Introduction to Apache Cassandra (Days)

This hands-on, project-based Apache Cassandra class is for people new to Cassandra. Students will learn how to install Cassandra, to define a database, to create tables, to insert, update, and delete data, to create an application, and to create and modify a multiple-node cluster.

Class Prerequisites

Experience in the following *would be useful* for this Apache Cassandra class:

- Experience with databases, application servers, and programming.

Goals

- Learn how to install Cassandra.
- Learn to define a database in Cassandra.
- Learn to create tables in Cassandra.
- Learn to insert, update, and delete data in a Cassandra database.
- Learn to create a an application in Cassandra.
- Learn to create and modify a multiple-node cluster.

Outline

The Purpose of Cassandra

- Understanding What Cassandra is
- Exercise: Access the Apache Cassandra website
- Learning What Cassandra is Being Used For
- Exercise: Learn what Cassandra is being used for
- Lab: Open the main course virtual machine

Getting Started with the Architecture

- Understanding that Cassandra is a Distributed Database
- Exercise: View Cassandra documentation
- Learning What Snitch is For
- Learning What Gossip is For
- Exercise: Check your understanding
- Learning How Data Gets Distributed and Replicated
- Exercise: Access related documentation
- Lab: View the course projects

Installing Cassandra

- Downloading Cassandra
- Exercise: Locate the tarball
- Ensuring Oracle Java 7 is Installed
- Exercise: Check that Oracle Java 7 is installed
- Installing Cassandra
- Exercise: Install Cassandra
- Viewing the Main Configuration File

- Exercise: View the cassandra.yaml file
- Providing Cassandra with Permission to Directories
- Exercise: Provide Cassandra with permission to directories
- Starting Cassandra
- Exercise: Start Cassandra
- Checking Status
- Exercise: Check status
- Accessing the Cassandra system.log File
- Exercise: Access the Cassandra system.log file

Communicating with Cassandra
- Understanding Ways to Communicate with Cassandra
- Exercise: View CQL commands
- Using cqlsh
- Exercise: Access cqlsh
- Lab: Use cqlsh

Creating a Database
- Understanding a Cassandra Database
- Exercise: View the existing keyspaces
- Defining a Keyspace
- Exercise: Define a keyspace
- Lab: Create a second database

Creating a Table
- Creating a Table
- Exercise: Enter a keyspace
- Defining Columns and Data Types
- Exercise: Define columns and data types
- Defining a Primary Key
- Exercise: Define a primary key
- Recognizing a Partition Key
- Exercise: Recognize a partition key
- Specifying a Descending Clustering Order
- Exercise: Specify a descending clustering order
- Lab: Create a second table

Inserting Data
- Understanding Ways to Write Data
- Using the INSERT INTO command
- Exercise: Use the INSERT INTO command
- Using the COPY command
- Exercise: Use the COPY command
- Seeing How Data is Stored in Cassandra
- Exercise: See how data is stored in Cassandra

- Seeing How Data is Stored on Disk
- Exercise: See how data is stored on disk
- Lab: Insert data

Modeling Data
- Understanding Data Modeling in Cassandra
- Understanding Secondary Indexes
- Exercise: Use a WHERE clause
- Creating a Secondary Index
- Exercise: Create a secondary index
- Defining a Composite Partition Key
- Exercise: Define a composite partition key

Creating an Application
- Understanding Cassandra Drivers
- Exercise: View the Cassandra drivers
- Exploring the DataStax Java Driver
- Exercise: Explore the DataStax Java driver
- Setting Up a Development Environment
- Exercise: Access a development environment
- Creating an Application Page
- Exercise: Create an application page
- Getting the DataStax Java Driver Classes
- Exercise: Get the DataStax Java driver classes
- Connecting to a Cassandra Cluster
- Exercise: Connect to a Cassandra cluster
- Executing a Query
- Exercise: Execute a query
- Displaying Query Results
- Exercise: Display query results
- Using an MVC Pattern
- Exercise: View an MVC pattern
- Lab: Create a second application

Updating and Deleting Data
- Updating Data
- Exercise: Update data
- Understanding How Updating Works
- Exercise: Get an inside view into how updating works
- Deleting Data
- Exercise: Delete data
- Understanding Tombstones
- Exercise: View a tombstone
- Using TTLs
- Exercise: Create a TTL

- Lab: Update and delete data

Selecting Hardware
- Understanding Hardware Choices
- Understanding RAM and CPU Recommendations
- Exercise: Check for RAM
- Selecting Storage
- Exercise: View available storage on your computer
- Deploying in the Cloud
- Exercise: View cloud recommendations

Adding Nodes to a Cluster
- Understanding Cassandra Nodes
- Exercise: Set up a second a node
- Having a Network Connection
- Exercise: Have a network connection
- Specifying the IP Address for a Node
- Exercise: Specify the IP address for a node
- Specifying Seed Nodes
- Exercise: Specify seed nodes
- Bootstrapping a Node
- Exercise: Bootstrap a node
- Cleaning Up a Node
- Exercise: Clean up a node
- Using cassandra-stress
- Exercise: Use cassandra-stress
- Lab: Add a third node

Monitoring a Cluster
- Understanding Cassandra Monitoring Tools
- Using nodetool
- Exercise: Use nodetool
- Using JConsole
- Exercise: Use JConsole
- Learning About OpsCenter
- Exercise: Learn about OpsCenter

Repairing Nodes
- Understanding Repair
- Exercise: Modify the replication factor of a keyspace
- Repairing Nodes
- Exercise: Repair nodes
- Understanding Consistency
- Exercise: Specify consistency
- Understanding Hinted Handoff

- Understanding Read Repair
- Exercise: View read repair and hinted handoff settings
- Lab: Repair nodes for a keyspace

Removing a Node
- Understanding Removing a Node
- Decommissioning a Node
- Exercise: Decommission a node
- Putting a Node Back into Service
- Exercise: Put a node back into service
- Removing a Dead Node
- Exercise: Remove a dead node
- Lab: Put a node back into service

Redefining a Cluster for Multiple Data Centers
- Redefining for Multiple Data Centers
- Changing Snitch Type
- Exercise: Change snitch type
- Modifying cassandra-rackdc.properties
- Exercise: Modify cassandra-rackdc.properties
- Changing Replication Strategy
- Exercise: Change replication strategy