

История развития СУБД

Предыстория



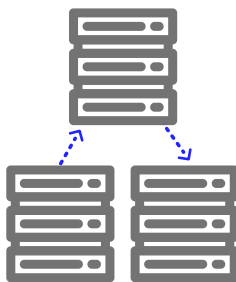
Важным этапом развития ИС явился переход к использованию централизованных систем управления файлами. Широкое распространение получили индексированные файлы. Файл является простым набором записей (record), которые содержат логически связанные данные. Каждая запись содержит логически связанный набор из одного или нескольких полей (field), каждое из которых представляет некоторую характеристику моделируемого объекта. С точки зрения прикладной программы файл – именованная область внешней памяти, в которую можно записывать и из которой можно считывать данные.

Множественные недостатки файловых систем послужили причиной разработки нового подхода к управлению информацией. Разработаны системы управления базами данных (СУБД).

Первое поколение СУБД

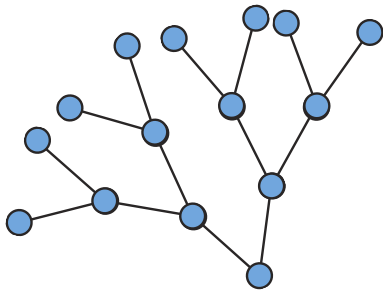
Инвертированные списки

К СУБД первого поколения можно отнести системы, основанные на инвертированных списках, иерархические и сетевые системы управления базами данных. Инвертированные списки представляют собой список таблиц и список индексов, позволяющих осуществлять доступ к данным, хранимым в таблицах. Общие правила определения целостности в системах на основе инвертированных списков отсутствуют. В некоторых системах поддерживаются ограничения уникальности значений некоторых полей, но в основном все возлагается на прикладную программу.



С помощью индексов осуществляется поиск по вторичному ключу. БД, организованная с помощью инвертированных списков, похожа на реляционную, но с тем различием, что индексы доступны.

Иерархические и сетевые СУБД



В 1968 году появилась первая версия базы данных Information Management System (IMS) фирмы IBM. Применяемая структура, напоминала перевернутое дерево и была названа иерархической структурой. Другим заметным достижением середины 1960-х годов было появление системы IDS (Integrated Data Store) фирмы General Electric. Развитие этой системы привело к созданию нового типа систем управления базами данных — сетевых СУБД. 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.

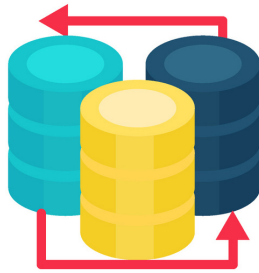
В 1971 году для утверждения Национальным институтом стандартизации США (American National Standards Institute — ANSI), группой DBTG был представлен стандарт баз данных, содержащий три компонента:

- 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 пользователями или приложениями.
- Язык управления данными — инструмент для определения характеристики структуры данных, а также для управления ими.

Второе поколение СУБД

Реляционные СУБД

В 1970 году Э. Ф. Кодд (E. F. Codd), работавший в исследовательской лаборатории корпорации IBM, опубликовал 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. Корпорацией IBM, расположенной в городе Сан-Хосе, штат Калифорния, созданной в конце 1970-х годов, был задуман проект с целью доказать практичность реляционной модели, что достигалось посредством реализации предусмотренных ею структур данных и требуемых функциональных возможностей.

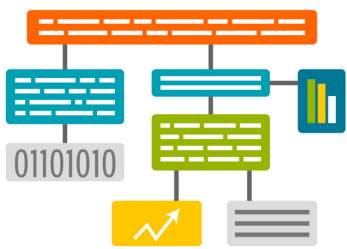


На основе этого проекта были получены важнейшие результаты.

- 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. Реляционные СУБД относятся к СУБД второго поколения.

Улучшенные реляционные СУБД



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 Питер Чен предложил модель "сущность-связь" (Entity-Relationship model — ER-модель), которая в настоящее время стала самой распространенной технологией проектирования баз данных и является основой методологии. 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-Driented 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.