

Basic Hadoop course



DELTA**FROG**
Technology

Big Data and Hadoop Course:--

1)Basic Understanding of Big Data and Hadoop

Course Objective:- In this class you will learn what bigdata is, and problem to process bigdata with existing(traditional) system and how Hadoop provide solution for this problem.

Topics:

- .Introduction to Big Data & Big Data Challenges**
- .Limitations & Solutions of Big Data Architecture**
- .Hadoop & its Features**
- .Hadoop Ecosystem**
- .Hadoop 2.x Core Components**
- .Hadoop Storage: HDFS (Hadoop Distributed File System)**
- .Hadoop Processing: MapReduce Framework**
- .Different Hadoop Distributions**

2)Hadoop Architecture and HDFS

Course Objective:-In this class you learn hadoop architecture and HDFS Command.

Topics:

- .Hadoop 2.x Cluster Architecture
- .Federation and High Availability Architecture
- .Hadoop Cluster in Production
- .Common Hadoop Shell Commands
- .Installation of Single Node Cluster & Multi-Node Cluster.
- .Basic Hadoop Administration

3)Hadoop MapReduce Framework

Class Objective:-In this class , you will learn how to store data in HDFS and process with the help of MapReduce framework.

Topics:

- .Traditional way vs MapReduce way
- .Why MapReduce
- .YARN Architecture
- .YARN MapReduce Application Execution Flow
- .Anatomy of MapReduce Program
- .Input Splits, Relation between Input Splits and HDFS Blocks
- .MapReduce: Combiner & Partitioner
- .Practical Demo of realTime Dataset

4)Apache Pig

Class Objective:- In this class , you will learn how to process data with the help pig.

Topics:

- .Introduction to Apache Pig
- .MapReduce vs Pig
- .Pig Components & Pig Execution
- .Pig Data Types & Data Models in Pig
- .Pig Latin Programs

- .Shell and Utility Commands
- .Pig UDF & Pig Streaming
- .Pig Demo on RealTime Dataset

5)Apache Hive

Class Objective:- In this class, we will help you in understanding Hive concepts, Hive Data types, loading and querying data in Hive, running hive scripts and Hive UDF.

Topics:

- .Introduction to Apache Hive
- .Hive vs Pig
- .Hive Architecture and Components
- .Hive Metastore
- .Limitations of Hive
- .Comparison with Traditional Database
- .Hive Data Types and Data Models
- .Hive Partition
- .Hive Bucketing
- .Hive Tables (Managed Tables and External Tables)
- .Importing Data
- .Querying Data & Managing Outputs
- .Hive Script & Hive UDF
- .HQL: Joining Tables,
- .Dynamic Partitioning
- .Custom MapReduce Scripts
- .Hive Indexes and views
- .Hive Query Optimizers
- .Hive UDF
- .Hive Demo on RealTime Dataset

6)Apache HBase

Class Objective:-In this class , we will help you to understand NoSQL database and storing and processing data using Hbase.

Topics:

- .Introduction to NoSQL Databases and HBase

- .CAP Theorem
- .HBase v/s RDBMS
- .HBase Components
- .HBase Architecture
- .HBase Configuration
- .HBase Data Model
- .HBase Shell
- .HBase Client API
- .Hive Data Loading Techniques
- .HBase Bulk Loading
- .Getting and Inserting Data
- .HBase Filters

7)Apache Sqoop

Class Objective:-In this class ,We will help you to import and export data into HDFS from RDBMS with the help of Sqoop tool.

Topics:

- .Downloading and Installing Sqoop
- .Importing Data
- .Incremental Import
- .Free-Form Query Import
- .Export:-Transferring Data from Hadoop
- .Importing Data Directly into Hive