

1 ВВЕДЕНИЕ

1.1 Жизненный цикл системы баз данных

1.1.1 Структура жизненного цикла

Проектирование базы данных (БД) — это один из этапов общего процесса проектирования системы баз данных организации. Поэтому, прежде чем говорить о процедурах проектирования БД, рассмотрим жизненный цикл системы баз данных в целом.

Под *жизненным циклом (ЖЦ)* системы баз данных (СБД) понимается непрерывный процесс, который начинается в момент принятия решения о создании СБД и заканчивается в момент полного изъятия системы из эксплуатации.

Этот процесс принято подразделять на ряд этапов (фаз). На каждом этапе выполняется определённый перечень работ, направленных на создание или эксплуатацию и развитие системы. Следующая схема представляет *идеальную* временную последовательность этапов ЖЦ СБД (рис. 1.1).

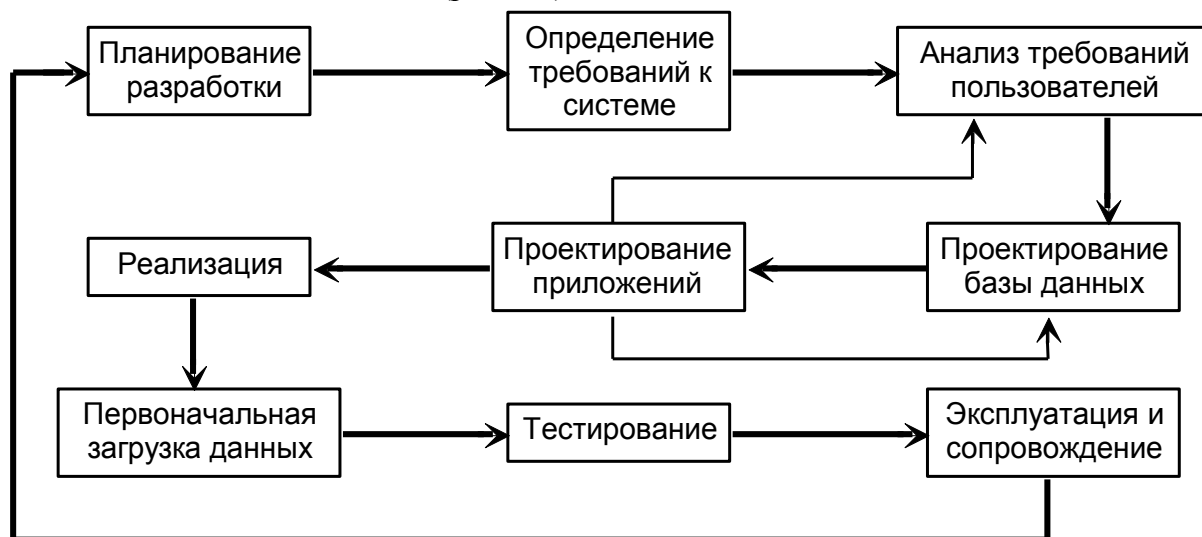


Рис. 1.1 — Жизненный цикл СБД

В идеале результаты очередного этапа должны являться исходными данными для следующего. **Однако** реальный жизненный цикл — это *итерационный процесс*. На каждом этапе могут возникнуть проблемы, которые потребуют пересмотра решений, принятых на любом из предшествующих. Поэтому приведённую здесь схему следует рассматривать как *перечень видов работ*, а не как строгую временную последовательность их исполнения.

Рассмотрим фазы ЖЦ подробнее.

1.1.2 Планирование разработки

Планирование разработки — это *подготовительный этап*. Здесь оцениваются

- требуемый объём работ;
- требуемые ресурсы;
- общая стоимость проекта.

Эти оценки приближённые. Они нужны руководству организации для принятия решения о создании (или радикальной переработки) системы. Обычно этот этап выполняется небольшой группой (2÷3 человека) высококвалифицированных и авторитетных специалистов. Возможно привлечение независимых экспертов.

1.1.3 Определение требований к системе

На этом этапе определяются диапазон действия системы (поддерживаемые бизнес-процессы), её границы и состав пользователей.

1.1.4 Анализ требований пользователей

Это наиболее ответственный этап ЖЦ. Недостаточно тщательно выполненный анализ приведёт к выходу за рамки бюджета проекта, превышению сроков разработки и, возможно, краху проекта.

На этом этапе собирается и анализируется информация о деятельности организации, необходимая для проектирования БД и приложений.

Выявляются

- функции пользователей системы,
- данные, необходимые для выполнения функций,
- бизнес-правила, ограничивающие функции.

Эта информация собирается различными способами. В частности

- путём опроса отдельных наиболее авторитетных сотрудников организации;
- посредством наблюдений за деятельностью организации;
- посредством изучения документов, в особенности тех, которые используются для сбора или представления информации;
- путём анкетирования широкого круга будущих пользователей системы.

Привлекается и опыт разработчиков в проектировании аналогичных систем.

Собранная информация по каждой области применения и каждой группе пользователей **должна включать**

- входную и выходную документацию организации;
- подробные сведения о выполняемых транзакциях;
- список требований пользователей с указанием приоритетов.

Получаемая аналитиками информация о деятельности организации **всегда плохо структурирована**.

Задача аналитиков — структурировать её и определить ***спецификации требований*** пользователей, т.е. дать точное и однозначное описание **функций** пользователей и их **потребностей в данных**.

1.1.5 Проектирование БД

Это этап создания информационного ядра системы.

Основные **цели этапа** следующие.

- Определить данные и связи между ними, необходимые для всех основных областей применения создаваемой СБД и любых существующих групп пользователей.
- Создать модель данных, способную поддерживать выполнение любых требуемых транзакций обработки данных.
- Разработать предварительный вариант проекта СБД, структура которого позволит удовлетворить основные требования к производительности системы.

Вообще говоря, эти цели **противоречивы**. Задача проектировщиков — найти удачный компромисс между универсальностью и производительностью системы.

Процесс проектирования БД мы обсудим подробно в следующих лекциях.

1.1.6 Проектирование приложений

Это процесс создания *интерфейса конечного пользователя и прикладных программ*, предназначенных для обработки данных.

Реально проектирование БД и приложений выполняется параллельно.

БД предназначена для поддержки приложений. Поэтому её проект должен учитывать нужды приложений, отражающие потребности локальных пользователей. Однако и приложения невозможно создавать, не зная структуры БД.

Поэтому между фазами проектирования БД и приложений идёт постоянный обмен информацией, отражённый на рисунке дугами обратных связей.

Проектирование приложений невозможно завершить до завершения проектирования БД.

Главная задача этой фазы — создание интерфейса конечного пользователя для всех функций системы, предусмотренных спецификациями требований пользователей.

1.1.7 Реализация

На этом этапе

- создаётся «пустая база данных»,
- определяются **все** специфические пользовательские представления,
- реализуются **все** приложения,
- реализуются **все** средства защиты данных и поддержания целостности.

Для реализации БД и внешних представлений создают их **описание** на языке определения данных (ЯОД) целевой системы управления базами данных (СУБД).

Это описание используется системой для **создания схем данных** всех уровней, и пустых файлов физической базы данных (ФБД) на устройствах внешней памяти.

Приложения реализуются на языках программирования высокого уровня (ЯВУ).

Транзакции обработки БД описываются на языке манипулирования данными (ЯМД) целевой СУБД. Они могут включаться непосредственно в текст программ, а могут быть реализованы в виде отдельного модуля.

Многие современные СУБД имеют собственные среды быстрого создания приложений.

Средства защиты данных и поддержания целостности частично описываются средствами ЯОД. Однако возможности СУБД в этом отношении не всегда достаточны. Поэтому часто приходится создавать специальные программы, выполняющие проверки правил целостности (*триггеры БД*), а также утилиты и специальные приложения, обеспечивающие безопасность данных.

1.1.8 Первоначальная загрузка

Это перенос любых существующих данных в новую БД и модификация всех существующих приложений с целью обеспечения их совместной работы с новой БД.

Работы на этом этапе зависят от того, есть ли в организации компьютерная информационная система. Если новая СБД заменяет старую, то файлы старой ФБД должны быть конвертированы в новые форматы. Каждая современная СУБД имеет утилиту загрузки существующих файлов в новую БД. Этой утилите нужно задать спецификации файлов.

Адаптация старых приложений к новой БД — гораздо более сложная задача. Нередко проще разработать приложение заново, чем адаптировать его.

Если ИС организации создаётся впервые, то администратор базы данных (АБД) должен

- выделить подмножество данных организации, которые будут перенесены с бумажных носителей в БД;
- продумать процедуру ввода этих данных и создать вспомогательные приложения, предназначенные исключительно для обеспечения корректного ввода;
- обучить операторов и постоянно контролировать их работу.

1.1.9 Тестирование

Это процесс выполнения прикладных программ с целью поиска ошибок.

Тестирование *не может доказать безошибочность* программ. С помощью тестов можно лишь обнаружить часть имеющихся ошибок. Если в результате тестирования не выявлено ошибок, то либо тесты были неудачными, либо программа с большой вероятностью работает в соответствии со спецификацией.

Существует множество стратегий и методик тестирования программ. Мы не будем обсуждать их здесь, поскольку это предмет технологии программирования, а не технологии БД.

1.1.10 Эксплуатация и сопровождение

Это наблюдение за системой и поддержание её нормального функционирования по окончании развёртывания.

На этом этапе осуществляется контроль производительности системы и сопровождение и модернизация приложений.

Производительность системы может падать со временем.

Основные причины снижения — увеличение объёма хранимых данных и увеличение числа одновременно работающих приложений.

Если среднее время реакции системы на запрос приложения становится недопустимо большим, то АБД должен произвести *дополнительную настройку* системы с целью увеличения производительности.

Например,
создать дополнительные индексы,
объединить или расщепить отдельные таблицы,
изменить структуры файлов и т.п.

Для того чтобы выполнить настройку, АБД должен располагать сведениями о различных показателях функционирования системы,
в частности

- о частотах обращений к тем или иным фрагментам БД,
- об эффективности системы блокировок,
- о выбираемых стратегиях выполнения запросов и т.п.

Сбор этой информации обеспечивают специальные утилиты администрирования БД. Как правило, они входят в состав типичной СУБД. Однако если СУБД не предоставляет такие вспомогательные службы, то АБД должен создать их сам, поскольку мониторинг системы необходимо осуществлять в течение всего её жизненного цикла.

В процессе эксплуатации СБД у пользователей появляются **новые потребности**. Они должны отслеживаться и удовлетворяться.

Новые требования могут относиться к
функциональности существующих приложений,
составу полей отдельных таблиц,
составу таблиц БД в целом и т.п.

Необходимость учёта этих требований возвращает разработчиков на начальный этап ЖЦ.

В конце концов, может наступить такой момент, когда нужно будет принять решение о проектировании **новой СБД**, поскольку существующая не в состоянии достичь требуемых эксплуатационных показателей.

Если новая система создана и введена в эксплуатацию, то старую **не следует** немедленно выводить из эксплуатации. Будет лучше, если некоторое время они будут работать параллельно. Это обеспечит подстраховку на случай непредвиденных проблем с новой системой.

1.2 Обзор процесса проектирования БД

1.2.1 Цели проектирования

Проектирование БД — это процесс создания БД, предназначенной для поддержки функционирования организации и способствующей достижению её целей.

Этот процесс направлен на достижение двух целей.

Цель 1. Создать структуры хранения, способные обеспечить накопление данных организации и выполнение всех требуемых видов обработки данных.

Цель 2. Создать все приложения, способные обеспечить интерфейс конечных пользователей с базой данных и специфическую обработку данных в соответствии с требованиями различных конечных пользователей.

1.2.2 Фазы проектирования БД

Как вы помните, одна из основных концепций технологии БД — концепция трёхуровневого представления данных.

Описания данных на внешнем и концептуальном уровнях — внешние и концептуальные схемы (ВС и КС) — не привязаны к реализации базы данных. Они описывают логические структуры данных, соответствующие объектам ПО и их отношениям, и не содержат ссылок на программно-техническую платформу проектируемой системы. На этапе разработки ВС и КС платформа ещё не определена. В основе схем лежат *представления конечных пользователей о предметной области*.

Внешняя схема соответствует локальному представлению конечного пользователя (КП), исполняющего в организации конкретные функции, например бухгалтера группы материального учёта, инспектора отдела кадров, кладовщика и т.п.

Концептуальная схема обобщает локальные представления. Её можно понимать как описание представлений *абстрактного пользователя*, исполняющего функции *всех* КП. Любая внешняя схема может быть выведена из концептуальной.

Внутренняя схема — это *реализация* концептуальной схемы на конкретной программно-технической платформе. Как и КС, она описывает представления обобщённого пользователя БД, но, кроме того, содержит определения физического уровня.

В соответствии со сказанным, выделяют три фазы проектирования БД (*слайд 9*):

- концептуальное моделирование,
- логическое моделирование,
- физическое проектирование.

Процесс создания внешних и концептуальной схем называют моделированием данных организации. Цель моделирования — понять, ЧТО нужно сделать.

Концептуальное моделирование не связано ни с какими соображениями реализации. Результатом этого процесса является документированная модель представлений конечных пользователей об информационных потребностях бизнеса.

На фазе *логического моделирования* концептуальная модель преобразуется в *логическую* с учётом выбранного способа структурирования данных. Логическая модель и есть описание концептуальной схемы БД.

Процесс создания внутренней схемы называют проектированием ФБД или физическим проектированием БД. Его цель — решить, КАК сделать то, что нужно.

На фазе физического проектирования создаётся описание реализации БД на внешних запоминающих устройствах. Определяются структуры хранимых файлов и методы доступа, обеспечивающие эффективную обработку данных.

Обсудим фазы проектирования БД подробнее.

1.2.3 Концептуальное моделирование

Концептуальное моделирование — это процесс анализа информационных потребностей конечных пользователей системы.

На этой фазе цель аналитика — сформировать в своей голове представления об информационных потребностях бизнеса, адекватные представлениям предполагаемых конечных пользователей системы. Эти представления формируются в процессе изучения деятельности организации. Используются следующие источники информации:

- документы организации (хозяйственные книги, счета, накладные, финансовые отчёты и т.п.),
- собственные наблюдения за работой организации,
- беседы с пользователями,
- опыт предыдущих разработок,
- собственные предположения об информационных потребностях бизнеса.

Обрабатывая полученную информацию, аналитик делает выводы о структуре и связях объектов предметной области. Эти выводы он документирует в модели данных и обсуждает с пользователем.

Цель обсуждений — согласование представлений аналитика и пользователя о предметной области системы.

Основная проблема согласования в том, что пользователь и аналитик мыслят разными категориями.

Пользователь знает, какие формы и отчёты и с какими данными ему нужны, но он ничего не может сказать о структурах и связях объектов, отражаемых в этих формах и отчётах.

Разработчика же интересуют именно структуры и связи объектов. Он вынужден *реконструировать* объекты и связи из форм и отчётов.

Эта задача, как любая обратная задача, имеет множество решений. В ходе многократных обсуждений аналитик должен выбрать такое, которое адекватно модели данных, *имеющейся в голове пользователя*.

В противном случае результат разработки — *готовая система* — вряд ли удовлетворит пользователя.

Процесс моделирования больше искусство, чем наука. Существует множество средств и способов моделирования данных. Можно изучить их все. Мы изучим некоторые. Но их использование — это искусство, требующее опыта и интуиции.

Результат концептуального моделирования — концептуальная модель данных пользователя — **однозначно описывает** структуры и связи *объектов* предметной области и *бизнес-правила* организации. Это описание *не зависит* от

- типа целевой СУБД,
- набора создаваемых программ,
- используемых языков программирования,
- типа вычислительной платформы,
- а также от любых других особенностей физической реализации.

Концептуальная модель служит источником информации для фазы логического моделирования.

1.2.4 Логическое моделирование

В процессе логического моделирования концептуальная модель уточняется и преобразуется в логическую с **учётом базовой модели данных целевой СУБД**. Например, если целевая СУБД базируется на РМД, то структуры и связи объектов, представленные в концептуальной модели, преобразуются в систему взаимосвязанных отношений, удовлетворяющих определённым формальным требованиям.

К началу логического моделирования должен быть определён **тип** целевой СУБД (реляционная, сетевая, объектно-ориентированная...). Однако *все прочие аспекты* целевой СУБД (физическая организация структур хранения, способы построения индексов и т.п.) *полностью игнорируются*.

Логическая модель содержит описания структур данных в терминах *выбранной модели данных*, а также описания ограничений целостности. Она должна

- быть корректной,
- соответствовать требованиям пользователей, зафиксированным в концептуальной модели,
- обеспечивать поддержку всех необходимых пользователям транзакций.

Логическая модель является источником информации для фазы физического проектирования БД и для проектирования приложений.

Кроме того, она играет важную роль **на этапе эксплуатации** и сопровождения СБД. При правильно организованном сопровождении она составляет основную часть системного каталога и автоматически обновляется при внесении изменений в логические структуры данных. Благодаря этому АБД всегда может точно представить любые вносимые изменения и оценить их влияние на существующие приложения.

Моделирование данных — наиболее ответственный и трудоёмкий этап разработки системы. Это практически бесконечный итеративный процесс исследования информационных потребностей предприятия. Каждая итерация вносит уточнения и улучшения в модели, что, в свою очередь, может потребовать изменений в других частях проекта.

Если созданные модели неадекватно отражают представления пользователей о предметной области, то будет очень трудно или даже невозможно определить все необходимые внешние схемы, организовать поддержку целостности данных и обработку необходимых транзакций. Опыт показывает, что успешные проекты используют около 2/3 бюджета времени **на моделирование данных**. Именно качественные модели являются залогом их успеха.

1.2.5 Физическое проектирование

На этой фазе принимаются решения о способах реализации БД. Она может начаться только после выбора конкретной целевой СУБД. Целью физического проектирования является описание способа физической реализации логической модели.

Например, если используется реляционная СУБД, то в её среде должен быть создан набор таблиц и ограничений, представленный в логической модели. Должны быть определены конкретные структуры хранения данных (файлы) и методы доступа к данным, которые обеспечат требуемую производительность системы. Кроме того, должны быть разработаны средства защиты системы.

В идеале фазы логического моделирования и физического проектирования **следует разделять**, однако **реально это невозможно**, поскольку решения, принимаемые на фазе **физического проектирования** с целью повышения производительности системы, **могут повлиять** на структуру логической модели.

Фазы моделирования и физического проектирования БД предъявляют различающиеся требования к знаниям и профессиональным навыкам проектировщиков. Для успешного моделирования нужно обладать навыками исследовательской работы и владеть методологиями моделирования. Для успешного проектирования ФБД нужно хорошо знать возможности целевой программно-технической платформы, владеть её инструментами и иметь навыки системного программирования.

В дальнейшем мы сосредоточимся на методологиях концептуального и логического моделирования и примерах их использования.