



Exam material for Microsoft 70-475 Designing and Implementing Big Data Analytics Solutions

By: [Daniel Calbimonte](#) | [Read Comments](#) | Related Tips: [More](#) > [Professional Development Certifications](#)

Attend a SQL Server Conference for FREE >> [click to learn more](#)

Problem

I love your Microsoft [certification series](#) and want to be certified for Big Data solutions. Does Microsoft offer a certification in Big Data?

Solution

Yes, Microsoft does offer a certification in Big Data and in this tip we will talk about exam 70-475. This exam is oriented to designing and implementing Big Data Analytics Solutions.

FAQ

Who should take this exam?

This exam is oriented to DBAs, Data Scientist, Data Architects, Data Analysts, Data Developers or other professionals who want to learn or who want to be certified in Big Data.

Is this exam a requirement for any Certification?

This exam is optional to get the **MCSA** in Azure ([Cloud Platform Microsoft Certified Solutions Associate](#)), it is also an optional exam for the [MCSE: Data Management and Analytics](#).

Is the exam easy?

No. If you do not have experience with Hadoop, Azure, Data Lake, Data Factory, Machine Learning, Lambda architecture, this exam is very difficult.

What is the relationship between Big Data and SQL Server?

Everything is migrating to Azure. SQL Server on-premises is gradually being migrated to Azure and unstructured data is being stored in the Cloud. You can query data from SQL Server on-premises or Azure SQL using PolyBase.

Which books would you recommend for this exam?

- [Data Lake for Enterprises](#)
- [Getting started with Azure Data Factory](#)
- [Introducing Microsoft Azure HDInsight](#)
- [Mastering Azure Analytics: Architecting in the Cloud with Azure Data Lake, HDInsight, and Spark](#)

- [Big Data Analytics with Microsoft HDInsight in 24 Hours, Sams Teach Yourself](#)
- [HDInsight Essentials - Second Edition](#)
- [Processing Big Data with Azure HDInsight: Building Real-World Big Data Systems on Azure HDInsight Using the Hadoop Ecosystem](#)
- [Microsoft Big Data Solutions](#)
- [HDInsight: Microsoft's Cloud Hadoop](#)
- [Pro Microsoft HDInsight: Hadoop on Windows](#)
- [70-475: Designing and Implementing Big Data Analytics Solutions: Study Guide](#)
- [HDInsight For Beginners](#)
- [HDInsight Jump Start](#)
- [Dive In HDInsight](#)

Are there some courses for this exam?

Yes, here are some useful courses.

- [Big Data Analytics with HDInsight: Hadoop on Azure](#)
- [Implementing Big Data Analysis](#)
- [Design and Implement Big Data & Advanced Analytics Solutions](#)
- [Creating your first HDInsight cluster and run samples](#)
- [Introducing Azure Data Lake](#)
- [Orchestrating Data and Services with Azure Data Factory](#)
- [Azure Data Factory makes it even easier and convenient to uncover insights from data when using Data Lake Store with SQL Data Warehouse](#)
- [Processing Real-Time Data Streams in Azure](#)
- [Hands-On with Azure Machine Learning](#)

Are there some practice tests?

Yes, the following exams are available:

- [70-475: Designing and Implementing Big Data Analytics Solutions Microsoft Official Practice Test](#)
- [Microsoft \(70-475\): Designing and Implementing Big Data Analytics Solutions - English](#)

Do you have some links to study, for this exam?

Yes, these links will be useful:

Design big data batch processing and interactive solutions

- Ingest data for batch and interactive processing
 - [Upload data for Hadoop jobs in HDInsight](#)
 - [Copy data to and from Data Lake Store by using Data Factory](#)
 - [Introduction to Azure Data Lake Store](#)
 - [Getting data into Azure Blob Storage](#)
 - [Hadoop - Command Reference](#)
 - [Azure Cosmos DB Documentation](#)
 - [Move data to SQL Server on an Azure virtual machine](#)
 - [Working with table, blob, queues and file storage in Azure](#)
 - [Load data from SQL Server into Azure SQL Data Warehouse \(SSIS\)](#)
 - [Azure Feature Pack for Integration Services \(SSIS\)](#)
 - [SQL Server 2016 – PolyBase tutorial](#)

- [What is Apache Storm on Azure HDInsight?](#)
- [Using Azure PowerShell with Azure Storage](#)
- [Realtime Event Processing with Azure WebJobs and Reactive Extensions](#)
- [What is the Azure SDK for .NET?](#)
- [Using Apache Flume with HDInsight](#)
- [APACHE SQOOP](#)
- Design and provision compute clusters
 - [Set up clusters in HDInsight with Hadoop, Spark, Kafka, and more](#)
 - [Introduction to Azure HDInsight, the Hadoop technology stack, and Hadoop clusters](#)
 - [Introduction to Spark on HDInsight](#)
 - [What is HBase in HDInsight: A NoSQL database that provides BigTable-like capabilities for Hadoop](#)
 - [What is Apache Storm on Azure HDInsight?](#)
 - [Introduction to R Server and open-source R capabilities on HDInsight](#)
 - [Introducing Apache Kafka on HDInsight \(preview\)](#)
 - [Use Interactive Hive in HDInsight \(Preview\)](#)
 - [An introduction to Hadoop security with domain-joined HDInsight clusters \(Preview\)](#)
 - [HDInsight Pricing](#)
 - [HOW TO CHOOSE THE RIGHT AZURE HDINSIGHT CLUSTER](#)
- Design for data security
 - [Azure security technical capabilities](#)
 - [Set up Azure Key Vault with end-to-end key rotation and auditing](#)
 - [Using shared access signatures \(SAS\)](#)
 - [Use Role-Based Access Control to manage access to your Azure subscription resources](#)
 - [Data Management Gateway](#)
- Design for batch processing
 - [Ambari Documentation](#)
 - [Oozie, Workflow Engine for Apache Hadoop](#)
 - [WebHCat](#)
 - [ZooKeeper: Because Coordinating Distributed Systems is a Zoo](#)
 - [Welcome to Apache Pig!](#)
 - [Apache Hive](#)
 - [A Complete Tutorial to Learn Data Science with Python from Scratch](#)
 - [Python for Big Data Programming](#)
 - [HBase, Sqoop, Flume and More: Apache Hadoop Defined](#)
 - [Mahout](#)
 - [Spark SQL, DataFrames and Datasets Guide](#)
 - [MapReduce Tutorial](#)
 - [Handling big data with PowerShell](#)
 - [Using SSRS With Large Datasets](#)
 - [Process large-scale datasets using Data Factory and Batch](#)

Design big data real-time processing solutions

- Ingest data for real-time processing
 - [Realtime event processing in Hadoop with Nifi, Kafka and Storm](#)
 - [7 Big Data Examples – Application of Big Data in Real Life](#)
 - [What is Event Hubs?](#)
 - [Chapter 6. HBase and Schema Design](#)
- Design and provision compute resources

- [Reference architecture: Real-time event processing with Microsoft Azure Stream Analytics](#)
- [Chapter 1. Enterprise Analytics Fundamentals](#)
- [Manage Hadoop clusters in HDInsight by using Azure PowerShell](#)
- [Running a Multi-Node Storm Cluster](#)
- Design for Lambda architecture
 - [The Lambda architecture: principles for architecting realtime Big Data systems](#)
 - [What is the Lambda Architecture?](#)
 - [Building a Lambda Architecture On Azure](#)
 - [Lambda Architecture](#)
- Design for real-time processing
 - [Azure Stream Analytics](#)
 - [What is Stream Analytics?](#)
 - [Help choosing a streaming analytics platform: Apache Storm comparison to Azure Stream Analytics](#)
 - [Big data and analytics](#)

Operationalize end-to-end cloud analytics solutions

- Create a data factory
 - [Introducing Data Factory: Orchestration on Big Data](#)
 - [Create Azure Data Factory](#)
 - [Azure Data Factory Documentation](#)
 - [Azure Data Factory: My Data Wandered Lonely as a Cloud with Julie Smith](#)
 - [Azure Data Factory Visual Studio Extension for authoring pipelines](#)
 - [Create, monitor, and manage Azure data factories using Azure Data Factory .NET SDK](#)
- Orchestrate data processing activities in a data-driven workflow
 - [Data Factory scheduling and execution](#)
 - [Datasets in Azure Data Factory](#)
 - [Azure Data Factory - Frequently Asked Questions](#)
 - [Pipelines and Activities in Azure Data Factory](#)
- Monitor and manage the data factory
 - [Troubleshoot Data Factory issues](#)
 - [Monitor and manage Azure Data Factory pipelines by using the Azure portal and PowerShell](#)
 - [Monitor and manage Azure Data Factory pipelines by using the Monitoring and Management app](#)
- Move, transform, and analyze data
 - [Use Pig with Hadoop on HDInsight](#)
 - [Use MapReduce in Hadoop on HDInsight](#)
 - [Create predictive pipelines using Azure Machine Learning and Azure Data Factory](#)
 - [Monitoring Azure Data Factory using PowerBI](#)
 - [What is Azure Machine Learning Studio?](#)
 - [Azure Machine Learning - Your first experiment](#)
 - [Beginning Azure ML Part 1 - Importing Data, Accessing, & Creating a New Experiment](#)
- Design a deployment strategy for an end-to-end solution
 - [Tutorial: Build your first Azure data factory using Azure PowerShell](#)
 - [Tutorial: Create a Data Factory pipeline that moves data by using Azure PowerShell](#)
 - [Tutorial: Create a pipeline with Copy Activity using Data Factory Copy Wizard](#)
 - [Tutorial: Use Azure portal to create a Data Factory pipeline to copy data](#)
 - [Tutorial: Use REST API to create an Azure Data Factory pipeline to copy data](#)
 - [Azure Data factories](#)
 - [Walkthrough Step 5: Deploy the Azure Machine Learning web service](#)

- [Tutorial: Create a data factory by using Visual Studio](#)

Next Steps

If you have any more questions, feel free to ask in the comments section.

For more information about this exam, refer to these links:

- [Exam 70-475](#)
- [Exam 70-475: Designing and Implementing Big Data Analytics Solutions](#)
- [Cert Exam Prep: Exam 70-475: Big Data and Analytics Solutions](#)

Last Update: 2017-07-10

About the author



Daniel Calbimonte is a Microsoft SQL Server MVP, Microsoft Certified Trainer and Microsoft Certified IT Professional.

[View all my tips](#)

Related Resources

- [More Career Tips...](#)

Copyright (c) 2006-2017 [Edgewood Solutions, LLC](#) All rights reserved

Some names and products listed are the registered trademarks of their respective owners.