

УДК 004.658.6

С.Г.Ульянов (4 курс, каф ИУС), А.А.Зотов (асп., каф ИУС),
С.А.Пархоменко (главный системный аналитик, ООО “Деловые консультации, СПб”)

РАЗРАБОТКА ШАБЛОНОВ РЕАЛИЗАЦИИ СОМ-ОБЪЕКТОВ MS DTS

Конец XX – начало XXI веков прошло под знаком развития информационных технологий, на каждом этапе которого появлялись новые способы хранения и представления информации. Сначала персональная СУБД сменила текстовые форматы, вскоре ее место занял сервер баз данных на несколько клиентов, но и на смену ему пришел полнофункциональный сервер БД. Сосуществование старых и новых форматов стало нормальным явлением. Причиной этому было не столько нежелание переходить на новые системы, сколько невозможность быстрого и безболезненного переноса данных.

Разработчики СУБД всегда уделяли значительное внимание процедурам доступа к данным, но только появление универсальных средств в составе популярных СУБД (MSSQL, Oracle, Sybase) позволило работать с данными любой природы, независимо от их формата и места хранения. Пакет Data Transformation Services (DTS), впервые появившийся в Microsoft SQL Server 7.0, является, по сути, универсальным средством как экспорта/импорта данных в/из SQL Server, так и трансформации данных, позволяя загружать или сохранять данные в любом из известных форматов, по необходимости изменяя их структуру.

Основой DTS является пакет (package) - полнофункциональный инструмент по трансформации данных. Основными компонентами пакета служат соединение (connection), задача (task) и шаг (step). Соединения указывают источники данных, задачи позволяют работать с этими данными, а шаги определяют последовательность выполнения задач, позволяя организовать ветвления при удачном или неудачном выполнении предыдущей задачи.

Соединение, как компонент пакета, представляет собой описание источника или назначения, использующихся при переносе данных или других задач. Соединение с источником осуществляется через OLE DB, следовательно, DTS может работать с источниками данных любого типа и формата, в том числе нереляционными, для которых существует OLE DB-провайдер.

DTS включает целый набор задач, позволяющий выполнять разнообразные действия, включающие выполнение ActiveX скрипта, SQL-запроса, запуск других программ, посылку e-mail сообщений и многое другое.

Наиболее сложной задачей, которую поддерживает DTS, является создание хранилищ данных. Множество баз данных, хранящихся в различных системах и представленные в самых разнообразных форматах, закачиваются в хранилище для более удобной работы. Хранилище данных СКАУТ OLAP представляет собой аналитическую подсистему, на котором с помощью MS SQL Analysis Services настроены OLAP кубы, которые позволяют отслеживать тенденции изменения данных, производить соответствующие выборки. При загрузке данных в хранилище из различных источников встала проблема несовместимости формата метаданных библиотеки СКАУТа и формата исходных данных. Как решение этой проблемы было выбрано создание собственной задачи, подключаемой в набор задач DTS Microsoft SQL Server и реализующей необходимые модели трансформации.

Для того чтобы задача могла быть подключена к DTS, она была оформлена в виде СОМ-объекта и наследована от интерфейса IDTSCustomTask – интерфейса, от которого, кстати, произведены все типы задач. Была реализована коллекция свойств, метод Execute,

т.е. непосредственно функциональное ядро задачи, и т.п. Для того чтобы в редакторе DTS Designer было удобнее работать с задачей, также использовался IDTSCustomTaskUI – интерфейс, поддерживающий графическое представление задач

Задача позволила перекачивать данные двумя соединениями, т.е. переносить множество записей соединения-источника во множество записей соединения-приемника – просто копируя их или выполняя определенные трансформации, т.к. поля таблицы-приемника могут быть абсолютно неидентичны полям источника. Универсализм задачи обеспечивается также возможностью написания SQL-запросов, позволяющих выполнять разнообразные операции над данными таблицы-источника и/или таблицы-приемника. Источник и приемник данных выбираются внутри самой задачи, что позволяет отказаться от использования соединений как таковых. Таким образом, созданная задача DTSPumpTask позволила загружать данные в хранилище, при необходимости трансформируя их. Подключаемая в набор задач, она позволила расширить возможности DTS и облегчить процесс заполнения хранилища.