

Training: Cloudera
Cloudera Training for Apache HBase


FORM OF TRAINING	MATERIALS	PRICE	DURATION
Traditional	Digital materials	1780 EUR	3 days
Traditional	CTAB Tablet	1930 EUR	3 days
Distance learning	Digital materials	1780 EUR	3 days
Distance learning	CTAB Tablet	1780 EUR	3 days

LOCATIONS

Krakow - 5 Tatarska Street, II floor, hours: 9:00 am - 4:00 pm
 Warsaw - 17 Bielska Street, hours: 9:00 am - 4:00 pm

TRAINING TERMS

2019-11-12 | 3 days | Warszawa

TRAINING GOALS:

This four-day training course for **Apache HBase** enables participants to store and access massive quantities of multi-structured data and perform hundreds of thousands of operations per second.

Apache HBase is a distributed, scalable, NoSQL database built on Apache Hadoop. HBase can store data in massive tables consisting of billions of rows and millions of columns, serve data to many users and applications in real time, and provide fast, random read/write access to users and applications.

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

- The use cases and usage occasions for HBase, Hadoop, and RDBMS
- Using the HBase shell to directly manipulate HBase tables
- Designing optimal HBase schemas for efficient data storage and recovery
- How to connect to HBase using the Java API, configure the HBase cluster, and administer an HBase cluster
- Best practices for identifying and resolving performance bottlenecks

CONSPECT:

- Introduction

- Introduction to Hadoop and HBase
 - What Is Big Data?
 - Introducing Hadoop
 - Hadoop Components
 - What Is HBase?
 - Why Use HBase?
 - Strengths of HBase
 - HBase in Production
 - Weaknesses of HBase
- HBase Tables
 - HBase Concepts
 - HBase Table Fundamentals
 - Thinking About Table Design
- The HBase Shell
 - Creating Tables with the HBase Shell
 - Working with Tables
 - Working with Table Data
- HBase Architecture Fundamentals
 - HBase Regions
 - HBase Cluster Architecture
 - HBase and HDFS Data Locality
- HBase Schema Design
 - General Design Considerations
 - Application-Centric Design
 - Designing HBase Row Keys
 - Other HBase Table Features
- Basic Data Access with the HBase API
 - Options to Access HBase Data
 - Creating and Deleting HBase Tables
 - Retrieving Data with Get
 - Retrieving Data with Scan
 - Inserting and Updating Data
 - Deleting Data
- More Advanced HBase API Features
 - Filtering Scans
 - Best Practices

- HBase Coprocessors
- HBase on the Cluster
 - How HBase Uses HDFS
 - Compactions and Splits
- HBase Reads and Writes
 - How HBase Writes Data
 - How HBase Reads Data
 - Block Caches for Reading
- HBase Performance Tuning
 - Column Family Considerations
 - Schema Design Considerations
 - Configuring for Caching
 - Dealing with Time Series and Sequential
- Data
 - Pre-Splitting Regions
- HBase Administration and Cluster Management
 - HBase Daemons
 - ZooKeeper Considerations
 - HBase High Availability
 - Using the HBase Balancer
 - Fixing Tables with hbck
 - HBase Security
- HBase Replication and Backup
 - HBase Replication
 - HBase Backup
 - MapReduce and HBase Clusters
- Using Hive and Impala with HBase
 - Using Hive and Impala with HBase
- Conclusion

REQUIREMENTS:

- This course is appropriate for developers and administrators who intend to use HBase.
- Prior experience with databases and data modeling is helpful, but not required.
- Prior knowledge of Java is helpful.
- Prior knowledge of Hadoop is not required, but Cloudera Developer Training for Apache Hadoop

provides an excellent foundation for this course.

Difficulty level



CERTIFICATE:

The participants will obtain certificates signed by Cloudera. This course helps prepare for a Cloudera Certified Specialist in Apache HBase (CCSHB) certification exam <https://www.cloudera.com/about/training.html>. Certification is a great differentiator; it helps establish you as a leader in the field, providing employers and customers with tangible evidence of your skills and expertise.

TRAINER:

Certified Cloudera Instructor.