

Бизнес интеллигентност (Business Intelligence - BI)

СКЛАДОВЕ ОТ ДАННИ
(DATA WAREHOUSE)

Бизнес интелигентност

- **BI** е набор от теории, методологии, архитектури и технологии, които трансформират суровите данни в смислена и полезна информация за бизнес цели.
- BI технологиите предоставят исторически, текущ и прогнозен поглед върху бизнес операциите.

ОСНОВНИ ПОНЯТИЯ

- Data Warehouse - Склад от данни
- Datawarehousing - Складиране на данни
- 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)

СКЛАД ОТ ДАННИ

- **Склад от данни** – база от данни, която съдържа данни, представляващи исторически архив на бизнеса в организацията
- Складът от данни съдържа както детайлни, така и силно обобщени данни
- Източници на данни за склада от данни – оперативни БД, файлове – изходи от други приложения

Складиране на данни

- Процес на интегриране на големи по обем корпоративни данни в единствено хранилище, наречено склад от данни.
- Обединяват се различни информационни източници, като се уеднаквяват, преформатират и почистват преди да се съхранят в склада от данни.

СКЛАД ОТ ДАННИ

- “Складът от данни е предметно-ориентиран, интегриран, приспособен към времето и непроменлив набор от данни, предназначен за подпомагане вземането на решения” [\[1\]](#).

[\[1\]](#) Bill Inmon (“The Father of Data Warehousing.”)

Предметна ориентираност

- Организиран около основни обекти в предметната област, като клиенти, продукти, продажби.
- Фокусира се върху моделирането и анализа на данните за вземане на решения, а не върху ежедневните операции или транзакции, като се изключват данните, които не са полезни за процеса на вземане на решения.

Интегрираност

- Интегрирани данни от множество хетерогенни източници на данни:
 - Релационни бази от данни, плоски файлове, файлове от други приложения.
- Прилагат се техники за пречистване и интегриране на данните
 - Осигурява се устойчивост в конвенциите за имената, структура на кодовете, мерни единици и т.н. на данните от различните източници;
 - При преместване на данните в склада от данни те се конвертират.

Свързаност с времето

- Времевият хоризонт на склада от данни е значимо по-дълъг от този при оперативните системи
 - Оперативни бази от данни: текущи стойности на данните.
 - Складове от данни: осигурява информация в историческа перспектива (напр. последните 5-10 години).
- Във всяка ключова структура на склада от данни се съдържа елемент време изрично или по подразбиране.
- Ключовете на оперативните данни могат да не съдържат елемент “време”.

Непроменливост

- Физически отделено хранилище на данни, трансформирани от оперативната среда.
- В склада от данни операция обновяване на данни не се прави.
- Необходими са само две операции при достъпа до склада от данни:
 - *Първоначално зареждане на данни;*
 - *Достъп до данните.*

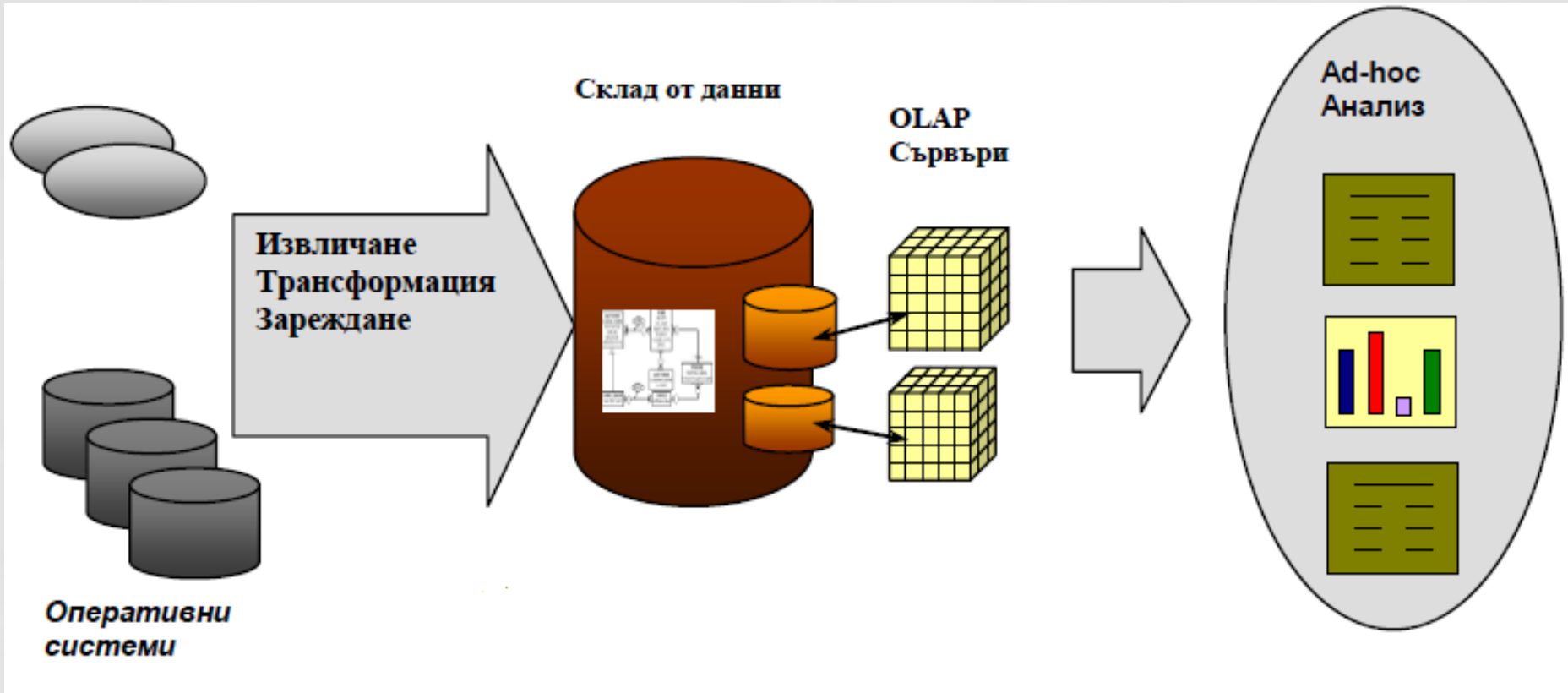
Преимущества на DW

- Интегриране на всички данни, свързани с дадена организация.
- Възможност за достъп до съхранени исторически данни.
- Ефективно управление както на детайлизирани, така и на обобщени данни.
- Повишаване ефективността при вземане на корпоративни решения.

Data Marts – магазини (витрини) от данни

- Съхраняват подмножества от обобщени данни.
- Предназначени за определен отдел от предприятието (маркетинг, финанси и т.н.).
- Могат да се разглеждат като локални складове от данни.

Архитектура на склада от данни



Физическа архитектура на склада от данни

- Трислоен клиент-сървър:
 - Сървър на склада от данни – обслужва базите от данни в склада; обикновено е релационна СУБД
 - Сървър на OLAP – с релационен или многомерен модел на данните
 - Клиенти - инструментални средства за анализ на данните (Ad-hoc)

Инструментални средства за анализ на данните

- Средства за генериране на стандартни отчети.
- Нерегламентирани заявки.
- Многомерен анализ - OLAP средства.
- Извличане на знания - Data mining tools.

OLAP (OnLine Analytical Processing)

- Технология , чрез която се използва информацията, съхранена в склада от данни.
- Съвкупност от концепции, принципи и изисквания, лежащи в основата на програмните продукти, осигуряващи достъп до данните за целите на техния анализ по няколко параметъра.

FASMI изисквания към OLAP средствата

- **Fast (Бърз)**. Минимално време за достъп до аналитичните данни.
- **Analysis (Анализ)**. Осигурява възможност на потребителя да осъществява числов и статистически анализ.
- **Shared (Разпределен достъп)**. Предоставя възможност за работа с информацията от много потребители едновременно.
- **Multidimensional (Многомерност)**.
- **Information (Информация)**. Дава възможност на потребителя да получи необходимата информация, независимо в кой склад от данни се намира.