

АРХІТЕКТУРА OLAP-СИСТЕМИ

Лекція 7

Автор Голуб Б.Л.



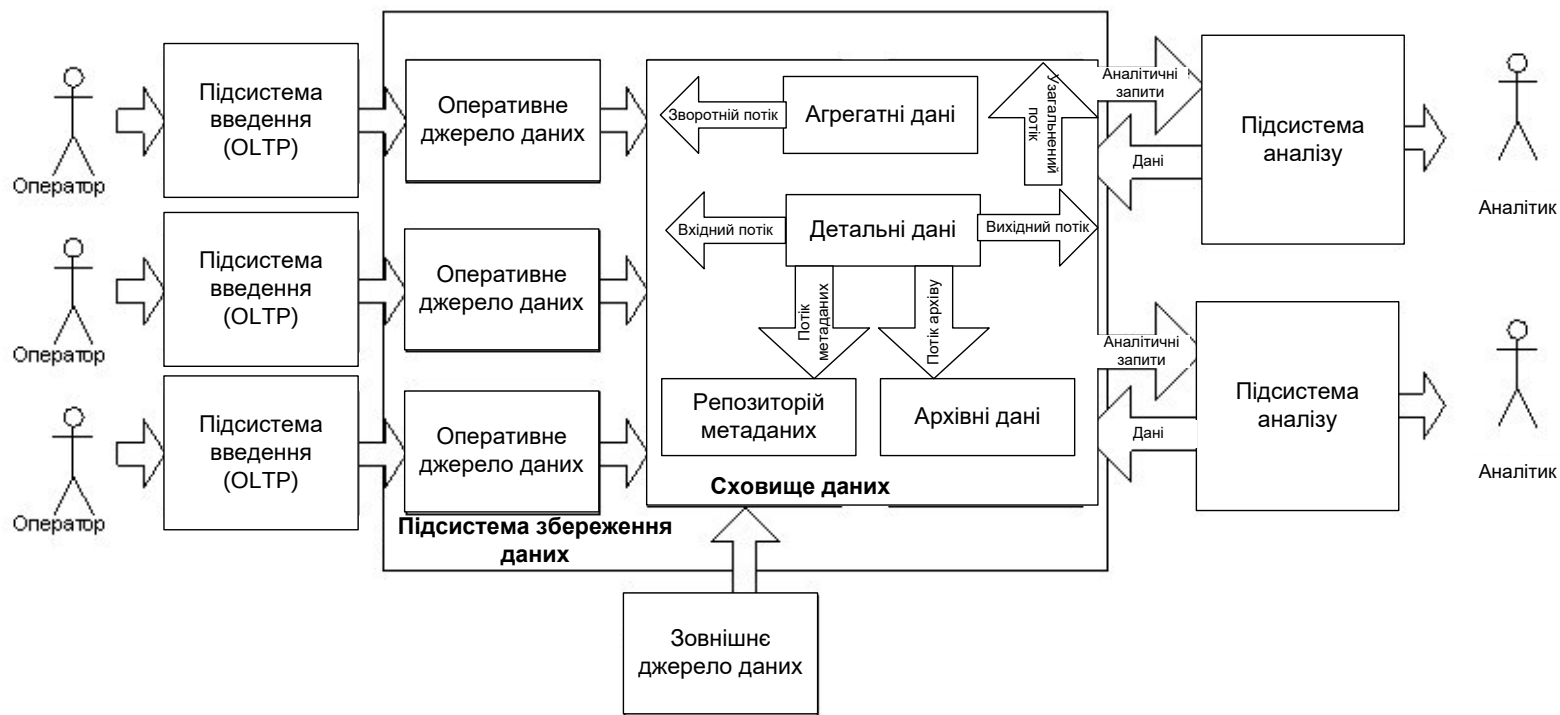
ПОНЯТТЯ OLAP



OLAP (*Online Analytical Processing*) - технологія оперативної аналітичної обробки даних, що використовує методи і засоби для збору, зберігання та аналізу багатовимірних даних з метою підтримки процесів прийняття рішень

- Основне призначення **OLAP**-систем - підтримка аналітичної діяльності, довільних запитів користувачів - аналітиків
 - Мета **OLAP**-аналізу - перевірка гіпотез
- 

КАТЕГОРІЇ ДАНИХ СД



Детальні дані

Детальні дані - дані, що переносяться безпосередньо з **OLTP**-підсистем. Відповідають елементарним подіям, фіксуються в **OLTP**-системах.

Поділяються на:

* *Виміри* - набори даних, необхідні для опису подій (товар, продавець, покупець, магазин, ...)

* *Факти* - дані, що відображають сутність події (кількість проданого товару, сума продажів, ...)



УЗАГАЛЬНЕНІ ДАНІ



*Агреговані (узагальнені)
дані* - дані, одержані на
підставі детальних
шляхом підсумовування за
певними вимірами,
спеціальної обробки
шляхом знаходження
середніх значень,
екстремальних значень
тощо

МЕТАДАНИ

Метадані - дані про дані, що містяться в СД. Можуть описувати:

* *Причини*
виконання різних дій над даними

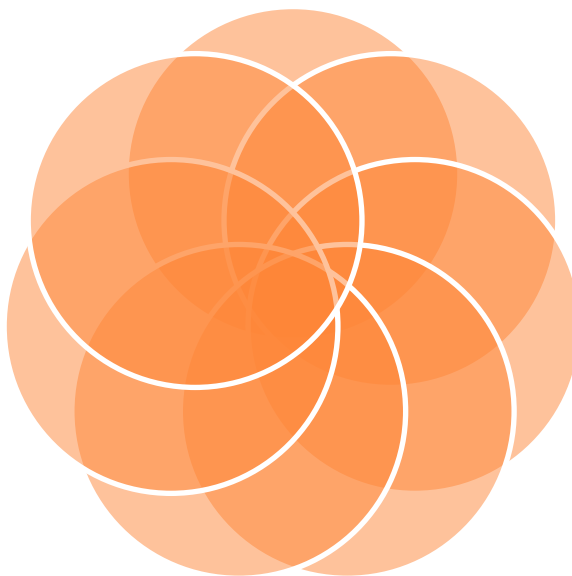
* *Час* виконання різноманітних дій над даними

* *Дії*, що виконуються над даними

* *Об'єкти*
предметної області, інформація про яких міститься в СД

* *Категорії*
користувачів, що використовують дані в СД

* *Місця і способи*
зберігання даних



ІНФОРМАЦІЙНІ ПОТОКИ В СД

- **Вхідний потік** - утворюється даними, копійованими з OLTP-систем в СД; дані при цьому часто очищаються і збагачуються шляхом додавання нових атрибутів
- **Потік узагальнення** - утворюється агрегуванням детальних даних та їх збереженням у СД
- **Архівний потік** - утворюється переміщенням детальних даних, кількість звернень до яких знизилася
- **Потік метаданих** - утворюється потоком інформації про дані в репозиторій даних
- **Вихідний потік** - утворюється даними, що отримують користувачі
- **Зворотний потік** - утворюється очищеними даними, які записуються в OLTP-системи



Вимоги

Повномасштабна **OLAP**-система повинна виконувати складні і різноманітні функції, що включають збір даних з різних джерел, їх узгодження, перетворення і завантаження в сховище, зберігання аналітичної інформації, регламентну звітність, підтримку довільних запитів, багатомірний аналіз та ін.



ШАРИ OLAP-СИСТЕМИ

У загальному вигляді архітектура корпоративної OLAP-системи описується схемою з трьома виділеними шарами:

- *Шар вилучення, перетворення і завантаження даних*
- *Шар зберігання даних*
- *Шар аналізу даних*



ОРГАНІЗАЦІЙНА ТОЧКА ЗОРУ

Шар включає підрозділи і структури організації всіх рівнів, що підтримують бази даних оперативного доступу.

Він являє собою низовий рівень генерації інформації, рівень внутрішніх і зовнішніх інформаційних джерел, що виробляють "сиру" інформацію. Ця інформація є робочою для повсякденної діяльності різних підрозділів, які її виробляють і використовують.

СИСТЕМНА ТЕХНІЧНА ТОЧКА ЗОРУ

З системно-технічної
точки зору шар
представлений ЛОС усіх
підрозділів усіх рівнів, до
яких підключені
спеціалізовані технічні
комплекси, що
зберігають інформацію
Найчастіше реалізовані у
вигляді реляційних
СУБД

ПРОЦЕДУРИ ЗАВАНТАЖЕННЯ

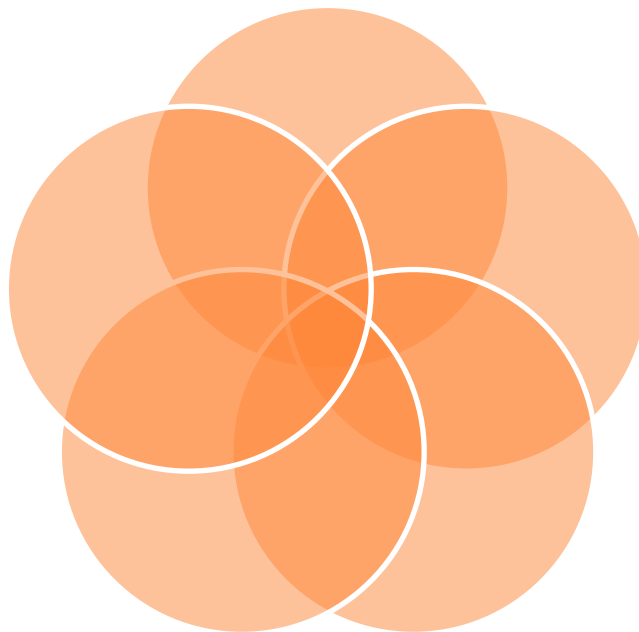
Завантаження
даних з джерел в
сховище
здійснюється
спеціальними
процедурами, що
дозволяють:

витягати дані з
різних баз даних,
текстових файлів;

перетворювати
дані при
переміщенні їх
від джерел до
сховища;

виконувати різні
типи узгодження і
очищення даних;

завантажувати
узгоджені і
"очищені" дані в
структуру
сховища.



ШАР ЗБЕРІГАННЯ ДАНИХ

Призначений
безпосередньо для
зберігання значущою,
перевіреної, узгодженої,
несуперечливої і
хронологічно цілісної
інформації, яку з досить
високим ступенем
впевненості можна
вважати достовірною



ШАР АНАЛІЗУ ДАНИХ

Для організації доступу аналітиків до даних СД і ВД використовуються спеціалізовані робочі місця, підтримують необхідні технології як для оперативного, так і довготривалого аналізу

Результати роботи аналітиків оформляються у вигляді звітів, графіків, рекомендацій, і зберігаються як на локальному комп'ютері, так і в загальнодоступному вузлу локальної мережі.

Основні види аналітичної діяльності:

- стандартна звітність;
- нерегламентовані запити;
- багатомірний аналіз (OLAP);
- вилучення знань (data mining).

