

Introduction to Apache Cassandra

This 3-day class cover the essentials of Cassandra DB for DBAs Architects and Developers

By the end of this course, you will master Apache Cassandra installation administration and development using the Datastax Java Driver for Cassandra.

Apache Cassandra is a free and open-source, distributed, wide column store, NoSQL database management system designed to handle large amounts of data across many commodity servers, providing high availability with no single point of failure. Cassandra offers robust support for clusters spanning multiple datacenters, with asynchronous master-less replication allowing low latency operations for all clients.

123Completed Experts assisted numerous customers in Architecting the Cassandra cluster and Data Schema for achieving zero data loss, maximize performance and gain availability
For more information www.123completed.com

AUDIENCE

Architects
DBAs
Developers

KNOWLEDGE REQUIREMENTS

SQL, RDBMS
Linux
Java/Python

LENGTH

3 Days

BONUS

Hands-on lab sessions

SYLLABUS

Day 1

Course Introduction
Cassandra Introduction
Cassandra Internals introduction
Data Cassandra Data modeling
Basics
Advanced Data Modeling and CQL
Data Manipulation
CQL Advanced Topics

Day 2

Cassandra Read and Write Path
Configuration Files
Simple Cassandra Cluster
Multi Data Center Cluster
Monitoring Cassandra

Day 3

Security
Backup and Recovery
Sizing and Performance planning
Summary Cassandra Backend
Cassandra Client Python/Java (One)
Summary Cassandra Client

HARDWARE AND SOFTWARE REQUIREMENTS

Computer Requirements

- RAM: minimum 10 GB of RAM required for exercises and platform to operate, 12 GB and up recommended.
- Disk Space: At least 40 GB of free disk space
- Internet connection

Supported Operating Systems

- Linux Ubuntu VM is provided. Virtual Machine provided for the labs is configured for 8GB RAM usage
- Machine should be able to run VMware Player and host the Ubuntu VM

Additional Software Requirements

- PDF Reader
- Zip software

Class HW required (If delivered on customer site)

- Projector 1024*768 minimum resolution (HD preferred)
- White Board
- Erasable Markers
- Desktops or Laptops (see HW Requirements)
- 12-24 ports Switch
- Internet connectivity
- Electricity outlets for all computers/monitors and other equipment.
- At least 3 electricity outlets next to instructor location.

DAY 1 - AGENDA

Lesson 1: Course Introduction

Duration: 1.0 hour Including Lab

- Course Introduction
- Courseware walkthrough
- Course environment
- Documentation
- Lab

Lesson 2: Cassandra Introduction

Duration: 1.5 hours Including Lab

- Cassandra DB introduction
- Getting started
- Configuration file's introduction
- Cassandra query language shell – CQLSH
- CQL introduction
- Lab

Lesson 3: Cassandra Internals introduction

Duration: 1.5 hours Including Lab

- Google big table
- Write path basics
- Read path basics
- Clustering introduction
- Nodetool introduction
- Lab

Lesson 4: Data Cassandra Data modeling Basics

Duration: 1.5 hours Including Lab

- RDBMS vs. Cassandra data modeling
- CAP theorem
- Consistency level
- Choosing primary key
- Denormalization
- Lab

Lesson 5: Advanced Data Modeling and CQL

Duration: 1.5 hours Including Lab

- Data Definition Language (DDL)
- Tables
- Static/frozen columns
- Indexes
- Data retrieval (Select)
- Simple Select
- Range Scans
- Filter non-PK Columns
- IN operator
- Sorting – order by
- Aggregation functions
- Date functions
- Lab

Lesson 6: Data Manipulation

- Insert, Update, Delete
- TTL
- Collections
- Set
- List
- Map
- Counters
- Light weight transactions
- Lab

Lesson 7: CQL Advanced Topics**Duration: 0.5 hours
Including Lab**

- Indexes should we use them
- Materialized views
- Aggregations and Cassandra

DAY 2 - AGENDA**Lesson 8: Cassandra Read and Write Path****Duration: 1 hours Including Lab**

- Write path
- Delete
- Compaction
- Read path
- Lab

Lesson 9: Configuration Files**Duration: 1 hours Including Lab**

- Installation directory
- cassandra.yaml
- cassandra-env.sh
- cassandra-env.ps1
- jvm.options
- cassandra-rackdc.properties
- Lab

Lesson 10: Simple Cassandra Cluster**Duration: 1 hours Including Lab**

- Simple cluster architecture and virtual nodes
- Automatic cluster consistency repair mechanisms
- Simple cluster configuration
- Experiencing CAP and consistency
- Lab

Lesson 11: Multi Data Center Cluster**Duration: 1 hours Including Lab**

- Multi datacenter cluster
- Multi datacenter configuration
- Cassandra cluster additional configurations for production environments
- Lab

Lesson 12: Monitoring Cassandra**Duration: 1 hours Including Lab**

- Monitoring Script
- Linux Useful Commands
- Nodetool Monitoring
- GC Monitoring
- Lab

DAY 3 - AGENDA

Lesson 13: Nodetool administration utility

Duration: 1 hours Including Lab

- Nodetool Utility
- Common Commands
- Adding a node to the cluster
- Removing a node from the cluster
- Nodetool repair
- Lab

Lesson 14: Security

Duration: 0.5 hours Including Lab

- Authentication
- Authorization
- Encryption
- Hardening
- Lab

Lesson 15: Backup and Recovery

Duration: 0.5 hours Including Lab

- Backup
- Restore

Lesson 16: Sizing and Performance planning (If time permits)

Duration: 1 hours Including Lab

- On Heap/Off Heap configuration
- Sizing Estimation
- Planning for Load and Load Testing

Lesson 17: Summary Cassandra Backend

Duration: 1 hours Including Lab

- Summary Cassandra Backend

Lesson 18: Cassandra Client - The Java Driver/Python Driver

Duration: 0.5 hour Including Lab

- Datastax Java/Python Driver
- Driver Classes
- Connecting to the cluster
- Address Resolution
- Authentication
- Session

Lesson 19: Simple Statements

Duration: 0.5 hour Including Lab

- Simple Statements
- Select
- Insert
- Update
- Delete
- What is JSON
- Querying JSON

Lesson 20: Prepared Statements

Duration: 0.5 hour Including Lab

- Prepared Statements

- Build Statement
- Batch Statement

Lesson 21: Paging

Duration: 0.5 hour including lab

- Paging
- Fetch Size
- Result Set

Lesson 22: Query Builder

Duration: 0.5 hour Including Lab

- Query builder
- CRUD with Query Builder
- Retry policy
- Additional properties

Lesson 23: Advanced Driver Topics

Duration: 0.5 hour Including Lab

- Load Balancing
- Failover and Retry policy
- Statement TTL, TIMESTAMP and LIMIT
- Consistency Level
- Lightweight Transactions

Lesson 24: In Memory Cache to Cassandra (If time permits)

Duration: 0.5 hour Including Lab

- Caching Cassandra in your App
- Ignite
- Redis

Lesson 27: Summary Cassandra Client

Duration: 0.5 hour Including Lab

- Summary