



# БАЗИ ОТ ДАННИ

Преподаватели:

*Доц. д-р Мария Кашева*

*[m.kasheva@ue-varna.bg](mailto:m.kasheva@ue-varna.bg)*

*Гл. ас. д-р Иван Куюмджиев*

*[ivan\\_ognyanov@ue-varna.bg](mailto:ivan_ognyanov@ue-varna.bg)*

# Основни теми

- Теоретични основи на базите от данни
  - Концепция за бази от данни.
  - Модел на данните “Същност-връзки”.
  - Релационен модел. Релационна алгебра. Нормализация на релационния модел.
  - Складове от данни. Многомерно моделиране на данните
  - СУБД – функции, обзор. MS SQL Server
- Structured Query Language SQL
  - Transact-SQL

# ОСНОВНИ ПОНЯТИЯ

- **Базата от данни (БД)** е множество от свързани данни, съхранявани заедно за нуждите на различни приложения в една организация (фирма, предприятие).
- БД представя някаква част от реалния свят, т.е. има определена **предметна област** и отразява установените в нея бизнес правила
- За създаване и управление на БД се използва специализиран софтуер, наречен система за управление на бази от данни (СУБД – **Data Base Management System DBMS**)

# Видове бази от данни

- Операционни бази от данни
  - Характеризират се с динамичност на съхраняваните данни
  - Използват се при OLTP (On-line transaction processing – он-лайн обработка на транзакции)
  - **Транзакция** - логическа неделима единица за работа с базата от данни, която включва една или няколко операции с данните (добавяне на нови данни, модифициране или изтриване на съществуващи данни).

# Online Transaction Processing (OLTP)

- **OLTP** е режим на обработка, който се характеризира с кратки транзакции, записващи бизнес събития
- Изисква висока надеждност, наличност и съгласуваност на данните и кратко време за реакция.
- Приложенията изискват на заявка за обслужване да се отговори в рамките на предвидим период, който се доближава до "реално време".

# Видове бази от данни

- Аналитични бази от данни (складове от данни)
  - Съхраняват статични данни
  - Данните са свързани с различни периоди от време и осигуряват информация в историческа перспектива (напр. последните 5-10 години).
  - Използват като основен източник операционните бази от данни
- Подпомагат вземането на решения данните
  - **OLAP (Online Analytical Prococesing)** – Аналитична интерактивна обработка на данни, при която данните се представят в различни разрези
  - **Data Mining** – Анализ на данни с цел откриване на неявни или непредполагаеми връзки и шаблони в тях (процес откриване на знания)

# Потребители на базата от данни

- Администратори на базата от данни
- Проектанти (дизайнери) на базата от данни
- Разработчици на приложения
- Системни аналитици
- Крайни потребители

# Основни характеристики на БД:

- Устойчивост на данните – относително постоянни във времето факти за дейността на организацията
- Минимално дублиране на данните
- Интегрираност на данните - съхранява се информация за връзките между данните
- Реализиране на различни модели на данните и различни потребителски представи



# Основни характеристики на БД:

- Интегритет (цялост) на данните – вярност, точност и непротиворечивост на данните.
- Поддържане описание на данните отделно от данните в т.н. речник на данните (съхраняват се данни за данните, наречени метаданни).
- Независимост на данните от програмите.

## Тристепенна ANSI-SPARC архитектура на базата от данни

- Приета през от 1975г. от ANSI/SPARC (American National Standards Institute/ Standards Planning and Requirement Committee)
- Моделът осигурява основно разбиране на основните функционалности на системите за управление на бази от данните.

# Тристепенна ANSI-SPARC архитектура на базата от данни

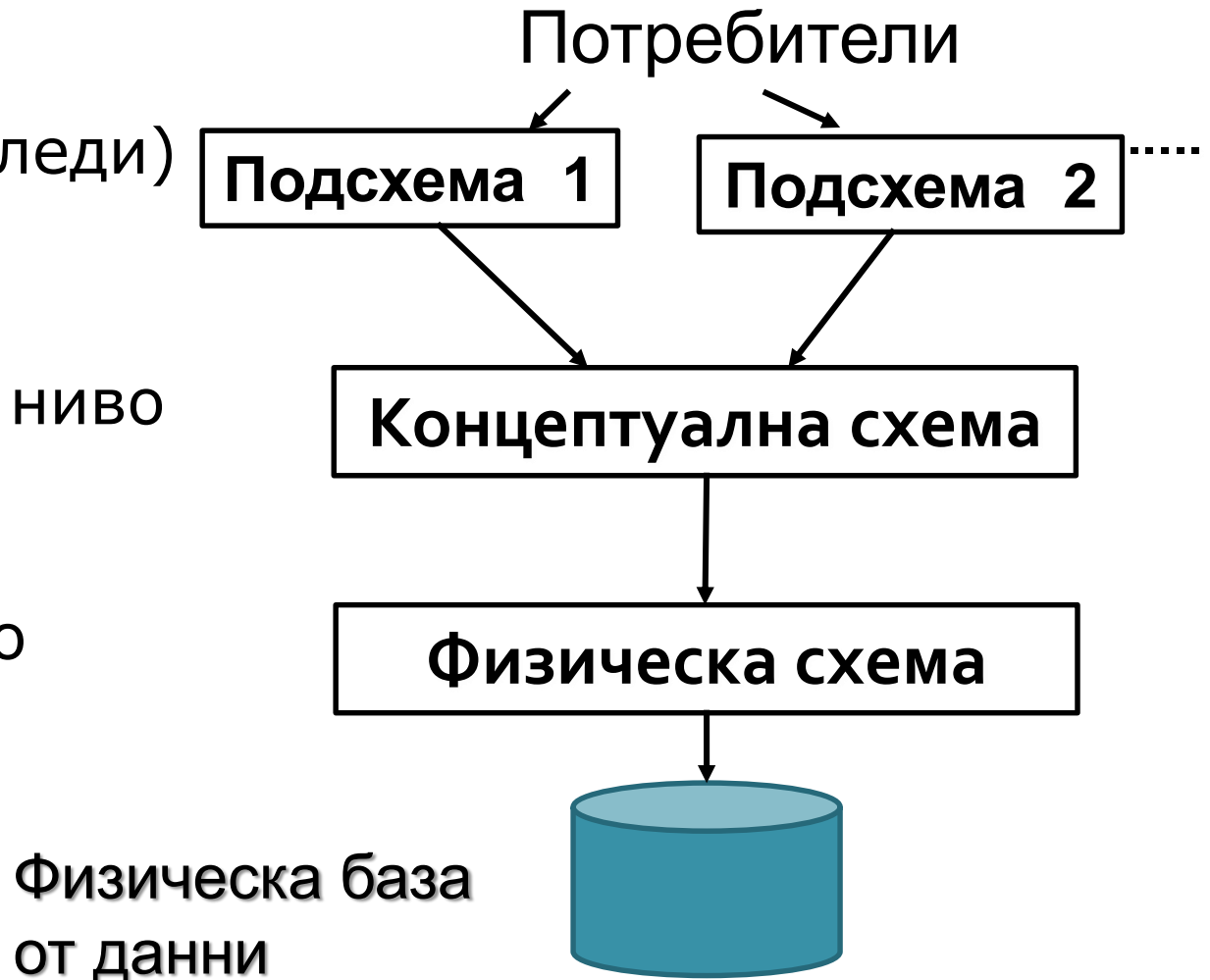
- Осигурява:
  - Независимост на данните от приложенията
  - Поддържане на различни потребителски представи за данните
- Постига се чрез представяне на БД на три нива:
  - **Външно ниво** (External Level) – вижданията за данните от различните потребители
  - **Концептуално ниво** (Conceptual Level) – представяне на данните, независимо от физическата реализация
  - **Вътрешно (физическо) ниво** (Internal Level) – организирането на данните върху физическата среда на съхранение

# Схеми на базата от данни

- Общото цялостно описание на базата от данни се нарича **схема на базата от данни**.
- Схемата се специфицира в процеса на проектиране на базата от данни и се очаква да не се променя често.
- Съществуват три различни типа схеми на базата от данни, които се дефинират съответно на трите нива на архитектурата.

# ANSI/SPARC архитектура на базата от данни

- Външно ниво (подсхеми/изгледи)
- Концептуално ниво
- Вътрешно ниво



## Външно ниво - подсхеми/изгледи

- На най-високо ниво се дефинират множество външни схеми, наречени подсхеми или изгледи.
- Подсхемите кореспондират на различните виждания (изгледи) за данните на различните потребители или група потребители.
- Различните подсхеми/изгледи могат да имат различни представяния на едни и същи данни.

# Пример:

## База от данни за университет University

- За потребителите от фронт офиса БД съдържа данни за Студенти, Преподаватели, Учебни дисциплини, Учебни планове, Изпити ...
- За потребителите от отдел „Човешки ресурси“ БД съдържа пълни данни за преподавателите и служителите и тяхната квалификация.
- За потребителите от отдел „Учебни дейности“ – Учебни планове, График на учебния процес, Преподаватели, Учебни групи, Зали ...

# Концептуално ниво

- Дава „общ поглед“ на данните, включващ описание на всички налични данни.
- Поддържа всеки външен изглед; наличните за даден потребител данни трябва да се съдържат или извличат от концептуалното ниво.
- Осигурява съответствие и желаната независимост между външното и вътрешното ниво.
- Не съдържа каквито и да е детайли за зависимост от физическата памет.



# Концептуална схема

- На концептуално ниво имаме **концептуална схема**, която описва всички същности, атрибути и връзки заедно с ограниченията за интегритет (цялост) на данните.
- Дадена съвкупност от данни притежава свойството **интегритет** (цялост), ако данните са логически съвместими, непротиворечиви и изпълняват определени ограничения.

# Вътрешно ниво - вътрешна схема

- На най-долното ниво имаме вътрешната схема, която е пълно описание на вътрешния модел.
- Вътрешната схема описва физическото представяне на базата от данни в компютъра. Описва как данните се съхраняват в базата от данни.
- Съдържа дефиниции на съхраняваните записи, на полетата с данни, организацията на файловете и на индексите, алгоритмите на хеширане и др.
- Имаме само една концептуална и една вътрешна схема за дадена база от данни.

# Независимост на данните

- Главна цел на тристепенната архитектура е осигуряване на независимост на данните, което означава, че по-горните нива не се влияят от промените в по-долните нива.
- Два вида независимост на данните:
  - **Логическа независимост на данните**
  - **Физическа независимост на данните**

# Логическа независимост на данните

- Логическата независимост на данните се отнася до имунитета (защитата) на външните схеми от промените в концептуаната схема.
- Промени в концептуалната схема, като добавяне или премахване на нови типове същности, атрибути или връзки е възможно без да налага промени в съществуващите външни подсхеми или пренаписване на приложните програми.
  - Напр. в БД University добавяте нови атрибути към таблица Students, но подсхемата за HRM отдела продължава да работи

# Физическа независимост на данните

- Физическата независимост се отнася до имунитета на концептуалната схема от промените във вътрешната схема.
- Промените във вътрешната схема, като използване на различна файлова организация, модифицирането на индексите или на алгоритъма за хеширане е възможно без да налага промени в концептуалната схема или във външните подсхеми.

# Дизайн на базата от данни

- Методологията на дизайна на базите от данни включва три основни фази:
  - Концептуален дизайн на базата от данни
  - Логически дизайн на базата от данни
  - Физически дизайн на базата от данни

# Модели на данните

- Моделите от данни описват структурите в базата от данни и връзките между тях.
- Класификация:
  1. High-Level Model
  2. Representation Model
  3. Low-Level Model

# High-Level Model

- Осигурява потребителските изисквания към данните. Независим от СУБД, която ще се използва за базата от данни.
- Чрез този модел се описват:
  - обектите, за които ще се съхраняват данни в базата от данни;
  - свойствата или характеристиките на обектите;
  - взаимоотношенията между обектите;
  - изискванията към данните.



# High-Level Model

- Използва се за концептуален дизайн на базата от данни – описание на потребителските представи за БД и концептуалната схема
- Разпространени модели от високо ниво за проектиране на бази от данни:
  - **Entity-Relationship Model (ERM);**
  - **Enhanced Entity–Relationship Model (EERM).**

# Representation Model (модел на представянето)

- Тези модели се използват за представяне на данните, съхранявани в базата от данни (логически дизайн на базата от данни).
- Структурите за съхраняване на данните се описва с терминологията на избраната СУБД.
- Описват логическите структури на базата от данни (напр., таблици, полета, изгледи)
- Класификация:
  - a. Йерархичен модел (Hierarchical model)
  - b. Мрежови модел (Network model)
  - c. Релационен модел (Relational model)

# Low-Level Model

- Моделите от т.н. ниско ниво описват физическите характеристики на базата от данни (физическия дизайн на базата от данни).
- Включва описание на файловете с данни, методи на организация, тип и размер на елементите от данни, размери на паметта за файловете и индексите, място на съхраняване, техники за компресиране и шифроване на данните ...

# Историческо развитие на технологията бази от данни

- Ранни приложения на базите от данни :
  - **Йерархични и Мрежови** модели – в средата на 1960-те и доминира през седемдесетте.
  - Information Management Systems (IMS) – 1968 г.
  - Integrated Data Store (IDS) (Ч. Бахман),
  - Комитет CODASYL DBTG – 1969, 1971 г.
- **Релационни бази от данни**
  - Оригиналният релационен модел е въведен през 1970 от специалистите на IBM (Е. F. Codd)
  - Релационните СУБД продукти се появяват през 1980-те години.

# Пострелационни бази от данни

- Обектно-ориентирани бази от данни OODBMS са въведени в края на 1980-те и началото на 1990-те за да задоволят нуждите от обработка на комплексни данни, CAD системите и други приложения.
- Обектно-релационни системи ORDBMS.
- Разширени релационни системи с допълнителни възможности за работа с мултимедийни данни, XML (eXtended Markup Language) и други типове данни
- **Web** бази от данни и приложения за е-бизнес
- **No-SQL** бази от данни

## Етапи на проектиране на базата от данни

- Планиране на БД и дефиниране на системата
- Анализ на предметната област
  - Формулиране на бизнес правилата и изискванията към данните
- Избор на СУБД
- Дизайн на БД
  - Концептуално проектиране
  - Логическо проектиране
  - Физическо проектиране

# Етапи на проектиране на базата от данни (продължение)

- Проектиране на приложенията
- Реализация
  - Създаване и инсталация на БД
  - Зареждане с данни
  - Тестване
- Развитие и поддържане

## *Жизнен цикъл на БД*

