Corporate Training Group

Microsoft GOLD CERTIFIED Partner | Learning Solutions Information Worker Solutions

Course Outline

Course 2795:

Designing an ETL Solution Architecture Using Microsoft SQL Server 2005 Integration Services

Course Length: 2 Days

Introduction

The purpose of this 2-day instructor-led course is to teach Business Intelligence (BI) professionals working in enterprise environments to design an extract, transform, and load (ETL) solution that supports their BI solution. Students will learn how to plan an ETL solution, and specifically how to design and implement a SQL Server Integration Services (SSIS) based ETL solution. They will also learn how to monitor, optimize, and deploy an SSIS solution.

The course focuses on the planning and design aspects of an ETL solution and does not teach students how to create SSIS packages or how to use the development tools provided with SQL Server 2005.

Audience:

This course is intended for experienced Business Intelligence (BI) professionals. The target students for this course already have experience of using the SQL Server 2005 tools to implement ETL functionality, but need to develop their understanding of design principles and best practices when planning, implementing, and deploying an ETL solution.

At Course Completion

After completing this course, students will be able to:

- Plan data transfer and staging solutions for an ETL operation.
- Plan an SSIS Solution.
- Design and implement data flows.
- Incorporate logging, error handling, and reliability into a package.
- Optimize an SSIS solution.
- Deploy and operate an SSIS solution.

Prerequisites

Before attending this course, students must:

- Have hands-on experience with database development tasks. For example:
  - Creating Transact-SQL queries
  - Writing and optimizing advanced queries (for example, queries that contain complex joins or subqueries)
  - Creating database objects such as tables, views, and indexes
- Be familiar with SQL Server 2005 features, tools, and technologies. In particular, they must have built an SSIS package.
- Have foundational conceptual understanding of data warehousing, data marts, and Business Intelligence. Students must be well-versed on the subjects of data warehousing, data marts, and BI, and preferably have read at least one book by Ralph Kimball or Bill Inmon.
- Have a conceptual understanding of ETL processes.
- Have foundational understanding of Microsoft Windows security. For example, how groups, delegation of credentials, and impersonation function in a security context.
- Have foundational understanding of Web-based architecture. For example, SSL, SOAP, and IIS-what they are and what their role is.

- Already know how to use:
  - Microsoft Office Visio
  - Microsoft SQL Server Business Intelligence Development Studio
  - Microsoft SQL Server Management Studio
  - Performance Monitor
  - Microsoft SQL Server Profiler

Course Outline

Module 1: Planning for ETL

In this module, Students will learn how to plan the way in which data will be transferred and staged in an ETL solution.

Lessons

- Identifying Data Sources and Destinations
- Evaluating Source Data
- Identifying Staging Requirements

Lab 1: Planning an ETL Solution

- Identifying ETL Requirements
- Examining Source Data

After completing this module, students will be able to:

- Plan data access and load operations.
- Evaluate source data.
- Design the ETL staging environment.

Module 2: Planning an SSIS Solution

In this module, students will learn how to plan an SSIS-based ETL solution. Specifically, they will learn how to plan SSIS packages, how to plan an effective package development process, and how to design the control flow within the packages.

Lessons

- Planning Packages
- Planning Package Development
- Designing Package Control Flow

Lab 2: Implementing SSIS Packages

- Designing an SSIS Solution
- Creating a Package Template
- Implementing SSIS Packages

After completing this module, students will be able to:

- Plan SSIS packages.
- Plan package development.
- Design package control flow.

Module 3: Designing Data Flow

In this module, students will learn how to design data flows that extract, transform, and load data. They will also learn about factors to consider when working with slowly changing dimensions (SCDs) and when implementing custom SCD transformations.

Lessons

- Understanding Data Flow
- Designing Data Flow Operations
- Handling Data Changes

Lab 3: Implementing Data Flow

- Designing Data Flow
- Implementing Data Flow

After completing this module, students will be able to:

- Describe how data flows work in SSIS.
- Design data flow operations
- Design data flows for updated data.

Module 4: Logging, Error Handling, and Reliability

In this module, students will learn how to incorporate logging, error handling, and reliability into your SSIS package designs. The module covers standard logging, custom logging, and log reporting. The module also teaches students how to implement error handling and how to handle bad data. Finally, the module shows students how to implement a transaction strategy, use checkpoints, and handle restarts and rollbacks. Lessons

- Logging ETL Operations
- Handling Errors in SSIS
- Implementing Reliable ETL Processes with SSIS

Lab 4: Implementing Reliable Packages

- Implementing Logging
- Implementing Error Handling
- Implementing Transactions and Checkpoints

After completing this module, students will be able to:

- Design and implement logging in an SSIS package.
- Design and implement error handling in an SSIS package.
- Design and implement reliability in an SSIS package.

Module 5: Optimizing an SSIS Solution

In this module, students will learn how to monitor and optimize SSIS performance. They will also learn how to scale-out SSIS packages Lessons

- Monitoring SSIS Performance
- Optimizing SSIS Packages
- Scaling Out SSIS Packages

Lab 5: Optimizing Packages

- Monitoring SSIS
- Optimizing an SSIS Package

After completing this module, students will be able to:

- Monitor SSIS packages.
- Optimize SSIS packages.
- Design scale-out solutions for SSIS packages.

Module 6: Deploying and Operating an SSIS Solution

This module describes considerations for deploying an SSIS solution into a production environment. It also discusses some of the issues related to operations and management that you must consider when planning an SSIS solution. Lessons

- Deploying SSIS Packages
- Operating an SSIS Solution

Lab 6: Deploying and Managing Packages

- Deploying an SSIS Solution
- Backing Up an SSIS Package

After completing this module, students will be able to:

- Deploy an SSIS solution.
- Operate an SSIS Solution.