

# Moderný dátový sklad a riešenia pre analýzu veľkých dát v Microsoft Azure

Kód kurzu: MOC DP-203

Cieľom kurzu je zoznámiť účastníkov s kľúčovými prístupmi a technológiami používanými pri tvorbe moderných dátových skladov a riešení pre analýzu veľkých dát v cloude Microsoft Azure. V rámci kurzu sa zoznámite so službou Azure Synapse Analytics, ktorá v sebe kombinuje relačný dátový sklad postavený na Microsoft SQL Serveri spolu s modernými prístupmi založenými na analýze veľkých dát s využitím Apache Spark. Na kurze sa naučíte vybrať a navrhnúť vhodné úložisko pre vaše dáta, dozviete sa, ako implementovať ETL procesy v cloude a ako následne uložené dáta dotazovať a analyzovať vrátane aplikácie strojového učenia a prezentácie výsledkov v Power BI.

## Pre koho je kurz určený

- Dátoví špecialisti, dátoví architekti a BI architekti, ktorí sa chcú naučiť navrhovať riešenia pre analýzu dát v cloude Microsoft Azure
- Dátoví analytici, ktorí budú pracovať s analytickými riešeniami v Microsoft Azure

## Čo Vás naučíme

- Zorientujete sa v ponuke služieb pre uloženie a spracovanie veľkých objemov dát v Microsoft Azure
- Zoznámite sa so službami Azure Synapse Analytics, Azure Data Lake Storage Gen2, Azure Stream Analytics a Azure Databricks
- Naučíte sa implementovať ETL/ELT procesy v cloude a načítať dáta z rôznych zdrojových systémov do úložiska v cloude
- Dozviete sa, ako ukladať a spravovať dáta v big data úložisku a ako navrhovať a používať relačné dátové sklady v cloude
- Zoznámite sa s možnosťami zberu a spracovania udalostí v reálnom čase

## Požadované vstupné znalosti

- Základné znalosti práce s prostredím Microsoft Azure a správou zdrojov v Microsoft Azure na úrovni kurzu AZ-900
- Základná orientácia v možnostiach uloženia a spracovania dát v Microsoft Azure na úrovni kurzu DP-900

## Metódy výučby

Odborný výklad s praktickými ukážkami, cvičenia na počítačoch

## Študijné materiály

Elektronická verzia originálnych študijných materiálov spoločnosti Microsoft v anglickom jazyku (Digital MOC). Študijné materiály obsahujú pracovnú knižku v elektronickej podobe s pretlačou prezentácií a doplňujúcimi informáciami a detailné návody na samostatné praktické cvičenia. Praktické cvičenia zaberajú zvyčajne aspoň tretinu času stráveného na kurze. Študijné materiály tiež reprezentujú platnú licenciu k účasti na originálnom školení spoločnosti Microsoft. Každý účastník musí mať, podľa licenčných podmienok, v priebehu kurzu k dispozícii vlastné študijné materiály. Keď materiály na daný kurz má už k dispozícii z minulosti, nemusí si ich nutne znovu zakupovať.

## Osnova kurzu

Modul 1: Úvod do spracovania dát v Microsoft Azure

- Pochopíte základné pojmy súvisiace so spracovaním dát v Microsoft Azure
- Zoznámite sa so službou Azure Synapse Analytics a úložiskom Azure Data Lake Storage

Modul 2: Spracovanie dát s využitím Azure Synapse Analytics serverless SQL pools

- Naučíte sa dotazovať a parsovať súbory uložené v data lake úložisku pomocou Azure Synapse Analytics serverless SQL pools
- Dozviete sa, ako používať Azure Synapse Analytics serverless SQL pools pre transformáciu uložených dát
- Zistíte, ako vytvárať ucelené databázy nad data lake úložiskom s využitím Azure Synapse Analytics serverless SQL pools

### GOPAS Praha

Kodaňská 1441/46  
101 00 Praha 10  
Tel.: +420 234 064 900-3  
[info@gopas.cz](mailto:info@gopas.cz)

### GOPAS Brno

Nové sady 996/25  
602 00 Brno  
Tel.: +420 542 422 111  
[info@gopas.cz](mailto:info@gopas.cz)

### GOPAS Bratislava

Dr. Vladimíra Clementisa 10  
Bratislava, 821 02  
Tel.: +421 248 282 701-2  
[info@gopas.sk](mailto:info@gopas.sk)



Copyright © 2020 GOPAS, a.s.,  
All rights reserved

# Moderný dátový sklad a riešenia pre analýzu veľkých dát v Microsoft Azure

## Modul 3: Spracovanie dát pomocou Azure Synapse Apache Spark Pools

- Zoznámite sa s technológiou Apache Spark a jej integráciou do služby Azure Synapse Analytics
- Naučíte sa analyzovať, transformovať a vizualizovať dáta pomocou Apache Spark v prostredí Azure Synapse Analytics
- Dozviete sa, ako používať Delta Lake formát pre tvorbu a uloženie tabuliek

## Modul 4: Práca s dátovými skladmi v Azure Synapse Analytics

- Zoznámite sa s princípmi návrhu tabuliek v relačných dátových skladoch
- Dozviete sa, aké sú špecifiká implementácie tabuliek v Azure Synapse Analytics
- Naučíte sa efektívne importovať dáta do tabuliek zo súborových formátov s využitím EXTERNAL TABLES či príkazu COPY INTO

## Modul 5: Prenosy a transformácie dát pomocou Azure Synapse Analytics Pipelines

- Pochopíte princípy implementácie ETL/ELT riešení pre extrakciu, transformáciu a načítanie dát s využitím Azure Synapse Analytics Pipelines
- Naučíte sa implementovať control flows a data flows v Azure Synapse Analytics Pipelines
- Zoznámite sa s možnosťami použitia Spark Notebooks v rámci Azure Synapse Analytics Pipelines

## Modul 6: Hybridné transakčné a analytické spracovanie dát v Azure Synapse Analytics

- Zoznámite sa s princípmi prístupu Hybrid transactional and analytical processing (HTAP)
- Naučíte sa nasadiť Azure Synapse Link v kombinácii s Azure Cosmos DB
- Vyskúšate si implementovať Azure Synapse Link for SQL

## Modul 7: Spracovanie streamovaných dát v Azure Stream Analytics

- Naučíte sa v reálnom čase spracovávať prúd udalostí v službe Azure Stream Analytics
- Dozviete sa, ako spracované udalosti trvalo uložiť v rámci služby Azure Synapse Analytics
- Zoznámite sa s možnosťami vizualizácie real-time udalostí v službe Power BI

## Modul 8: Riadenie dát naprieč organizáciou

- Pochopíte, prečo je dôležité riešiť centrálnu katalogizáciu a riadenie dát vo vašej organizácii
- Dozviete sa, ako využiť službu Microsoft Purview pre objavenie, katalogizáciu a riadenie dát vo vašej organizácii
- Naučíte sa integrovať Microsoft Purview a Azure Synapse Analytics

## Modul 9: Spracovanie dát v Azure Databricks

- Zoznámite sa s možnosťami služby Azure Databricks
- Naučíte sa pracovať s Apache Spark v prostredí Azure Databricks
- Dozviete sa, ako spúšťať Azure Databricks notebooks v službe Azure Data Factory

### Príprava na certifikačné skúšky

Pri certifikačných skúškach Microsoft platí, že okrem certifikácie MCM nie je účasť na oficiálnom MOC kurze nutnou podmienkou pre zloženie skúšky.

Oficiálne kurzy MOC spoločnosti Microsoft aj naše vlastné kurzy GOC sú vhodnou súčasťou prípravy na certifikačné skúšky spoločnosti Microsoft ako sú MTA, MCP, MCSA, MCSE alebo MCM. Primárnym cieľom kurzu však nie je priamo príprava na certifikačné skúšky, ale zvládnutie teoretických princípov a osvojenie si praktických zručností potrebných na efektívnu prácu s daným produktom. MOC kurzy zvyčajne pokrývajú takmer všetky oblasti požadované na príslušných certifikačných skúškach. Ich prebratie na kurze ale nebýva daný vždy presne rovnaký čas a dôraz, ako vyžaduje certifikačná skúška.

Ako ďalšiu prípravu na certifikačné skúšky je možné využiť napríklad knihy od MS Press (tzv. Self-paced Training Kit) a elektronický self-test softvér.

#### GOPAS Praha

Kodaňská 1441/46  
101 00 Praha 10  
Tel.: +420 234 064 900-3  
[info@gopas.cz](mailto:info@gopas.cz)

#### GOPAS Brno

Nové sady 996/25  
602 00 Brno  
Tel.: +420 542 422 111  
[info@gopas.cz](mailto:info@gopas.cz)

#### GOPAS Bratislava

Dr. Vladimíra Clementisa 10  
Bratislava, 821 02  
Tel.: +421 248 282 701-2  
[info@gopas.sk](mailto:info@gopas.sk)



Copyright © 2020 GOPAS, a.s.,  
All rights reserved