

<https://www.ituonline.com/>

Microsoft SQL Server - Big Data

Module 1: What are Big Data Clusters?

- 1.1 Introduction
- 1.2 Linux, PolyBase, and Active Directory
- 1.3 Scenarios

Module 2: Big Data Cluster Architecture

- 2.1 Introduction
- 2.2 Docker
- 2.3 Kubernetes
- 2.4 Hadoop and Spark
- 2.5 Components
- 2.6 Endpoints

Module 3: Deployment of Big Data Clusters

- 3.1 Introduction
- 3.2 Install Prerequisites
- 3.3 Deploy Kubernetes
- 3.4 Deploy BDC
- 3.5 Monitor and Verify Deployment

Module 4: Loading and Querying Data in Big Data Clusters

- 4.1 Introduction
- 4.2 HDFS with Curl
- 4.3 Loading Data with T-SQL
- 4.4 Virtualizing Data
- 4.5 Restoring a Database

Module 5: Working with Spark in Big Data Clusters

- 5.1 Introduction
- 5.2 What is Spark
- 5.3 Submitting Spark Jobs
- 5.4 Running Spark Jobs via Notebooks
- 5.5 Transforming CSV
- 5.6 Spark-SQL
- 5.7 Spark to SQL ETL

Module 6: Machine Learning on Big Data Clusters

- 6.1 Introduction
- 6.2 Machine Learning Services
- 6.3 Using MLeap
- 6.4 Using Python
- 6.5 Using R

Module 7: Create and Consume Big Data Cluster Apps

- 7.1 Introduction
- 7.2 Deploying, Running, Consuming, and Monitoring an App
- 7.3 Python Example - Deploy with azdata and Monitoring
- 7.4 R Example - Deploy with VS Code and Consume with Postman
- 7.5 MLeap Example - Create a yaml file
- 7.6 SSIS Example - Implement scheduled execution of a DB backup

Module 8: Maintenance of Big Data Clusters

- 8.1 Introduction
- 8.2 Monitoring
- 8.3 Managing and Automation
- 8.4 Course Wrap Up