

## Professional Training for BigData and Apache Hadoop

**While watching we promise you will say WOW! At least once**

Accelerate Your and Organization Hadoop Education

Apache Hadoop is increasingly being adopted in a wide range of industries and as a result, Hadoop expertise is more valuable than ever for you and your organization. Using Hadoop, organizations can consolidate and analyze data in ways never before possible. Businesses can capture, manage and process information that they used to throw away. You can leverage years of Hadoop experience with training from [www.HadoopExam.com](http://www.HadoopExam.com)

The course/training is designed specifically for CEO, CTO to Managers, Software Architect to an Individual Developers and Testers to enhance their skills in BigData world. You will learn when the use of Hadoop is appropriate, what problems Hadoop addresses, how Hadoop fits into your existing environment, and what you need to know about deploying Hadoop.

Learn the basics of the Hadoop Distributed File System (HDFS) and MapReduce framework and how to write programs against its API, as well as discuss design techniques for larger workflows. This training also covers advanced skills for debugging MapReduce programs and optimizing their performance, plus introduces participants to related projects in Distribution for Hadoop such as Hive, Pig, and Oozie. After completing the training, attendees can leverage our Hadoop Certification Exam Simulator for Developer as well as Administrator to clear the Hadoop Certification. Since launch 467+ attendees already cleared the exam with the help of our simulator.

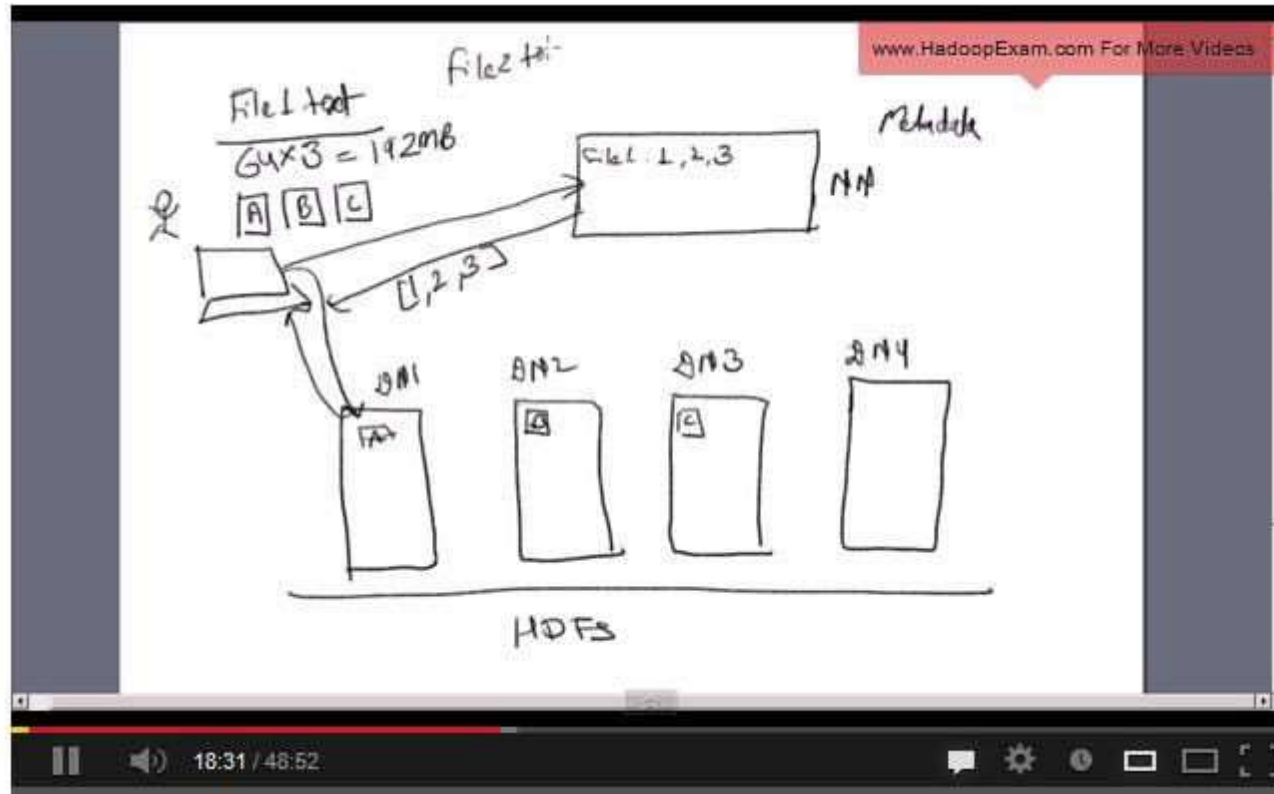
**Industry Where Hadoop is Being Used:** Energy & Utilities, Financial Services ,Government, Healthcare & Life Sciences, Media & Entertainment, Retail, eCommerce Consumer, Product, Technology, Telecommunications, **start ups**(They are trying their each & every resource should have Hadoop knowledge)

### Features of HadoopExam Learning Resources Training

1. **Very Convenient (24/7 access)** : This training will be available online for three months and any time you can access as per your comfort, so no need to wait as soon as access provided to you, you can start watching the trainings.
2. **No to PPT at ALL, As effective as classroom trainings** : We have taken the advanced step for providing online training and avoided using PPTs and created recordings as white board sessions.
3. **No travel, No Leaves & No Weekend classes**: This training is created keeping in mind that professionals cannot take leave so easily to attend the trainings. Because these sessions are going to be used by CEO, CTO, Solution Architect, Managers, Individual Developer and Tester. So no to travel, no to leave and No to weekend classes.
4. **Immediate and Cost-effective (Equal to price of a Book)**: It is so cost effective that it cost you equal to buying a book on BigData and Hadoop.

5. [In Just two Weeks 3200+ views on Youtube](#) : We provide Module 2 as free of cost for the Demo and after launch in just two weeks we got **3200+** views and all positive appreciation you can see in below testimonial sections.

Module 2 is freely available for Demo [Watch Now](#) (While watching you definitely says WOW!)



## Complete List of BigData/Hadoop Training Sessions

[Syllabus and Completed Hadoop Training is Below](#)

### Module 1 : Introduction to BigData, Hadoop (HDFS and MapReduce) : Available (Length 35 Minutes)

1. BigData Introduction
2. Hadoop Introduction
3. HDFS Introduction
4. MapReduce Introduction

Video URL : <http://www.youtube.com/watch?v=R-qjyEn3bjs> ([View Demo](#))

### Module 2 : Deep Dive in HDFS : Available (Length 48 Minutes)

1. HDFS Design
2. Fundamental of HDFS (Blocks, NameNode, DataNode, Secondary Name Node)
3. Rack Awareness
4. Read/Write from HDFS
5. HDFS Federation and High Availability ([Hadoop 2.x.x](#))
6. Parallel Copying using DistCp
7. HDFS Command Line Interface

Video URL : <http://www.youtube.com/watch?v=PK6lm7tBWow> ([View Demo](#))

### Module 2A : HDFS File Operation Lifecycle (Supplementary) : Available (Length 45 Minutes)

1. File Read Cycle from HDFS
  - DistributedFileSystem
  - FSDataInputStream
2. Failure or Error Handling When File Reading Fails
3. File Write Cycle from HDFS
  - FSDataOutputStream
4. Failure or Error Handling while File write fails

Video URL : <http://www.youtube.com/watch?v=Wu2EGfQY-i4> ([View Demo](#))

**Module 3 : Understanding MapReduce : Available (Length 60 Minutes)**

1. JobTracker and TaskTracker
2. Topology Hadoop cluster
3. Example of MapReduce  
Map Function  
Reduce Function
4. Java Implementation of MapReduce
5. DataFlow of MapReduce
6. Use of Combiner

**Video URL :** [Watch Private Video](#)

**Module 4 : MapReduce Internals -1 (In Detail) : Available (Length 57 Minutes)**

1. How MapReduce Works
2. Anatomy of MapReduce Job (MR-1)
3. Submission & Initialization of MapReduce Job (What Happen ?)
4. Assigning & Execution of Tasks
5. Monitoring & Progress of MapReduce Job
6. Completion of Job
7. Handling of MapReduce Job
  - Task Failure
  - TaskTracker Failure
  - JobTracker Failure

**Video URL :** [Watch Private Video](#)

**Module 5 : MapReduce-2 (YARN : Yet Another Resource Negotiator [Hadoop 2.x.x](#)) : Available (Length 52 Minutes)**

1. Limitation of Current Architecture (Classic)
2. What are the Requirement ?
3. YARN Architecture
4. JobSubmission and Job Initialization
5. Task Assignment and Task Execution
6. Progress and Monitoring of the Job

7. Failure Handling in YARN

- Task Failure
- Application Master Failure
- Node Manager Failure
- Resource Manager Failure

**Video URL :** [Watch Private Video](#)

**Module 6 : Advanced Topic for MapReduce (Performance and Optimization) : Available (Length 58 Minutes)**

1. Job Scheduling
2. In Depth Shuffle and Sorting
3. Speculative Execution
4. Output Committers
5. JVM Reuse in MR1
6. Configuration and Performance Tuning

**Video URL :** [Watch Private Video](#)

**Module 7 : Advanced MapReduce Algorithm : Available (Length 87 Minutes)**

File Based Data Structure

- Sequence File
- MapFile

Default Sorting In MapReduce

- Data Filtering (Map-only jobs)
- Partial Sorting

Data Lookup Strategies

- In MapFiles

Sorting Algorithm

- Total Sort (Globally Sorted Data)
- InputSampler
- Secondary Sort

**Video URL :** [Watch Private Video](#)

**Module 8 : Advanced MapReduce Algorithm -2 : Available : Private (Length 67 Minutes)**

1. MapReduce Joining
  - Reduce Side Join
  - MapSide Join
  - Semi Join
2. MapReduce Job Chaining
  - MapReduce Sequence Chaining
  - MapReduce Complex Chaining

**Video URL :** [Watch Private Video](#)

**Module 9 : Features of MapReduce : Available : Private (Length 61 Minutes)**

Introduction to MapReduce Counters

- Types of Counters
- Task Counters
- Job Counters
- User Defined Counters
- Propagation of Counters

Side Data Distribution

- Using JobConfiguration
- Distributed Cache
- Steps to Read and Delete Cache File

**Video URL :** [Watch Private Video](#)

**Module 10: MapReduce DataTypes and Formats : Available : Private (Length 77 Minutes)**

1. Serialization In Hadoop
2. Hadoop Writable and Comparable
3. Hadoop RawComparator and Custom Writable
4. MapReduce Types and Formats
5. Understand Difference Between Block and InputSplit
6. Role of RecordReader
7. FileInputFormat
8. ComineFileInputFormat and Processing whole file Single Mapper
9. Each input File as a record

10. Text/KeyValue/NLine InputFormat
11. BinaryInput processing
12. MultipleInputs Format
13. DatabasInput and Output
14. Text/Biinary/Multiple/Lazy OutputFormat MapReduce Types

**Video URL :** [Watch Private Video](#)

**Module 11 : Apache Pig : Available (Length 52 Minutes)**

1. What is Pig ?
2. Introduction to Pig Data Flow Engine
3. Pig and MapReduce in Detail
4. When should Pig Used ?
5. Pig and Hadoop Cluster
6. Pig Interpreter and MapReduce
7. Pig Relations and Data Types
8. PigLatin Example in Detail
9. Debugging and Generating Example in Apache Pig

**Video URL :** [Watch Private Video](#)

**Module 12 : Fundamental of Apache Hive Part-1 : Available (Length 60 Minutes)**

1. What is Hive ?
2. Architecture of Hive
3. Hive Services
4. Hive Clients
5. how Hive Differs from Traditional RDBMS
6. Introduction to HiveQL
7. Data Types and File Formats in Hive
8. File Encoding
9. Common problems while working with Hive

Video URL : [Watch Private Video](#)

**Module 13 : Apache Hive : Available (Length 73 Minutes )**

1. HiveQL
2. Managed and External Tables
3. Understand Storage Formats
4. Querying Data
  - Sorting and Aggregation
  - MapReduce In Query
  - Joins, SubQueries and Views
5. Writing User Defined Functions (UDFs)
3. Data types and schemas
4. Querying Data
5. HiveODBC
6. User-Defined Functions

**Video URL :** [Watch Private Video](#)

**Module 14 : Hands On : Single Node Hadoop Cluster Set Up In Amazon Cloud : Available (Length 60 Minutes Hands On Practice Session)**

1. ♦ How to create instance on Amazon EC2
  2. ♦ How to connect that Instance Using putty
  3. ♦ Installing Hadoop framework on this instance
  4. ♦ Run sample wordcount example which come with Hadoop framework.
- In 30 minutes you can create Hadoop Single Node Cluster in Amazon cloud, does it interest you ?

**Video URL :** [Watch Private Video](#)

**Module 15 : Hands On : Implementation of NGram algorithm : Available (Length 48 Minutes Hands On Practice Session)**

1. Understand the NGram concept using (Google Books NGram )
2. Step by Step Process creating and Configuring eclipse for writing MapReduce Code
3. Deploying the NGram application in Hadoop Installed in Amazon EC2
4. Analyzing the Result by Running NGram application (UniGram, BiGram, TriGram etc.)

**Video URL :** [Watch Private Video](#)



7000+ Learners already benefited from [www.HadoopExam.com](http://www.HadoopExam.com)

**Module 16 : Hands On : Hadoop MultiNode Cluster Setup and Running a BigData Example : Available (Length 70 Minutes) New**

1. Hadoop MultiNode Cluster
2. Setup Three Node Hadoop cluster
3. Running NGram Application on cluster
4. Analyze the cluster using
  - NameNode UI (Multiple Blocks and effect of Replication Factor)
  - JobTracker UI (Multiple MapTask running on different Nodes)
5. SettingUp Replication Factor

**Video URL :** [Watch Private Video](#)

**Module 17 : NOSQL Introduction and Implementation : Available (Length 56 Minutes) New**

1. What is NoSQL ?
2. NoSQL Characterstics or Common Traits
3. Catgories of NoSQL DataBases
  - Key-Value Database
  - Document DataBase
  - Column Family DataBase
  - Graph DataBase
4. Aggregate Orientation : Perfect fit for NoSQL
5. NOSQL Implementation
6. Key-Value Database Example and Use
7. Document DataBase Example and Use
8. Column Family DataBase Example and Use
9. What is Polyglot persistence ?

**Video URL :** [Watch Private Video](#)

**Module 18 : HBase Introduction : : Available (Part-1 Length 48 Minutes and Part-2 Length-37 Minutes) New**

1. Fundamentals of HBase
2. Usage Scenerio of HBase

3. Use of HBase in Search Engine
4. HBase DataModel
  - Table and Row
  - Column Family and Column Qualifier
  - Cell and its Versioning
  - Regions and Region Server
5. HBase Designing Tables
6. HBase Data Coordinates
7. Versions and HBase Operation
  - Get/Scan
  - Put
  - Delete

**Video URL :** Watch Private Video [Part-1](#) and [Part-2](#)

**Module 19 : Hands On Hadoop Setup and Running MapReduce Example on Virtual Machine on Windows : : Available (Length 51 Minutes) New**

1. Installation VM (Virtual Machine)
2. Creating Hadoop Environment in VM
3. Creating MapReduce Program
4. Running MapReduce Job
5. Analyzing JobTracker and NameNode UI

**Video URL :** [Watch Private Session](#)

**Module 20 : Apache Cassandra : Available (Length 63 Minutes) New**

1. BigData and Apache Cassandra
2. Why Cassanra is so Popular
3. Cassandra as a Distributed DataBase
4. Cassandra and High Availability
5. Cassandra and Replication Mechanism
6. Cassandra's Elastic Scalability

7. Tuneable consistency
  - Strict Consistency
  - Casual Consistency
  - Weak Consistency
8. Brewer's CAP Theorem
9. Cassandra as a Scema Free DataBase
10. Where should we use Cassandra
11. Who and why using the Cassandra

**Video URL :** [Watch Private Video](#)

**Module 21: Hands On MRUnit (MapReduce Testing Framework) : Available (Length 48 Minutes) New**

1. Practice Basic MapReduce Without Installing Hadoop Framework
2. Mapper Testing
3. Reducer Testing
4. Counter Testing
5. Full MapReduce Job Testing

**Video URL :** [Watch Private Video](#)

**Module 22 : Apache Sqoop (SQL To Hadoop) : Available (Length 66 Minutes) New**

1. Sqoop Tutorial
2. How does Sqoop Work
3. Sqoop JDBC Driver and Connectors
4. Sqoop Importing Data
5. Various Options to Import Data
  - Table Import
  - Binary Data Import
  - SpeedUp the Import
  - Filtering Import
  - Full DataBase Import Introduction to Sqoop

**Video URL :** [Watch Private Video](#)

**Module 23 : Apache Flume : Available (Length 28 Minutes) New**

1. Data Acquisition : Apache Flume Introduction
2. Apache Flume Components
3. POSIX and HDFS File Write
4. Flume Events
5. Interceptors, Channel Selectors, Sink Processor

**Video URL :** [Watch Public Video](#)

**Module 24 : Advanced Apache Flume : Available (Length 48 Minutes) New**

1. Sample Twiteer Feed Configuration
2. Flume Channel
  - Memory Channel
  - File Channel
3. Sinks and Sink Processors
4. Sources
5. Channel Selectors
6. Interceptors

**Video URL :** [Watch Public Video](#)

**Module 25 : YARN Introduction (Length 52 Mins) Available [Hadoop 2.x. YARN Training](#)**

1. Why to think Beyond MapReduce
2. New Components of YARN
3. Revisit Hadoop 1.0
4. How YARN fits in Hadoop Framework
5. Hadoop MR1 Components Revisit
6. Need for Non-MapReduce
7. YARN Components Introduction

**Module 26 : Fundamental Overview of YARN (Length 40 Mins) Available [Hadoop 2.x. YARN Training](#)**

1. YARN Functional Component
2. YARN Architecture Overview
3. Claiming and Re-claiming Resources
4. Functional Properties of Resource Manager

7000+ Learners already benefited from [www.HadoopExam.com](http://www.HadoopExam.com)

- Node Manager
- Application Master
- 5. YARN Scheduling Component
- 6. Introduction to FIFO Scheduler
- 7. Introduction to Capacity Scheduler

**Module 27 : Powerfull Hadoop 2.0 Framework (Length : 27 Mins) Available [Hadoop 2.x. YARN Training](#)**

- 1. HDFS 1.0 Versus Hadoop 2.0
- 2. Resource Manager - Subcomponent
- 3. Details About Fair Share Scheduler
- 4. Hierarchical Queues in Scheduler
- 5. Containers
- 6. Node Manager and Its Responsibility
- 7. Role of Application Master while submitting Jobs

**Module 28 : Submitting the Application to YARN Hadoop Cluster (Length : 27 Mins) Available [Hadoop 2.x. YARN Training](#)**

- 1. Submitting the Application to YARN Hadoop Cluster
- 2. Managing Application Dependencies
- 3. Writing a YARN Application : Birdseye View

**Module 29 : LocalResources of the Application Available [Hadoop 2.x. YARN Training](#)**

- 1. Understanding of YARN Application/Jobs Dependencies
- 2. Types of LocalResource
- 3. Visibilities of Local Resources
- 4. Lifetime of Local Resources
- 5. Good and Bad Local Resources
- 6. Target Directories of Local Resources

**Module 30 : Deep Dive in Capacity Scheduler (Length 39 Mins) Available [Hadoop 2.x. YARN Training](#)**

- 1. Introduction and Enabling Capacity Scheduler
- 2. Setting Up Queues in the CS
- 3. Access Control List Setup
- 4. Managing Cluster Capacity in with Queues
- 5. Resource Distribution Workflow Example

**Module 31 : Managing Capacity Scheduler (Length 39 Mins) Available [Hadoop 2.x. YARN Training](#)**

- 1. Managing Capacity with Queues
- 2. Resource Distribution Example
- 3. Understanding User Limits
- 4. Application Reservation
- 5. Understanding the Preemption

**Module 32 : Hadoop Security : Kerberos Authentication (Length 23 Mins) Available [Hadoop Security Training](#)**

- 1. Kerberos Authentication
- 2. Important entity of Kerberos Authentication

3. How Kerberos Process works

**Module 33 : Apache Spark : Introduction to Apache Spark (Length 48 Mins) Available 100 Time Faster Data Processing**

1. Introduction to Apache Spark
2. Features of Apache Spark
3. Apache Spark Stack
4. Introduction to RDD's
5. RDD's Transformation
6. What is Good and Bad In MapReduce
7. Why to use Apache Spark

To buy the training [Click Here](#) or Visit [www.HadoopExam.com](http://www.HadoopExam.com)

## Testimonials: (Many More By emails)



[Raghavendra R](#) 2 weeks ago

Awesome explanation sir.... if u upload for pig sqoop hive and hbase it wil be a very usefull too all. thankq



[Siva kumar](#) 2 weeks ago

nice training flow... Thanks.



[kantha penikelapatti](#) 2 weeks ago

very good explanation about hdfs



[Kalyan Chakravarthy](#) 2 weeks ago

Nice explanation. tks



[Java Jadhav](#) 2 weeks ago

Nice explanation about HDFS....ThanQ and keep it up. Hope to see more tutorials in future.



[Anju Singh](#) 4 weeks ago

VERY NICE TUTORIALS FOR BEGINNERS, I ENJOYED THIS VIDEO.

Website Search Help Tags: Hadoop Training, BigData Training, HBase Training, Cassandra Training, MongoDB Training, Apache Spark Training, NoSQL Training, SAS Training, Hortonworks Hadoop Training, Salesforce Training, AWS Training

## Other Certification Material we have

Cloudera Certification	Amazon WebService Certification	SAS Professional Certification	Salesforce Certification
<ol style="list-style-type: none"> <li><a href="#">CCD-410 Hadoop Developer Simulator : 430 Questions (New Syllabus &amp; Pattern ) + 15 Page Revision Notes</a></li> <li><a href="#">CCA410 : Hadoop Administrator Simulator : 307 Questions (New Syllabus &amp; Pattern) + 15 Page Revision Notes</a></li> <li><a href="#">CCB400 : 262 HBase Certification Practice Questions (CCSHB or CCB-400)</a></li> </ol>	<ol style="list-style-type: none"> <li><a href="#">424+ Practice Questions for Amazon AWS Certified Solutions Architect : Associate Level/Professional</a></li> <li><a href="#">424+ Practice Questions for Amazon AWS Certified Solutions Architect : Associate Level / Professional</a></li> <li><a href="#">Practice Questions for Amazon AWS Certified SysOps : Associate Level / Professional</a></li> </ol>	<ol style="list-style-type: none"> <li><a href="#">A00-211 : 490+ SAS Base Certification Practice Questions (Includes latest 9.3 edition as well)</a></li> <li><a href="#">A00-212 : 356+ SAS Advanced (A00-212) Certification Practice Questions (Includes latest 9.3 edition as well)</a></li> <li><a href="#">A00-240 : 86+ SAS Certified Statistical Business Analyst Using SAS 9: Regression and</a></li> </ol>	<ol style="list-style-type: none"> <li><a href="#">Total 405 Q&amp;A + 60 Pages Study Cum Revision Notes for Salesforce.com Certified Force.com Developer (Dev401 Exam)</a></li> </ol>

993+ Learners Already Benefited from [www.HadoopExam.com](http://www.HadoopExam.com)

<p>4. <a href="#">DS200: 235 Data Science Certification Practice Questions (CCP:DS200) + 150 Pages Revision Study Notes</a></p> <p>5. <a href="#">CCA500 &amp; 505 : 200 Practice Questions for Cloudera CDH5 (CCA500 &amp; CCA505) Certification Practice Questions + Revision Notes of 28 Pages</a></p>		<p><a href="#">Modeling Credential (A00-240) Certification Practice Questions (Includes latest 9.3 edition as well)</a></p> <p>4. <a href="#">A00-250 : 85+ SAS Certified Platform Administrator 9 (A00-250) Certification Practice Questions (Includes latest 9.3 &amp; 9.4 edition as well)</a></p>	
---	--	---	--

To buy the training [Click Here](#) or Visit [www.HadoopExam.com](http://www.HadoopExam.com)

**[For Corporate Bulk License Please Contact Us](#)**

Email: [admin@hadoopexam.com](mailto:admin@hadoopexam.com), [hadoopexam@gmail.com](mailto:hadoopexam@gmail.com)

**Phone: 022-42669636**

**Mobile: +91-8879712614**

**HadoopExam Learning Resource**

**B-902 Shah Trade Centre**

**Mumbai 400097**