

BDSCP MODULE 07

Fundamental Big Data Engineering

Arcitura®
CERTIFIED
Big Data Engineer

OVERVIEW

This course covers engineering-related concepts, techniques and technologies for the processing and storage of Big Data datasets. It highlights the unique challenges faced when processing and storing large, volatile and disparate sets of data. NoSQL is covered and the MapReduce data processing engine is explained in detail as a base framework for high-volume batch data processing.

The following primary topics are covered:

- Big Data Engineering Techniques and Challenges
- Big Data Storage, including Sharding, Replication, CAP Theorem, ACID and BASE
- Master-Slave, Peer-to-Peer Replication, Combining Replication with Sharding
- Big Data Storage Requirements, Scalability, Redundancy and Availability
- Fast Access, Long-term Storage, Schema-less Storage and Inexpensive Storage
- On-Disk Storage, including Distributed File System and Databases
- Introduction to NoSQL and NewSQL
- NoSQL Rationale and Characteristics
- NoSQL Database Types, including Key-Value, Document, Column-Family and Graph Databases
- Big Data Processing Engines
- Distributed/Parallel Data Processing, Schema-less Data Processing
- Multi-Workload Support, Linear Scalability and Fault-Tolerance
- Big Data Processing Requirements, including Batch, Cluster and Realtime Modes
- MapReduce for Big Data Processing, including Map, Combine, Partition, Shuffle and Sort and Reduce
- MapReduce Algorithm Design
- Task Parallism, Data Parallism

Duration: 1 Day

TRAINING + CERTIFICATION

This course is part of a training and accreditation program through which official certification can be achieved and for which official training and certification badges can be issued.



www.arcitura.com/bdscp/engineer
www.arcitura.com/bdscp/consultant

eLEARNING + PRINTED STUDY KITS

This course is available as part of a study kit that includes course materials and supplemental self-study materials. The study kit is available in full-color printed format and eLearning format. Both types of study kits include the contents listed below.

- Workbook
- Self-Study Guide
- Mind Map Poster
- Flashcards
- Audio Tutor Recordings



Arcitura®

Pearson | VUE | OnVUE | Arcitura On-Site EXAM PROCTORING | cclaIm | Credly

BDSCP MODULE 07

Fundamental Big Data Engineering

	Certified Big Data Professional*	Certified Big Data Science Professional	Certified Big Data Scientist	Certified Big Data Consultant	Certified Big Data Engineer	Certified Big Data Architect	Certified Big Data Governance Specialist
MODULE 01 Fundamental Big Data	●	●	●	●	●	●	●
MODULE 02 Big Data Analysis & Technology Concepts	●	●	●	●	●	●	●
MODULE 03 Big Data Analysis & Technology Lab		●		●			
MODULE 04 Fundamental Big Data Analysis & Science			●	●			
MODULE 05 Advanced Big Data Analysis & Science			●				
MODULE 06 Big Data Analysis & Science Lab			●				
MODULE 07 Fundamental Big Data Engineering				●	●		
MODULE 08 Advanced Big Data Engineering					●		
MODULE 09 Big Data Engineering Lab					●		
MODULE 10 Fundamental Big Data Architecture						●	
MODULE 11 Advanced Big Data Architecture						●	
MODULE 12 Big Data Architecture Lab						●	
MODULE 13 Fundamental Big Data Governance							●
MODULE 14 Advanced Big Data Governance							●
MODULE 15 Big Data Governance Lab							●

* The Certified Big Data Professional designation is automatically issued when achieving any other BDSCP certification. It can also be achieved by receiving passing grades on Exams B90.01 + B90.02.

Arcitura Big Data School (www.arcitura.com/big-data-school)