Iverson Associates Sdn Bhd (303330-M)
Suite T113 – T114, 3rd Floor, Centrepoint, Lebuh Bandar Utama
Bandar Utama, 47800 Petaling Jaya, Selangor Darul Ehsan
Tel: 03-7726 2678   Fax: 03-7727 9737   Website: www.iverson.com.my

Course Outline :: Cloudera Essentials for Apache Hadoop ::

Module Title	:	**Cloudera Essentials for Apache Hadoop**
Duration	:	**1 day**

Overview

This one-day essentials course presents an overview of Apache Hadoop and how it can help decision-makers meet business goals, providing a fundamental introduction to the main components of Hadoop and its use cases in various industries. This course is a good starting point for any role or set of objectives and is part of the data analyst learning path

Hands-On Hadoop

Through instructor-led discussion and interactive, hands-on exercises, participants will navigate the Hadoop ecosystem, learning topics such as:

- The occasions on which Hadoop is the appropriate platform for managing and analyzing data
- The ways organizations in different verticals are using Hadoop to maximize the value of data
- How Hadoop fits into the existing hardware, software, and data environment
- What you need to know about choosing Hadoop for your company
- Which resources are necessary to effectively deploy Hadoop in development and production

Audience & Prerequisites

This course is best suited to anyone who wants the big picture of Hadoop in the enterprise: CTOs, CIOs, technical managers, architects, data analysts, developers, administrators. A basic level of technical proficiency is valuable. Prior knowledge of Apache Hadoop is not required.

Course Outline

Introduction

The Motivation for Hadoop

- Problems with traditional large---scale systems
- Requirements for a new approach
- Introducing Hadoop

Iverson Associates Sdn Bhd (303330-M)
Suite T113 – T114, 3rd Floor, Centrepoint, Lebuh Bandar Utama
Bandar Utama, 47800 Petaling Jaya, Selangor Darul Ehsan
Tel: 03-7726 2678   Fax: 03-7727 9737   Website: www.iverson.com.my

Course Outline :: Cloudera Essentials for Apache Hadoop ::

Hadoop: Basic Concepts

- What is Hadoop?
- The Hadoop Distributed File System (HDFS)
- How MapReduce works
- Anatomy of a Hadoop cluster

Hadoop Solutions

- The most common problems Hadoop can solve
- Real---world Hadoop use cases

The Hadoop Ecosystem

- Introduction
- Hive and Pig
- HBase
- Flume
- Other ecosystem components

Hadoop in the Data Center

- How Hadoop integrates with other infrastructure components
- What physical resources are needed?
- Managing a Hadoop cluster
- Cloudera Manager

Managing the Elephant in the Room

- What people or resources are required?
- How to hire or cross---train
- How can Cloudera help?