

Advanced EDB PostgreSQL Administration, Performance Tuning and High Availability

Duration: 16 Days

Related technologies: EDB Postgres Advanced Server (EPAS), PEM, EFM, HAProxy and ClickHouse

Day 1: EDB Architecture and Multi-Cluster Management

Module 1: EDB Postgres Advanced Server Architecture

- EPAS architecture and enterprise capabilities
- Database instance, schema and tablespace architecture
- Server processes and background workers
- Memory, storage and transaction architecture

Module 2: Multi-Cluster Architecture

- PostgreSQL cluster concepts
- Running multiple clusters on the same host
- Port, service, data-directory and configuration isolation
- Cluster naming and administration standards

Module 3: Multi-Cluster Design and Governance

- Resource allocation across clusters
 - Workload separation strategies
 - OLTP, reporting and maintenance workload isolation
 - Configuration, patching and ownership standards
-

Day 2: Operating System and Resource Tuning

Module 1: Database Workload Assessment

- CPU-, memory-, disk- and network-bound workloads
- Workload characterization
- Performance baseline collection
- Resource-bottleneck identification

Module 2: Linux and Memory Optimization

- CPU scheduling and processor considerations
- Memory allocation, swapping and huge pages
- File-descriptor and process limits
- PostgreSQL-related kernel parameters

Module 3: Storage and Resource Isolation

- Storage design for data, WAL and temporary files
- IOPS, throughput and latency analysis
- Resource allocation between clusters

- Preventing one cluster from affecting other clusters
-

Day 3: EDB Memory and Configuration Tuning

Module 1: PostgreSQL Memory Architecture

- Shared and local memory
- Shared buffers and operating-system cache
- Work memory and maintenance memory
- Memory consumption across concurrent sessions

Module 2: Memory Parameter Tuning

- shared_buffers
- work_mem
- maintenance_work_mem
- effective_cache_size
- temp_buffers, wal_buffers and huge_pages

Module 3: Connection and Configuration Management

- Connection limits and session memory
 - Idle and long-running sessions
 - Connection pooling concepts
 - Instance-, database-, role- and session-level configuration
-

Day 4: SQL Workload Analysis

Module 1: Query Processing Architecture

- SQL parsing, rewriting, planning and execution
- Cost-based optimization
- Statistics used by the optimizer
- Cardinality and selectivity estimation

Module 2: Execution Plan Analysis

- Understanding EXPLAIN and EXPLAIN ANALYZE
- Scan, join, sort and aggregate operations
- Estimated versus actual rows
- Cost, execution time, buffers and loops

Module 3: Expensive SQL Identification

- Using pg_stat_statements
 - High-duration and high-frequency queries
 - CPU- and I/O-intensive queries
 - Query-performance regression analysis
-

Day 5: Advanced SQL and Query Tuning

Module 1: Advanced Indexing Strategies

- B-tree, Hash, GIN, GiST and BRIN indexes
- Composite and covering indexes
- Partial and expression indexes
- Detecting unused and duplicate indexes

Module 2: Join and Subquery Optimization

- Nested loop, hash and merge joins
- Join ordering and selectivity
- Correlated subqueries
- CTE, EXISTS and IN optimization

Module 3: Sorting, Aggregation and Parallel Query Tuning

- In-memory and disk-based sorting
 - Hash and group aggregation
 - Window-function performance
 - Parallel scans, joins and aggregations
-

Day 6: Large Database and Space Management

Module 1: PostgreSQL Physical Storage

- Heap pages and tuple storage
- TOAST storage
- Free Space Map and Visibility Map
- Table and index storage organization

Module 2: Tablespace and Large-Object Management

- Tablespace design and administration
- Separating tables and indexes
- Moving objects between tablespaces
- Storage tiering and space monitoring

Module 3: Large Database Capacity Planning

- Data, index and WAL growth
 - Temporary-space requirements
 - Backup-storage requirements
 - Capacity forecasting for databases approaching 1 TB
-

Day 7: Partitioning, Vacuum and Bloat Management

Module 1: Large Table Partitioning

- Range, list and hash partitioning

- Partition-key selection
- Partition pruning
- Historical data archiving and retention

Module 2: Vacuum and Autovacuum Tuning

- MVCC and dead tuples
- Standard vacuum and full vacuum
- Autovacuum thresholds and scale factors
- Worker and cost-based tuning

Module 3: Bloat and Transaction ID Management

- Table and index bloat
 - Bloat detection and remediation
 - Reindexing strategies
 - Transaction ID wraparound and freeze management
-

Day 8: WAL Architecture and Physical Replication

Module 1: Write-Ahead Logging Architecture

- WAL generation and processing
- WAL records and segments
- WAL writer and checkpoints
- WAL archiving and retention

Module 2: WAL and Checkpoint Tuning

- wal_level
- max_wal_size and min_wal_size
- checkpoint_timeout
- checkpoint_completion_target
- WAL compression considerations

Module 3: Physical Streaming Replication

- Primary and standby architecture
 - WAL sender and receiver processes
 - Synchronous and asynchronous replication
 - Replication slots and WAL retention
-

Day 9: Advanced Replication and Disaster Recovery

Module 1: Standby Server Management

- Warm and hot standby
- Read-only workloads on standby
- Cascading replication
- Multiple-standby architectures

Module 2: Logical Replication

- Publications and subscriptions
- Initial data synchronization
- Replication identity
- Table-level replication and monitoring

Module 3: Replication Monitoring and Troubleshooting

- Send, write, flush and replay lag
 - Missing WAL segments
 - Replication-slot growth
 - Timeline divergence and standby recovery
-

Day 10: EDB Failover Manager and High Availability

Module 1: EFM Architecture and Configuration

- EDB Failover Manager components
- Primary, standby and witness roles
- EFM agent configuration
- Cluster communication and security

Module 2: Automatic Failover Management

- Failure detection
- Standby selection and promotion
- Failover conditions
- Quorum and split-brain prevention

Module 3: Switchover, Failover and Reintegration

- Planned switchover
 - Unplanned failover
 - Rejoining the old primary
 - Post-failover health and consistency checks
-

Day 11: PEM Monitoring and Performance Diagnostics

Module 1: PEM Architecture and Administration

- PEM server and agent architecture
- Monitored hosts and database servers
- Agent registration
- Monitoring permissions and topology

Module 2: Database Performance Monitoring

- Database and server dashboards
- Session and transaction monitoring

- Wait-event analysis
- Lock and blocking-session analysis

Module 3: Query and Infrastructure Diagnostics

- Query-performance monitoring
 - CPU, memory, disk and network metrics
 - Correlating SQL with infrastructure events
 - Root-cause analysis workflow
-

Day 12: PEM Alerting and Capacity Planning

Module 1: Alerts and Notifications

- Alert templates
- Warning and critical thresholds
- Notification and escalation policies
- Reducing duplicate and non-actionable alerts

Module 2: Capacity Monitoring and Forecasting

- Database and tablespace growth
- Table and index growth
- WAL-generation trends
- Connection and storage utilization

Module 3: Custom Monitoring and Reporting

- Custom probes and metrics
 - Custom alerts and dashboards
 - Availability and performance reports
 - Management-level capacity reports
-

Day 13: Proxy and HAProxy Integration

Module 1: Database Proxy Architecture

- Purpose of a database proxy
- Application connection abstraction
- Read-write and read-only endpoints
- Connection-routing strategies

Module 2: HAProxy Configuration

- Frontend and backend configuration
- Database health checks
- Routing based on database role
- Connection and timeout settings

Module 3: HAProxy Integration with EFM

- Detecting the current primary
 - Redirecting connections after failover
 - Connection-pool recovery
 - Application connectivity during switchover and failover
-

Day 14: Database Security and Backup Strategy

Module 1: EDB Security Architecture

- Role and privilege management
- Principle of least privilege
- Authentication methods
- Host-based access control

Module 2: Data and Connection Security

- SSL/TLS connectivity
- Password and certificate management
- Auditing and activity tracking
- Encryption considerations

Module 3: Enterprise Backup Architecture

- Logical and physical backups
 - Full, incremental and differential strategies
 - WAL archiving
 - Backup retention and validation
-

Day 15: Recovery, Switchover and Failover Operations

Module 1: Database Recovery Techniques

- Instance and crash recovery
- Physical backup restoration
- Point-in-time recovery
- Recovery target and timeline selection

Module 2: Controlled Switchover and Failover

- Pre-switchover readiness checks
- Standby promotion
- Application connection redirection
- Post-switchover validation

Module 3: Disaster-Recovery Governance

- RPO and RTO definition
- Failover decision matrix
- Failback planning
- Recovery documentation and post-incident review

Day 16: Production Simulation and ClickHouse Reporting

Module 1: Production Incident Simulation

- Primary database failure
- Replication-lag incident
- Storage-capacity issue
- Blocking and long-running SQL scenarios

Module 2: End-to-End HA/DR Response

- PEM alert identification
- EFM failover decision
- Standby promotion
- HAProxy connection redirection
- Application and database validation

Module 3: EDB-to-ClickHouse Reporting Architecture

- Extracting data from multiple EDB clusters
- Centralized reporting architecture
- Incremental data ingestion patterns
- Data consistency and reconciliation
- Separating transactional and analytical workloads