Training: Cloudera
Cloudera Training for Apache HBase

CLOUDERA Training Partner

TRAINING GOALS:

Take your knowledge to the next level with Cloudera Training for Apache HBase. Cloudera Educational Services' three-day training course enables participants to store and access massive quantities of multi-structured data and perform hundreds of thousands of operations per second.

What You Will Learn

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 to insert and retrieve data in real time
- Best practices for identifying and resolving performance bottlenecks

CONSPECT:

- Introduction to Hadoop and HBase
  - Introducing Hadoop
  - Core 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
- HBase Shell
  - Creating Tables with the HBase Shell

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

- 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 Write Path
  - HBase Write Path
  - Compaction
  - Splits
- HBase Read Path
  - How HBase Reads Data
  - Block Caches for Reading
- HBase Performance Tuning
  - Column Family Considerations
  - Schema Design Considerations
  - Configuring for Caching
  - Memory Considerations
  - Dealing with Time Series and Sequential Data
  - Pre-Splitting Regions

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

- 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
  - How to Use Hive and Impala to Access HBase
- Conclusion
- Appendix A: Accessing Data with Python and Thrift
  - Thrift Usage
  - Working with Tables
  - Getting and Putting Data
  - Scanning Data
  - Deleting Data
  - Counters
  - Filters

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. Knowledge of Java is assumed. Prior knowledge of Hadoop is not required, but Cloudera Developer Training for Spark and Hadoop provides an excellent foundation for this course.

Difficulty level



CERTIFICATE:

The participants will obtain certificates signed by Cloudera (course completion).

Upon completion of the course, attendees are encouraged to continue their study and register for the

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

CDP Data Developer exam
https://www.cloudera.com/about/training/certification/cdp-datadev-exam-cdp-3001.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

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007