# Learning Apache Cassandra

**Section 1: Introduction to Cassandra**

**1.1 The Course Overview**

This video provides an overview of the entire course.

**1.2 What Is Apache Cassandra?**

This video explains what Apache Cassandra is.

* When Cassandra is a great fit
* Learn the pros of Cassandra

**1.3 Key Space, Table Schema, Partition Key, and Clustering Key**

In this video, we will understand key space and the various keys used in Cassandra.

* Learn about the partition key
* Understand the table schema
* Learn about the clustering key

**1.4 Start a Single Node Cassandra Database**

In this video, we will install Cassandra.

* Start Cassandra
* Use the instructions to install Cassandra

**1.5 Introduction to Cqlsh Command Line Client**

In this video, we will start cqlsh.

* Create simple Cassandra table
* Create key space

**1.6 Loading and Reading Data**

In this video, we will insert elements to a table

* Query the table
* Learn the process of loading and reading data

**Section 2: Cassandra Distributed Architecture**

**2.1 Node and Ring Structure**

In this video, we will learn token ring.

* Learn virtual nodes
* Learn partitioning

**2.2 Replication and Consistency Model**

In this video, we will see what replication is.

* Learn consistency
* Learn replication factor

**2.3 Racks and Datacenters**

In this video, we will see what datacenter is.

* Learn racks
* Understand how Cassandra is aware of DC and racks

**2.4 CAP Theorem**

This video explains CAP theorem.

* See the consistency
* See the availability

**2.5 Gossip**

This video explains the gossip protocol.

* Learn how Cassandra achieves co-ordination between nodes
* How nodes stays in sync

**2.6 Read Repair, Hinted Handoff**

This video explains read repair.

* Explain hinted hand-off
* Read path

**Section 3: Diagnostics**

**3.1 Understanding Files in the Data Directory**

This video explains SSTables.

* Learn Cassandra write path
* Understand how data is written

**3.2 Use Nodetool to Examine Performance Statistics**

This video uses nodetool.

* Get statistics about Cassandra
* Learn how to use nodetool

**3.3 System and Output Logs**

In this video, we will examine Cassandra logs.

* Examine debug logs

**3.4 JMX to Monitor Metrics**

In this video, we will connect to Cassandra with JMX.

* Examine threads
* Examine memory

**3.5 Choosing the Appropriate Compaction Strategy**

In this video, we will see what the compaction is.

* Learn how compaction works
* Learn how to choose proper compaction strategy

**Section 4: Data Modelling Principles**

**4.1 Primary Key and Cluster Ordering**

In this video, we will see what a primary key is.

* Learn what a clustering key is
* How to model data using primary and clustering key

**4.2 Denormalization and Design for the Read Performance**

In this video, we will see how Cassandra is write-effective.

* Understand that data model can have duplicates
* Spread load evenly in the cluster

**4.3 Optimizing for BlindWrites**

In this video, we will see how to optimize performance of Cassandra writes.

* Understand how replication factor affects performance
* Learn how compaction strategy affects write performance

**Section 5: Data Modelling in Cassandra**

**5.1 Collection Types**

* In this video, we will learn the different collection types.
* Learn about Set
* Get to know about map

**5.2 Static Columns**

In this video, we will learn about static column.

* Share static column

**5.3 Indexes, Materialized Views**

* This video explains Cassandra index.
* Learn about Casandra materialized view

**5.4 Data Aggregation**

In this video, we will learn about data aggregation.

* Learn about min
* Learn about max
* Learn about avg and sum

**5.5 compareAndSet**

In this video, we will explain concurrent access using lightweight transactions.

* Explain compareAndSet
* Explain Cassandra IF EXISTS syntax

**5.6 Counter Type**

In this video, we will learn about the different counter types.

* Learn about read repair
* Learn about hinted handoff
* Learn about read path

**Section 6: Optimization of Data**

**6.1 The Impact of Frequent Updates and Delete**

This video shows what a Tombstone is.

* Learn how delete works

**6.2 Wide Rows and Primary Key Considerations**

This video shows what a wide row is.

* Learn how it affects primary key

**6.3 Load Testing with CQL Stress**

These video tests write performance.

* Test read performance

**6.4 Logged and Unlogged Batching**

This video shows what the unlogged batching is.

* Learn about logged batching
* Understand when to use which

**Section 7: Integrating Cassandra Database with Your Application**

**7.1 A Maven Project Using the Java Driver**

This video shows how to setup Cassandra Java project.

* Add Cassandra to dependencies

**7.2 Connection Information for the Driver**

This video shows how to start Cassandra node.

* Establish connection from java driver to Cassandra server

**7.3 Basic Statements**

In this video, we will use Select query from Java driver.

* Load data from Cassandra

**7.4 Using Prepared Statements**

In this video, we will create prepared statements.

* Use prepared statements to insert data to Cassandra

**7.5 Understanding Errors**

In this video, we will examine possible exceptions.

* Handle exceptions

**Section 8: Overview of Apache Spark**

**8.1 What Is Apache Spark and Spark Architecture**

In this video, we will see Spark architecture.

* Learn Resilient Distributed Dataset (RDD)
* Learn partitioning
* Learn transformations and actions and learn about laziness of transformations

**8.2 Get Started with Spark**

In this video, we will create a project with Spark.

* Start Spark context
* Learn how to create your own project with Spark

**8.3 Working with Spark’s Data Structures – RDD, Data Frame, and Dataset**

In this video, we will create data frame.

* Create a dataset
* Create a RDD
* Get to know about the API of DF, DS, and RDD

**8.4 Setting Up the Spark Connector**

In this video, we will add a Spark connector to the Spark project.

* Import implicit connector method in the RDD context

**Section 9: Connecting Spark with Cassandra**

**9.1 Writing Data to Cassandra from Spark**

In this video, we will save data to Cassandra using connector.

* Create a Spark project with Cassandra
* Integrate with Cassandra database

**9.2 Reading Data from Cassandra Using Spark RDD**

In this video, we will read data from Cassandra.

* Test, write, and retrieve data using Spark
* Use implicit functions from Cassandra driver to perform select function

**9.3 Join, Aggregate Data Using Spark Data Frame API and Spark SQL**

In this video, we will join data from two Cassandra tables.

* Join RDD with Cassandra table
* Use implicit functions from Cassandra driver to perform join function

**9.4 Cassandra Aware Partitioning in Spark**

In this video, we will learn about Cassandra aware partitioning.

* Join RDD with Cassandra using aware partitioning
* Use implicit functions from Cassandra driver to perform aware function

**Section 10: Integrate Cassandra with Spark Streaming**

**10.1 Use Cases for Near Real Time Stream Processing Using Spark Streaming**

In this video, we will see why use streaming.

* Learn Spark streaming architecture
* Learn DStream

**10.2 Advanced Stream Receiver Using Kafka Connectors**

In this video, we will create DStream.

* Use KafkaUtils to create stream based on Kafka topic

**10.3 Stateless and Stateful Transformations**

In this video, we will create stateless processing.

* Create stateful processing
* See the difference between stateless and stateful processing

**10.4 Persistence of Live Stream on to Cassandra**

In this video, we will use mapWithStat() method.

* Save Spark streaming state to Cassandra
* Load Spark streaming state from Cassandra
* Define check pointing