springpeople

COURSE NAME

Cassandra Intermediate (L1)

OBJECTIVE

Cassandra is a massively scalable NoSQL datastore capable of handling very high operations per second. It originated from Facebook and derives inspiration from Google's BigTable and Amazon DynamoDB. It is used by some of the iconic businesses around the world like Apple, Netflix, eBay,Walmart, Instagram etc, it has become a product of choice because of the high availability and masterless architecture.

DURATION

2 Days

TRAINEE PRE-REQUISITES

- Applicable for Developers / Designers / Technical enthusiasts
- With understanding of RDBMS and SQL
- And having basic familiarity of distributed systems
- Having Experience > 3 years

LEARNING OBJECTIVES

1. Understand the different use cases of using NoSQL databases
2. Realize the benefits offered by the NoSQL databases
3. Compare different categories of NoSQL databases
4. Master the architecture of Cassandra cluster design
5. Understand how does Cassandra provided great scalability , availabity and performance for massive datasets
6. Build application to store and access data from Cassandra

LAB REQUIREMENTS DETAILS

1. Intel I3 processor or equivalent or above
2. 8GB RAM or above + 500 MB free space or more Software (please consider 64bit/32bit installers depending on your hardware, 32bit not preferred)
3. Ubuntu desktop with internet access For both Windows and Mac download VirtualBox
from https://www.virtualbox.org/wiki/Downloads and install Ubuntu 16.04.x LTS, download
from https://www.ubuntu.com/download/desktop
4. Oracle JDK 8u92+ (download and install)
5. Python 2.7.10+ (download and install if not available)
6. Apache Cassandra tarball (latest stable release - download only, configuration is a hands on session)

Environment setup JAVA_HOME needs to be set. Java and Python should be in the PATH.

COURSE CONTENT

Day1

☐Explain the differences between NoSQL and RDBMS databases,

☐Explain what the various NoSQL databases are,

☐Topics - Quick Review of RDBMS:

Transactions, ACIDity, Schema, Two Phase Commit, Sharding and Share

Nothing Architecture, Feature Based, Key Based, Lookup Table Based,

☐NoSQL Databases, Brewers CAP Theorem, Cassandra Definition and Features,

Distributed and Decentralised,

☐Elastic Scalability, High Availability and Fault Tolerance,

☐Tuneable Consistency,

☐Strict Consistency,

☐Casual Consistency,

☐Weak (Eventual Consistency),

☐Column Orientation, Schema Free, High Performance, USE Cases for Cassandra, Cassandra Installation.

Introduction to Cassandra

☐Understanding What Cassandra is

☐What Cassandra is Being Used For

☐Explain the various Cassandra features, Explain why Cassandra scores over other NoSQL

databases,

☐Distinguish between use cases when Cassandra is a strong choice and when it is not,

Understand the use cases where Cassandra is implemented.

☐System Requirements

Cassandra Architecture

☐Cassandra - a Distributed Database

☐Introduction to Snitch

☐Introduction to Gossip

☐Introduction to Data Distribution

☐Introduction to Replication

☐Introduction to Virtual Nodes

Installing Cassandra

☐Downloading Cassandra

☐Installing Cassandra

☐Viewing the Main Configuration File

☐Providing Cassandra with Permission to Directories

☐Starting Cassandra

☐Checking Status

☐Accessing the Cassandra system.log File

Communicating with Cassandra

☐Ways to Communicate with Cassandra

☐Using Cqlsh

Creating a Database

☐Understanding a Cassandra Database

☐Defining a Keyspace

☐Deleting a Keyspace

Creating a Table

☐Creating a Table

☐Defining Columns and Data Types

☐Defining a Primary Key

☐Recognizing a Partition Key

☐Specifying a Descending Clustering Order

Inserting Data

☐Understanding Ways to Write Data

☐Using the Insert Into Command

☐Using the Copy Command

☐How Data is Stored in Cassandra

☐How Data is Stored on Disk

Hands On:

1. Installing Cassandra cluster using NetworkTopology to demonstrate how to create real

time cluster

2. Adding and removing into existing cluster

3. Creating keyspace, table , CRUD operations , data types , custom types and key modeling

exersie to understand demoralization concept.

Day2

Modeling Data

☐Understanding Data Modeling in Cassandra

☐Using a Where Clause

☐Understanding Secondary Indexes

☐Creating a Secondary Index

☐Defining a Composite Partition Key

☐Importing data in Cassandra from RDBMS

Creating An Application

☐Understanding Cassandra Drivers

☐Exploring the DataStax Java Driver

☐Setting Up a Development Environment

☐Creating An Application Page

☐Acquiring the DataStax Java Driver Files

☐Getting the DataStax Java Driver Files Through Maven

☐Providing the DataStax Java Driver Files Manually

☐Connecting to a Cassandra Cluster

☐Executing a Query

☐Displaying Query Results

☐Using An MVC Pattern

Hands On:

1. Using Java client for creating keyspace, table , CRUD operations , data types , custom types and key

modeling exersie to understand demoralization concept.

Updating and Deleting Data

- Updating Data

- Understanding How Updating Works

- Deleting Data

- Understanding Tombstones

- Using TTLs

- Updating a TTL

- Different ways of running CQL scripts

- CQL Types

- Numeric Data Types

- Textual Data Types

- Time and Identity Data Types

- Other Simple Data Types

- Collections

- User-Defined Types

- Secondary Indexes

Functions and aggregation