



KIMBALL
UNIVERSITY

Kimball University: Course Description

Microsoft Data Warehouse in Depth

Duration: 4 days

Why Attend

With SQL Server 2005 and 2008, Microsoft provides a complete suite of tools for you to build data warehouse and business intelligence systems using software from a single vendor. This course prepares you to deal with the many facets of developing, deploying, operating, and growing your Microsoft data warehouse system. This course applies to both SQL Server 2005 and 2008. Examples and demos are based on SQL Server 2008.

Taught by Joy Mundy and Warren Thornthwaite, co-authors of the best selling Microsoft Data Warehouse Toolkit, this lecture and demo course will provide you a detailed technical introduction to building a Microsoft data warehouse / business intelligence system that meets the needs of your business users. The class offers Microsoft-specific detailed guidance for working through the data warehouse lifecycle, from requirements gathering and design through the ETL system, relational data warehouse, OLAP and data mining applications, to reporting and other BI applications. We'll discuss issues around deploying, operating, and securing the Microsoft data warehouse system. This course covers a lot of material in a short time.

Who should attend

We've designed this course to appeal to all major roles on a data warehouse project on the Microsoft platform. It's for data warehouse team managers, system architects, ETL system architects and developers, data warehouse operational staff, and BI application designers and developers.

About This Course

The focus of the course is architectural—how should you design the components of the system. We expect our attendees to be able to read documentation and follow Microsoft's generally straightforward user interfaces. Our goal is to teach you the hard stuff: not which button to push, but how to design and build a successful Microsoft DW/BI system. The pace of the course permits only demos by the instructor; do not expect hands on tutorials during class time. You should be familiar with the SQL Server product family, including the BI Studio, SQL Manager Studio, SSIS, SSAS, Reporting Services, and the relational database, at least at the tutorial level. You should also have a basic understanding of the principles of dimensional modeling; the class only covers dimensional modeling topics at a high level. These principles are summarized in the first four articles of the Fundamentals series on this web page.

COURSE OUTLINE

Day 1

Introduction to the Business Dimensional Lifecycle

- Roadmap for creating the data warehouse/business intelligence system

Project Planning and Management

- Assess readiness
- Define and plan the project
- Manage the project

Defining Business Requirements

- Gathering business requirements
- Requirements prioritization session
- Tips and tactics for achieving milestones and avoiding pitfalls
- Exercise: Translating requirements into the DW Bus Matrix

Designing the Business Process Dimensional Model

- Basic dimensional modeling concepts
- Conformed dimensions
- Slowly changing dimensions
- Additional concepts in dimensional modeling
 - Hierarchies and snowflaking
 - Degenerate and junk dimensions
 - Many to many dimensions
- The dimensional model in Analysis Services
- The dimensional modeling process and design spreadsheet

Microsoft Data Warehouse/Business Intelligence System Architecture

- Common components of DW/BI system architecture
- Mapping Microsoft components to the general DW/BI architecture
- Process: How to create an architecture plan
- The conformed data warehouse
- Exercise/discussion: Develop and present summary architecture for attendees' systems

System Setup

- System configuration
 - Distributing SQL Server components across servers
 - Memory
 - Storage
- Rough system sizing
- Installing SQL Server
 - Choosing the edition of SQL Server
 - Issues to consider during installation and configuration
 - Planning early for development, test, and production systems

Day 2

SQL Server Relational Data Warehouse, Physical Design

- Column names and data types
- Primary and foreign keys
- Initial index plan
- Fact table partitioning

ETL System Design

- Develop a high-level map
- Exercise: High-level map for MDWT_AdventureWorksDW
- Develop standard strategies for common tasks
- Develop table-level details
- The ETL system specification

ETL System Development Using Integration Services

- Introduction to Integration Services and the BI Studio tools
 - Control Flow elements
 - Data Flow elements
 - Debugging within the BI Studio development environment
- General SSIS design techniques
 - Modularize packages

- Variables, expressions, and configurations
- Save extracted data before transformation
- Template packages
- Populating dimension tables
 - Using the slowly changing dimension wizard
 - Avoiding the slowly changing dimension wizard
- Populating fact tables
 - Basic fact table processing
 - The surrogate key pipeline
 - Advanced issues in fact table processing
- A simple Audit system

Day 3

Business Intelligence Applications

- Basic BI application concepts
- The reporting system design process
 - Standard report template
 - Report specifications
 - Reporting system navigation design
- The BI application development process

Delivering BI Applications with Reporting Services

- Reporting Services overview
- Designing a report
 - Sourcing from the relational database
 - Sourcing from Analysis Services
- Deploying a report
 - Report Manager
 - The BI portal
- Reporting Services metadata
- Report Builder

Adding Business Value with Data Mining

- Data Mining overview
- The SQL Server 2005 and 2008 data mining architecture
- The data mining process
- Extended demo: Creating a data mining model
 - Developing the input data set
 - Selecting algorithms
 - The data mining designer
 - Iterating
- Validating the model
- Using the data mining model in production
- Data mining metadata and maintenance

The Metadata Morass

- Defining metadata
- Managing metadata
- Metadata in SQL Server 2005 and 2008
- A simple business metadata data model

The Analysis Services OLAP Database: Designing Dimensions

- Why Use Analysis Services OLAP?
 - Aggregation design and management
 - Query performance
 - Complex calculation logic

- Getting Ready
 - Partially populate the relational DW
 - Create a data source and data source view
- Designing Dimensions
 - The vocabulary of dimensions
 - The Dimension Wizard
 - The Dimension Editor

The Analysis Services OLAP Database: Designing Cubes

- Designing cubes
 - Analysis Services cube terminology
 - Cube structure
 - Dimension usage
 - Process and browse the cube
 - Calculations, KPIs, and Actions
 - Translations and Perspectives
- Physical design considerations
 - Physical design terminology
 - Storage mode
 - Aggregations
 - Partitions
- Processing cubes

Day 4

Securing the Microsoft Data Warehousing and Business Intelligence System

- Getting ready
- Developing a plan
 - Providing open access
 - Itemizing sensitive data
 - Securing various types of data access
- Securing SQL Server components
 - Relational database
 - Analysis Services
 - Reporting Services

Deployment: The Great Unveiling

- System deployment
 - Pre-deployment testing
 - Testing the deployment process
 - Deploying the relational database
 - Deploying Integration Services packages
 - Deploying Analysis Services databases
 - Deploying Reporting Services reports
- User readiness
 - DW/BI documentation
 - User training
 - User support
 - Desktop readiness and configuration

Operations and Maintenance

- Providing ongoing user support
- Executing SSIS packages in production
- Monitoring the DW/BI system
 - General operating system monitoring
 - Monitoring the relational data warehouse
 - Monitoring Integration Services
 - Analysis Services

-Reporting Services

- Performance tuning
- Backup and recovery

Growing the Data Warehousing and Business Intelligence System

- Technical work: Iterating the Lifecycle
- Organizational work: Marketing and managing expectations
- System interconnections

Real Time Business Intelligence

- Defining real-time BI
- Making the case for (and against) real-time
- Alternative approaches to providing real-time data
 - Executing reports in real-time
 - Loading the DW/BI system in real-time
 - Using Analysis Services with real-time data