

MOC 20767 C: Implementing a SQL Data Warehouse

Course Summary

Description

This course describes how to implement a data warehouse platform to support a BI solution. Students will learn how to create a data warehouse with Microsoft SQL Server with Azure SQL Data Warehouse, to implement ETL with SQL Server Integration Services, and to validate and cleanse data with SQL Server Data Quality Services and SQL Server Master Data Services. This course provides students with the knowledge and skills to provision a Microsoft SQL Server database. The course covers SQL Server provision both on-premise and in Azure, and covers installing from new and migrating from an existing install.

Objectives

By the end of this course, students will be able to :

- Describe the key elements of a data warehousing solution
- Describe the main hardware considerations for building a data warehouse
- Implement a logical design for a data warehouse
- Implement a physical design for a data warehouse
- Create columnstore indexes
- Implementing an Azure SQL Data Warehouse
- Describe the key features of SSIS
- Implement a data flow by using SSIS
- Implement control flow by using tasks and precedence constraints
- Create dynamic packages that include variables and parameters
- Debug SSIS packages
- Describe the considerations for implement an ETL solution
- Implement Data Quality Services
- Implement a Master Data Services model
- Describe how you can use custom components to extend SSIS
- Deploy SSIS projects
- Describe BI and common BI scenarios

Topics

- Introduction to Data Warehousing
- Planning Data Warehouse Infrastructure
- Designing and Implementing a Data Warehouse
- Columnstore Indexes
- Implementing an Azure SQL Data Warehouse
- Creating an ETL Solution
- Implementing Control Flow in an SSIS Package
- Debugging and Troubleshooting SSIS Packages
- Implementing a Data Extraction Solution
- Enforcing Data Quality
- Using Master Data Services
- Extending SQL Server Integration Services (SSIS)
- Deploying and Configuring SSIS Packages
- Consuming Data in a Data Warehouse

Audience

The primary audience for this course is database professionals who need to fulfil a Business Intelligence Developer role. They will need to focus on hands-on work creating BI solutions including Data Warehouse implementation, ETL, and data cleansing.

MOC 20767 C: Implementing a SQL Data Warehouse

Course Summary (cont.)

Prerequisite:

In addition to their professional experience, students who attend this training should already have the following technical knowledge:

- Basic knowledge of the Microsoft Windows operating system and its core functionality.
- Working knowledge of relational databases.
- Some experience with database design.

Duration

Five Days

MOC 20767 C: Implementing a SQL Data Warehouse

Course Outline

I. *Introduction to Data Warehousing*

This module describes data warehouse concepts and architecture consideration.

- A. Overview of Data Warehousing
- B. Considerations for a Data Warehouse Solution

Lab : Exploring a Data Warehouse Solution

- Exploring data sources
- Exploring an ETL process
- Exploring a data warehouse

II. *Planning Data Warehouse Infrastructure*

This module describes the main hardware considerations for building a data warehouse.

- A. Considerations for data warehouse infrastructure.
- B. Planning data warehouse hardware.

Lab : Planning Data Warehouse Infrastructure

- Planning data warehouse hardware

III. *Designing and Implementing a Data Warehouse*

This module describes how you go about designing and implementing a schema for a data warehouse.

- A. Designing dimension tables
- B. Designing fact tables
- C. Physical Design for a Data Warehouse

Lab : Implementing a Data Warehouse Schema

- Implementing a star schema
- Implementing a snowflake schema
- Implementing a time dimension table

IV. *Columnstore Indexes*

This module introduces Columnstore Indexes.

- A. Introduction to Columnstore Indexes
- B. Creating Columnstore Indexes
- C. Working with Columnstore Indexes

Lab : Using Columnstore Indexes

- Create a Columnstore index on the FactProductInventory table
- Create a Columnstore index on the FactInternetSales table
- Create a memory optimized Columnstore table

V. *Implementing an Azure SQL Data Warehouse*

This module describes Azure SQL Data Warehouses and how to implement them.

- A. Advantages of Azure SQL Data Warehouse
- B. Implementing an Azure SQL Data Warehouse
- C. Developing an Azure SQL Data Warehouse
- D. Migrating to an Azure SQ Data Warehouse
- E. Copying data with the Azure data factory

Lab : Implementing an Azure SQL Data Warehouse

- Create an Azure SQL data warehouse database
- Migrate to an Azure SQL Data warehouse database
- Copy data with the Azure data factory

VI. *Creating an ETL Solution*

At the end of this module you will be able to implement data flow in a SSIS package.

- A. Introduction to ETL with SSIS
- B. Exploring Source Data
- C. Implementing Data Flow

Lab : Implementing Data Flow in an SSIS Package

- Exploring source data
- Transferring data by using a data row task
- Using transformation components in a data row

VII. *Implementing Control Flow in an SSIS Package*

This module describes implementing control flow in an SSIS package.

- A. Introduction to Control Flow
- B. Creating Dynamic Packages
- C. Using Containers
- D. Managing consistency.

Lab : Implementing Control Flow in an SSIS Package

- Using tasks and precedence in a control flow
- Using variables and parameters
- Using containers

Lab : Using Transactions and Checkpoints

- Using transactions
- Using checkpoints

MOC 20767 C: Implementing a SQL Data Warehouse

Course Outline (cont.)

VIII. *Debugging and Troubleshooting SSIS Packages*

This module describes how to debug and troubleshoot SSIS packages.

- A. Debugging an SSIS Package
- B. Logging SSIS Package EventsHandling Errors in an SSIS Package

Lab : Debugging and Troubleshooting an SSIS Package

- Debugging an SSIS package
- Logging SSIS package execution
- Implementing an event handler
- Handling errors in data flow

IX. *Implementing a Data Extraction Solution*

This module describes how to implement an SSIS solution that supports incremental DW loads and changing data.

- A. Introduction to Incremental ETL
- B. Extracting Modified Data
- C. Loading modified data
- D. Temporal Tables

Lab : Extracting Modified Data

- Using a datetime column to incrementally extract data
- Using change data capture
- Using the CDC control task
- Using change tracking

Lab : Loading a data warehouse

- Loading data from CDC output tables
- Using a lookup transformation to insert or update dimension data
- Implementing a slowly changing dimension
- Using the merge statement

X. *Enforcing Data Quality*

This module describes how to implement data cleansing by using Microsoft Data Quality services.

- A. Introduction to Data Quality
- B. Using Data Quality Services to Cleanse Data
- C. Using Data Quality Services to Match Data

Lab : Cleansing Data

- Creating a DQS knowledge base
- Using a DQS project to cleanse data
- Using DQS in an SSIS package

Lab : De-duplicating Data

- Creating a matching policy
- Using a DS project to match data

XI. *Using Master Data Services*

This module describes how to implement master data services to enforce data integrity at source.

- A. Introduction to Master Data Services
- B. Implementing a Master Data Services Model
- C. Hierarchies and collections
- D. Creating a Master Data Hub

Lab : Implementing Master Data Services

- Creating a master data services model
- Using the master data services add-in for Excel
- Enforcing business rules
- Loading data into a model
- Consuming master data services data

XII. *Extending SQL Server Integration Services (SSIS)*

This module describes how to extend SSIS with custom scripts and components.

- A. Using scripting in SSIS
- B. Using custom components in SSIS

Lab : Using scripts

- Using a script task

XIII. *Deploying and Configuring SSIS Packages*

This module describes how to deploy and configure SSIS packages.

- A. Overview of SSIS Deployment
- B. Deploying SSIS Projects
- C. Planning SSIS Package Execution

Lab : Deploying and Configuring SSIS Packages

- Creating an SSIS catalog
- Deploying an SSIS project
- Creating environments for an SSIS solution
- Running an SSIS package in SQL server management studio
- Scheduling SSIS packages with SQL server agent

MOC 20767 C: Implementing a SQL Data Warehouse

Course Outline (Cont.)

XIV. Consuming Data in a Data Warehouse

This module describes how to debug and troubleshoot SSIS packages.

- A. Introduction to Business Intelligence
- B. An Introduction to Data Analysis
- C. Introduction to reporting
- D. Analyzing Data with Azure SQL Data Warehouse

Lab : Using a data warehouse

- Exploring a reporting services report
- Exploring a PowerPivot workbook
- Exploring a power view report