



Implementing a Data Warehouse with Microsoft SQL Server

Course 20463C: 4 days; Instructor-Led

Introduction

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 2014, implement ETL with SQL Server Integration Services, and validate and cleanse data with SQL Server Data Quality Services and SQL Server Master Data Services.

Note: This course is designed for customers who are interested in learning SQL Server 2012 or SQL Server 2014. It covers the new features in SQL Server 2014, but also the important capabilities across the SQL Server data platform.

At Course Completion

After completing this course, students will be able to:

- Describe data warehouse concepts and architecture considerations.
- Select an appropriate hardware platform for a data warehouse.
- Design and implement a data warehouse.
- Implement Data Flow in an SSIS Package.
- Implement Control Flow in an SSIS Package.
- Debug and Troubleshoot SSIS packages.
- Implement an ETL solution that supports incremental data extraction.
- Implement an ETL solution that supports incremental data loading.
- Implement data cleansing by using Microsoft Data Quality Services.
- Implement Master Data Services to enforce data integrity.
- Extend SSIS with custom scripts and components.
- Deploy and Configure SSIS packages.
- Describe how BI solutions can consume data from the data warehouse.

Audience

This course is intended for database professionals who need to create and support a data warehousing solution. Primary responsibilities include:

- Implementing a data warehouse.
- Developing SSIS packages for data extraction, transformation, and loading.
- Enforcing data integrity by using Master Data Services.
- Cleansing data by using Data Quality Services.

Prerequisites

This course requires that you meet the following prerequisites:

- At least 2 years' experience of working with relational databases, including:
- Designing a normalized database.
- Creating tables and relationships.
- Querying with Transact-SQL.
- Some exposure to basic programming constructs (such as looping and branching).

- An awareness of key business priorities such as revenue, profitability, and financial accounting is desirable.

Course Outline

Module 1: Introduction to Data Warehousing

This module provides an introduction to the key components of a data warehousing solution and the high-level considerations you must take into account when you embark on a data warehousing project.

Lessons

- Overview of Data Warehousing
- Considerations for a Data Warehouse Solution

Lab: Exploring a Data Warehousing Solution

- Exploring Data Sources
- Exploring and ETL Process
- Exploring a Data Warehouse

After completing this module, you will be able to:

- Describe the key elements of a data warehousing solution
- Describe the key considerations for a data warehousing project

Module 2: Planning Data Warehouse Infrastructure

This module discusses considerations for selecting hardware and distributing SQL Server facilities across servers.

Lessons

- Considerations for Data Warehouse Infrastructure
- Planning Data Warehouse Hardware

Lab: Planning Data Warehouse Infrastructure

- Planning Data Warehouse Hardware

After completing this module, you will be able to:

- Describe key considerations for BI infrastructure.
- Plan data warehouse infrastructure.

Module 3: Designing and Implementing a Data Warehouse

This module describes the key considerations for the logical design of a data warehouse, and then discusses best practices for its physical implementation.

Lessons

- Data Warehouse Design Overview
- Designing Dimension Tables
- Designing Fact Tables
- Physical Design for a Data Warehouse

Lab: Implementing a Data Warehouse

- Implement a Star Schema
- Implement a Snowflake Schema
- Implement a Time Dimension

After completing this module, you will be able to:

- Describe a process for designing a dimensional model for a data warehouse
- Design dimension tables for a data warehouse

- Design fact tables for a data warehouse
- Design and implement effective physical data structures for a data warehouse

Module 4: Creating an ETL Solution with SSIS

This module discusses considerations for implementing an ETL process, and then focuses on Microsoft SQL Server Integration Services (SSIS) as a platform for building ETL solutions.

Lessons

Introduction to ETL with SSIS

- Exploring Data Sources
- Implementing Data Flow

Lab: Implementing Data Flow in an SSIS Package

- Exploring Data Sources
- Transferring Data by Using a Data Flow Task
- Using Transformations in a Data Flow

After completing this module, you will be able to:

- Describe the key features of SSIS.
- Explore source data for an ETL solution.
- Implement a data flow by using SSIS

Module 5: Implementing Control Flow in an SSIS Package

This module describes how to implement ETL solutions that combine multiple tasks and workflow logic.

Lessons

- Introduction to Control Flow
- Creating Dynamic Packages
- Using Containers
- 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

After completing this module, you will be able to:

- Implement control flow with tasks and precedence constraints
- Create dynamic packages that include variables and parameters
- Use containers in a package control flow
- Enforce consistency with transactions and checkpoints

Module 6: Debugging and Troubleshooting SSIS Packages

This module describes how you can debug packages to find the cause of errors that occur during execution. It then discusses the logging functionality built into SSIS that you can use to log events for troubleshooting purposes. Finally, the module describes common approaches for handling errors in control flow and data flow.

Lessons

- Debugging an SSIS Package
- Logging SSIS Package Events
- Handling 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 a Data Flow

After completing this module, you will be able to:

- Debug an SSIS package
- Implement logging for an SSIS package
- Handle errors in an SSIS package

Module 7: Implementing a Data Extraction Solution

This module describes the techniques you can use to implement an incremental data warehouse refresh process.

Lessons

- Planning Data Extraction
- Extracting Modified Data

Lab: Extracting Modified Data

- Using a Datetime Column
- Using Change Data Capture
- Using the CDC Control Task
- Using Change Tracking

After completing this module, you will be able to:

- Plan data extraction
- Extract modified data

Module 8: Loading Data into a Data Warehouse

This module describes the techniques you can use to implement data warehouse load process.

Lessons

- Planning Data Loads
- Using SSIS for Incremental Loads
- Using Transact-SQL Loading Techniques

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

After completing this module, you will be able to:

- Describe the considerations for planning data loads
- Use SQL Server Integration Services (SSIS) to load new and modified data into a data warehouse
- Use Transact-SQL techniques to load data into a data warehouse

Module 9: Enforcing Data Quality

This module introduces Microsoft SQL Server Data Quality Services (DQS), and describes how you can use it to cleanse and deduplicate data.

Lessons

- Introduction to Data Quality
- Using Data Quality Services to Cleanse Data
- Using Data Quality Services to Cleanse Data

Lab: Cleansing Data

- Creating a DQS Knowledge Base
- Using a DQS Project to Cleanse Data
- Using DQS in an SSIS Package

After completing this module, you will be able to:

- Describe how Data Quality Services can help you manage data quality
- Use Data Quality Services to cleanse your data
- Use Data Quality Services to match data

Module 10: Master Data Services

Master Data Services provides a way for organizations to standardize data and improve the quality, consistency, and reliability of the data that guides key business decisions. This module introduces Master Data Services and explains the benefits of using it.

Lessons

- Introduction to Master Data Services
- Implementing a Master Data Services Model
- Managing Master Data
- 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

After completing this module, you will be able to:

- Describe key Master Data Services concepts
- Implement a Master Data Services model
- Use Master Data Services tools to manage master data
- Use Master Data Services tools to create a master data hub

Module 11: Extending SQL Server Integration Services

This module describes the techniques you can use to extend SSIS. The module is not designed to be a comprehensive guide to developing custom SSIS solutions, but to provide an awareness of the fundamental steps required to use custom components and scripts in an ETL process that is based on SSIS.

Lessons

- Using Scripts in SSIS
- Using Custom Components in SSIS

Lab: Using Custom Scripts

- Using a Script Task

After completing this module, you will be able to:

- Include custom scripts in an SSIS package
- Describe how custom components can be used to extend SSIS

Module 12: Deploying and Configuring SSIS Packages

In this module, students will learn how to deploy packages and their dependencies to a server, and how to manage and monitor the execution of deployed packages.

Lessons

- Overview of SSIS Deployment
- Deploying SSIS Projects
- Planning SSIS Package Execution

Lab: Deploying and Configuring SSIS Packages

- Creating an SSIS Catalog
- Deploying an SSIS Project
- Running an SSIS Package in SQL Server Management Studio
- Scheduling SSIS Packages with SQL Server Agent

After completing this module, you will be able to:

- Describe considerations for SSIS deployment.
- Deploy SSIS projects.
- Plan SSIS package execution.

Module 13: Consuming Data in a Data Warehouse

This module introduces business intelligence (BI) solutions and describes how you can use a data warehouse as the basis for enterprise and self-service BI.

Lessons

- Introduction to Business Intelligence
- Enterprise Business Intelligence
- Self-Service BI and Big Data

Lab: Using a Data Warehouse

- Exploring an Enterprise BI Solution
- Exploring a Self-Service BI Solution

After completing this module, you will be able to:

- Describe BI and common BI scenarios
- Describe how a data warehouse can be used in enterprise BI scenarios
- Describe how a data warehouse can be used in self-service BI scenarios