767
Implementing a SQL Data Warehouse

Exam number: 70-767

Exam title: Implementing a SQL Data Warehouse

Publish date:

GUID:

Language(s) this exam will be available in:

Audience (IT professionals, Developers, Information workers, etc.)**:**

Technology:

Credit type (example: MCSA)**:**

Exam provider (VUE, Certiport, or both)**:**

Exam Design

This document shows changes to objectives and functional groupings. These changes are effective as of March 2017.

Audience Profile

This exam is intended for extract, transform, and load (ETL) and data warehouse developers who create business intelligence (BI) solutions. Their responsibilities include data cleansing as well as ETL and data warehouse implementation.

Skills measured

Design, ~~and~~ implement, and maintain a data warehouse (35-40%)

- Design and implement dimension tables
  - Design shared and conformed dimensions; determine support requirements for slowly changing dimensions; determine attributes; design hierarchies; determine star or snowflake schema requirements; determine the granularity of relationship by using fact tables; determine auditing or lineage requirements; determine keys and key relationships for a data warehouse; implement dimensions; implement data lineage of a dimension table
- Design and implement fact tables
  - Identify measures; identify dimension table relationships; create composite keys; design a data warehouse that supports many-to-many relationships; implement semi-additive measures; implement non-additive measures
- Design and implement indexes for a data warehouse workload
  - Design an indexing solution; select appropriate indexes; implement clustered, non-clustered, filtered, and columnstore indexes
- Design storage for a data warehouse
  - Design an appropriate storage solution, including hardware, disk, and file layout
- Design and implement partitioned tables and views
  - Design a partition structure to support a data warehouse; implement sliding windows; implement partition elimination; design a partition structure that supports the quick loading and scale-out of data

~~Manage and maintain a SQL Data Warehouse~~

- ~~Manage queries by using labels; manage statistics; manage partition distribution; scale out the data warehouse; grow, shrink, and pause the data warehouse~~

Extract, transform, and load data (40-45%)

- Design and implement an extract, transform, and load (ETL) control flow by using a SQL Server Integration Services (SSIS) package
  - Design and implement ETL control flow elements, including containers, tasks, and precedence constraints; create variables and parameters; create checkpoints, sequence and loop containers, and variables in SSIS; implement data profiling, parallelism, transactions, logging, and security
- Design and implement an ETL data flow by using an SSIS package
  - Implement slowly changing dimension, fuzzy grouping, fuzzy lookup, audit, blocking, non-blocking, and term lookup transformations; map columns; determine the appropriate transform object for a given task; determine appropriate scenarios for Transact-SQL joins vs. SSIS lookup; design table loading by using bulk loading or standard loading; remove extra rows or bad rows by using de-duplication
- Implement an ETL solution that supports incremental data extraction
  - Design fact table patterns; enable Change Data Capture; create a SQL MERGE statement

Implement an ETL solution that supports incremental data loading

Design a control flow to load change data; load data by using Transact-SQL Change Data Capture functions; load data by using Change Data Capture in SSIS

Debug SSIS packages

Fix performance, connectivity, execution, and failed logic issues by using the debugger; enable logging for package execution; implement error handling for data types; implement breakpoints; add data viewers; profile data with different tools; perform batch clean-up

Deploy and configure SSIS packages and projects

Create an SSIS catalog; deploy packages by using the deployment utility, SQL Server, and file systems; run and customize packages by using DTUTIL

~~**Integrate solutions with cloud data and big data**~~

~~Integrate external data sources with SQL Server by using Polybase~~

~~Integrate with Hadoop; integrate with text files stored in the Azure Blob service; manage external tables; access data in Hadoop databases with Transact-SQL; access data in the Azure Blob service by using Transact-SQL; import data from Hadoop or blobs as regular SQL Server tables; export data to Hadoop or the Azure Blob service~~

~~Extract, transform, and load data from SQL Data Warehouse by using Polybase~~

~~Integrate Azure SQL Data Warehouse with on-premises data warehouses; implement bi-directional data synchronization between Azure and on-premises systems; load data into SQL Data Warehouse from Polybase; design an incremental load strategy by using Polybase and the Azure Blob service~~

~~Design and implement an Azure SQL Data Warehouse~~

~~Create a new Azure SQL Data Warehouse database by using the Azure portal; create an Azure SQL Data Warehouse database by using Transact-SQL; select the appropriate method to load data into Azure SQL Data Warehouse~~

~~Manage and maintain a SQL Data Warehouse~~

~~Manage queries by using labels; manage statistics; manage partition distribution; scale out the data warehouse; grow, shrink, and pause the data warehouse~~

Build data quality solutions (15-20%)

Create a knowledge base

Create a Data Quality Services (DQS) knowledge base; determine appropriate use cases for a DQS knowledge base; perform knowledge discovery; perform domain management

Maintain data quality by using DQS

Add matching knowledge to a knowledge base; prepare a DQS for data deduplication; create a matching policy; clean data by using DQS knowledge; clean data by using the SSIS DQS task; install DQS

Implement a Master Data Services (MDS) model

Install MDS; implement MDS; create models, entities, hierarchies, collections, and attributes; define security roles; import and export data; create and edit a subscription; implement entities, attributes, hierarchies, and business rules

Manage data by using MDS

Use MDS tools; use the Master Data Services Configuration Manager; create a Master Data Manager database and web application; deploy a sample model using MDSModelDeploy.exe; use the Master Data Services web application; use the Master Data Services Add-in for Excel; create a Master Data Management hub; stage and load data; create subscription views