

Apache Cassandra Administration

Duration

4 Days

Training Mode

ILT with 50% Theory & 50% Handson

Contents

Chapter 1: BigData Overview

- What is BigData?
- What is NoSQL?
- CAP Theorem
- Understanding data consistency
- Why Cassandra?
- Cassandra Terminologies
- Use cases for Cassandra

Chapter 2: Data Modeling using Cassandra

- Designing differences between Relational & Cassandra data models
- Building Data models
- Clusters, Keyspaces
- Column Families & Columns
- Super Columns

Chapter 3: Configuring Cassandra

- Defining Keyspaces
- Partitioning data
- Data partitioning strategies
- Partitioner
- Replication factor
- Replication strategies
- Snitches

Chapter 4: Managing Data in Cassandra

- Query differences between SQL & CQL
- Basic write properties
- Basic read properties
- Using API
- Cassandra Query Language
- Cassandra Client Interface
- Performance Tuning Strategies

Chapter 5: The Write Path

- Log Structured Storage
- Memtables
- Flushing

Chapter 6: The Read Path

- SSTables
- Row Merging
- Cache (Key, Row)
- Compaction
- Distributed Deletes
- Memory Mapped Files
- Evolving Applications

Chapter 7: Partition Tolerance

- Data partitioners
- Replication strategies
- Snitches
- Hinted Handoff

Chapter 8: Availability

- How Cassandra handles failure of one or more nodes
- What to do in the face of failure
- Elastic scalability

Chapter 9: Tunable Consistency

- Consistency
- Coordinators
- Read Repair
- Phi Accrual
- Hinted Handoff
- Anti-entropy Service

Chapter 10: Monitoring Cassandra

- Monitoring CPU
- Disk
- Network
- Goals for Sizing
- Logging

Chapter 11: Backup & Recovery

- Backup

- Recovery

Chapter 12: Storage Strategies

- Specific concerns for cloud hosting
- Logical and Physical disk configuration
- Local disks vs. network mounted/shared drives

Chapter 13: Security

- Authentication
- Authorization
- Physical Security