

ClickHouse System & Storage Administration

Day 1 - ClickHouse Architecture for Infrastructure Engineers

Topics

- ClickHouse overview and use cases
- Server architecture
- Client/server communication
- Processes and threads
- Ports and interfaces
- Deployment models
- Single-node vs cluster
- Linux prerequisites
- Hardware sizing concepts

Day 2 - Installation and Linux Service Administration

Topics

- ClickHouse packages
- Installing a specific ClickHouse version
- Directory structure
- systemd and systemctl
- Start/stop/restart
- Boot-time enablement
- Version verification
- Package hold for training consistency

Day 3 - Server Configuration and Networking

Topics

- config.xml
- Configuration fragments
- Network configuration
- Listening interfaces
- Native TCP and HTTP ports
- Hostnames and name resolution
- Memory-related settings overview
- Logging configuration
- Safe configuration changes

Day 4 - ClickHouse Storage Architecture

Topics

- ClickHouse data directory
- Metadata directories
- Data parts

- Background merges
- Disk layout
- Storage volumes
- Storage policies
- Filesystem considerations
- Permissions and ownership

Day 5 - Linux Storage Administration for ClickHouse

Topics

- Capacity planning
- df, du and lsblk
- Filesystem utilization
- Mount points
- Linux file permissions
- I/O monitoring
- vmstat and iostat
- Basic RAID/LVM concepts relevant to ClickHouse
- Storage growth and bottleneck identification

Day 6 - Monitoring, Metrics and Logs

Topics

- system.metrics
- system.events
- system.processes
- system.query_log
- system.parts
- system.merges
- Linux top, free, vmstat and iostat
- ClickHouse server logs
- journalctl
- Basic server-health methodology

Day 7 - ClickHouse Cluster and Distributed Architecture

Topics

- Nodes, shards and replicas
- Cluster topology
- Horizontal scaling
- Cluster configuration
- Local and Distributed tables overview
- Network considerations
- system.clusters
- Basic failure domains

Day 8 - ClickHouse Keeper and Replication

Topics

- Replication fundamentals
- ReplicatedMergeTree overview
- ClickHouse Keeper architecture
- Keeper configuration
- Replica synchronization
- Replication queues
- Basic high availability
- Failure and recovery concepts

Day 9 - Backup, Restore and Operational Resilience

Topics

- Backup concepts
- Table and database backup
- Restore operations
- Backup destinations
- Recovery validation
- Operational maintenance
- Upgrade planning
- Version compatibility
- Basic disaster-recovery concepts

Day 10 - Troubleshooting and Infrastructure Capstone

Topics

- Server startup failures
- Configuration errors
- Port conflicts
- Permission issues
- Disk-full conditions
- Memory pressure
- Slow storage
- Cluster connectivity issues
- Keeper problems
- Replication problems
- Infrastructure health-check methodology