

ЛЕКЦІЯ. ВСТУП ДО БАЗ ДАНИХ

План:

1. Технології баз даних.
2. Компоненти банків даних.
3. Компоненти систем баз даних.

Література

1. Гайна Г.А. Основи проектування баз даних: Навчальний посібник. К.: КНУБА, 2005. – 204 с.
2. Гайна Г.А. Організація баз даних і знань. Мови баз даних: Конспект лекцій.–К.:КНУБА, 2002. – 64 с.
3. Гайна Г.А., Попович Н.Л. Організація баз даних і знань. Організація реляційних баз даних: Конспект лекцій.–К.:КНУБА, 2000. – 76 с.

1. Технології баз даних

Виникнення технологій баз даних припадає на початок 60-х років. Їх швидкому розвитку сприяли потреби в обробці інформації, досягнення в суміжних областях інформаційних технологій таких, як операційні системи, мови програмування, технічне забезпечення. Спочатку зароджувалися певні ідеї щодо управління ресурсами даних, формувалися основи методології побудови систем баз даних. Із самого початку було зрозуміло, що цей напрям має самостійне значення і буде відігравати одну з ключових ролей у побудові інформаційних систем різного призначення.

Інформаційні задачі на відміну від обчислювальних мають такі особливості:

- збереження даних складної структури;
- відносно прості алгоритми обробки;
- великі обсяги оброблюваної інформації.

Інформаційна система виконує функції збирання, зберігання, розповсюдження і обробки інформації. Під *інформацією* розуміють будь-які відомості про будь-яку подію, сутність, процес і т.ін., які є об'єктом певних операцій: передачі, перетворення, зберігання або використання. *Дані* можна визначити, як інформацію зафіксовану у певній формі, яка придатна для

подальшої обробки, зберігання і передачі (рис. 1.1). *Предметна область* – область застосування конкретної бази даних.

Інформації відповідає *інфологічне представлення*, яке розглядає питання пов'язані зі змістом інформації, незалежно від представлення її в пам'яті комп'ютера. *Даталогічне представлення* розглядає питання представлення даних в пам'яті комп'ютера.

Функції управління інформацією в інформаційних системах виконують системи управління базами даних.

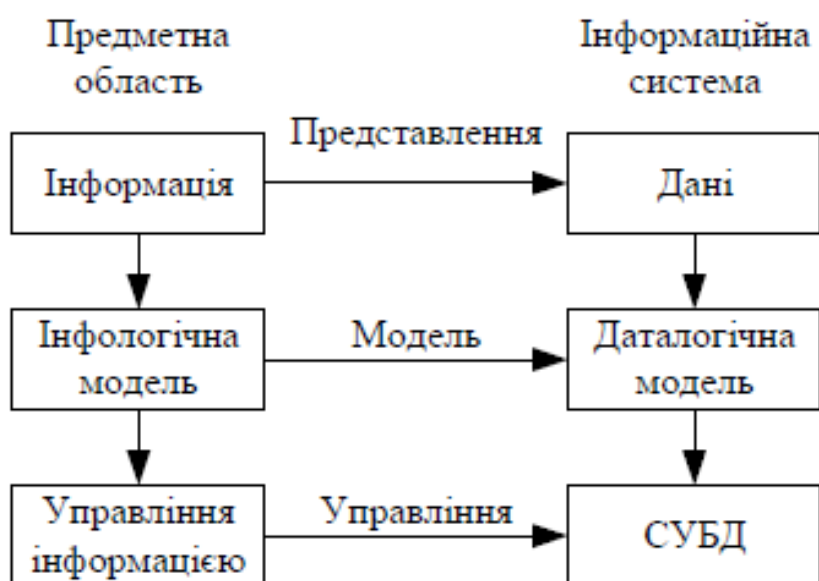


Рис. 1.1. Зв'язок між інформацією і даними

Спочатку для збереження даних застосовувалися *файлові системи* (рис. 1.2). Для цих систем були характерні такі особливості:

–□структура запису файла даних була відома тільки прикладній програмі, яка з ним працювала, а система управління файлами її не знала; кожна програма, яка працювала з файлом даних, повинна була мати у себе структуру даних, яка відповідала цьому файлу (така ситуація характеризувалась, як *залежність програм від даних*);

–□система управління файлами повинна була забезпечити авторизацію доступу до файлів для різних користувачів, але були відсутні централізовані методи управління доступом до інформації;

–□ для організації паралельної роботи користувачів з даними необхідне узгоджене управління, а система управління файлами при цьому працювала надто повільно.

Для подолання цих недоліків практично паралельно почалися роботи над ієрархічними і мережними базами даних (БД) і над відповідними системами управління цими БД СУБД. В основі цих БД лежали відповідні моделі даних: ієрархічні і мережні (рис. 1.2).

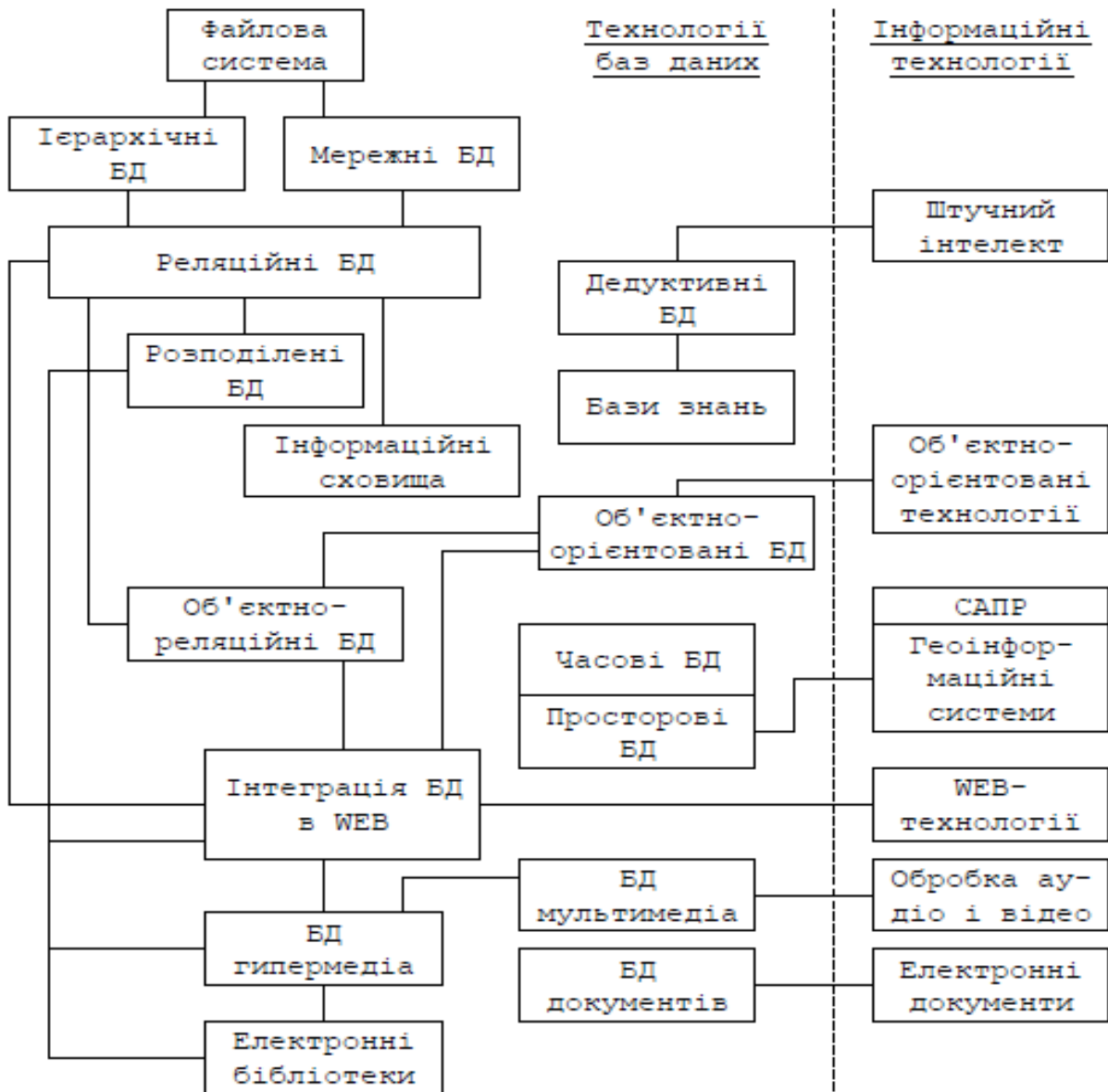


Рис. 1.2. Етапи розвитку баз даних

Ці моделі належать до теоретико-графових моделей. До переваг цих БД можна віднести ефективне використання пам'яті комп'ютера, швидке виконання основних операцій над даними.

Недоліками цих БД є таке:

- □ труднощі при обробці інформації з достатньо складними зв'язками;
- □ складність розуміння змісту звичайним користувачем, залежність від фізичної реалізації.

На початку 70-х років були сформовані теоретичні основи сучасних технологій БД і остаточно сформувався самостійний напрям інформаційних технологій – наука о базах даних. Головною подією цього періоду стало виникнення *реляційних* баз даних. В основі реляційної моделі лежить поняття відношення. Реляційна модель є простою і зрозумілою, а її фізичне представлення добре реалізується на комп'ютері. До недоліків реляційних моделей можна віднести складність реалізації ієрархічних і мережних зв'язків.

У 70-х роках почали формуватися підходи в БД, які пов'язані з використанням апарата логіки в якості моделі даних. Ці роботи привели до створення *дедуктивних* БД. Розвиток цього напрямку дозволяє створювати *бази знань*.

У 80-х роках реляційні БД набули домінуючого положення. Однак уже тоді існувало і постійно розширювалося коло застосувань, для яких ця технологія була неадекватною. Це стосується мультимедійних застосувань, систем, які оперують просторовими даними і т.ін. В середині 80-х років успіхів досягло об'єктно-орієнтоване програмування. Під його впливом і в зв'язку з необхідністю реалізації нетрадиційних застосувань, вимоги яких погано узгоджуються з можливостями реляційних систем, почались роботи з практичної реалізації *об'єктно-орієнтованих* БД. Логічна структура об'єктно-орієнтованої БД зовні схожа на структуру ієрархічної БД, але вона доповнена об'єктно-орієнтованими механізмами. Об'єктно-орієнтовані БД у порівнянні з реляційними БД мають можливість відображати інформацію складних взаємозв'язках об'єктів, визначати функції обробки окремих записів. До недоліків об'єктно-орієнтованих моделей належить висока понятійна складність, низька швидкість виконання запитів, незручність обробки даних.

Бажання розробників реляційних БД зберегти лідируючі позиції і підвищити ефективність реляційної моделі, привели до розробки об'єктно-реляційної моделі, на основі якої почали розроблятися *об'єктно-реляційні* БД. В

цих моделях припускається застосування багатозначних полів – полів, які складаються з підзначень. Ці БД дозволяють забезпечити високу наглядність представлення інформації і підвищити ефективність її обробки. До недоліків об'єктно-реляційних моделей можна віднести складність рішення забезпечення цілісності і несуперечливості даних.

На початку 90-х років почало збільшуватися значення інформаційного забезпечення систем підтримки прийняття рішень. Це привело до необхідності створення багатомірних СУБД і на їх основі розробки *інформаційних сховищ*. Ці системи використовують історичну інформацію, яка представляється в агрегованому вигляді. На основі цієї інформації виконується аналітична обробка, прогнозування даних, а також інтелектуальний аналіз даних.

В цей час також велика увага приділялася розвитку розподілених систем. Успіхи в розробці комп'ютерних мереж стимулювали дослідження в технологіях *розподілених БД*, були розроблені архітектурні концепції *клієнт – сервер*.

Одним з найбільших досягнень 90-х років в області інформаційних технологій стало створення відкритої глобальної розподіленої неоднорідної гіпермедійної інформаційної системи, яка використовує комунікаційну мережу Internet. Ця система отримала назву WWW (Web). З самого початку виконувались спроби інтегрувати системи БД у Web. Одним з напрямів роботи є *інтеграція структурованих даних БД і слабкоструктурованих даних Web*, проводяться роботи зі створення БД на мові XML.

У даний час в багатьох комп'ютерних компаніях здійснюються роботи зі створення *цифрових бібліотек* – інформаційних систем, які призначені для зберігання, пошуку, обробки, аналізу і розповсюдження інформаційних ресурсів різної природи – структурованих і слабо-структурованих даних, мультимедійної інформації, повнотекстових документів. Для створення таких систем використовуються Web-технології, розробляються методи інтеграції неоднорідних ресурсів, розпізнавання і пошуку інформаційних ресурсів.

2. Компоненти банків даних

Банк даних – це система спеціальним чином організованих даних (баз даних), програмних, мовних, технічних, організаційно-методичних засобів призначених для підтримки інформаційної моделі предметної області з метою забезпечення інформаційних потреб користувачів (рис. 1.3).

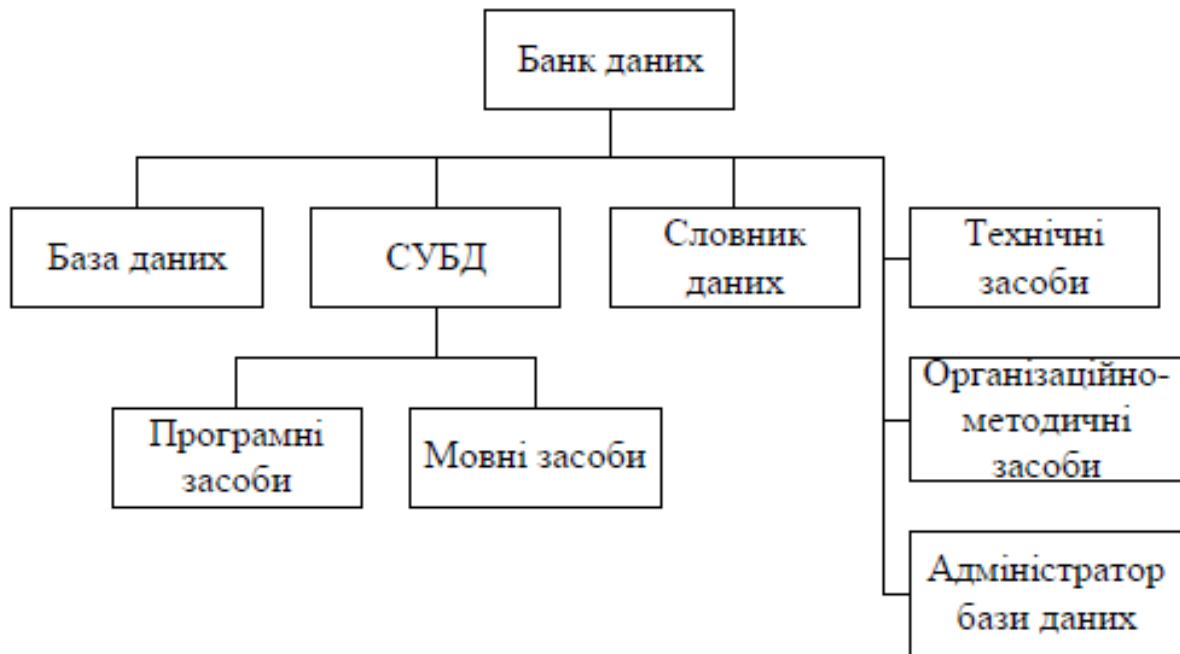


Рис. 1.3. Банк даних

База даних – поїменована сукупність взаємозв'язаних даних, які знаходяться під управлінням СУБД. В БД зберігаються дані, логічно пов'язані між собою. До головних властивостей БД належать такі:

- □ *цілісність* означає, що в будь-який момент часу відомості в БД повинні бути несуперечливі;
- □ *безпека* означає, що виконується захист даних від санкціонованого і несанкціонованого доступу;
- □ *відновленість* означає можливість відновлення БД після збоїв роботи системи.

Система управління базами даних (СУБД) – сукупність мовних і програмних засобів, які призначені для створення, ведення і сумісного використання БД багатьма користувачами.

До головних функцій СУБД належать такі:

–□ управління даними у зовнішній пам'яті і буферами оперативної пам'яті;

–□ управління транзакціями і паралельним доступом;

–□ відновлення БД;

–□ підтримка мов БД;

–□ контроль доступу до даних;

–□ підтримка цілісності даних;

–□ підтримка незалежності даних;

–□ підтримка обміну даними.

Склад БД містить не тільки дані, що зберігаються, але і опис БД. Опис БД належить до *метаінформації*, тобто інформації про інформацію. Опис БД часто називають *схемою*.

Централізоване сховище метаінформації називається словником даних.

Словник даних (каталог даних) – використовується для централізованого накопичення і опису ресурсів даних. Словник даних відповідає за визначення всіх елементів даних:

–□ імена, типи і розміри елементів даних;

–□ імена зв'язків;

–□ обмеження даних по підтримці цілісності;

–□ схеми БД (зовнішня, концептуальна і внутрішня), а також відображення між ними;

–□ імена користувачів і їх права доступу до даних;

–□ статистична інформація.

Програмні засоби БД включають в свій склад ядро СУБД, транслятори, утіліти, прикладні програми.

Мовні засоби поділяються на мови опису даних (МОД) і мови маніпулювання даними (ММД). МОД призначені для опису схеми БД або її частини. З її допомогою виконується опис типів даних, їх структур і зв'язків між собою. Відповідно до отриманого опису СУБД знаходить в програмі необхідні дані, перетворює їх і передає в прикладну програму. ММД виконує

функції вибірки з БД даних за певними умовами, зміну даних, додавання даних, вилучення даних і т.ін.

Адміністратор даних – людина, яка відповідає за управління даними (планування БД, розробку і супроводження стандартів, прикладних алгоритмів і ділових процедур), а також за концептуальне і логічне проектування БД.

Адміністратор БД – людина, яка відповідає за фізичну реалізацію БД (фізичне проектування і втілення проекту), за забезпечення безпеки і цілісності даних, за супроводження операційної системи, а також за забезпечення максимальної продуктивності застосувань і користувачів.

Адміністратор даних і адміністратор БД виконують функції: управління структурою БД, управління паралельною обробкою, розподіл прав і обов'язків при обробці, забезпечення безпеки БД, відновлення БД, управління СУБД, підтримка репозиторія даних.

Адміністрування даними і БД передбачає управління інформаційними ресурсами, проектування БД, управління реалізацією застосувань, підтримку цілісності даних, захист даних, спостереження за поточною продуктивністю системи, а також реорганізацію БД при необхідності.

Переваги і недоліки застосування СУБД наведені в табл. 1.1.

Переваги і недоліки застосування СУБД

Переваги СУБД	Недоліки СУБД
Мінімізація збитковості даних	Використання значної частини ресурсів на потреби СУБД, а не на прикладну задачу
Несуперечливість даних і контроль їх цілісності	Вартість СУБД
Незалежність прикладних програм від даних	Підвищені вимоги до технічного і програмного забезпечення
Підвищена безпека	Продуктивність
Розвинені засоби резервного копіювання і відновлення	Підвищені вимоги до кваліфікації робітників
Багатокористувацький режим роботи	Наслідки збоїв

3. Компоненти системи баз даних

Компонентами системи баз даних є БД, СУБД і прикладні програми, з якими працюють як розробники, так і користувачі.

В СУБД входять такі компоненти (рис. 1.4): ядро СУБД, підсистема засобів проектування і підсистема засобів обробки.

Ядро СУБД – містить сукупність базових механізмів СУБД, які використовуються при будь-яких варіантах конфігурації системи. Ядро СУБД виконує функцію посередника між підсистемами засобів проектування і обробки і даними. Сучасні БД у більшості представляють користувачу дані у вигляді таблиць. Ядро СУБД отримує запити від інших компонентів в термінах таблиць, стовпців, рядків і перетворює ці запити в команди операційної системи, які виконують запис і читання з фізичних носіїв інформації.

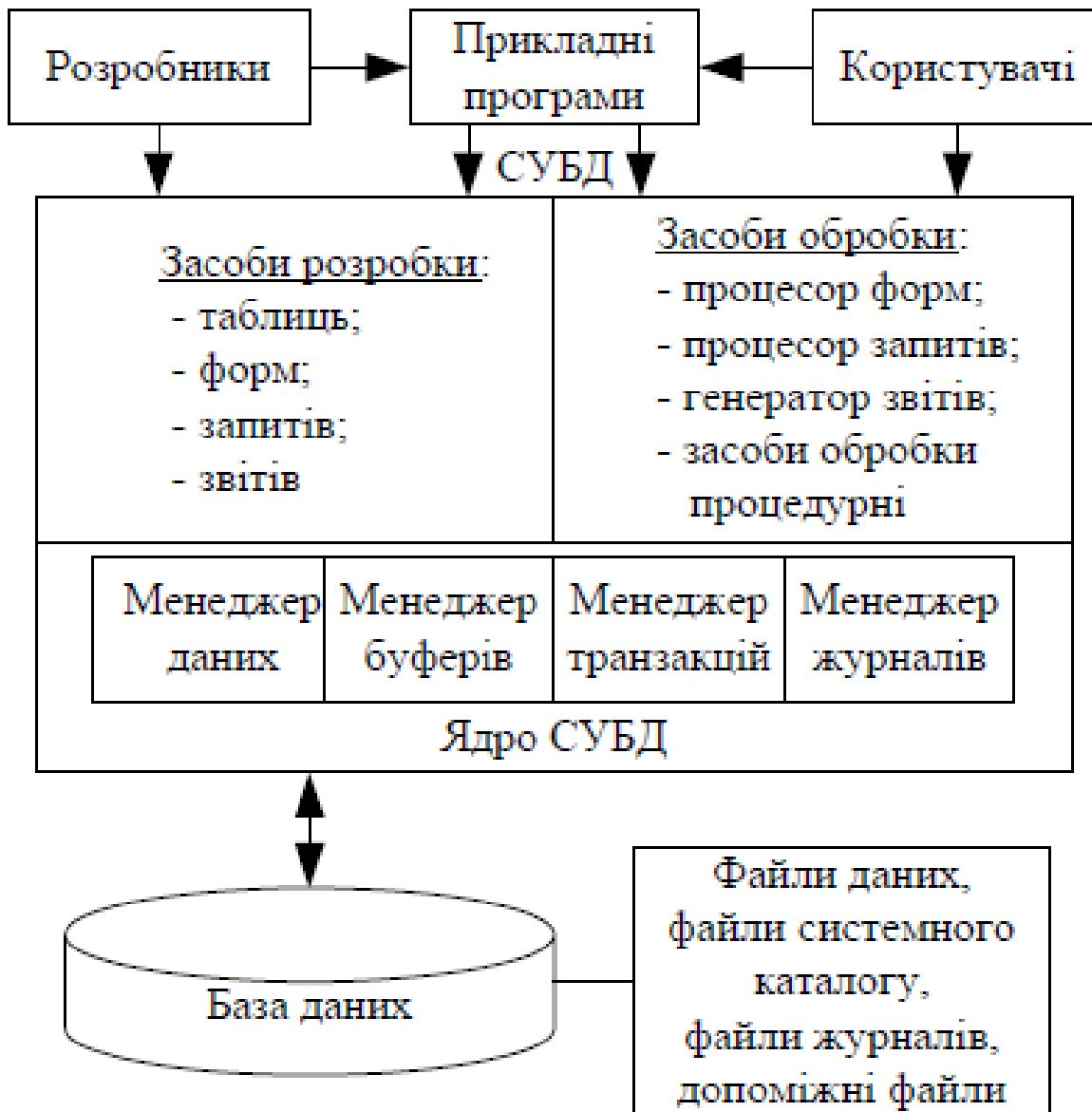


Рис. 1.4. Компоненти системи бази даних

Крім того, ядро СУБД задіяне в управлінні транзакціями, блокуванні, резервному копіюванні і відновленні. В ядро СУБД входять менеджери буферів, даних, транзакцій, журналів.

Менеджер буферів – призначений для рішення задач ефективної буферизації оперативної пам'яті.

Менеджер даних – призначений для управління зовнішньою пам'яттю, забезпечення створення структур для даних, що зберігаються і допоміжних структур (індекси і т.ін.).

Менеджер транзакцій – підтримує механізми фіксації і відкату транзакцій, пов'язаний з менеджером буферів оперативної пам'яті і забезпечує зберігання всієї інформації, яка потрібна після збоїв системи.

Менеджер журналів – забезпечує реєстрацію відомостей про виконання транзакцій, про працюючих користувачів, про виконання застосування, про доступи до різних структур даних і т.ін.

Підсистема засобів проектування являє собою набір інструментів, які спрощують проектування і реалізацію баз даних і їх застосувань. Як правило, цей набір містить засоби для створення таблиць, форм, запитів й звітів. В СУБД є також мови програмування і інтерфейси до них.

Підсистема обробки здійснює обробку компонентів застосування, які створені за допомогою засобів проектування.

Застосування БД складається з форм, запитів, звітів, меню і прикладних програм. Форми, запити і звіти можна створювати за допомогою засобів, що постачаються у комплекті з СУБД. Прикладні програми повинні бути написані або на вхідній мові СУБД, або на одній зі стандартних мов, а потім за допомогою СУБД з'єднані з БД.

Контрольні запитання

1. Дати визначення таких термінів: *інформація, інформаційна система, дані, предметна область.*
2. Дати визначення *бази даних і СУБД.*
3. Пояснити, чому база даних є моделлю. Яка існує різниця між реальністю і моделлю реальності?
4. Перелічити основні функції СУБД.
5. Перелічити переваги і недоліки СУБД.
6. Дати визначення словника даних.
7. Назвати основні компоненти системи бази даних і пояснити функцію кожної з них.
8. Перелічити основні етапи розвитку технологій баз даних.
9. Що таке банк даних?
10. Які складові частини містить банк даних?
11. Кого називають адміністратором бази даних?
12. Що таке ядро СУБД?
13. Які підсистеми входять в СУБД?