

29.07.2015

MS-10777 Внедрение хранилищ данных в Microsoft SQL Server 2012

Продолжительность курса: 40 академических часов

Аттестация: удостоверение о повышении квалификации
установленного образца (или сертификат ТПУ)

Обучение: вечернее

Краткая аннотация курса

В ходе занятий по этому курсу Вы овладеете реализацией BI-платформы для поддержки аналитики для информационных работников. Вы узнаете, как создать хранилище данных в SQL Server 2012, внедрить решение ETL со службами SQL Server Integration Services, проверить и очистить данные с помощью SQL Server Data Quality Services и SQL Server Master Data Services. Успешное окончание курса поможет Вам подготовиться к экзамену *Microsoft 70-463*.

Аудитория

Основной аудиторией данного курса являются специалисты баз данных, которые выполняют роль разработчика бизнес-аналитики. Их обязанности фокусируются на практической работе по созданию BI-решений, включая внедрение хранилищ данных, ETL, и очистке данных. Основные обязанности:

- Реализация хранилища данных
- Разработка пакетов SSIS для извлечения данных и загрузки / передачи / преобразования
- Обеспечение целостности данных, с использованием Master Data Services
- Очистка данных с помощью служб Data Quality

Предварительная подготовка

Как минимум 2 года работы с реляционными базами данных, включая:

- Написание запросов с использованием Transact-SQL
- Навыки разработки реляционных баз данных
- Знания проектирования и нормализации баз данных

Чему вы научитесь

По окончании курса Вы будете уметь:

- Представлять концепцию и архитектуру хранилища данных.
- Выбирать подходящую аппаратную платформу для хранилищ данных.
- Планировать и создавать хранилища данных.
- Реализовывать потоки данных в SSIS пакетах.
- Производить отладку и устранение неполадок SSIS пакетов.
- Реализовывать SSIS решение, которое поддерживает инкрементную загрузку хранилища и изменение данных.
- Осуществлять интеграцию данных из облака в хранилище данных.
- Реализовывать очистку данных с помощью службы Microsoft Data Quality.
- Реализовывать Master Data Services для обеспечения целостности данных.
- Расширять пакеты SSIS пользовательскими скриптами и компонентами.
- Развертывать и настраивать пакеты SSIS.
- Представлять, как информационные работники могут использовать данные из хранилища данных.

МЫ ДЕЛАЕМ ЗНАНИЯ БЛИЖЕ!



Институт
кибернетики

ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Учебный центр ТПУ&Softline
634034, г.Томск ул.Советская 84/3 оф.209
Тел. (3822) 60-60-03
e-mail: ituniversity@tpu.ru
<http://it-university.tpu.ru>

Программа курса

Модуль 1. Введение в хранилища данных (4 часа)

- Описание концепции и архитектуры хранилища данных
- Решение для хранилища данных
- **Лабораторная работа: Изучение решения для хранилища данных**
 - Изучение источников данных
 - Изучение процессов ETL
 - Изучение хранилищ данных

Модуль 2. Аппаратное обеспечение для хранилища данных (2 часа)

- Сложности построения хранилищ данных
- Архитектура хранилищ данных
- Оборудование для хранилищ данных

Модуль 3. Проектирование и реализация хранилищ данных (2 часа)

- Логический дизайн хранилищ данных
- Физический дизайн хранилищ данных
- **Лабораторная работа: Реализация схемы хранилища данных**
 - Реализация схемы Звезда
 - Реализация схемы Снежинка
 - Реализация таблицы времени

Модуль 4. Разработка и внедрение схемы для хранилища данных (4 часа)

- Введение в ETL с SSIS
- Изучение источников данных
- Реализация потока данных
- **Лабораторная работа: Реализация потока данных в пакете служб SSIS**
 - Изучение источников данных
 - Передача данных с помощью задачи потока данных
 - Использование преобразования в потоке данных

Модуль 5. Реализация потока управления в пакете SSIS (2 часа)

- Введение в поток управления
- Создание динамических пакетов
- Использование контейнеров
- Управление согласованностью
- **Лабораторная работа: Реализация потока управления в пакете SSIS**
 - Использование задач и очередность в потоке управления
 - Использование переменных и параметров
 - Использование контейнеров
- **Лабораторная работа: Использование транзакций и контрольных точек**
 - Использование транзакций
 - Использование контрольных точек

Модуль 6. Отладка и устранение неполадок служб SSIS (4 часа)

- Отладка пакетов служб SSIS
- Ведение журнала событий пакетов служб SSIS
- Обработка ошибок в пакете служб SSIS
- **Лабораторная работа: Отладка и устранение неполадок пакетов служб SSIS**
 - Отладка пакетов служб SSIS

МЫ ДЕЛАЕМ ЗНАНИЯ БЛИЖЕ!

- Ведение журнала выполнения пакетов служб SSIS
- Реализация обработчиков событий
- Обработка ошибок в потоке данных

Модуль 7. Реализация инкрементных ETL процессов (4 часа)

- Введение в инкрементные процессы ETL
- Извлечение изменения данных
- Загрузка измененных данных
- **Лабораторная работа: извлечение измененных данных**
 - Использование столбца типа DateTime для поэтапного извлечения данных
 - Использование отслеживания изменений
- **Лабораторная работа: загрузка инкрементных изменений**
 - Использование задачи Lookup task, чтобы вставить данные измерений
 - Использование задачи Lookup task, чтобы вставить или обновить данные измерений
 - Реализация медленно изменяющихся измерений
 - Использование инструкции MERGE для загрузки фактических данных

Модуль 8. Включение данных из облака в хранилище данных (2 часа)

- Обзор «облачных» источников данных
- SQL Server Azure
- Рынок данных Azure
- **Лабораторная работа: Включение данных из облака в хранилище данных**
 - Извлечение данных из SQL Azure
 - Получение данных из рынка данных Azure

Модуль 9. Обеспечение качества данных (4 часа)

- Введение в очистку данных
- Использование Data Quality Services для очистки данных
- Использование Data Quality Services для соответствия данных
- **Лабораторная работа: Очистка данных**
 - Создание базы знаний DQS
 - Очистка данных с помощью проекта DQS
 - Использование DQS в SSIS пакете
- **Лабораторная работа: удаление повторяющихся данных**
 - Создание политики соответствия
 - Использование Data Quality Services для соответствия данных

Модуль 10. Использование Master Data Services (2 часа)

- Основные понятия Master Data Services
- Реализация модели Master Data Services
- Использование надстройки Excel с Master Data Services
- **Лабораторная работа: Реализация Master Data Services**
 - Создание базовой модели MDS
 - Редактирование моделей MDS с помощью Excel
 - Загрузка данных в MDS
 - Обеспечение бизнес-правил
 - Получение Master Data Services данных

Модуль 11. Расширение SSIS (2 часа)

- Использование пользовательских компонентов в SSIS
- Использование сценариев в SSIS

МЫ ДЕЛАЕМ ЗНАНИЯ БЛИЖЕ!

- **Лабораторная работа: использование скриптов и пользовательских компонентов**
 - Использование пользовательских компонентов
 - Использование Задачи сценарий (Script Task)

Модуль 12. Развертывание и настройка пакетов служб SSIS (4 часа)

- Обзор развертывания
- Развертывание проектов служб SSIS
- Планирование выполнения пакетов служб SSIS
- **Лабораторная работа: Развертывание и настройка пакетов служб SSIS**
 - Создание каталога SSIS
 - Развертывание SSIS проекта
 - Создание среды для SSIS решений
 - Запуск пакетов служб SSIS в SQL Server Management Studio
 - Запуск пакетов SSIS по расписанию агентом SQL Server

Модуль 13. Использование данных в хранилище (4 часа)

- Использование Excel для анализа данных в хранилище
- Введение в PowerPivot
- Введение в Crescent
- **Лабораторная работа: использование хранилищ данных**
 - Использование PowerPivot для создания запроса к хранилищу данных
 - Визуализация данных с помощью Crescent

МЫ ДЕЛАЕМ ЗНАНИЯ БЛИЖЕ!