



Virtuelt kursus

3 dage

Nr. 87923 P

DKK 17.499

ekskl. moms

Dato

23-10-2020

Sted

Virtuelt kursus

Virtual classroom: Big Data on AWS

Big Data on AWS giver dig indsigt i cloud-baseret big data løsninger gennem forskellige Amazon systemer. På kurset bliver der arbejdet med forskellige Amazon features der kan bruges til at skabe big data miljøer, samt hvorledes man skaber en big data løsning, der er sikker og omkostningseffektiv.



Lær i dine egne
omgivelser



Lær med kursister
fra hele verden



Spar tid og udgifter
på transport



Adgang til udvidet
undervisningsmateriale

Virtuelt kursus

Dette virtuelle kursus foregår på din egen computer live via GoToMeeting med en engelsktalende underviser. Under kurset har du mulighed for at stille spørgsmål, deltage i diskussioner, se whiteboard på din skærm og lave lab øvelser.

Get introduced to AWS big data platform

Big Data on AWS introduces you to cloud-based big data solutions such as Amazon Elastic MapReduce (EMR), Amazon Redshift, Amazon Kinesis and the rest of the AWS big data platform. In this course, we show you how to use Amazon EMR to process data using the broad ecosystem of Hadoop tools like Hive and Hue. We also teach you how to create big data environments, work with Amazon DynamoDB, Amazon Redshift, and Amazon Kinesis, and leverage best practices to design big data environments for security and cost-effectiveness.

Prerequisites

We recommend that attendees of this course have the following prerequisites:

- Basic familiarity with big data technologies, including Apache Hadoop, MapReduce, HDFS, and SQL/NoSQL querying
- Students should complete the Big Data Technology Fundamentals web-based training or have equivalent experience
- Working knowledge of core AWS services and public cloud implementation
- Students should complete the AWS Essentials course or have equivalent experience
- Basic understanding of data warehousing, relational database systems, and database design

Target Audience

This course is intended for:

- Individuals responsible for designing and implementing big data solutions, namely Solutions Architects and SysOps Administrators
- Data Scientists and Data Analysts interested in learning about big data solutions on AWS

Content

This course teaches you how to:

- Fit AWS solutions inside of a big data ecosystem
- Leverage Apache Hadoop in the context of Amazon EMR
- Identify the components of an Amazon EMR cluster
- Launch and configure an Amazon EMR cluster
- Leverage common programming frameworks available for Amazon EMR including Hive, Pig, and Streaming
- Leverage Hue to improve the ease-of-use of Amazon EMR
- Use in-memory analytics with Spark and Spark SQL on Amazon EMR
- Choose appropriate AWS data storage options
- Identify the benefits of using Amazon Kinesis for near real-time big data processing
- Define data warehousing and columnar database concepts
- Leverage Amazon Redshift to efficiently store and analyze data
- Comprehend and manage costs and security for Amazon EMR and Amazon Redshift deployments
- Identify options for ingesting, transferring, and compressing data
- Use visualization software to depict data and queries
- Orchestrate big data workflows using AWS Data Pipeline

The course content is distributed over 3 days as following:

Day 1

- Overview of Big Data
- Ingestion, Transfer, and Compression
- Storage Solutions
- Storing and Querying Data on DynamoDB
- Big Data Processing and Amazon Kinesis



- Introduction to Apache Hadoop and Amazon EMR
- Using Amazon Elastic MapReduce

Day 2

- Hadoop Programming Frameworks
- Processing Server Logs with Hive on Amazon EMR
- Processing Chemistry Data Using Hadoop Streaming on Amazon EMR
- Streamlining Your Amazon EMR Experience with Hue
- Running Pig Scripts in Hue on Amazon EMR
- Spark on Amazon EMR
- Interactively Creating and Querying Tables with Spark and Spark SQL on Amazon EMR
- Managing Amazon EMR Costs
- Securing your Amazon EMR Deployments

Day 3

- Data Warehouses and Columnar Datastores
- Amazon Redshift and Big Data
- Optimizing Your Amazon Redshift Environment
- Big Data Design Patterns
- Visualizing and Orchestrating Big Data
- Using Tibco Spotfire to Visualize Big Data

[Læs mere om vores virtuelle kurser og se svar på dine spørgsmål \(FAQ\).](#)

Søgte du et andet virtuelt kursus?

Vi tilbyder virtuelle kurser inden for mange forskellige områder. Kontakt os på tlf. 72203000 eller kurser@teknologisk.dk, så vi kan hjælpe med at imødekomme dit behov.

Har du faglige spørgsmål så kontakt



Charlotte Heimann
+45 72203147
chhn@teknologisk.dk