Module1: UNIX & Basic Sun Solaris OS concepts:

Basic Primitives of UNIX:

Origin of Unix, Overview of System Administration

Unix basic commands

Working with editors – vi and sed

Linking files with symbolic link and hard link

Scheduling Jobs using Crontab and at

Locating a process and monitoring processes with ps and prstat

Killing Process in different ways

Installation:

Planning system resources

Solaris installation on workstations

Managing Software Packages and Patches

User Management and Security:

Creation of groups, users and modifying their accounts

Changing passwords, set password aging Assigning environmental variables to users, administration

of Initialization of files

Change ownership of files and directories

Setting advanced file permissions like setuid, setgid, stickybit

Disk and File System Management:

Managing disks formatting, labeling, viewing volume table of contents Bringing disks to O/S control

dynamically

File Systems types, creating, mounting, unmounting checking and repairing File Systems.

Module2: CORE JAVA

Basic Java Programming

Oops

Abstraction

Inheritance

Polymorphism

Interfaces

Java.lang package

String, StringBuffer & StringBuilder

Wrapper classes

Generics

TypeSafty

Type Casting

File I/O and Serialization

Collection Framework

Module3: HADOOP DEVELOPER COURSE DETAILS

Understand Big Data & Hadoop: Basic Concepts

What is BigData

Characteristics of BigData

Challenges with Traditional Systems

Problems with BigData

Handling BigData

HADOOP Core Concepts

Problems with Existing Distributed Systems to deal Big Data

Why Hadoop and An Overview and History of Hadoop

Requirements of NewApproach

The Hadoop Project and Hadoop Components

Hadoop Distributed File System – HDFS

What is HDFS, Why it is required for running Map-Reduce.

How it differs from other distributed file systems.

Design of HDFS & Concepts

Command Line Interface, Hadoop File Systems, Java Interface

Data Flow (Anatomy of a File Read, Anatomy of a File Write)

Hadoop Archives.

Hands-on Exercise on HDFS

Understanding - Map-Reduce Basics and Map-Reduce Types and Formats

Describe Map-Reduce framework works & Flow

Why Map-Reduce is tightly coupled with HDFS.

What are the different types of Input and Output formats and why they are required?

Hadoop Data Types,

Concept of Mappers & Reducers

Concept of Partitioners & Combiners

MapReduce Types

Input Formats (Input Splits and Records, Text Input, Binary Input, Multiple Inputs, Database Input

and Output)

Output Formats (TextOutput, BinaryOutPut, Multiple Outputs, Databaseoutput).

Hands-on Exercise

Setting up Hadoop Cluster

Cluster introduction

How to setup Hadoop Cluster

Different configuration files which need to be edited for Cluster setup.

Cluster Specification,

Cluster Setup and Installation

SSH Configuration

Hadoop Configuration

Important Hadoop Daemon Properties

Hadoop Daemon Addresses and Ports

The Five Daemons working

Other Hadoop Properties,

User Account Creation.

Install Pseudo cluster

Install Multi node cluster

Hands-on Exercise

Developing Map Reduce Programs

Setting up Eclipse Development Environment.

Eclipse integration with HADOOP for Rapid Application Development

Understanding HADOOP API

Creating Map Reduce Projects,

Writing MapReduce Drivers, Mappers and Reducers in Java

Driver Code

Mapper Code

Reducer Code

Map Reduce Code

Differences Between the Old and New MapReduce APIs

Hands-on Exercise

Understanding ToolRunner

More about ToolRunner

Combiner

Reducer

configure and close methods

Hands-on Exercise

Common MapReduce Algorithms

Sorting

Searching

Indexing

TF-IDF

Word\_CoOccurance

Hands-on Exercise

HADOOP EcoSystem

Flume

Sqoop

Importing data from RDBMS using sqoop

Hive

Introduction

Introduction to Hadoop and Hive

Getting Data into Hive

Manipulating Data with Hive

Partitioning and Bucketing Data

Pig

Introduction to pig

Different modes of pig

when to use hive and when to use pig

Hadoop Developer Programme

HBASE

Basics of HBASE

Hands-on Exercise

Advanced MapReduce Programming

Developing custom Writable

Developing custom Writable Comparable

Understanding Input Output formats

Hands-on Exercise

Introduction to Ooziee