.
- W prompts the pg_dump command to prompt for the password on adminrole, adding some security when including this command in a script.
- F specifies the output file format. In this case, the t stands for plain-text SQL script file. The alternative is -F d to create a TAR file (an archive format similar to ZIP).
- mydb is the database that we want to back up.
- > is like an arrow that “points” to the destination for the backup.
- c:\backup\mydb.sql is the output backup file name and path that we are backing up to. We could optionally omit the path and just provide the file name, to output the backup file to the location from which we are running the command.

Here are some additional options:

a	Backs up only the data, not the schema. This option is useful only if backing up the data in plain text format.
s	Backs up only the schema, not the data.
c	Cleans the database objects using DROP statements first before creating them.
t <tables> </tables>	Backs up only specific tables that match the table name that is passed. For example, -t employee would only back up the employee table.

© 2024 SOPHIA Learning, LLC. SOPHIA is a registered trademark of SOPHIA Learning, LLC.

T <tables> </tables>	Backs up all tables other than those listed.
C	Starts the output with a command to create the database and reconnect to the created database. Using this option, you don't have to create the database first.

HINT

The pg_dumpall command backs up all of the databases on a server. This is not commonly done, as we typically want to back up specific databases. It exports all of the databases into a single file, so restoring from this can be unreliable. It uses the same options as the pg_dump command. The only difference is that the -W option is not used, as you wouldn't want to have to type the password for each individual database.

TERMS TO KNOW

pg_dump
A utility for backing up a single PostgreSQL database.

pg_dumpall
A utility for backing up all the databases on a server.

SUMMARY

In this lesson's **introduction**, you learned that there are different options for backing up a database, including both command line and graphical interface options.

Next, you learned about two **PostgreSQL backup tools**, pg_dump and pg_dumpall.; pg_dump is a command line utility that enables users to back up their databases. The program does not actually perform the backup, but rather it generates a text file containing SQL commands that can be used as a script to recreate the database's schema and data. Database migration, disaster recovery, and data preservation are all made easier with the help of this utility.

Depending on the database connection details, users can tailor their backup process by specifying various options, including the backup file name, the user role initiating the backup (and its password), and the file format. You can also add options that back up only the data or only the schema, or back up only certain tables.

Source: THIS TUTORIAL WAS AUTHORED BY DR. VINCENT TRAN, PHD (2020) AND FAITHE WEMPEN (2024) FOR SOPHIA LEARNING. PLEASE SEE OUR TERMS OF USE.

TERMS TO KNOW

© 2024 SOPHIA Learning, LLC. SOPHIA is a registered trademark of SOPHIA Learning, LLC.

Dumping

The process of extracting data or information from a system or application and storing it in a file or memory space for backup, analysis, or debugging.

pg_dump

A utility for backing up a single PostgreSQL database.

pg_dumpall

A utility for backing up all the databases on a server.

© 2024 SOPHIA Learning, LLC. SOPHIA is a registered trademark of SOPHIA Learning, LLC.