

Формализация информации и Big Data

«02.03.03 - Математическое обеспечение и администрирование информационных систем
направленность разработка и администрирование информационных систем»

<http://vikchas.ru>

Лекция 12 «СУБД Adabas Big Data»

Часовских Виктор Петрович

д-р техн. наук, профессор кафедры ШИиКМ

ФГБОУ ВО «Уральский государственный экономический
университет»

Екатеринбург 2023

Удобство эксплуатации

Структура БД, используемая в ADABAS, значительно уменьшает трудозатраты, связанные с проектированием и эксплуатацией БД, — пользователю предоставляется язык описания данных, при котором структура каждого файла может быть определена независимо от других файлов.

Для большинства приложений при моделировании объектов предметной области достаточно тех понятий, которые допустимы на уровне схемы файла: простая группа, периодическая группа, простые и множественные атрибуты.

Кроме того, для определения ассоциативных связей между записями в схеме БД некоторые из атрибутов можно объявить дескрипторами. Для поисковых атрибутов (дескрипторов) в ассоциаторе автоматически строятся инвертированные списки.

На любом этапе существования БД ее логическая структура может быть модифицирована без перезагрузки или реорганизации БД. Объектами модификации могут быть как логические структуры файлов, так и связи между ними. Тем самым предоставляется возможность отражать в БД изменения, которые происходят в модели предметной области.

Структура файлов может быть расширена путем включения простых и множественных атрибутов и групп. В схеме БД возможно добавление нового или удаление старого файла, а также определение новой или удаление старой связи.

После расширения логической структуры значения новых атрибутов могут присутствовать во вновь вводимых записях.

В старые записи значения новых атрибутов добавляются с помощью команд корректировки.

Для каждой категории пользователей ADABAS предоставляет простые и непроцедурные языковые средства. Система располагает диалоговым языком программирования, который существенно сокращает трудозатраты по реализации прикладных задач.

Описания данных (длина, тип значения, информация для редактирования, имена колонок отчетов и т. д.) хранятся в словаре данных. Эти описания доступны всем пользователям ADABAS и не зависят от формата и соответствующего описания, используемого в БД, что обеспечивает определение логических объектов, отражающих специфику решаемой задачи.

Средства конструирования интерактивных технологий позволяют администратору базы данных работать с естественными понятиями его предметной области и создавать стандартные технологии сопровождения БД, обеспечивающие легкость и простоту ее эксплуатации.

Стандартные технологические процессы позволяют автоматизировать ввод, контроль и преобразование данных, загрузку и корректировку данных, организацию и ведение словаря данных и т. д.

Анализ функциональных возможностей ADABAS показывает, что эта система может служить эффективным инструментом для реализации информационных систем, существенно сокращающим трудоемкость разработки прикладных задач.

Базовая модель и структура данных системы

ADABAS разработана как высоконадежная и производительная СУБД для создания и эксплуатации больших баз данных на мейнфреймах.

В настоящее время она представляет собой систему управления **сверхбольшими базами** данных и позволяет строить не только традиционные системы обработки структурированных данных, но и текстовые информационно-поисковые СУБД ADABAS системы, географические и экспертные системы, системы обработки изображений и т. д.

Это становится возможным благодаря тому, что ADABAS обеспечивает поддержку следующих моделей и типов данных.

1. Не первая нормальная форма (NF2 — Non-First Normal Form). Эта модель (традиционная) данных характерна для начальной версии ADABAS.

2. Традиционная реляционная модель данных. Эта модель соответствует ANSI/ISO стандарту SQL.

3. Модель данных сущность-связь (E/R модель). В ADABAS предусмотрено расширение до модели Entity-Relationship (или E/R модели) для управления сложными структурами данных с высокой степенью связности. Поддерживаются также рекурсивные структуры данных.

Объединяя эту модель с другими моделями ADABAS, можно строить мощные интегрированные базы данных и, соответственно, прикладные системы. Предпочтительные области для применения этих моделей — системы представления знаний, моделирование поведения сложных технических и биологических систем, расчеты потребностей, планирование материальных ресурсов различного вида и назначения (Bills of Materials).

4. Обработка и управление произвольными текстами. Этот тип данных (и соответствующие средства манипулирования ими) обеспечивает доступ к документам, обрабатываемым ADABAS, как к произвольным текстам. Возможно сочетать разного рода обработку текстовых данных со стандартными процедурами ADABAS для работы с форматированными данными. Например, это могут делать библиотеки, юридические конторы, службы новостей, телекомпании, газеты и журналы, правительственные организации — все, кто имеет и будет иметь данные в виде так называемых «свободных» или неструктурированных текстов.

5. Обработка и управление географическими данными. В отличие от традиционных географических (картографических) систем, которые используют структуры хранения данных (реляционные и др.) независимо от собственно географической информации, обеспечивая их совместное использование, как правило, только на уровне SQL-интерфейса, подход ADABAS характеризуется глубокой интеграцией на всех уровнях хранения и обработки.

Картографическая информация, координатные привязки объектов, форматированные данные и текстовая информация **неограниченного объема хранятся в рамках единой БД.**

Точно так же в рамках одной программы и даже **одного оператора** (среда Natural) интегрирована обработка этой комплексной информации.

Важно, что при этом обеспечивается не только высокая производительность геоинформсистем, но за счет поддержания логики транзакций гарантируется целостность и безопасность хранения БД и другие ее неотъемлемые характеристики.

Структура системы

Многоуровневая архитектура

Для современной концепции баз данных характерно многоуровневое представление данных, при котором каждый уровень играет определенную роль в процессе проектирования, эксплуатации и использования БД.

Многоуровневая архитектура обеспечивает независимость данных и способность информационных систем, построенных на основе СУБД, к эволюции. Наибольшее признание в настоящее время получила трехуровневая архитектура, включающая внешний, концептуальный и внутренний уровни.

Внешний уровень поддерживает частные представления о данных отдельных задач пользователей.

Концептуальный уровень является глобальным представлением о предметной области вне зависимости от того, как данные хранятся и используются приложениями.

Внутренний уровень связан с представлением данных в памяти ЭВМ.

Каждый уровень представления данных в архитектуре СУБД является, по существу, самостоятельным уровнем абстракции данных и определяется соответствующей моделью данных. Поэтому современные СУБД с многоуровневой архитектурой характеризуются не одной, а несколькими моделями данных.

Трехуровневая архитектура СУБД обладает высокой степенью адаптации к происходящим изменениям в процессе эволюции информационных систем возможность независимой модификации внешних и внутренних представлений о БД благодаря существованию стабильной концептуальной схемы и действию соответствующих отображений.

ADABAS является СУБД с трехуровневой архитектурой.

Она предоставляет в распоряжение пользователей концептуальный и внутренний уровни, ориентированные на АД, и внешний уровень, ориентированный на конечных пользователей, администратора прикладных задач и прикладных программистов. Кроме того, прикладные программисты могут обращаться к БД на языке манипулирования данными концептуального уровня. Однако в системе ADABAS отсутствует четкая граница между концептуальным и внутренним уровнями. Описанные выше функциональные возможности системы ADABAS реализуются совокупностью программных компонентов, объединенных архитектурой, отраженной на рис. 1.

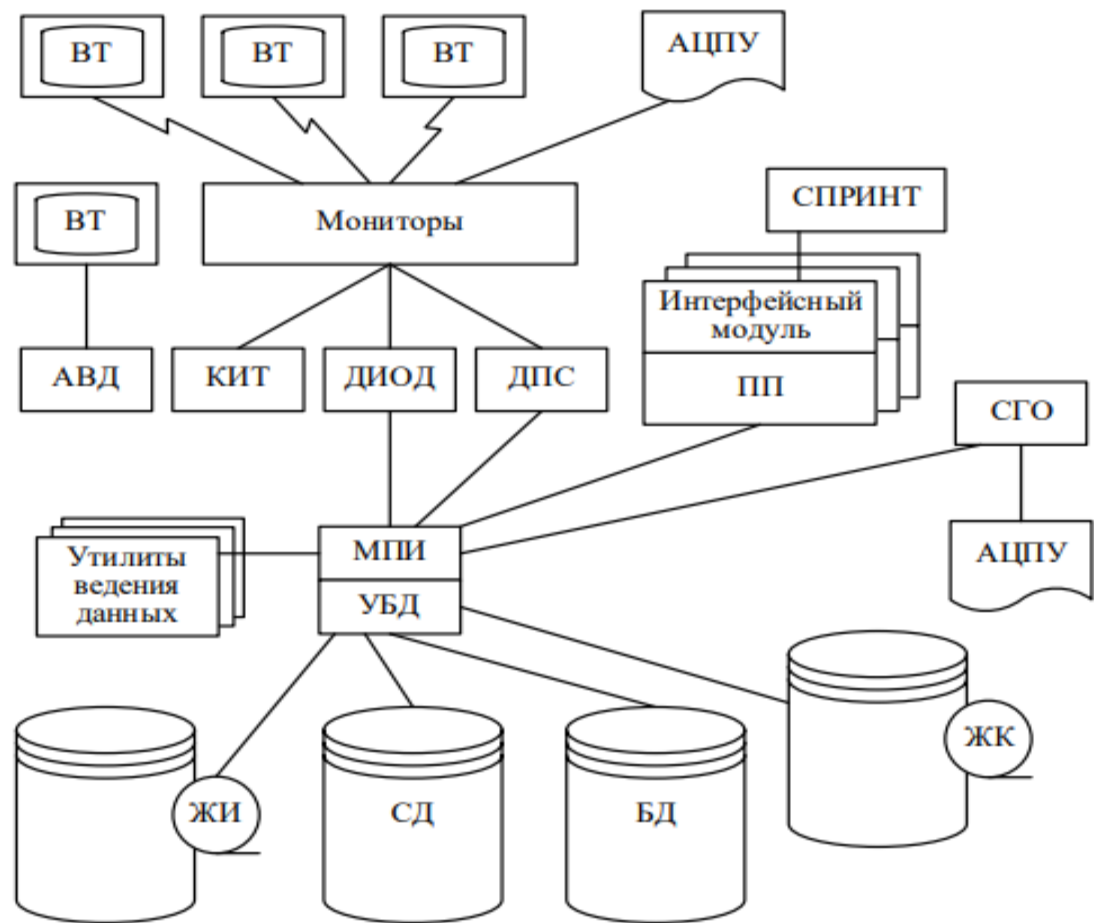


Рис. 1 Архитектура ADABAS:

VT — видеотерминал; АЦПУ — алфавитно-цифровое печатающее устройство;
 ПП — прикладные программы; ЖК — журнал команд;
 ЖИ — журнал изменений; СД — словарь данных; БД — база данных;
 АВД — автономное ведение данных; КИТ — конструирование интерактивных технологий; ДИОД — диалоговая обработка данных; ДПС — диалоговая подготовка справок; МПИ — мультипрограммный интерфейс;
 УБД — управление базой данных; СГО — средства генерации отчетов;
 СПРИНТ — средства программирования интерфейсных модулей

В состав ADABAS входят программные средства, обеспечивающие:

- управление видеотерминалами (МВТ);
- конструирование интерактивных технологий (КИТ);
- диалоговую обработку данных (ДИОД);
- диалоговую подготовку справок (ДПС);
- программирование интерфейсных модулей (СПРИНТ);
- генерацию отчетов (СГО);
- управление базой данных (УБД);
- мультипрограммный интерфейс (МПИ);
- автономное ведение данных (АВД);
- ведение данных.

Концептуальный и внутренний уровни

Концептуальный и внутренний уровни реализуются программой УБД и утилитами ведения данных.

Утилиты ведения данных включают утилиты создания и ведения базы данных, утилиты ведения словаря данных и утилиту обеспечения сохранности данных. В интерактивном режиме они получают параметры от средств АВД и КИТ. Могут запускаться также в пакетном режиме работы.

Утилиты ведения словаря данных осуществляют накопление, обновление и выдачу сведений о данных, программах, пользователях и их связях с данными и программами.

Программа УБД и утилиты ведения данных обеспечивают механизм отображения модели данных концептуального уровня на модель данных, поддерживаемую на внутреннем уровне.

На концептуальном уровне ADABAS с помощью языка описания данных даст представление о БД как о совокупности взаимосвязанных файлов, состоящих из однотипных записей, включающих атрибуты (простые и повторяющиеся), которые моделируют реальные информационные объекты предметной области, их свойства и взаимосвязи.

Операции манипулирования БД обеспечивают возможность создания, модификации и доступа к записям файлов с учетом их взаимосвязей и объявленных в схеме БД ассоциаций.

На концептуальном уровне ADABAS допускается возможность обращения к БД из прикладных программ, написанных на языках ассемблера COBOL, FORTRAN и др. Особенности операционной среды, под управлением которой выполняются прикладные программы, учитываются с помощью специализированных интерфейсов.

Модель данных, используемая в ADABAS на концептуальном уровне, определяется в системе как базовая.

На внутреннем уровне дается представление о БД с использованием таких понятий среды хранения, как набор данных, экстент, блок, хранимая запись, индекс и т. д.

Механизм отображения концептуального представления БД на ее внутреннее представление обеспечивает преобразование различных типов данных, определенных на них ограничений и операций базовой модели в понятия, ограничения и операции внутренней модели.

Внешний уровень

Внешний уровень в архитектуре ADABAS представлен рядом компонентов, в число которых входят программные средства ДИОД, СГО, ДПС, СПРИНТ и утилита форматированного ввода данных.

Каждый из перечисленных компонентов внешнего уровня поддерживает некоторый уровень абстракции БД, ориентированный на соответствующую категорию пользователей системы, и может рассматриваться как механизм отображения внешнего уровня на концептуальный.

Представление данных, поддерживаемое каждым из компонентов внешнего уровня, характеризуется структурными свойствами этого представления, совокупностью операций над элементами соответствующей структуры и ограничениями целостности.

Языковые средства описания структуры данных и манипулирования ими для каждого из представлений данных внешнего уровня позволяют определить внешний уровень ADABAS как управляемый.

Модель данных, поддерживаемая средствами ДИОД, СГО и ДПС и соответствующими языками, практически совпадает с моделью данных, используемой на концептуальном уровне.

Близость представлений данных, применяемых в средствах ДИОД, ДПС и СГО, к представлениям данных концептуального уровня обеспечивает сравнительно простое и эффективное отображение внешних моделей данных, поддерживаемых этими компонентами, на базовую модель данных концептуального уровня.

Благодарю за внимание!

