

К ВОПРОСУ ПОДГОТОВКИ СПЕЦИАЛИСТОВ В ОБЛАСТИ РАЗРАБОТКИ БАЗ ДАННЫХ

Медведева О.А.

Кыргызско-Российский Славянский университет. Кыргызстан. olgamedv@rambler.ru

В настоящее время системы баз данных являются одним из лидеров на рынке программного обеспечения. Это обусловлено огромной потребностью в данной технологии коммерческих и правительственных структур.

В последнее десятилетие бурно развивается направление распределенной обработки данных, и на сегодняшний день это одно из наиболее динамично растущих направлений. Распределенная обработка – это обработка с использованием множества вычислительных ресурсов. Под распределенной базой данных понимают хранение таблиц (или частей таблиц) в различных узлах вычислительной сети. Фундаментальным принципом распределенных баз данных, по определению К. Дейта, является то, что для конечного пользователя распределенная система должна выглядеть так же, как нераспределенная [5].

В распределенных системах важной задачей становится задача тиражирования (репликации) данных между несколькими системами управления базами данных (СУБД). В докладе об изменении профиля корпоративных данных, который был озвучен на саммите АРАС, посвященном хранилищам данных и состоявшемся в 2007г. в Хошимине, Рик Вилларс (Rick Villars, Head of Investment Technical Services HSBC) показал, что объем тиражируемых данных увеличивается ежегодно на 43.9% и к 2009 г. сравняется с объемом традиционных данных, размещаемых в хранилищах [3]. Все это требует высокой квалификации специалистов как по созданию программного обеспечения (ПО), так и по его обслуживанию.

Важной характеристикой современного программного обеспечения является его сложность. Сложность коммерческого ПО (за исключением высокоспециализированных программных пакетов) заключается в разнообразии взаимосвязей между многочисленными объектами и категориями объектов (классами) в предметной области задач производства. В связи с этим вопрос подготовки специалистов в области разработки баз данных и коммерческого программного обеспечения с их использованием является важной и актуальной задачей.

Курс «Базы данных» входит в федеральный компонент государственного образовательного стандарта ГОСТ 2000 [4] и является одной из важных дисциплин по подготовке специалистов по специальности «Программное обеспечение вычислительной техники и автоматизированных систем».

В Кыргызско-Российском Славянском университете курс по базам данных разбит на две дисциплины: «Базы данных» и «Распределенные базы данных», изучаемые студентами в 6 и 7 семестрах соответственно. В настоящее время в качестве инструментальных средств при обучении этих дисциплин используются соответственно СУБД Microsoft Access и Microsoft SQL Server.

Использование в преподавании курса баз данных СУБД Microsoft Access имеет ряд недостатков. Данная СУБД является «вещью в себе» – структура БД и связи между данными глубоко скрыты внутри предлагаемого пользователю интерфейса, за встроенными инструментами разработки. Диалект языка структурированных запросов (Structured Query Language, SQL) данной СУБД является урезанным и не имеет ряд конструкций. Язык SQL на сегодня является стандартом де-факто при работе с реляционными базами данных. Достоинством этой СУБД является простота её освоения и возможность быстрого создания интерфейса пользователя и отчетов. Однако это, в свою очередь, является и недостатком, поскольку не дает четкого представления о разделении функций между клиентом и сервером, об отделении бизнес-логики от интерфейса пользователя.

В качестве альтернативы СУБД MS Access при обучении студентов базам данных некоторые авторы предлагают использовать СУБД Mysql [6, 9]. В этих работах отмечается популярность данной СУБД в программной индустрии, а также ее кросс-платформенность и открытость исходных кодов. Однако, такие серверные объекты баз данных как триггеры и хранимые процедуры появились лишь в последних версиях данной СУБД, и не полностью соответствуют стандартам языка SQL[10]. Кроме того, в Mysql написание распределенных запросов находится в зачаточном состоянии: начиная с версии 5.0 в запросах появилась возможность объединять таблицы с разных серверов Mysql (другие СУБД участвовать не могут), при этом отсутствует поддержка распределенных транзакций.

В настоящем докладе предлагается построить курс по базам данных полностью на базе СУБД MS SQL Server [11]. Данная СУБД является достаточно популярной на рынке разработки программного обеспечения (входит в тройку «big tree»: Oracle, MS Sql Server и DB2 компании IBM). В данной СУБД хорошо развит диалект языка SQL – Transact-SQL, что позволит научить студентов писать запросы различной сложности на этом языке. Кроме того, по сравнению с Mysql, данная СУБД позволяет успешно работать с такими объектами баз данных как хранимые процедуры и триггеры. Также одним из достоинств MS SQL Server является его хорошая совместимость со средой разработки Microsoft Visual Studio .NET, которую предлагается использовать в качестве базовой среды для разработки клиентских приложений.

MS SQL Server позволяет создавать распределенные запросы для доступа к данным из нескольких разнородных источников данных и обрабатывать распределенные транзакции. Каждый распределенный запрос может ссылаться на несколько связанных серверов и выполнять операции обновления или считывания отдельно на каждом из связанных серверов. В качестве источников данных для распределенного запроса могут выступать данные, хранящиеся не только в экземплярах SQL Server, но и в различных реляционных и нереляционных источниках данных, доступ к которым осуществляется с использованием поставщика OLE DB. На рис. 1 показаны соединения между клиентским компьютером, экземпляром SQL Server и поставщиком OLE DB.

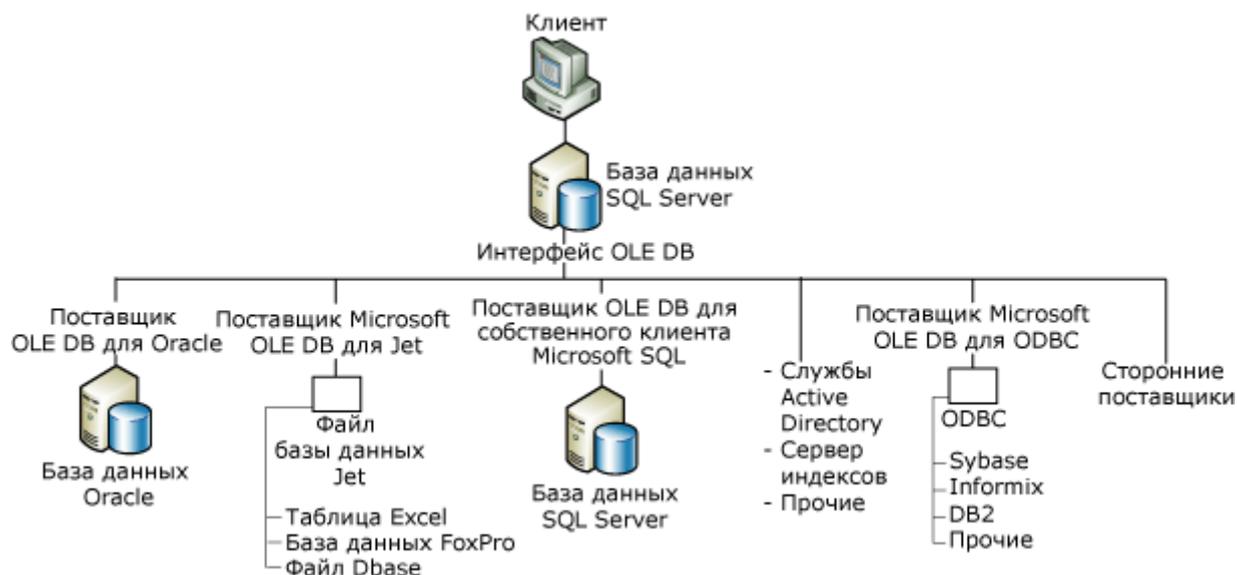


Рис. 1. Конфигурация связанных(linked) серверов в MS SQL Server 2008

Основной целью предлагаемого автором курса по базам данных является приобретение студентами практических навыков проектирования и реализации распределенных информационных систем типа клиент-сервер. Курс дисциплины рассчитан на два семестра. В теоретической части курса предлагается выделить следующие разделы:

1. Введение в базы данных (основные понятия и определения, функции СУБД, классификация моделей данных).
2. Основные положения теории реляционных баз данных (базовые понятия, реляционная алгебра, функциональные зависимости и нормализация отношений).
3. Модели архитектуры «клиент-сервер».
4. Использование MS SQL Server в качестве сервера баз данных и языка SQL при разработке программных приложений (создание таблиц и связей, представлений, реализация бизнес-логики приложения в виде хранимых процедур и функций пользователя, использование триггеров).
5. Базовые объектные архитектуры распределенных систем. Технологии .NET, (D)COM+, CORBA, EJB. Реализация клиентского GUI-приложения на языках высокого уровня. Использование технологий объектно-реляционного отображения (ORM).
6. Распределенные базы данных, транзакции и технологии тиражирования баз данных.
7. Способы защиты данных в современных СУБД.

Первые четыре из перечисленных разделов предлагается рассматривать в первом семестре курса дисциплины «Базы данных», последние три – во втором семестре курса дисциплины. При рассмотрении теоретических основ моделей данных предлагается уделить основное внимание даталогическому подходу [2], поскольку в КРСУ инфологический подход детально рассматривается в курсе «Технология разработки программного обеспечения».

При разработке реальных программных приложений важно уметь отделять бизнес-логику работы приложения от пользовательского интерфейса и уровня хранения данных для повышения степени повторного использования кода и его читабельности. В связи с этим в курс предлагается включить обзорный материал по использованию некоторых шаблонов объектно-ориентированного проектирования, в частности шаблон Model View Presenter.

Шаблоны проектирования (design pattern) – это многократно применяемая архитектурная конструкция, предоставляющая решение общей проблемы проектирования в рамках конкретного контекста и описывающая значимость этого решения [1]. Главная польза каждого отдельного шаблона состоит в том, что он описывает решение целого класса абстрактных проблем.

Курс лабораторных работ предлагается разделить на три большие части:

1. Работа с серверными объектами баз данных (создание БД, таблиц и отношений, написание запросов к БД на языке SQL, разработка представлений, хранимых процедур, функций пользователя, использование триггеров и курсоров).
2. Проектирование структуры базы данных.
3. Принципы и технологии разработки интерактивных N-уровневых клиентских приложений.

Лабораторные работы первого семестра обучения дисциплине (первые два раздела) посвящены практическим аспектам построения серверной части приложения, что позволит закрепить теоретические знания по проектированию структуры баз данных и использованию языка SQL (создание таблиц, связей между ними, представлений, написание различных запросов, а также применение процедурного языка Transact-SQL).

Немаловажной частью курса баз данных является приобретение практических навыков проектирования структуры базы данных, нормализации и денормализации таблиц.

Лабораторный практикум второго семестра по базам данных предполагает обучение студентов теории проектирования и реализации клиентских приложений распределенных информационных систем, а также реализации бизнес-логики на стороне сервера баз данных в виде хранимых процедур и функций пользователя, написания распределенных запросов.

Студенты приобретают навык работы со средами разработки GUI-приложений по части работы с базами данных.

В качестве среды для выполнения лабораторных работ предлагается использовать Microsoft Visual Studio .NET [8,12], поскольку платформа Microsoft .NET Framework является очень популярной и востребованной в области коммерческой разработки программного обеспечения. Она предоставляет множество классов для создания приложений, работающих с локальными и удаленными реляционными базами данных (технология ADO.NET).

ADO .NET (ActiveX Data Objects .NET) является набором классов, реализующих программные интерфейсы для облегчения подключения к базам данных из приложения независимо от особенностей реализации конкретной системы управления базами данных и от структуры самой базы данных, а также независимо от места расположения этой самой базы — в частности, в распределенной среде (клиент-серверное приложение) на стороне сервера [7].

В рамках выполнения лабораторных работ студентам предлагается познакомиться с принципами работы в соединенном (connected) и отсоединенном (disconnected) режимах технологии ADO.NET, а также познакомиться с технологией объектно-реляционного отображения (ORM).

В процессе обучения в вузе студенты должны получить представление обо всех стадиях создания ПО: от определения задачи и сбора требований до внедрения и сопровождения. При выполнении лабораторных работ студентам необходимо объяснять, что программирование – лишь малая часть процесса создания программного обеспечения, занимающая небольшое количество времени. На этапе разработки архитектуры системы (и структуры БД в частности) могут быть использованы CASE-средства – программные пакеты, помогающие при проектировании программного обеспечения: Rational Rose, Power Designer, Visual Paradigm. (Изучение CASE-средств выходит за рамки курса по базам данных и детально рассматривается в дисциплине «Технология разработки программного обеспечения»). Кроме того, немаловажным этапом создания программного обеспечения является его тестирование. На данный момент существуют различные пакеты и средства, позволяющие автоматизировать процесс тестирования. При выполнении курсовой работы студенты могут применять эти пакеты и средства, однако, детальное их изучение – предмет внимания дисциплины «Тестирование и оценка качества программного обеспечения».

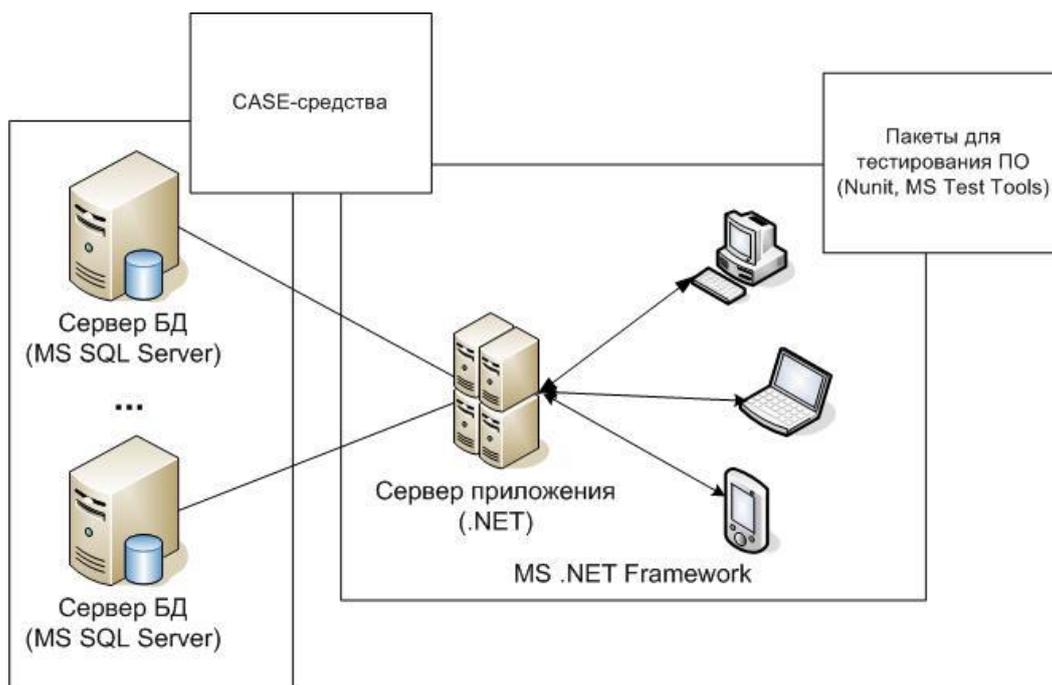


Рис. 2. Структура 3-х уровневой распределенной системы и средства ее реализации

Для закрепления практических навыков студентам предлагается выполнить курсовую работу в виде проекта по разработке трехуровневой распределенной информационной системы (от анализа предметной области, разработки архитектуры ПО и построения структуры базы данных до практической реализации системы). На рис. 2 показана структура 3-х уровневое клиент-серверного приложения, а также инструментарий, используемый при ее реализации.

Следует также отметить, что курс по базам данных необходимо рассматривать во взаимосвязи с другими дисциплинами. В частности, для успешного обучения базам данных студент должен обладать знаниями из дисциплин «Объектно-ориентированное программирование», «CASE-технологии», «Сети ЭВМ» и др. В свою очередь, знания, полученные студентами в курсе по базам данных, могут быть в дальнейшем использованы при изучении дисциплины «Тестирование и оценка качества программного обеспечения».

Таким образом, в докладе излагается новая концепция преподавания курса по базам данных, в результате изучения которого студенты будут иметь представление о современном состоянии вопроса и быть востребованными специалистами на рынке разработки коммерческого программного обеспечения.

Литература

1. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений на C++. – М.: Бинوم, 1998.
2. Гаврилец Е.З., Медведева О.А. Теория баз данных: учебно-методическое пособие. Бишкек: Изд-во КРСУ, 2010.
3. Гладченко А.Ю. Основы репликации SQL Server 2008. – <http://msmvps.com/blogs/gladchenko/archive/2009/01/15/1662288.aspx>
4. Государственный образовательный стандарт высшего профессионального образования. Направление 654600 – Информатика и вычислительная техника. – М: Министерство образования РФ, 2000.
5. Дейт К. Дж. Введение в системы баз данных. – М., 2001.
6. Карпукова А.Л. Методология преподавания баз данных // Труды Всероссийской научно-практической конференции Информационные технологии в образовании и науке «ИТОН-2006». – Москва, 2006.
7. Марченко А.Л. Основы ADO.NET. – <http://www.intuit.ru/department/se/msadonet/>
8. Морган С., Райан Б., Хорн Ш., Бломсма М. Разработка распределенных приложений на платформе Microsoft .NET Framework: Учебный курс Microsoft. – СПб.: «Питер», 2008.
9. Рюмшина О.А. Общий курс «Базы данных и экспертные системы» на специальности «Прикладная математика и информатика» – http://2006.edu-it.ru/docs/3/09_03_Ryumshina.doc.
10. MySQL Reference Manuals. – <http://dev.mysql.com/doc/index.html>.
11. Sikha Saha Bagui, Richard Walsh Earp. Learning SQL on SQL Server 2005. – O'Reilly, 2006.
12. Troelsen A. Pro C# 2008 and the .NET 3.5 Platform. – Apress, 2007.