

Big Data on AWS (BDAWS)

ID AW-BDAWS Preço a solicitar Duração 3 dias

Quem deve assistir

Esse curso destina-se a:

- Indivíduos responsáveis pelo design e implementação de soluções de Big Data, nomeadamente Arquitetos de Soluções
- Cientistas de dados e analistas de dados interessados em aprender sobre soluções de Big Data na AWS

Pré-requisitos

Recomendamos que os participantes deste curso tenham os seguintes pré-requisitos:

- Familiaridade básica com tecnologias de big data, incluindo consultas ao Apache Hadoop, MapReduce, HDFS e SQL / NoSQL
- Os alunos devem concluir o treinamento na Web baseado em Big Data Technology Fundamentals ou ter experiência equivalente
- Conhecimento prático dos principais serviços da AWS e implementação de nuvem pública
- Os alunos devem concluir o curso técnico da **AWS Essentials (AWSE)** ou ter experiência equivalente
- Conhecimento básico de armazenamento de dados, sistemas de banco de dados relacionais e design de banco de dados

Objetivos do curso

Este curso ensina como:

- Adapte soluções da AWS dentro de um ecossistema de big data
- Aproveite o Apache Hadoop no contexto do Amazon EMR
- Identifique os componentes de um cluster do Amazon EMR
- Iniciar e configurar um cluster do Amazon EMR
- Aproveite estruturas de programação comuns disponíveis para o Amazon EMR, incluindo Hive, Pig e Streaming
- Aproveite o Hue para melhorar a facilidade de uso do Amazon EMR
- Use a análise na memória com o Spark e o Spark SQL no Amazon EMR
- Escolha as opções apropriadas de armazenamento de dados da AWS
- Identifique os benefícios do uso do Amazon Kinesis para processamento de big data em tempo quase real

- Definir conceitos de data warehousing e banco de dados de colunas
- Aproveite o Amazon Redshift para armazenar e analisar dados com eficiência
- Compreender e gerenciar custos e segurança para implantações do Amazon EMR e Amazon Redshift
- Identificar opções para ingerir, transferir e compactar dados
- Use o software de visualização para descrever dados e consultas
- Orquestrar fluxos de trabalho de big data usando o AWS Data Pipeline

Conteúdo do curso

Dia 1

- Visão geral de big data
- Ingestão, Transferência e Compressão
- Soluções de armazenamento
- Armazenando e consultando dados no DynamoDB
- Big Data Processing e Amazon Kinesis
- Introdução ao Apache Hadoop e ao Amazon EMR
- Usando o Amazon Elastic MapReduce

Dia 2

- Frameworks de programação do Hadoop
- Processando logs do servidor com o Hive no Amazon EMR
- Processamento de dados de química usando o Hadoop Streaming no Amazon EMR
- Simplificando sua experiência Amazon EMR com o Hue
- Executando Scripts Pig em Matiz no Amazon EMR
- Spark no Amazon EMR
- Criando e consultando interativamente tabelas com o Spark e o Spark SQL no Amazon EMR
- Gerenciando custos do Amazon EMR
- Protegendo suas implantações do Amazon EMR

Dia 3

- Armazéns de dados e datastores colunares
- Amazon Redshift e Big Data
- Otimizando seu ambiente do Amazon Redshift
- Padrões de Design de Big Data
- Visualizando e Orquestrando Big Data
- Usando o Tibco Spotfire para visualizar Big Data

Centros de Entrenamiento Mundial

