

INTRODUCTION TO BIG DATA AND NOSQL TRAINING (1 Day)

We live in the information age where business success is grounded on the ability of organizations to convert raw data coming from various sources into high-grade business information.

Many organizations are overwhelmed by the sheer volume of information they have to process in order to stay competitive. Traditional database systems may become either prohibitively expensive to handle the exponential growth of data volumes or found unsuitable for the job. At this point, the data gets mystically morphed into the Big Data.

This course provides an introduction to Big Data as well as NoSQL (Not Only SQL) database systems. The fundamental concepts of and ideas behind Big Data / NoSQL technologies are methodically explored and many buzzwords demystified. The course is supplemented by hands-on labs that help attendees reinforce their theoretical knowledge of the subject.

TOPICS

- Defining Big Data
- Big Data Stores Overview
- NoSQL
- Big Data Business Intelligence and Analytics
- Real World Case Studies
- Adopting NoSQL

AUDIENCE

General audience including business and technology team leadership

PRE-REQUISITES

Basic programming skills, some knowledge of SQL

CHAPTER 1. DEFINING BIG DATA

- Transforming Data into Business Information
- Gartner's Definition of Big Data
- More Definitions of Big Data
- Challenges Posed by Big Data
- The Cloud and Big Data
- The Business Value of Big Data
- Big Data: Hype or Reality?

CHAPTER 2. NOSQL AND BIG DATA SYSTEMS OVERVIEW

- Limitations of Relational Databases
- What are NoSQL (Not Only SQL) Databases?
- NoSQL Past and Present
- NoSQL Database Properties
- NoSQL Benefits
- NoSQL Database Storage Types
- The CAP Theorem
- Limitations of NoSQL Databases
- Big Data Sharding
- Sharding Example
- Amazon S3
- Amazon Storage SLAs
- Amazon Glacier
- Amazon S3 Security
- Data Lifecycle Management with Amazon S3
- Amazon S3 Cost Monitoring
- OpenStack
- Object Store (Swift)
- Components of Swift
- Google BigTable
- BigTable-based Applications
- BigTable Design
- Google App Engine
- Google App Engine Billing
- Google Cloud Storage
- Hadoop
- Hadoop's Core Components
- Hadoop Distributed File System
- Accessing HDFS
- HBase
- HBase design
- MemcacheDB
- MongoDB
- MongoDB Operational Intelligence
- MongoDB Use Cases

CHAPTER 3. BIG DATA BUSINESS INTELLIGENCE AND ANALYTICS

- Comparison with other systems
- NoSQL Data Querying and Processing
- MapReduce programming model
- Example of Map & Reduce Operations using JavaScript
- Analyzing Big Data with Hadoop
- Hadoop's MapReduce
- Hadoop Streaming
- Making things simpler with Hadoop Pig Latin
- Example of a Pig Script in Batch Mode
- Amazon Elastic MapReduce
- Big Data in Google App Engine
- Example of Google AppEngine Java Datastore API
- MongoDB Data Model
- MongoDB Query Language (QL)
- The
- find
- and
- findOne
- methods
- A MongoDB QL Example
- What is Hive
- Interfacing with Hive
- Business analytics with Hive
- The UnQL Specification

CHAPTER 4. BIG DATA REAL WORLD CASE STUDIES

- Hadoop @ Yahoo
- Yahoo for Hadoop
- Yahoo!!
- Big Data @ Facebook
- Hive @ Facebook
- Mailtrust (Rackspace's mail division)

CHAPTER 5. ADOPTING NOSQL

- Hype Cycle and Technology Adoption Model
- Barriers to Adoption
- Dismantling Barriers to Adoption
- Use Cases for NoSQL Database Systems
- Example Applications
- Industry trends
- Enterprise Big Data / NoSQL Offerings
- NoSQL Technology Adoption Action Plan