

Introduction to Hadoop for Managers

Course Summary

Description

The Introduction to Hadoop for Managers training course is designed to equip managers and other decision makers with an understanding of Hadoop.

To fully understand the impact Hadoop has on your organization's infrastructure, it is critical to understand its business benefits, required resources and potential uses. In this course, you will gain a thorough understanding of what the technology is and how it would impact their organizations. You will learn how to use the technology to gain competitive advantage, save time and money and solve data problems.

Objectives

By the end of this course, students will be able to:

- Understand reasons to use Hadoop
- Understand the Hadoop platform
- Know how to work with Hadoop in your existing environment
- Know how to manage a Hadoop cluster, including determining what costs and other resources are needed

Topics

- Why Hadoop?
- The Hadoop Platform
- The Big Picture: How Hadoop Fits into Your Architecture
- Operations: Managing a Hadoop Cluster

Prerequisites

There are no prerequisites for this course.

Duration

One day

Introduction to Hadoop for Managers

Course Outline

I. Why Hadoop?

- A. The Motivation for Hadoop
- B. Use Cases and Case Studies About Hadoop

II. The Hadoop Platform

- A. MapReduce and HDFS
- B. Hive
- C. Pig
- D. HBase
- E. Sqoop
- F. Flume
- G. Hue
- H. Hadoop Distributions and How to Choose One

III. The Big Picture: How Hadoop Fits into Your Architecture

- A. How Hadoop Can Augment Your Existing Environment
 - 1. Relational Databases
 - 2. SANs
 - 3. OLAP

IV. Operations: Managing a Hadoop Cluster

- A. What People Resources are Required
- B. What Training Will be Required
- C. What Physical Resources are Required
- D. What This Will Cost Your Organization
- E. How It Scales for Growth