HADOOP SYLLABUS

- Big Data

- Hadoop

- Hadoop Tools

- MapReduce

- SQOOP

- FLUME

- PIG

- HIVE

- IMPALA

- HBASE

- Ooziee

- SPARK Basics

- Real Time Scenario – Perform Twitter Analysis

  YouTube Data Analysis

HADOOP SYLLABUS

Big Data:

What is Big Data?

How big data arrives?

How it works in real world?

Big data Market in future.

Difference between Big data and Data?

Hadoop:

What is Hadoop?

Hadoop history?

Venders in Hadoop?

How Hadoop works?

Hadoop Installation.

Hadoop Ecosystem:

Introduction to Hadoop Ecosystem

Tools available in Hadoop Ecosystem
(Hive,pig,sqoop,flume,impala,oozie,HBase)

How tools used to solve the real time problems?

Introduction to Hadoop Architecture

How data stored in HDFS

Namenode,Datanode,JobTracker,Tasker

HADOOP SYLLABUS

MapReduce:

What is MapReduce and how it works?

What is Map and what is Reduce?

File Format and Data types available in Hadoop

What is combiner?

What is shuffle and sort?

What is input split?

What is Record reader and Record writer?

Introduction to YARN

How YARN works?

Example

SQOOP:

What is SQOOP and its configuration?

How SQOOP works?

Import and Export from RDBMS

Incremental Load in RDBMS

Connect from SQLSERVER to HDFS

Real time scenarios in SQOOP

FLUME:

What is Flume and its Configuration?

What is Flume agents?

What is Source?

What is Channel?

What is Sinks?

Real time scenarios in FLUME

HIVE & IMPALA:

What is hive?

Different types of Tables?

Function and Query in hive (all)

Partition

Buckets

Example in hive

Introduction to IMPALA

Impala Overview

PIG:

Introduction to PIG

Pig Data Model

Inbuilt function in PIG

Real time Examples.

Oozie:

Introduction to Oozie

How to use action and controls

How to schedule a job using Oozie

How workflow,Co-ordinator,Bundles work

HADOOP SYLLABUS

Example using Oozie

HBase and Zookeeper

Introduction to HBase

HBase overview

Introduction to Zookeeper

Zookeeper overview

SPARK:

Introduction to spark

How it Works and RDD in spark

Basic of spark.

HADOOP SYLLABUS