Masterclass: Secure Data Engineering in the Azure Cloud

Duration: 3 Days **Course Code: DAC** **Delivery Method: Virtual Learning**

Overview:

In the modern world, every organization relies on data more every day. Businesses tend to deal with large amount of data which could be spread across various sources and to effectively analyze it – it is crucial to consolidate it into one key location. The modern era posed us with many challenges – especially in the Data sector as it is getting harder to connect all the dots together. Microsoft's Azure Cloud is a great solution that makes it easier to manage and analyze the data that we deal with across various sources.
This intense hands-on training if perfect for anyone that has experience with Windows Servers, SQL Servers and deals with or is interested in data management and analysis. During the training, you will gain hands-on experience on how to choose and work with the compute and storage options for data engineering workloads, deep-dive into Azure Data Lake and Databricks. Moreover, you will also go through the key aspects of Azure Synapse Analytics and you will have the chance to become the master of Hybrid Transactional Analytical Processing. By the end of the training you will also gain practical knowledge on data transformation and data movement orchestration using Azure Data Factory and Realtime stream processing!
The training is though and authored by one of the best Experts in the data field – Microsoft's Data Platform MVP Damian Widera.
The course is an intense workshop! During these 3 days we recommend a good cup of coffee – this course is really intense and in order not to miss a thing you MUST stay awake!

Virtual Learning

This interactive training can be taken from any location, your office or home and is delivered by a trainer. This training does not have any delegates in the class with the instructor, since all delegates are virtually connected. Virtual delegates do not travel to this course, Global Knowledge will send you all the information needed before the start of the course and you can test the logins.

Target Audience:

Data engineers, developers, IT professionals, database administrators, security consultants and other people responsible for creating, implementing and maintaining cloud data platform solutions.
To attend this training you should have good hands-on experience in administering Microsoft SQL Server infrastructure. At least 2 years in the field is recommended.

Prerequisites:

We recommend good hands-on experience in administering Microsoft SQL Server infrastructure with least 2 years in the field.

Testing and Certification

Content:

MODULE 1: How to choose and work with the compute and storage options for data engineering workloads?

- Microsoft Azure services for data engineering
- Reference architecture for the data engineering workloads
- Azure Storage account configuration
- Secure data in the Azure Storage

MODULE 2: Explore and design Data Lake

- Why do we need a Data Lake?
- Security design of a Data Lake O Data Lake performance

MODULE 3: Benefits and challenges of the Delta Lake

- Why do we need to create a Delta Lake?
- Usage patterns of the Delta Lake in data engineering workloads
- Challenges of the Delta Lake - transaction handling

MODULE 4: Azure Data bricks

- Create and configure the cluster
- Perform data engineering tasks - work with data frames, perform ETL operations in Scala and Python
- Integrate Databricks with CosmosDB, Azure SQL Database and Azure Storage
- Working with the Delta Lake

MODULE 5: Azure Synapse Analytics - serverless pool

- Prepare for the explorative analytics
- Working with the metadata
- Secure data and manage users

MODULE 6: Azure Synapse Analytics - Apache Spark pool

- Ingest data with Apache Spark
- Data transformation
- Connection to SQL pool

MODULE 7: Azure Synapse Analytics - dedicated pool

- Configure Azure Synapse Analytics with the Dedicated Pool
- Security considerations during the data ingestion
- Security considerations during data processing and retrieval

MODULE 8: Understand the Hybrid Transactional Analytical Processing

- Configure Azure Synapse Link with Azure CosmosDB
- Query Azure CosmosDB with Apacke-Spark and serverless pool

MODULE 9: Data transformation and data movement orchestration using Azure Data Factory

- Create connections to data services
- Create data flows and pipelines
- Perform data transformation and enrichment
- Secure the solution

MODULE 10: Realtime stream processing

- With Azure Stream Analytics
- With Event Hubs and Azure Databricks

Additional Information:

Unique exercises:
All exercises are based on Azure Cloud. with additional Linux and virtual network appliances that covers common scenarios. After the workshop, you will receive PowerPoint slides, tools and lab instructions.

Platform and Technical Requirements:
To participate in the course you need a Stable internet connection. For best learning experience we also need you to have a webcam, headphones and a microphone. Open RDP port 3391 for the connection to the Lab environment is needed as well. We will setup a secure Zoom classroom for every day of the course – we will send you a safe link to join the conference by e-mail.

Further Information:

For More information, or to book your course, please call us on Head Office 01189 123456 / Northern Office 0113 242 5931

info@globalknowledge.co.uk

www.globalknowledge.com/en-gb/

Global Knowledge, Mulberry Business Park, Fishponds Road, Wokingham Berkshire RG41 2GY UK