

Cassandra for Developers

Course Outline:

Section 1: Introduction to Big Data / NoSQL

- NoSQL overview
- CAP theorem
- When is NoSQL appropriate
- Columnar storage
- NoSQL ecosystem

Section 2 : Cassandra Basics

- Design and architecture
- Cassandra nodes, clusters, datacenters
- Keyspaces, tables, rows and columns
- Partitioning, replication, tokens
- Quorum and consistency levels
- Labs : interacting with cassandra using CQLSH

Section 3: Data Modeling – part 1

- introduction to CQL
- CQL Datatypes
- creating keyspaces & tables
- Choosing columns and types
- Choosing primary keys

- Data layout for rows and columns
- Time to live (TTL)
- Querying with CQL
- CQL updates
- Collections (list / map / set)
- Labs : various data modeling exercises using CQL ; experimenting with queries and supported data types

Section 4: Data Modeling – part 2

- Creating and using secondary indexes
- composite keys (partition keys and clustering keys)
- Time series data
- Best practices for time series data
- Counters
- Lightweight transactions (LWT)
- Labs : creating and using indexes; modeling time series data

Section 5 : Data Modeling Labs : Group design session

- multiple use cases from various domains are presented
- students work in groups to come up designs and models
- discuss various designs, analyze decisions
- Lab : implement one of the scenario

Section 6: Cassandra drivers

- Introduction to Java driver

- CRUD (Create / Read / Update, Delete) operations using Java client
- Asynchronous queries
- Labs : using Java API for Cassandra

Section 7 : Cassandra Internals

- understand Cassandra design under the hood
- sstables, memtables, commit log
- read path / write path
- caching
- vnode

Section 8: Administration

- Hardware selection
- Cassandra distributions
- Cassandra best practices (compaction, garbage collection,)
- troubleshooting tools and tips
- Lab : students install Cassandra, run benchmarks

Section 9: Bonus Lab (time permitting)

- Implement a music service like Pandora / Spotify on Cassandra