

PostgreSQL 备份与迁移工具

PostgreSQL 提供命令行工具用于逻辑备份和数据库迁移。这些工具为可执行程序，非 SQL 命令，通过生成 SQL 语句或其他格式文件保存数据库结构和数据。以下详细介绍 PostgreSQL 的 pg_dump 和 pg_dumpall。

pg_dump

- 功能: pg_dump 用于备份单个 PostgreSQL 数据库，生成可用于恢复的 SQL 脚本或其他格式文件。
- 输出格式:
 - 纯文本 SQL 脚本（默认，使用 psql 恢复）。
 - 自定义格式（压缩二进制，需 pg_restore 恢复）。
 - 目录格式（支持并行导出）。
 - tar 格式（便携性强）。
- 特点:
 - 支持 -j 选项并行导出，加速大数据库备份。
 - 可通过 --exclude-table 排除特定表。
 - 包括表、视图、触发器等对象，需确保目标服务器支持相关功能。
- 示例:
 - pg_dump -U username database_name > backup.sql

备份 database_name 到 backup.sql 文件。

- 恢复:
 - psql -U username -d template1 -f backup.sql

或使用 pg_restore 恢复自定义格式：

pg_restore -U username -d database_name backup.dump

pg_dumpall

- 功能: pg_dumpall 备份 PostgreSQL 集群中的所有数据库及全局对象（如角色、表空间），生成单一 SQL 脚本。
- 特点:
 - 调用 pg_dump 逐个备份数据库，并添加全局对象定义。
 - 需超级用户权限以访问所有数据库。
 - 支持 -g 选项仅备份全局对象。
- 示例:
 - pg_dumpall -U username > all_databases.sql

备份所有数据库到 all_databases.sql。

- 恢复:
 - psql -U username -f all_databases.sql

功能比较

工具	数据库系统	备份范围	输出格式	特殊功能
pg_dump	PostgreSQL	单个数据库	SQL、自定义、目录、tar	并行导出（-j）
pg_dumpall	PostgreSQL	所有数据库及全局对象	SQL	备份全局对象（-g）

- 备份范围: pg_dump 专注于单数据库，pg_dumpall 覆盖整个集群。

- **输出格式:** `pg_dump` 提供多种格式, 适合不同场景; `pg_dumpall` 仅生成 SQL。
- **性能:** `pg_dump` 的并行导出适合大数据库

使用注意事项

- **权限:**
 - `pg_dump` 和 `pg_dumpall` 需数据库用户权限, `pg_dumpall` 通常需超级用户权限。
- **安全性:**
 - 使用 `.pgpass` (PostgreSQL) 或 `.my.cnf` (MySQL) 存储密码, 避免命令行暴露。
 - 避免在命令历史中记录敏感信息。
- **性能:**
 - 大数据库备份可能消耗大量 CPU 和 I/O, 建议在低负载时段运行。
 - `pg_dump` 的 `-j` 选项加速备份, 但增加服务器负载。
- **恢复:**
 - PostgreSQL: SQL 脚本用 `psql`, 自定义格式用 `pg_restore`。
 - 恢复前需检查目标服务器兼容性 (如字符集、版本)。
- **压缩:**
 - `pg_dump` 支持自定义格式压缩, 或通过 `gzip`:
 - `pg_dump -U username database_name | gzip > backup.sql.gz`
- **验证:**
 - 定期验证备份文件完整性, 如用 `pg_restore -v` (PostgreSQL)。
 - 检查备份文件大小是否合理。
- **调度:**
 - 使用 `cron` 调度定期备份, 结合脚本管理备份轮换。
 - 示例 `cron` 任务 (PostgreSQL):
 - `0 2 * * * pg_dump -U username database_name | gzip > /backups/db_$(date +%F).sql.gz`

高级功能

- **PostgreSQL:**
 - `pg_dump` 支持 `--exclude-table` 排除表, `-b` 包括大对象。
 - `pg_dumpall` 的 `-g` 选项仅备份角色和表空间。
- **跨版本迁移:**
 - PostgreSQL 备份文件通常跨版本兼容, 但需检查功能支持。
- **其他工具:**
 - PostgreSQL 的 `pg_basebackup` 用于物理备份, 适合大数据库。

总结

`pg_dump` 和 `pg_dumpall` 是 PostgreSQL 的核心逻辑备份工具, 适用于备份、恢复和迁移。`pg_dump` 适合单数据库, 而 `pg_dumpall` 覆盖整个集群。合理使用这些工具, 结合权限管理、压缩和定期验证, 可确保数据安全和高效迁移。