

## **Этапы проектирования БД, их характеристика. Цели проектирования**

**Цель:** сформировать знание основных этапов проектирования базы данных

Процесс проектирования баз данных длительный, он требует обсуждений с заказчиком, со специалистами в предметной области. При разработке серьезных корпоративных информационных систем проект базы данных является тем фундаментом, на котором строится вся система в целом.

Проектирование баз данных состоит в построении комплекса взаимосвязанных моделей данных. Причем процесс построения модели данных неотделим от процессов обработки и манипулирования данными.

Проектирование базы данных состоит из двух основных фаз: логического и физического моделирования. Во время фазы логического моделирования разработчик собирает требования к разрабатываемой БД, составляет описание предметной области и разрабатывает модель, не зависящую от конкретной СУБД. Во время фазы физического моделирования разработчик создает модель, оптимизированную для СУБД и конкретных приложений пользователей.

При создании базы данных для информационной системы наиболее важными являются задачи, связанные с разработкой правильной логической структуры данных, обеспечивающей выполнение всего требуемого набора функций информационной системы. Плохо продуманная база данных оказывается, как правило, неэффективной и даже бесполезной.

Разработка базы данных является достаточно сложной задачей. Зачастую к ней предъявляется много противоречивых требований. Создание правильной логической структуры предусматривает комплексный анализ всех факторов, влияющих на формирование и обработку данных.

Хорошо спроектированная база данных:

- удовлетворяет всем требованиям пользователей к содержимому базы данных. Перед проектированием базы необходимо провести исследование требований пользователей к функционированию базы данных;
- гарантирует непротиворечивость и целостность данных;
- обеспечивает естественное, легкое для восприятия структурирование информации. Качественное построение базы позволяет делать запросы к базе более «прозрачными» и легкими для понимания; следовательно, снижается вероятность внесения некорректных данных и улучшается качество сопровождения базы;
- удовлетворяет требованиям пользователей к производительности базы данных. При больших объемах информации вопросы сохранения производительности начинают играть главную роль, сразу «высвечивая» все недочеты этапа проектирования.

Задача проектировщика состоит в учете всех этих факторов с целью разработки наиболее оптимальной базы данных. Основные задачи проектирования баз данных можно сформулировать следующим образом:

- обеспечение хранения в БД всей необходимой информации;
- обеспечение возможности получения данных по всем необходимым запросам;
- сокращение избыточности и дублирования данных; и
- обеспечение целостности базы данных.

То, как решаются эти задачи, будет рассмотрен ниже. Но прежде нужно вспомнить определение некоторых терминов.

**Предметная область** – часть реального мира, данные о которой необходимо отразить в базе данных. Например, в качестве предметной области можно выбрать студенческую библиотеку, бухгалтерию предприятия, отдел кадров, банк, магазин и т.д.

Предметная область содержит как существенные понятия и данные, так и малозначащие или вообще не значащие. Так, если в качестве предметной области выбрать учет товаров на складе, то понятие «накладная» является существенным. В то же время количество детей сотрудника, принимающего накладные для учета товаров, является несущественной информацией. Но для расчета заработной платы данные о наличии детей являются необходимыми. Таким образом, любая предметная область имеет свои границы рассмотрения и при проектировании БД необходимо выделить информационные объекты внутри границ предметной области и абстрагироваться от информации вне границ предметной области.

**Модель предметной области** – это формализованные знания предметной области, выраженные при помощи каких-либо средств. В качестве таких средств могут выступать текстовые описания предметной области (должностные инструкции, описание документооборота, примеры первичных документов, выходных отчетов и т.п.). Гораздо более информативными и полезными при работе баз данных являются описания предметной области, выполненные при помощи специализированных графических нотаций. От того, насколько правильно смоделирована предметная область, зависит успех дальнейшей разработки информационной системы.

Процесс проектирования БД представляет собой последовательность переходов от неформального словесного описания информационной структуры предметной области к формализованному описанию объектов предметной области в терминах некоторой **модели**.

В общем случае можно выделить следующие этапы проектирования.

1. Системный анализ и словесное описание информационных объектов предметной области.

2. Проектирование концептуальной (инфологической) модели предметной области – частично формализованное описание объектов предметной области в терминах некоторой семантической модели.

3. Даталогическое или логическое проектирование БД, т. е. описание БД в терминах принятой логической модели данных.

4. Физическое проектирование БД, т. е. выбор эффективного размещения БД на внешних носителях для обеспечения наиболее эффективной работы приложения.

На каждом из этих этапов разрабатывается та или иная модель данных.

Если учесть, что между вторым и третьим этапами необходимо принять решение, с использованием какой стандартной СУБД будет реализовываться проект, то условно процесс проектирования БД можно представить последовательностью выполнения соответствующих этапов как представлено на рисунке 1.



Рисунок 1

*Концептуальное (инфологическое) проектирование* – построение семантической модели предметной области, т. е. информационной модели наиболее высокого уровня абстракции. Такая модель создается без ориентации на какую-либо конкретную СУБД и модель данных. Термины «семантическая модель», «концептуальная модель» и «инфологическая модель» являются синонимами. Кроме того, в этом контексте равноправно могут использоваться слова «модель базы данных» и «модель предметной области» (например, «концептуальная модель базы данных» и «концептуальная модель предметной области»), поскольку такая модель является как образом реальности, так и образом проектируемой базы данных для этой реальности.

Конкретный вид и содержание концептуальной модели базы данных определяются выбранным для этого формальным аппаратом. Обычно используются графические нотации, подобные ER- диаграммам.

Чаще всего концептуальная модель базы данных включает в себя:

- описание информационных объектов, или понятий предметной области и связей между ними;
- описание ограничений целостности, т. е. требований к допустимым значениям данных и к связям между ними.

*Логическое (даталогическое) проектирование* – создание схемы базы данных на основе конкретной модели данных, например, реляционной. Для реляционной модели данных логическая модель – набор схем отношений, обычно с указанием первичных ключей, а также «связей» между отношениями, представляющих собой внешние ключи. Преобразование

концептуальной модели в логическую модель, как правило, осуществляется по формальным правилам. Этот этап может быть в значительной степени автоматизирован.

На этапе логического проектирования учитывается специфика конкретной модели данных, но может не учитываться специфика конкретной СУБД.

*Физическое проектирование* – создание схемы базы данных для конкретной системы управления базами данных. Специфика конкретной СУБД может включать в себя ограничения на именование объектов базы данных, ограничения на поддерживаемые типы данных и т.п. Кроме того, специфика конкретной СУБД при физическом проектировании включает выбор решений, связанных с физической средой хранения данных (выбор методов управления дисковой памятью, разделение БД по файлам и устройствам, методов доступа к данным), создание индексов и т.д.

К области проектирования и разработки баз данных используются различные средства моделирования предметной области и БД, причем даже в рамках одной конкретной системы необходим целый комплекс моделей разного назначения.

### **Анализ предметной области**

Анализ предметной области предшествует созданию любой информационной системы и является частью ее разработки. На этом этапе определяются потребности пользователей в информации, которые в свою очередь определяют структуру и содержание базы данных для будущей системы. Предметная область – некоторая совокупность реальных объектов. Каждый из этих объектов обладает определенным набором свойств (атрибутов). Между объектами предметной области могут существовать связи, имеющие различный содержательный смысл. Создавая базу данных для информационной системы, пользователь стремится упорядочить информацию по различным признакам. Это делается для того, чтобы по необходимости извлекать нужную совокупность данных, получать выборку с желаемым сочетанием признаков. Осуществить это возможно только тогда, когда данные структурированы, т. е. отвечают соглашениям о способах представления данных.

Модель предметной области имитирует структуру и/или функционирование исследуемой предметной области и должна быть адекватной этой области. Особую роль модели предметной области играют на стадии формирования требований к будущей информационной системе при ее создании. Предварительное моделирование предметной области позволяет сократить время и сроки проведения проектировочных работ и получить более эффективный и качественный проект. Без проведения моделирования предметной области велика вероятность допущения большого количества ошибок к решению стратегических вопросов, приводящих к экономическим потерям и высоким затратам на последующее перепроектирование системы.

Исследование предметной области необходимо проводить в целом для разрабатываемой системы, частью которой является база данных. При этом

модель данных может быть создана только в случае, если выявлены все объекты системы, логика их взаимодействия, потоки передаваемой информации. База данных является хранилищем передаваемых данных, которые используются системой при работе. Можно сказать, что база данных это фундамент системы, следовательно, к ее созданию нужно подходить очень серьезно. Очень много ошибок при создании базы данных происходит по причине недостаточной продуманности ее структуры и некачественного выполнения этапа проектирования, который и начинается с исследования предметной области.

Очень важно на этапе проектирования достичь взаимопонимания как между разработчиками системы, так и между экспертами предметной области, заказчиками, так как каждый имеет свое видение проекта. При этом работа сводится к поэтапному выделению объектов, значимых функций системы, информационных потоков и системы их взаимосвязей. Точки зрения участников разработки по определенным проблемам могут совпадать, при этом формы их представления могут быть различными, что ведет к усложнению совместной работы над одним проектом. Важным инструментом в данном случае является использование единого языка представления проектных решений — языка моделирования. Нужно определить систему обозначений, правил описания процессов, объектов, явлений и их взаимосвязи, позволяющую всем участникам проекта понимать друг друга («говорить на одном языке»).

*Язык моделирования* это набор графических нотаций, которые используются для описания моделей в процессе проектирования.

*Нотация* представляет собой совокупность графических объектов, используемых в модели, и является синтаксисом языка моделирования. Язык моделирования, с одной стороны, должен делать решения проектировщиков понятными пользователю, с другой предоставлять проектировщикам средства достаточно формализованного и однозначного определения проектных решений, подлежащих реализации в виде программных комплексов, образующих целостную систему.

Комплексность подхода и использование единой нотации очень важно, не только на этапе моделирования предметной области, но и на последующих этапах разработки программной системы.

Изучение предметной области складывается из непосредственного наблюдения протекающих в ней процессов, изучения документов, циркулирующих в системе, а также интервьюирования процессов.

План разработки модели предметной области определяют некоторые границы, в пределах которых можно развивать логическую модель данных, т.е. моделировать объекты, не выходящие за пределы рассматриваемой предметной области. При этом все второстепенные детали опускаются, чтобы чрезмерно не усложнять процесс анализа и исследования полученной модели.

Результатам проведения исследования предметной области должен стать перечень системных требований, спецификаций, информационных

потоков и их описание. Очень часто для этого **применяются** стандартные способы описания предметной области (использованием моделей DFD (диаграмма потоков данных), UML (унифицированный язык моделирования)).

### **Вопросы**

1. Назовите основные фазы проектирования базы данных.
2. Перечислите требования, которым удовлетворяет хорошо спроектированная база данных.
3. Перечислите основные задачи проектирования.
4. Дайте определение понятиям «предметная область» и «модель предметной области». Приведите примеры предметных областей.
5. Назовите и поясните основные этапы проектирования базы данных.
6. Дайте определение понятию «язык моделирования».
7. Дайте определение понятию «нотация».
8. Перечислите основные результаты исследования предметной области.

Преподаватель

С.В. Банцевич