

DATABASES

INTERMEDIATE

4 DAYS

PostgreSQL Administration

A four-day operational course for teams running PostgreSQL in production. Covers architecture, configuration, roles and authentication, MVCC and autovacuum, write-ahead logging, backup and point-in-time recovery, streaming and logical replication, monitoring, security hardening and major version upgrades. Delivered on PostgreSQL 18 with notes for versions 14 through 17, and every module ends with a lab on a live cluster rather than a walkthrough.

Why PostgreSQL Administration Training?

Built around failure, not happy paths

Participants deliberately break clusters and recover them, including a full point-in-time restore and a replica promotion.

Current on PostgreSQL 18

Includes the asynchronous I/O subsystem, OAuth authentication and upgrade improvements introduced in version 18.

Autovacuum taught properly

Bloat and transaction ID wraparound cause more PostgreSQL incidents than anything else, and both get a dedicated module.

Cloud and self-managed

Covers self-hosted clusters alongside the operational differences on RDS, Aurora, Azure Database and Cloud SQL.

What you'll be able to do

- ✓ Explain the PostgreSQL process architecture, memory structures and on-disk layout
- ✓ Install, initialise and configure a cluster for a defined workload
- ✓ Manage roles, authentication methods and object level privileges securely
- ✓ Diagnose and control bloat, long transactions and autovacuum behaviour
- ✓ Prevent transaction ID wraparound before it becomes an outage
- ✓ Design and execute a backup strategy including continuous archiving
- ✓ Perform a point-in-time recovery to a specified timestamp
- ✓ Build streaming replication with failover and understand the trade-offs of synchronous modes
- ✓ Use logical replication for selective distribution and near-zero-downtime upgrades
- ✓ Instrument a cluster and diagnose performance problems from evidence
- ✓ Harden a cluster against the controls your auditors will test
- ✓ Plan and execute a major version upgrade with a rollback path

Who this is for

- Database administrators new to PostgreSQL or moving from Oracle, SQL Server or MySQL
- DevOps, platform and site reliability engineers who operate PostgreSQL

- Backend engineers responsible for their own database infrastructure
- Cloud engineers managing RDS, Aurora, Azure Database or Cloud SQL instances
- Technical leads accountable for database availability and recovery objectives

Curriculum

01 PostgreSQL Architecture and Process Model

- Postmaster, backend processes and the background worker set
- Shared buffers, WAL buffers and per-backend memory areas
- The data directory layout and what each subdirectory holds
- Client connection lifecycle and the cost of a connection
- PostgreSQL 18 asynchronous I/O and what changed operationally
- **Hands-on:** inspect a running cluster's processes, memory and directories

02 Installation, Initialisation and Cluster Management

- Packaged installs, PGDG repositories and containerised deployments
- initdb options, locale, encoding and checksums set at creation time
- Running multiple clusters and instances on one host
- Service management, start and stop modes and controlled shutdown
- Directory ownership, permissions and filesystem choices
- **Hands-on:** build a cluster from scratch with checksums and a chosen locale

03 Configuration and Parameter Management

- postgresql.conf, include files and ALTER SYSTEM precedence
- Parameter contexts: which settings need a reload and which need a restart
- Memory parameters: shared_buffers, work_mem, maintenance_work_mem and effective_cache_size
- Connection limits, and why a connection pooler is usually required
- PgBouncer pooling modes and their application constraints
- **Hands-on:** tune a default install for a defined workload and measure the change

04 Roles, Authentication and Connection Security

- Roles, login attributes, role inheritance and group design
- pg_hba.conf rule evaluation order and the mistakes that lock people out
- scram-sha-256, certificate, LDAP and OAuth authentication methods
- Retiring legacy md5 password hashes
- TLS configuration, certificate management and enforcing encrypted connections
- **Hands-on:** implement a role hierarchy and a hardened host-based access policy

05 Databases, Schemas and Privilege Management

- Databases, schemas, search_path and the public schema defaults
- GRANT and REVOKE across tables, columns, functions and sequences
- Default privileges for objects created in future
- Row level security policies and their operational implications
- Auditing effective privileges for a given role
- **Hands-on:** design least-privilege access for application, reporting and admin roles

06 Storage Internals and Space Management

- Heap files, pages, tuples and the fill factor
- TOAST storage for oversized values and its performance impact
- Tablespaces for separating workloads across storage
- Measuring table, index and database size accurately
- Free space map, visibility map and their role in performance
- **Hands-on:** locate the largest objects in a cluster and account for their space

07 MVCC, Bloat and Autovacuum

- Multi-version concurrency control and why deleted rows persist
- Dead tuples, table bloat and index bloat and how to measure each
- Autovacuum thresholds, scale factors and per-table overrides
- Long running transactions and idle-in-transaction sessions as root causes
- VACUUM, VACUUM FULL, pg_repack and the availability trade-offs
- **Hands-on:** generate bloat deliberately, diagnose it and reclaim the space

08 Transaction ID Wraparound and Freezing

- Transaction ID space, freezing and why wraparound halts a cluster
- Monitoring datfrozenxid and age across databases
- Aggressive and anti-wraparound vacuum behaviour
- Multixact IDs and their separate wraparound risk
- Alerting thresholds and the runbook for a cluster approaching the limit
- **Hands-on:** build wraparound monitoring and rehearse the response

09 Indexes and Plan Reading for Administrators

- B-tree, GIN, GiST, BRIN and hash indexes and their operational profiles
- CREATE INDEX CONCURRENTLY and safe index maintenance in production
- Detecting unused, duplicate, bloated and invalid indexes
- Reading EXPLAIN ANALYZE output well enough to triage an incident
- Statistics, ANALYZE and the effect of stale statistics on plans
- **Hands-on:** audit an index set and rebuild without blocking traffic

10 Write-Ahead Logging, Checkpoints and Durability

- How WAL guarantees durability and crash recovery
- Checkpoints, their I/O cost and tuning checkpoint behaviour
- WAL levels, retention and the causes of runaway WAL growth
- synchronous_commit, fsync and the durability versus latency trade-off
- Replication slots and the disk risk an inactive slot creates
- **Hands-on:** tune checkpoints and observe the effect on write throughput

11 Backup Strategy, Archiving and Point-in-Time Recovery

- Recovery point and recovery time objectives as design inputs
- pg_dump and pg_dumpall formats, parallelism and selective restores
- pg_restore options for partial and reordered recovery
- pg_basebackup for physical backups and its limitations
- Backup verification, retention and offsite storage requirements
- **Hands-on:** take and verify both a logical and a physical backup
- WAL archiving configuration and archive command reliability
- Base backups combined with archived WAL for arbitrary recovery targets

12 Streaming Replication and High Availability

- Physical streaming replication setup and standby configuration
- Asynchronous, synchronous and quorum commit modes and their cost
- Hot standby, read scaling and query conflict handling
- Replication lag monitoring and the causes of a falling-behind standby
- Failover, promotion and split-brain prevention
- Automated cluster management with Patroni or repmgr
- **Hands-on:** build a standby, force a failure and promote it under time pressure

13 Logical Replication and Data Distribution

- Publications, subscriptions and selective table replication
- Logical decoding and how it differs from physical streaming
- Cross-version replication and near-zero-downtime migration patterns
- Conflict handling, sequence gaps and DDL limitations
- Monitoring subscription lag and recovering a broken subscription
- **Hands-on:** replicate a subset of tables between two clusters on different versions

14 Monitoring, Logging and Diagnostics

- pg_stat_activity, pg_stat_database and the statistics views that matter
- pg_stat_statements for workload level query analysis
- Wait events and separating CPU, I/O, lock and network bound problems
- Logging configuration, slow query logging and log_line_prefix
- Prometheus exporters, Grafana dashboards and pgBadger reporting
- Alert thresholds worth setting and the ones that only cause noise
- **Hands-on:** instrument a cluster and triage an unknown performance incident

15 Security Hardening and Compliance

- Network exposure, listen addresses and firewall placement
- Encryption at rest options and transparent data encryption alternatives
- pgaudit for statement level audit trails
- Secret management, password policies and credential rotation
- Data masking and anonymisation for non-production environments
- Mapping controls to ISO 27001, PCI DSS and regional data protection requirements
- **Hands-on:** harden a cluster against a supplied audit checklist

16 Upgrades, Extensions and Day-Two Operations

- Minor version patching and the release cadence to plan around
- Major version upgrades with pg_upgrade, including the version 18 improvements
- Logical replication as a near-zero-downtime upgrade path
- Version support lifecycle and planning off end-of-life releases
- Managing extensions: pg_stat_statements, pgcrypto, PostGIS, pgvector and others
- Operational differences on RDS, Aurora, Azure Database and Cloud SQL
- **Hands-on:** capstone upgrade of a replicated cluster with a documented rollback plan