**Physical Backups with pg\_basebackup in PostgreSQL**

**What is a Physical Backup?**

A physical backup is a **binary copy of the PostgreSQL data directory**, including all table data, indexes, transaction logs (WAL), configuration files, and more. It is typically used for:

* Disaster recovery
* Setting up streaming replication
* Cloning production environments

The main tool for physical backups in PostgreSQL is pg\_basebackup.

**Requirements for Using pg\_basebackup**

1. PostgreSQL server must allow replication connections.
2. The user running pg\_basebackup must have the REPLICATION privilege.
3. WAL archiving or streaming must be enabled to ensure backup consistency.

**Basic Syntax**

pg\_basebackup -h <host> -p <port> -U <replication\_user> -D <backup\_dir>

* -h: Host of the PostgreSQL server
* -p: Port (default is 5432)
* -U: Replication user
* -D: Destination directory for the backup

**Example: Simple Physical Backup**

pg\_basebackup -h localhost -p 5432 -U replicator -D /var/lib/postgresql/backup -Fp -Xs -P -v

**Explanation of Options**

* -Fp: Write the backup in plain format (just a copy of the data directory)
* -Xs: Include WAL files by streaming them
* -P: Show progress
* -v: Verbose output

**Formats**

* -Fp: Plain format (default), exact copy of the data directory
* -Ft: Tar format, creates a tar archive instead of copying files directly

**Tar format example**

pg\_basebackup -h localhost -U replicator -Ft -D /backups/db.tar

**Creating a Replication User**

CREATE ROLE replicator WITH REPLICATION LOGIN PASSWORD 'securepass';

Update pg\_hba.conf to allow replication connections:

# Allow replication connections

host replication replicator 127.0.0.1/32 md5

Then reload the configuration:

pg\_ctl reload

**Backup Destination Directory**

The destination directory (-D) must be:

* Empty or non-existent
* Writable by the user running pg\_basebackup

**Streaming WAL Files**

To make the backup consistent, PostgreSQL needs to include **Write-Ahead Logs (WAL)**. Options:

* -Xf: Fetch WAL files at the end
* -Xs: Stream WAL files during the backup (recommended)
* -Xnone: Do not include WAL (not safe for recovery)

**Backup Verification**

After taking a backup, you can check that it is complete:

* The directory should contain all files from the data directory
* There should be a backup\_label and pg\_wal folder
* It can be used to start a new PostgreSQL instance

**Restoring from a Physical Backup**

To restore:

1. Stop the PostgreSQL server.
2. Replace the current data directory with the backup.
3. Remove any stale files (postmaster.pid, etc.).
4. Start the server.

Example:

systemctl stop postgresql

rm -rf /var/lib/postgresql/15/main

cp -r /backups/main /var/lib/postgresql/15/main

chown -R postgres:postgres /var/lib/postgresql/15/main

systemctl start postgresql