**История развития СУБД**

Предшественницами СУБД были файловые системы. Однако появление СУБД не привело к полному исчезновению файловых систем. Для выполнения некоторых специализированных задач файловые системы используются до сих пор. Считается, что развитие СУБД началось еще в 1960-е годы, когда разрабатывался проект запуска корабля Apollo на Луну. Этот проект был начат по инициативе президента США Кеннеди, поставившего задачу осуществить пилотируемый полет и высадку человека на Луну к концу десятилетия. В то время не существовало никаких систем, способных обрабатывать или как-либо управлять тем огромным количеством данных, которое было необходимо для реализации этого проекта.

В результате специалисты основного подрядчика — компании North American Aviation (NAA) (которая теперь называется Rockwell International) — разработали программное обеспечение под названием GUAM (Generalized Update Access Method). Основная идея GUAM была построена на том, что малые компоненты объединяются вместе как части более крупных компонентов до тех пор, пока небудет собран воедино весь проект. Применяемую при этом структуру, напоминающую перевернутое дерево, часто называют иерархической структурой (hierarchical structure). В середине 1960-х годов корпорация IBM присоединилась к фирме NAA для совместной работы над GUAM, в результате чего была создана система IMS (Information Management System). Причина, по которой корпорация IBM ограничила функциональные возможности IMS только управлением иерархиями записей, заключалась в том, что необходимо было обеспечить работу с устройствами хранения с последовательным доступом, а именно с магнитными лентами, которые были в то время основным типом носителя. Спустя некоторое время это ограничение удалось преодолеть. Несмотря на то что IMS является самой первой из всех коммерческих СУБД, она до недавнего времени оставалась основной иерархической СУБД, используемой на большинстве крупных мэйнфреймов.

Другим заметным достижением середины 1960-х годов было появление системы IDS (Integrated Data Store) фирмы General Electric. Работу над ней возглавлял один из пионеров исследований в области систем управления базами данных — Чарльз Бачман (Charles Bachmann). Развитие этой системы привело к созданию нового типа систем управления базами данных — сетевых (network) СУБД, — что оказало существенное влияние на информационные системы того поколения. Сетевая СУБД создавалась для представления более сложных взаимосвязей между данными, чем те, которые можно было моделировать с помощью иерархических структур, а также для формирования стандарта баз данных. Для создания таких стандартов в 1965 году на конференции организации CODASYL (Conference on Data Systems Languages), проходившей при участии представителей правительства США и бизнесменов, была сформирована рабочая группа List Processing Task Forte, переименованная в 1967 году в группу Data Base Task Group (DBTG). В компетенцию группы DBTG входило определение спецификаций среды, которая допускала бы разработку баз данных и управление данными. Предварительный вариант отчета этой группы был опубликован в 1969 году, а первый полный вариант — в 1971 году. Предложения группы DBTG содержали три компонента.

• *Сетевая схема* — это логическая организация всей базы данных в целом (с точки зрения АБД), которая включает определение имени базы данных, типа каждой записи и компонентов записей каждого типа.

• *Подсхема* — это часть базы данных, как она видится пользователям или приложениям.

*• Язык управления данными* — инструмент для определения характеристик и структуры данных, а также для управления ими. Группа DBTG также предложила стандартизировать три языка.

• *Язык определения данных для схемы* (Data Definition Language — DDL), который позволяет АБД ее описать.

• *Язык определения данных для подсхемы* (также DDL), который позволяет определять в приложениях те части базы данных, доступ к которым будет необходим.

• *Язык манипулирования данными* (Data Manipulation Language — DML), предназначенный для управления данными.

Несмотря на то, что этот отчет официально не был утвержден Национальным институтом стандартизации США (American National Standards Institute - ANSI), большое количество систем было разработано в полном соответствии с этими предложениями группы DBTG. Теперь они называются CODASYL-системами, или DBTG-системами. CODASYL-системы и системы на основе иерархических подходов представляют собой СУБД первого поколения. Более подробно они рассматриваются в материалах, представленных на сопровождающем Web-узле (URL которого приведен во введении к данной книге). Однако этим двум моделям присущи перечисленные ниже недостатки.

• Даже для выполнения простых запросов с использованием переходов и доступом к определенным записям необходимо создавать достаточно сложные программы.

• Независимость от данных существует лишь в минимальной степени.

• Отсутствие общепризнанных теоретических основ,

В 1970 году Э. Ф. Кодд (Е. F. Codd), работавший в исследовательской лаборатории корпорации IBM, опубликовал очень важную и весьма своевременную статью о реляционной модели данных, позволявшей устранить недостатки прежних моделей. Вслед за этим появилось множество экспериментальных реляционных СУБД, а первые коммерческие продукты появились в 1970-1980-х годах. Особенно следует отметить проект System R, разработанный в исследовательской лаборатории корпорации IBM, расположенной в городе Сан-Хосе, штат Калифорния, созданный в конце 1970-х годов [9]. Этот проект был задуман с целью доказать практичность реляционной модели, что достигалось посредством реализации предусмотренных ею структур данных и требуемых функциональных возможностей. На основе этого проекта были получены важнейшие результаты.

• Был разработан структурированный язык запросов SQL, который с тех пор стал стандартным языком любых реляционных СУБД.

• В 1980-х годах были созданы различные коммерческие реляционные СУБД — например, DB2 или SQL/DS корпорации IBM или Oracle корпорации Oracle Corporation.

В настоящее время существует несколько сотен различных реляционных СУБД для серверов и персональных компьютеров, хотя во многих из них определение реляционной модели трактуется слишком широко. В качестве примеров многопользовательских СУБД могут служить система INGRES II фирмы Computer Associates и система Informix фирмы Informix Software, Inc. Примерами реляционных СУБД для персональных компьютеров являются Access и FoxPro фирмы Microsoft, Paradox фирмы Corel Corporation, InterBase и BDE фирмы Borland, а также R:Base фирмы R:Base Technologies. Реляционные СУБД относятся к СУБД второго поколения.

Однако реляционная модель обладает также некоторыми недостатками — в частности, ограниченными возможностями моделирования. Для решения этой проблемы был выполнен большой объем исследовательской работы. В 1976 году Чен (Chen) предложил модель "сущность-связь" (Entity-Relationship model — ER-модель), которая в настоящее время стала самой распространенной технологией проектирования баз данных. В 1979 году Кодд сделал попытку устранить недостатки собственной основополагающей работы и опубликовал расширенную версию реляционной модели — RM/T (1979), затем еще одну версию — RM/V2 (1990). Попытки создания модели данных, позволяющей более точно описывать реальный мир, неформально называют семантическим моделированием данных (semantic data modeling).

В ответ на все возрастающую сложность приложений баз данных появились две новые системы: объектно-ориентированные СУБД, или ООСУБД (Object-Oriented DBMS — OODBMS), и объектно-реляционные СУБД, или ОРСУБД (Object-Relational DBMS — ORDBMS). Однако, в отличие от предыдущих моделей, действительная структура этих моделей не совсем ясна. Попытки реализации подобных моделей представляют собой СУБД третьего поколения