Návrh a implementace moderního datového skladu a ETL/ELT procesů v Microsoft Fabric

Kód kurzu: GOC683

Kurz naučí účastníky navrhovat a implementovat moderní datové sklady v prostředí Microsoft Fabric a budovat nad nimi robustní ETL/ELT procesy s pomocí Data Factory. Kurz kombinuje osvědčenou teorii dimenzionálního modelování (star schema, typy dimenzí a faktových tabulek, pomalu se měnící dimenze) s její praktickou implementací ve Fabric Data Warehouse s důrazem na jazyk T-SQL, který je pro Fabric Warehouse primárním nástrojem práce s daty. Naučíte se navrhnout vrstvenou architekturu skladu, implementovat jeho schéma v jazyce T-SQL a zabezpečit jej na úrovni workspace, databázových objektů i jednotlivých řádků a sloupců. Významný prostor je věnován pokročilým scénářům v Data Factory – parametrizovaným a metadaty řízeným pipelines, ošetření chyb, auditování běhů a zejména návrhovým vzorům pro inkrementální načítání dat z transakčních systémů. Závěr kurzu se věnuje provozu, ladění výkonu, nasazování změn a migraci stávajících řešení z Microsoft SQL Serveru a SQL Server Integration Services, jejichž koncepty kurz průběžně srovnává s jejich protějšky v Microsoft Fabric. Kurz je cíleně zaměřen na technologie kolem jazyka T-SQL; na vybraných scénářích však ukazuje, kde je krátký notebook v Pythonu nejjednodušší cestou k cíli (systematicky se Pythonu a PySpark věnují kurzy GOC681 a GOC685). Vzhledem k rozsahu témat je kurz veden ve svižném tempu a neopakuje obecná témata platformy Microsoft Fabric – vstupní znalosti na úrovni kurzu GOC680 jsou proto důrazně doporučeny.

Co Vás naučíme

- Seznámíte se s architekturou moderních datových skladů a jejich implementací v Microsoft Fabric
- Naučíte se navrhovat dimenzionální modely – faktové tabulky, dimenze včetně pomalu se měnících dimenzí (SCD) a volbu granularity dat
- Vyzkoušíte si implementovat kompletní schéma datového skladu ve Fabric Warehouse v jazyce T-SQL
- Naučíte se budovat parametrizované a metadaty řízené pipelines v Data Factory včetně ošetření chyb a auditování
- Osvojíte si transformace dat v T-SQL včetně načítání dimenzí typu SCD 1 a SCD 2 a faktových tabulek
- Naučíte se navrhovat a implementovat inkrementální načítání dat z transakčních systémů
- Dozvíte se, jak datový sklad zabezpečit, monitorovat a ladit jeho výkon
- Zjistíte, jak nasazovat změny řešení mezi prostředími a jak migrovat stávající řešení z Microsoft SQL Serveru a SSIS

Pro koho je kurz určen

- Vývojářům a architektům datových skladů, kteří chtějí budovat řešení v prostředí Microsoft Fabric.
- Specialistům na datové sklady na Microsoft SQL Serveru a vývojářům ETL procesů v SQL Server Integration Services, kteří chtějí své know-how přenést do prostředí Microsoft Fabric.
- BI vývojářům a datovým inženýrům, kteří preferují práci v jazyce T-SQL.
- Zájemcům o ucelený a prakticky zaměřený úvod do návrhu a implementace datových skladů.

Požadované vstupní znalosti

- Důrazně doporučeny znalosti platformy Microsoft Fabric na úrovni kurzu GOC680 – kurz neopakuje obecná témata platformy (kapacity, licencování, práce s Lakehouse a Apache Spark) a bez těchto vstupních znalostí bude tempo kurzu velmi náročné.
- Znalost jazyka T-SQL na úrovni tvorby dotazů, ideálně na úrovni kurzu GOC278; zkušenost s tvorbou uložených procedur výhodou.
- Zkušenost s datovými sklady nebo SQL Server Integration Services výhodou, nikoliv podmínkou.

Osnova kurzu

1 Moderní datový sklad a seznámení s Fabric Data Warehouse

- Pochopíte roli datového skladu v moderní datové platformě a jeho místo mezi komponentami Microsoft Fabric
- Seznámíte se s klasickými architekturami datových skladů a přístupy Kimball a Inmon
- Dozvíte se, jak mapovat medallion architekturu (bronze/silver/gold) na schémata Warehouse či kombinaci Lakehouse a Warehouse
- Pochopíte posun od ETL k ELT a jeho dopad na návrh procesů ve srovnání s myšlením v SSIS

GOPAS Praha
Na Strži 2097/63
140 00 Praha 4 - Krč
Tel.: +420 226 201 390
info@gopas.cz

GOPAS Brno
Nové sady 996/25
602 00 Brno
Tel.: +420 530 513 590
info@gopas.cz

GOPAS Bratislava
Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 902 903 132
info@gopas.sk

Copyright © 2026 GOPAS, a.s.,
All rights reserved

- Seznámíte se s klíčovými vlastnostmi Fabric Data Warehouse a naučíte se rozhodovat mezi Lakehouse a Warehouse
- Vyzkoušíte si možnosti práce s Warehouse od webového editoru po SQL Server Management Studio
- Zjistíte, v čem se Fabric Warehouse koncepčně liší od datového skladu na Microsoft SQL Serveru a Azure Synapse

2 Dimenzionální modelování datového skladu

- Naučíte se postup návrhu dimenzionálního modelu – identifikaci business procesů, volbu granularity a sestavení bus matrix
- Pochopíte rozdíly mezi star a snowflake schématem a kdy které použít
- Naučíte se navrhovat dimenze včetně náhradních klíčů, dimenze data a pomalu se měnících dimenzí (SCD Type 0 až 3)
- Seznámíte se se speciálními typy dimenzí – role-playing, junk, degenerované a mini-dimenze
- Naučíte se navrhovat faktové tabulky různých typů (transakční, snapshoty, factless) a pochopíte aditivitu metrik
- Dozvíte se, jak na úrovni návrhu ošetřit neznámé členy dimenzí a pozdě přicházející data
- Vyzkoušíte si navrhnout kompletní dimenzionální model pro zadanou obchodní doménu

3 Implementace schématu datového skladu ve Fabric Warehouse

- Dozvíte se, v čem se T-SQL ve Fabric Warehouse liší od Microsoft SQL Serveru – datové typy, collation, omezení DDL
- Naučíte se organizovat vrstvy skladu pomocí databázových schémat a dotazovat data napříč databázemi
- Pochopíte roli NOT ENFORCED constraints a osvojíte si strategie generování náhradních klíčů
- Vyzkoušíte si zero-copy klonování tabulek a time travel dotazy pro vývoj, testování a audit
- Dozvíte se, jak fungují statistiky a fyzické uložení dat ve formátu Delta-parquet
- Vyzkoušíte si implementovat schéma navrženého dimenzionálního modelu v jazyce T-SQL včetně naplnění dimenze data

4 Seznámení s Data Factory v Microsoft Fabric

- Seznámíte se s komponentami Data Factory a jejich rolemi – Pipelines, Dataflows Gen2 a Copy Job
- Pochopíte koncept aktivit pipeline a naučíte se spravovat připojení k datovým zdrojům
- Zjistíte silné a slabé stránky Dataflows Gen2 oproti transformacím v T-SQL a kdy sáhnout po kterém nástroji
- Dozvíte se, jak se připojit k on-premises datovým zdrojům přes data gateway
- Naučíte se plánovat spouštění a sledovat běhy v Monitor hub
- Pochopíte, jak se koncepty SSIS mapují na jejich protějšky v Data Factory

5 Pokročilá práce s Pipelines

- Naučíte se parametrizovat pipelines a používat výrazový jazyk pro dynamický obsah
- Osvojíte si aktivity řízení toku (ForEach, If Condition, Switch, Until) a aktivity Lookup, Get Metadata a Script
- Pochopíte vzor metadaty řízeného načítání dat s řídicí tabulkou zdrojů
- Naučíte se ošetřovat chyby, nastavovat retry politiky a odesílat notifikace přes Teams a Outlook
- Vyzkoušíte si vytvořit metadaty řízenou pipeline pro plnění staging vrstvy včetně ošetření chyb

6 Techniky načítání dat do datového skladu

- Poznáte do hloubky Copy activity – konektory, mapování sloupců a výkonnostní nastavení
- Naučíte se načítat soubory z OneLake v jazyce T-SQL pomocí příkazu COPY INTO a funkce OPENROWSET
- Dozvíte se, jak načítat data dotazem napříč databázemi nad SQL analytics endpointem Lakehouse
- Seznámíte se s Copy Job a zrcadlením (Mirroring) databází jako téměř real-time zdrojem dat
- Zjistíte, ve kterých scénářích je Python notebook nejjednodušší cestou – volání REST API či zpracování komplexních JSON dat
- Vyzkoušíte si načíst data ze souborů v OneLake do staging tabulek a naučíte se volit správnou techniku načítání

7 Transformace dat v T-SQL a načítání dimenzí a faktů

GOPAS Praha
Na Strži 2097/63
140 00 Praha 4 - Krč
Tel.: +420 226 201 390
info@gopas.cz

GOPAS Brno
Nové sady 996/25
602 00 Brno
Tel.: +420 530 513 590
info@gopas.cz

GOPAS Bratislava
Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 902 903 132
info@gopas.sk

Copyright © 2026 GOPAS, a.s.,
All rights reserved

- Naučíte se implementovat ELT transformace pomocí uložených procedur, CTAS, INSERT ... SELECT a příkazu MERGE
- Pochopíte transakce a úroveň izolace ve Fabric Warehouse a naučíte se psát idempotentní procedury
- Vyzkoušíte si implementovat načítání dimenzí typu SCD 1 a SCD 2 včetně detekce změn
- Naučíte se načítat faktové tabulky včetně vyhledání náhradních klíčů a ošetření nenalezených členů dimenzí
- Dozvíte se, jak čistit a validovat data v T-SQL a jak vybudovat auditní rámec načítání

8 Inkrementální načítání dat

- Pochopíte, proč plné načítání nestačí a jaké jsou vzory inkrementálního zpracování
- Naučíte se implementovat načítání s vodoznakem (watermark) s řídicí tabulkou a dynamickými dotazy
- Dozvíte se, jak využít Change Tracking a Change Data Capture na straně SQL Server zdrojů včetně zachycení mazání záznamů
- Zjistíte, jak inkrementálně kopírovat data pomocí Copy Job a jak využít zrcadlenou databázi jako zdroj změn
- Naučíte se v cíli implementovat upsert, ošetřit pozdě přicházející data a zajistit opakovatelnost běhů
- Vyzkoušíte si implementovat kompletní inkrementální načítání transakční tabulky

9 Orchestrace, monitoring a ladění výkonu

- Naučíte se navrhovat master orchestraci s závislostmi mezi dimenzemi a fakty, paralelizací a restartovatelností
- Dozvíte se, jak spouštět zpracování plánovaně i na základě událostí, například příchodu souboru do OneLake
- Naučíte se monitorovat řešení pomocí Monitor hub, pohledů Query Insights a Dynamic Management Views
- Zjistíte, jak ladit výkon datového skladu – statistiky, caching výsledků, V-Order a kompakce dat
- Dozvíte se, jak sledovat a optimalizovat spotřebu kapacity ETL procesů
- Vyzkoušíte si sestavit master pipeline pro kompletní denní načtení skladu a analyzovat její běh

10 Zabezpečení datového skladu a zpřístupnění dat konzumentům

- Pochopíte model zabezpečení Fabric Warehouse – role workspace, sdílení položek a jejich skládání s právy v T-SQL
- Naučíte se nastavovat granulární oprávnění v T-SQL na úrovni databáze, schématu, objektu a sloupce
- Vyzkoušíte si implementovat Row-Level Security a Dynamic Data Masking pro ochranu citlivých údajů
- Dozvíte se, jak zpřístupnit data konzumentům přes SQL analytics endpoint a externí nástroje
- Seznámíte se s doporučeními pro návrh gold vrstvy jako zdroje pro sémantické modely a reporting

11 Nasazování řešení a migrace z Microsoft SQL Serveru a SSIS

- Naučíte se spravovat životní cyklus řešení pomocí Git integrace a nasazovat změny přes deployment pipelines
- Dozvíte se, jak pracovat s Warehouse jako databázovým projektem a řídit změny schématu
- Zjistíte, jak posoudit a provést migraci datového skladu z Microsoft SQL Serveru včetně využití Migration Assistant
- Pochopíte, jak namapovat SSIS procesy na komponenty Fabric a zvolit strategii postupné migrace
- Vyzkoušíte si připojit workspace na Git a nasadit změnu schématu do dalšího prostředí
- Seznámíte se s doporučenými postupy a nejčastějšími chybami při stavbě datového skladu ve Fabric

GOPAS Praha
Na Strži 2097/63
140 00 Praha 4 - Krč
Tel.: +420 226 201 390
info@gopas.cz

GOPAS Brno
Nové sady 996/25
602 00 Brno
Tel.: +420 530 513 590
info@gopas.cz

GOPAS Bratislava
Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 902 903 132
info@gopas.sk

Copyright © 2026 GOPAS, a.s.,
All rights reserved