

BigData and Hadoop Training Course

Java Fundamentals

Basic java concepts (Object,Constructor,Inheritance,Overriding)
Multi-threading
File I/O –Java. IO
Collections –Java.Util.*, Java.Math, Java.Lang
Java Serialization
Eclipse IDE – Java Development.

Hadoop Fundamentals

What is Big Data? Why Big Data?
Hadoop Architecture & Components
Hadoop Processing – Map Reduce, Spark Frameworks

Hadoop Architecture And HDFS

Hadoop1.x (HDFS Basis)
Hadoop2.x (YARN , Federation)
File Storage
Fault Tolerance, High Availablity
Hadoop Configuration File
Single Node Cluster / Multinode Cluster
HDFS Commands

Map Reduce

What Is MapReduce?
Basic MapReduce Concepts
Hadoop 2.x MapReduce Architecture and Components
Input Splits , Relation Between Input Splits and HDFS Blocks
Concepts of Mappers, Reducers, Combiners and Partitioning
Demo (Word Count , Weather DataSet)

Advance MapReduce

Counters , Distributed Cache
MapSide Join , Reduce Side Join
Inputs and Output formats to MR Program
MRUnit

Spark

What Is Spark?
Basic Spark Concepts
How Spark differs from Map Reduce?
Spark Ecosystem
Working With RDD
Spark SQL
Saprk Streaming
Spark GraphX
SparkMlib

Hive

What is Hive, why we need it and its importance?

How Hive is different from Traditional RDBMS

Metastore in Hive

Hive Data Types , Modeling in Hive, creating Hive structures and data load process.

Concepts of Blocks, Hashing, External Tables etc.

Concepts of serialization, deserialization

Different Hive data storage formats including ORC, RC

Introduction to HiveQL and examples.

Joining Table

Concepts of Partitioning, Bucketing, Indexing

Hive as an ELT tool and difference between Pig and Hive

Writing and mastering Hive UDFs and Thrift Server

Pig and Latin

Basics of Pig and Why Pig?

Grunt

Pig's Data Model (Data Types)

Apache Pig Architecture

Installation

Relational Operators (Group Operator, COGROUP Operator, Joins , Union, Diagnostic Operators)

Writing Evaluation (Programming Structure in Pig)

Filter

Load & Store Functions

Benefits of Pig over SQL language

Built In Functions

Execution of xml file using Pig

HBase

Introduction to NoSQL

HBase – Introduction (HBASE v/s RDBMS)

When to use HBase

Hbase Architecture

HBase Families & Components

HBase Data Model

Data Storage and Distribution

Zookeeper

HBase Master

Oozie

Flume and Sqoop Demo

Oozie Components, Oozie Workflow,

Scheduling with Oozie, Demo on Oozie Workflow,

Oozie for MapReduce, PIG, Hive, and Sqoop

Hadoop Project Demo,

Projects in Hadoop will also be covered in this training