



Лекція 1 - Огляд баз даних

Огляд модулів:



Лекції: 32 год.

Лабораторні: 32 год.

Консультації: 2 год.

Іспити: 2 год.

Дані VS Інформація

Говорячи про системи управління базами даних, майже неминуче стикаємося з поняттями інформації та даних, які часто використовуються як взаємозамінні. Це відбувається, навіть якщо вони мають ключові відмінності, що їх відрізняють.

Дані є необробленим, неаналізованим, неорганізованим, не пов'язаним, безперервним матеріалом, який використовується для отримання інформації після аналізу.

Дані: називаються факти щодо об'єктів та подій, що використовуються та обробляються для довідки чи аналізу, які можна записати та зберегти на комп'ютерних носіях.

Дані не є інформацією - вони стають інформацією, коли їх обробляють, інтерпретують, організовують, структурують або представляють.



Дані VS Інформація



Інформація: це дані, які були оброблені таким чином, щоб підвищити знання особи, яка використовує дані.

Як ми вже заявляли, інформація походить із наборів даних, які обробляються значущим чином через контекст, актуальність та мету. Інформація надає значення та покращує надійність даних.

Основа для порівняння	Дані	ІНФОРМАЦІЯ
Значення	Дані є незміщеними фактами і цифрами і використовуються як вхідні дані для комп'ютерної системи.	Інформація - це вихід оброблених даних.
Характеристики	Дані - це окремі одиниці, які містять сировину і не мають ніякого значення.	Інформація - це продукт і група даних, які разом несуть логічний сенс.
Залежність	Це не залежить від інформації.	Вона спирається на дані.
Особливість	Нечіткі	Конкретні.
Вимірювальний блок	Вимірюється в бітах і байтах.	Вимірюється в значущих одиницях, таких як час, кількість і т.д.

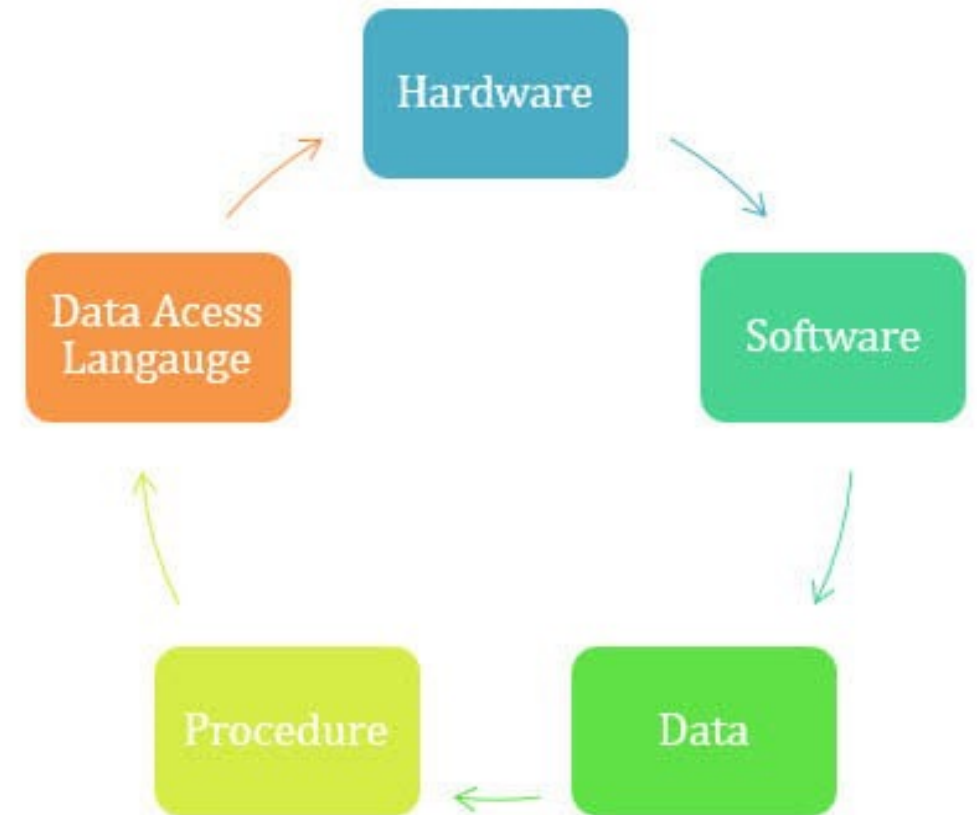
Що таке база даних?

- База даних - це організований набір даних.
- База даних - це сукупність інформації, яка організована таким чином, що до неї можна легко отримати доступ, керувати та оновлювати її.
- База даних — це набір пов'язаних даних, а дані — це набір фактів і цифр, які можна обробити для отримання інформації.

Переважно дані представляють факти, які можна записати.

Дані допомагають створювати інформацію, яка базується на фактах.

Наприклад, якщо у нас є дані про оцінки, отримані всіма студентами, ми можемо зробити висновок про найкращі та середні оцінки.



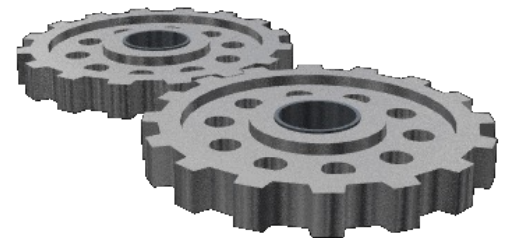
Дизайн Бази Даних

Для чого нам
дизайн?



Дизайн Бази Даних

- Визначення системних параметрів;
- Визначення робочих процесів;
- Побудова концептуальної моделі даних;
- Підготовка схеми бази даних;
- Проектування інтерфейсу користувача.



Дизайн Бази Даних

Метою концептуального проекту є отримання належного представлення корпоративних ресурсів даних, незалежно від рівня реалізації, а також конкретних потреб кожного користувача або додатка.

Завдання логічного проектування полягає у перетворенні концептуальної схеми шляхом її адаптації до моделі даних, що реалізує СУБД, яка буде використовуватися (як правило, реляційна).

Загальні міркування щодо проектування

- Користувачі;
- Системи / Дані;
- Вимоги до продукту;



Дизайн Бази Даних

Типова СУБД має користувачів з різними правами та дозволами, які використовують її для різних цілей. Деякі користувачі отримують дані, а деякі створюють їх резервні копії.



Користувачі:

Хто вони (рівень)?

- Адміністративний
- Науковий
- Технічний

Вплив

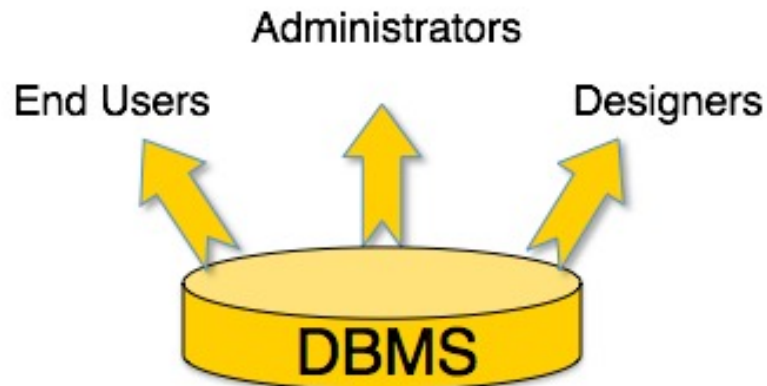
- Контроль доступу
- Інтерфейси
- Рівні обслуговування

Дизайн Бази Даних

Користувачі:

Адміністратори – адміністратори обслуговують СУБД і відповідають за адміністрування бази даних. Вони відповідають за його використання та за те, хто його має використовувати. Вони створюють профілі доступу для користувачів і застосовують обмеження для підтримки ізоляції та забезпечення безпеки. Адміністратори також піклуються про такі ресурси СУБД, як ліцензія на систему, необхідні інструменти та інше обслуговування, пов'язане з програмним і апаратним забезпеченням.

Дизайнери – це група людей, які фактично працюють над проектуванням бази даних. Вони уважно стежать за тим, які дані потрібно зберігати та в якому форматі. Вони ідентифікують і проектують увесь набір сутностей, відносин, обмежень і представлень.



Кінцеві користувачі – це ті, хто фактично отримує переваги від наявності СУБД. Кінцеві користувачі можуть варіюватися від простих глядачів, які звертають увагу на журнали чи ринкові ставки, до досвідчених користувачів, таких як бізнес-аналітики.

Дизайн Бази Даних

Системи / Дані



Які системи наразі діють?

Звідки беруться дані?

Як вони генерується?

У якому форматі здійснюється зберігання даних?

Для чого використовуються дані?

Які частини системи повинні залишатися статичними?

П'ять основних компонентів бази даних:

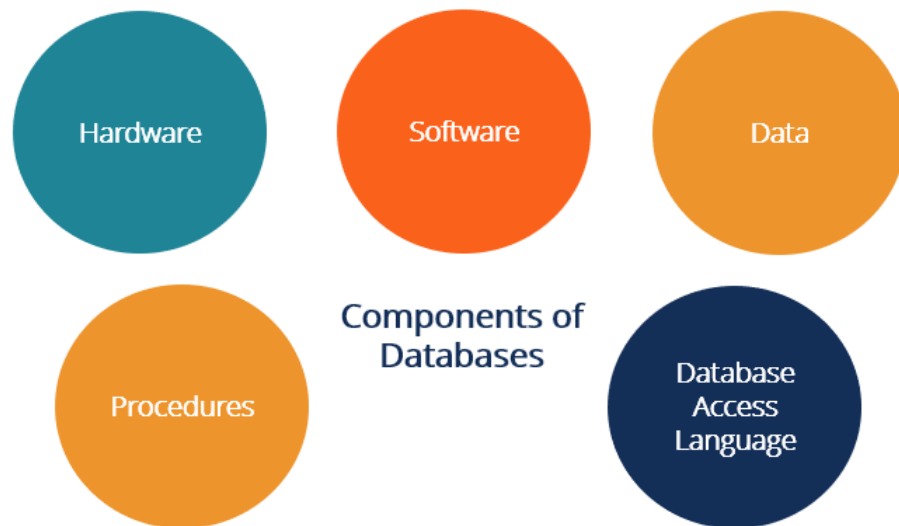
1. Обладнання;
2. Програмне забезпечення
3. Дані;
4. Процедури;
5. Мова доступу до бази даних;

1. Обладнання

Апаратне забезпечення стосується фізичних електронних пристроїв, таких як комп'ютери та жорсткі диски, які пропонують інтерфейс між комп'ютерами та системами реального світу.

2. Програмне забезпечення

Це набір програм, які використовуються для управління та контролю бази даних і включають програмне забезпечення бази даних, операційну систему, мережеве програмне забезпечення, яке використовується для обміну даними з іншими користувачами, і програми, які використовуються для доступу до даних.

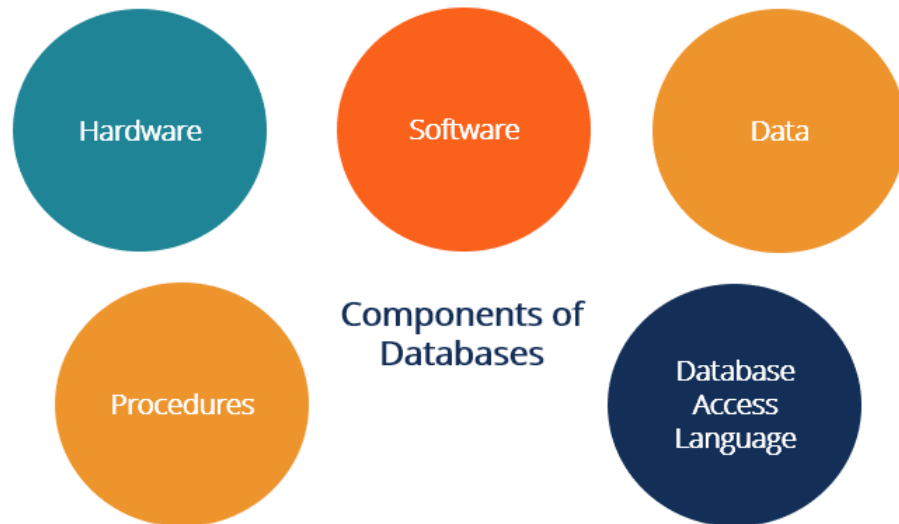


3. Дані

Дані — це необроблені факти та інформація, які потрібно впорядкувати й обробити, щоб зробити їх більш значущими. Словники бази даних використовуються для централізації, документування, контролю та координації використання даних в організації. База даних — це сховище інформації про базу даних (її також називають метаданими).

4. Процедури

Процедури стосуються інструкцій, які використовуються в системі керування базами даних, і охоплюють усе, починаючи від інструкцій щодо налаштування та інсталяції, входу та виходу з системи, керування повсякденними операціями, створення резервних копій даних і створення звітів.



5. Мова доступу до бази даних

Мова доступу до бази даних — це мова, яка використовується для написання команд для доступу, оновлення та видалення даних, що зберігаються в базі даних. Користувачі можуть писати команди за допомогою Database Access Language перед тим, як надсилати їх до бази даних для виконання. За допомогою мови користувачі можуть створювати нові бази даних, таблиці, вставляти та видаляти дані.

Дизайн Бази Даних

Вимоги до продукту

Яка база даних?

Аналітична обробка OnLine (OLAP) - це інтерактивна система що дозволяє переглядати різні підсумки по багатовимірних даних.

Онлайн транзакційна обробка (OLTP) - Спосіб організації БД, при якому система працює з невеликими за розмірами транзакціями, що йдуть великим потоком, і при цьому клієнту потрібний від системи максимально швидкий час відповіді.

Бюджет

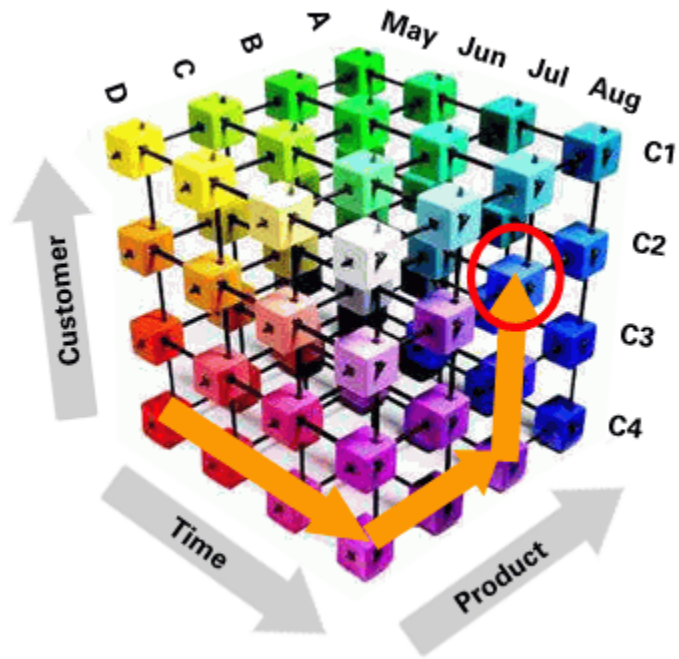
Платформа / Постачальник

Робочий процес?

порядок операцій

обробка помилок

звітність



Типи баз даних

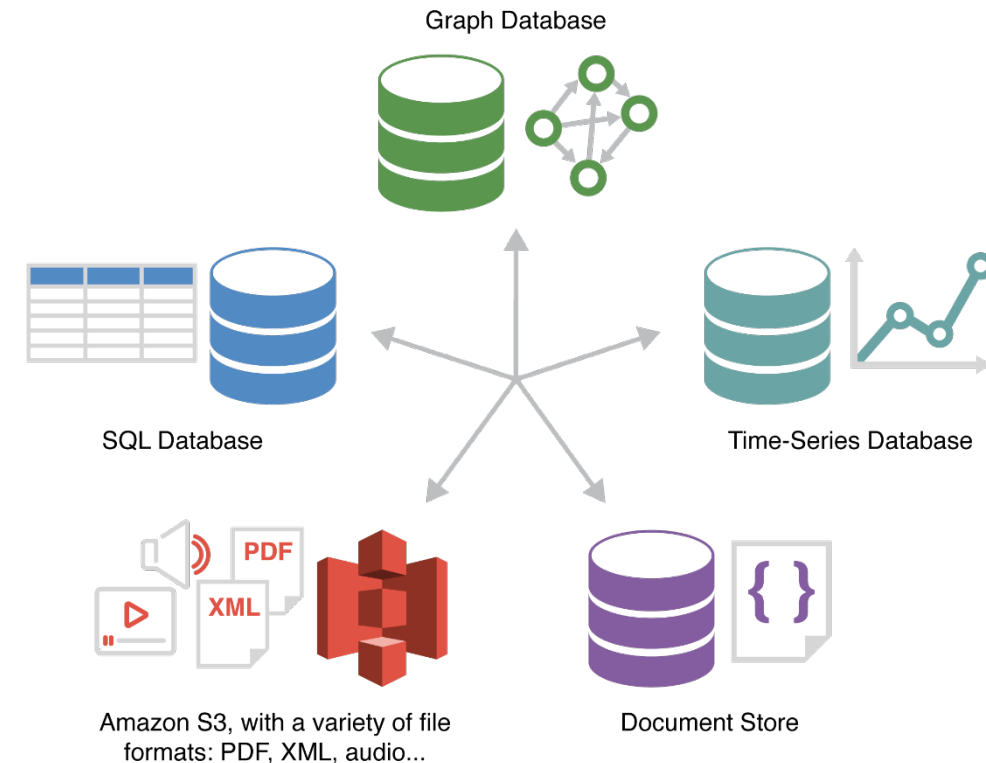
Бази даних еволюціонували з моменту їх створення в 1960-х роках, починаючи з ієрархічних та мережових баз даних, до 1980-х років з об'єктно-орієнтованими базами даних, а сьогодні з базами даних SQL та NoSQL та хмарними базами даних.

Розподілена база даних

Розподілена база даних - це база даних, у якій частини бази даних зберігаються в декількох фізичних місцях і в яких обробка розподіляється або тиражується між різними точками мережі.

Розподілені бази даних можуть бути однорідними або неоднорідними. Усі фізичні розташування в однорідній системі розподілених баз даних мають однакове базове обладнання та запускають однакові операційні системи та додатки баз даних.

Апаратне забезпечення, операційні системи або програми баз даних у різномірній розподіленій базі даних можуть бути різними в кожному з розташувань.

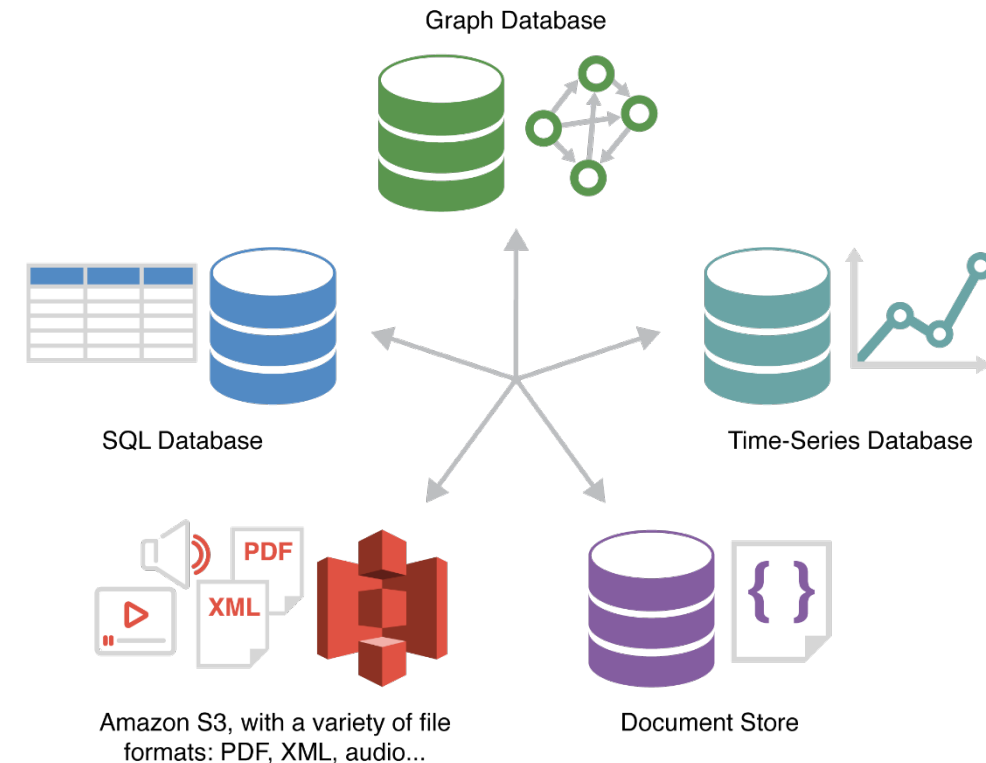


Типи баз даних

Об'єктно-орієнтована база даних

Елементи, створені з використанням об'єктно-орієнтованих мов програмування, часто зберігаються в реляційних базах даних, але для цих елементів краще підходять об'єктно-орієнтовані бази даних.

Об'єктно-орієнтована база даних організована навколо об'єктів, а не дій, даних - а не логіки. Наприклад, мультимедійний запис у реляційній базі даних може бути визначений об'єктом даних, на відміну від буквено-цифрового значення.



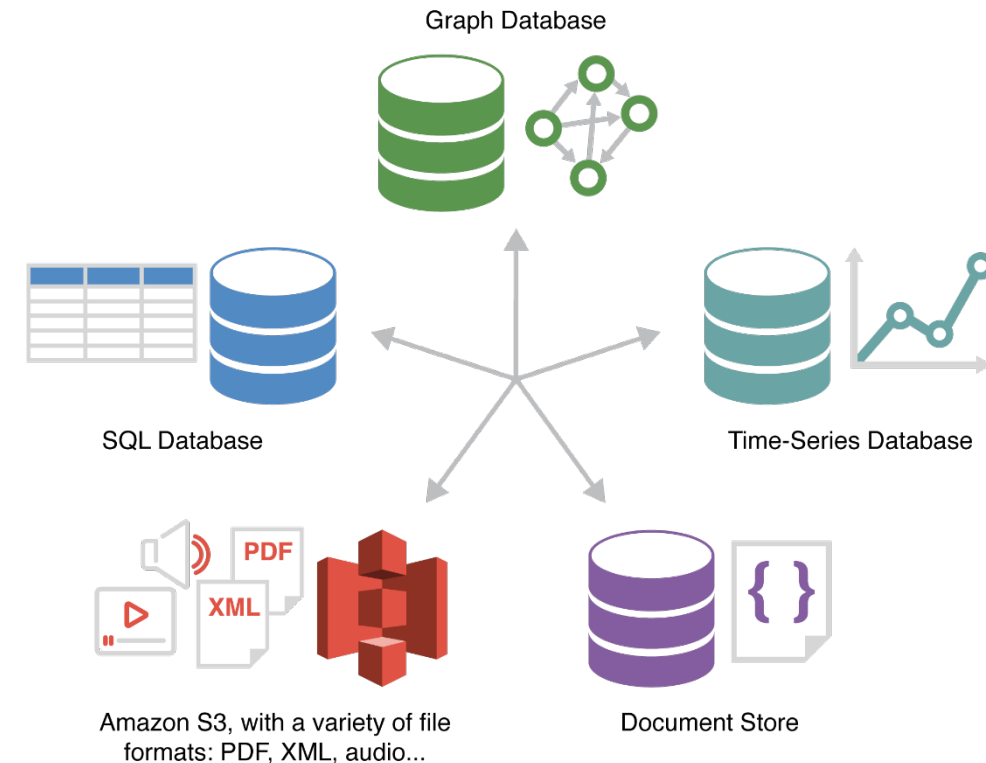
Типи баз даних

Хмарна база даних

Хмарна база даних - це база даних, яка була оптимізована або побудована для віртуалізованого середовища в гібридній хмарі, загальнодоступній хмарі чи приватній хмарі.

Хмарні бази даних забезпечують такі переваги, як можливість платити за об'єм і пропускну спроможність на основі кожного використання, і вони забезпечують масштабованість на вимогу, поряд з високою доступністю.

Хмарна база даних також надає підприємствам можливість підтримувати бізнес-додатки при розгортанні програмного забезпечення як послуги (software-as-a-service).



Типи баз даних

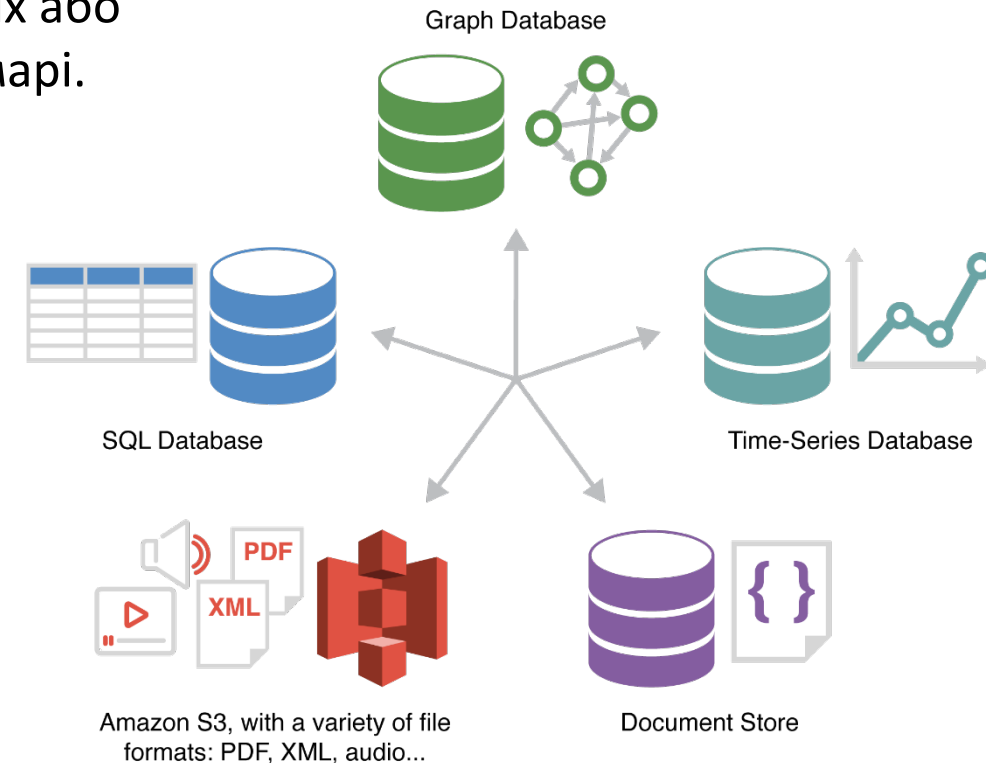
База даних NoSQL

Бази даних NoSQL корисні для великих наборів розподілених даних.

Бази даних NoSQL ефективні для вирішення проблем, пов'язаних з великими даними. Вони найбільш ефективні, коли організація повинна аналізувати великі фрагменти неструктурованих даних або даних, які зберігаються на декількох віртуальних серверах у хмарі.

База даних графів

База даних, орієнтована на графіки, або база даних графів - це тип бази даних NoSQL, який використовує теорію графів для зберігання, відображення та взаємозв'язку. Бази даних графіків - це в основному колекції вузлів і ребер, де кожен вузол представляє сутність, а кожне ребро - зв'язок між вузлами.



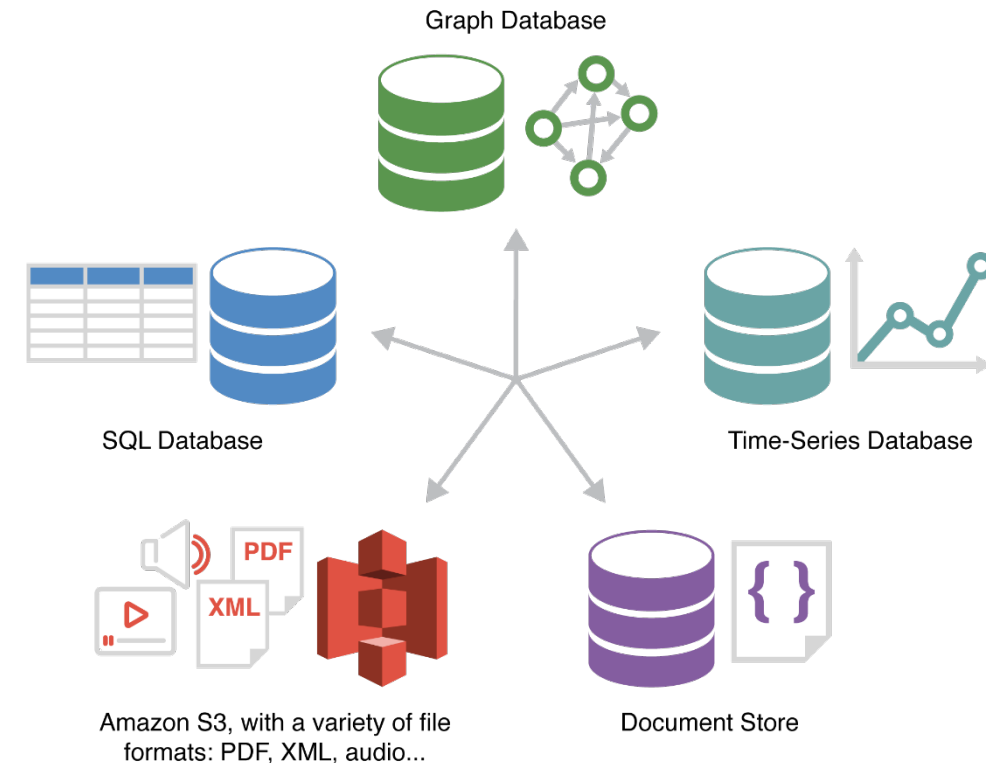
Типи баз даних

Реляційна база даних

Реляційна база даних, винайдена Е.Ф. Коддом в IBM в 1970 році, є табличною базою даних, де дані визначаються таким чином, що їх можна реорганізувати та отримати до них доступ різними способами.

Реляційні бази даних складаються з набору таблиць з даними, які входять у заздалегідь визначену категорію. Кожна таблиця має принаймні одну категорію даних у стовпці, і кожен рядок має певний екземпляр даних для категорій, визначених у стовпцях.

Мова структурованих запитів (SQL) - це стандартний інтерфейс користувача та прикладної програми для реляційної бази даних.



Що таке СУБД?

- Система управління базами даних (СУБД) - це спеціальна програмна програма, яка взаємодіє з користувачем, іншими програмами та самою базою даних для збору та аналізу даних.
- Наприклад: MySQL, PostgreSQL, Microsoft SQL Server, Oracle тощо.



TOP 10 Database Management Systems

DB-Engines Ranking класифікує системи управління базами даних за їх популярністю.

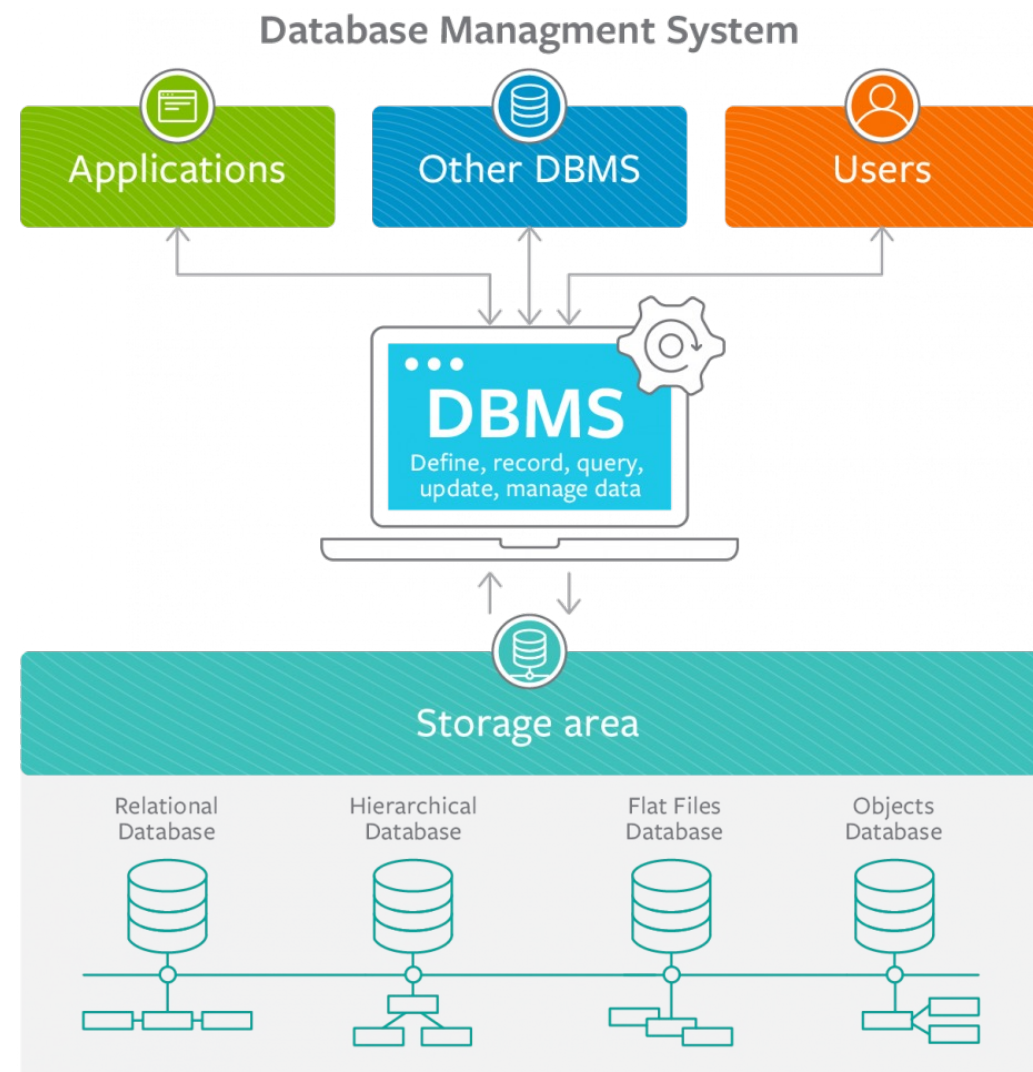
395 systems in ranking, September 2022									
Rank			DBMS	Database Model	Score				
Sep 2022	Aug 2022	Sep 2021			Sep 2022	Aug 2022	Sep 2021		
1.	1.	1.	Oracle +	Relational, Multi-model ⓘ	1238.25	-22.54	-33.29		
2.	2.	2.	MySQL +	Relational, Multi-model ⓘ	1212.47	+9.61	-0.06		
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model ⓘ	926.30	-18.66	-44.55		
4.	4.	4.	PostgreSQL +	Relational, Multi-model ⓘ	620.46	+2.46	+42.95		
5.	5.	5.	MongoDB +	Document, Multi-model ⓘ	400.64	+11.07	-6.07		
6.	6.	6.	Redis +	Key-value, Multi-model ⓘ	163.68	+0.72	-17.79		
7.	↑ 8.	↑ 8.	Elasticsearch	Search engine, Multi-model ⓘ	138.98	-0.94	-12.46		
8.	↓ 7.	↓ 7.	IBM Db2	Relational, Multi-model ⓘ	136.72	-2.52	-14.67		
9.	9.	↑ 11.	Microsoft Access	Relational	128.56	-1.78	-11.47		
10.	10.	↓ 9.	SQLite +	Relational	129.20	-0.72	-9.62		
11.	11.	↓ 10.	Cassandra +	Wide column	902.22	-18.60	-24.08		
12.	12.	12.	MariaDB +	Relational, Multi-model ⓘ	620.75	+0.37	+0.29		

422 systems in ranking, September 2023									
Rank			DBMS	Database Model	Score				
Sep 2023	Aug 2023	Sep 2022			Sep 2023	Aug 2023	Sep 2022		
1.	1.	1.	Oracle +	Relational, Multi-model ⓘ	1240.88	-1.22	+2.62		
2.	2.	2.	MySQL +	Relational, Multi-model ⓘ	1111.49	-18.97	-100.98		
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model ⓘ	902.22	-18.60	-24.08		
4.	4.	4.	PostgreSQL +	Relational, Multi-model ⓘ	620.75	+0.37	+0.29		
5.	5.	5.	MongoDB +	Document, Multi-model ⓘ	439.42	+4.93	-50.21		
6.	6.	6.	Redis +	Key-value, Multi-model ⓘ	163.68	+0.72	-17.79		
7.	7.	7.	Elasticsearch	Search engine, Multi-model ⓘ	138.98	-0.94	-12.46		
8.	8.	8.	IBM Db2	Relational, Multi-model ⓘ	136.72	-2.52	-14.67		
9.	↑ 10.	↑ 10.	SQLite +	Relational	129.20	-0.72	-9.62		
10.	↓ 9.	↓ 9.	Microsoft Access	Relational	128.56	-1.78	-11.47		

Всі ми знаємо, що в світі технологій баз даних існує два основних напрямки: SQL і NoSQL, реляційні та нереляційні бази даних. Відмінності між ними полягають в тому, як вони спроектовані, які типи даних підтримують, як зберігають інформацію.

Призначення СУБД

- Щоб правильно зберігати дані
- Забезпечити одночасний доступ до даних для багатьох користувачів
- Розмежувати доступ до даних для різних користувачів
- Щоб запобігти втраті даних



RDBMS

Реляційна система керування базами даних (RDBMS) - це програмна система, що забезпечує доступ до реляційної бази даних, винайденої британським вченим з IBM Едгаром Коддом.

Програмна система являє собою набір програмних додатків, які можуть використовуватися для створення, обслуговування, управління і використання бази даних.

«Реляційна база даних» - це база даних, структурована за «реляційною» моделлю. Дані зберігаються і представлені в табличному форматі, організованому в рядках і стовпцях з одним записом в рядку.

RDBMS дозволяє створювати, оновлювати і адмініструвати реляційну базу даних. Більшість систем керування базами даних використовують SQL для доступу до бази даних. Реляційні бази даних часто замінюють застарілі ієрархічні і мережні БД, так як вони простіше у використанні.



RDBMS

Традиційно дані були організовані у форматах файлів. Тоді СУБД була новою концепцією, і всі дослідження проводилися, щоб подолати недоліки традиційного стилю управління даними.

Сучасна БД має такі характеристики:

Об'єкт реального світу – сучасна БД є більш реалістичною та використовує об'єкти реального світу для розробки своєї архітектури. Він також використовує сутності та атрибути. Наприклад, шкільна база даних може використовувати учнів як сутність, а їхній вік – як атрибут.

Таблиці на основі зв'язків – БД дозволяє сутностям і зв'язкам між ними формувати таблиці. Користувач може зрозуміти архітектуру бази даних, просто подивившись на назви таблиць.



СУБД - RDBMS

Традиційно дані були організовані у форматах файлів. Тоді СУБД була новою концепцією, і всі дослідження проводилися, щоб подолати недоліки традиційного стилю управління даними.

Сучасна БД має такі характеристики:

Ізоляція даних і програми – система управління базою даних повністю відрізняється від своїх даних. База даних є активною сутністю, тоді як дані називаються пасивними, на яких база даних працює та організовує. СУБД також зберігає метадані, які є даними про дані, щоб полегшити власний процес.

Менше надлишковості – RDBMS дотримується правил нормалізації, яка розділяє відношення, якщо будь-який з його атрибутів має надлишкові значення. Нормалізація — це математично багатий і науковий процес, який зменшує надмірність даних.



СУБД - RDBMS

Традиційно дані були організовані у форматах файлів. Тоді СУБД була новою концепцією, і всі дослідження проводилися, щоб подолати недоліки традиційного стилю управління даними.

Сучасна БД має такі характеристики:

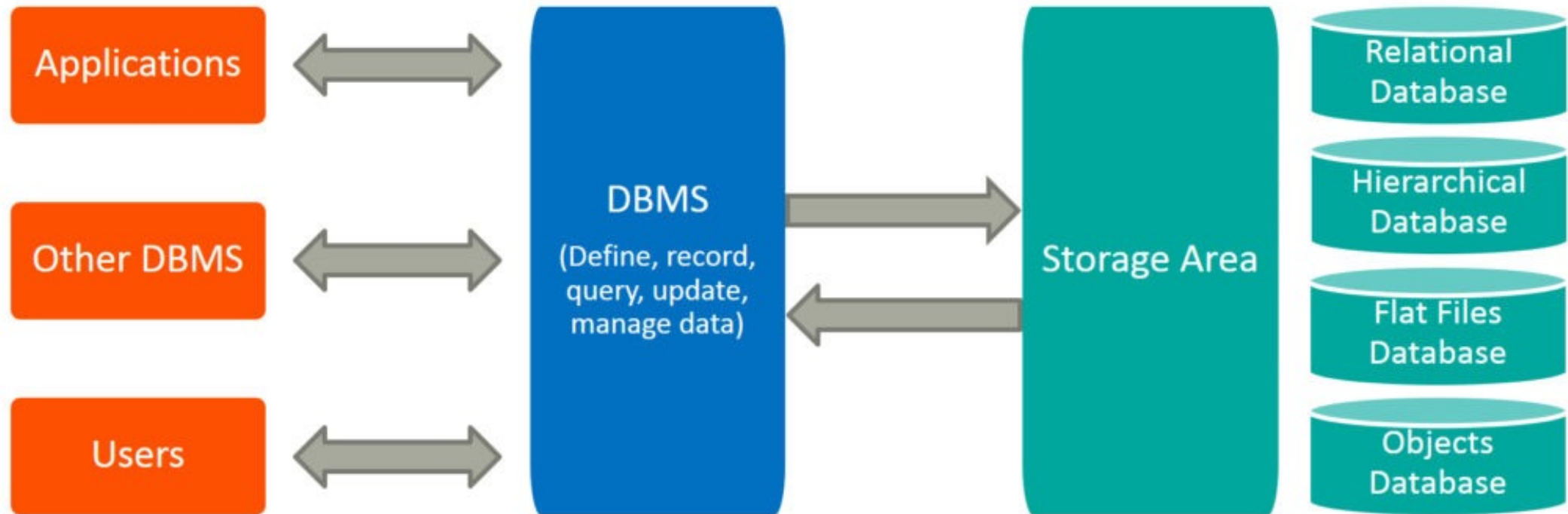
Узгодженість – це стан, коли кожне відношення в базі даних залишається узгодженим. Існують методи та техніки, які можуть виявити спробу залишити базу даних у неузгодженому стані. СУБД може забезпечити більшу узгодженість порівняно з попередніми формами програм зберігання даних, такими як системи обробки файлів.

Мова запитів – СУБД оснащена мовою запитів, що робить її більш ефективним отримання та маніпулювання даними. Користувач може застосовувати стільки різних параметрів фільтрації, скільки потрібно для отримання набору даних. Традиційно це було неможливо там, де використовувалася система обробки файлів.



Призначення СУБД

Database Management System



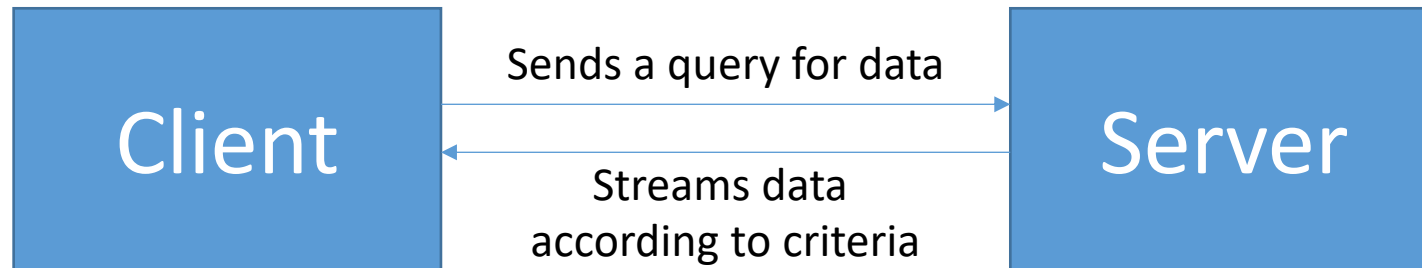
Типи архітектури СУБД

База даних зберігає багато важливої інформації для швидкого та безпечного доступу до даних.

Тому для ефективного керування даними важливо вибрати правильну архітектуру.

Архітектура СУБД допомагає користувачам виконувати свої запити під час підключення до бази даних.

Ми обираємо архітектуру бази даних залежно від кількох факторів, таких як розмір бази даних, кількість користувачів і відносини між користувачами.



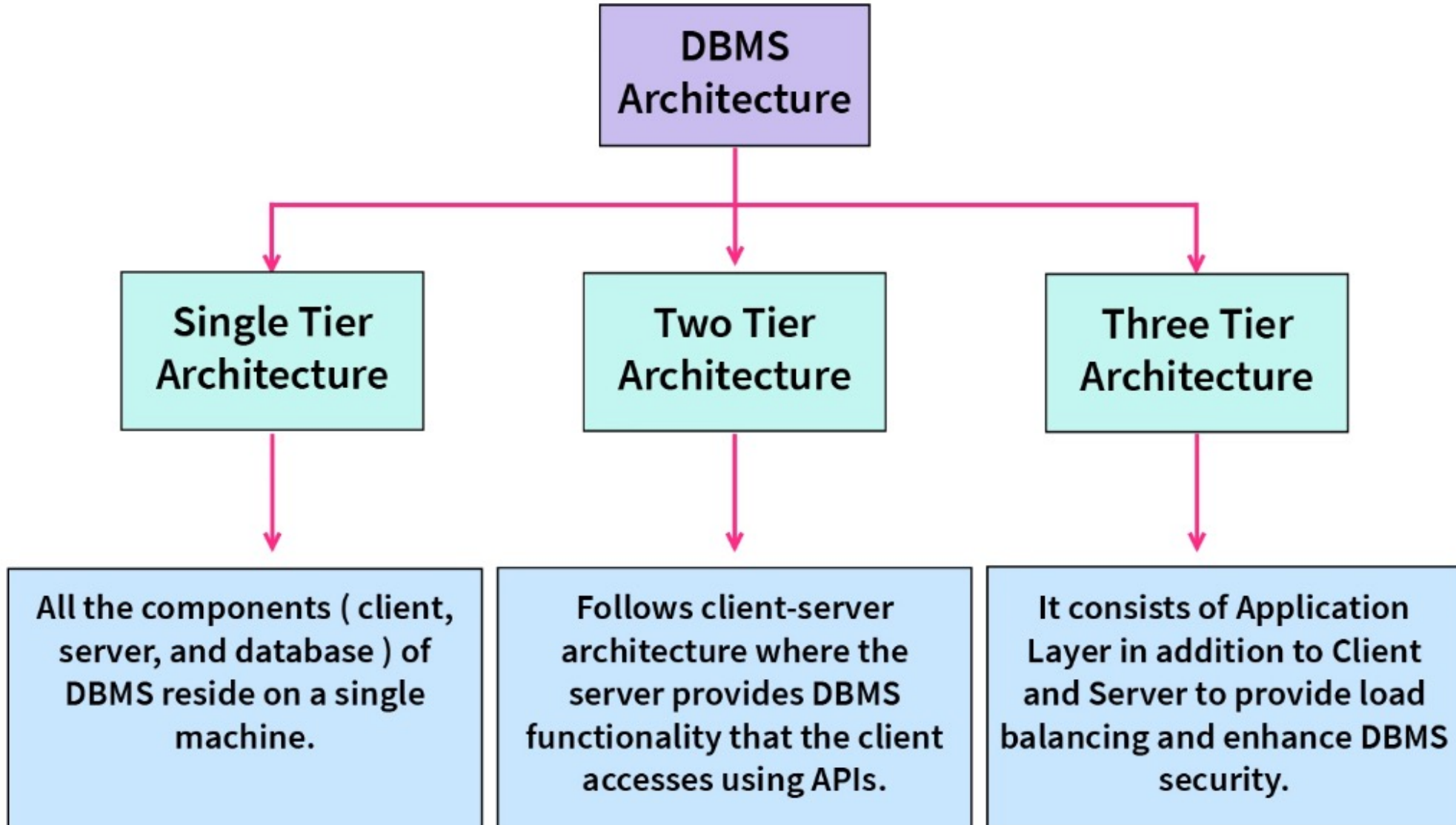
Типи архітектури СУБД

Типи архітектури СУБД

Існує кілька типів архітектури СУБД, які ми використовуємо відповідно до вимог використання.

- 1-рівнева архітектура - Single Tier Architecture (One-Tier Architecture)
- 2-рівнева архітектура - Two-Tier Architecture
- 3-рівнева архітектура - Three-Tier Architecture

Типи архітектури СУБД



Типи архітектури СУБД

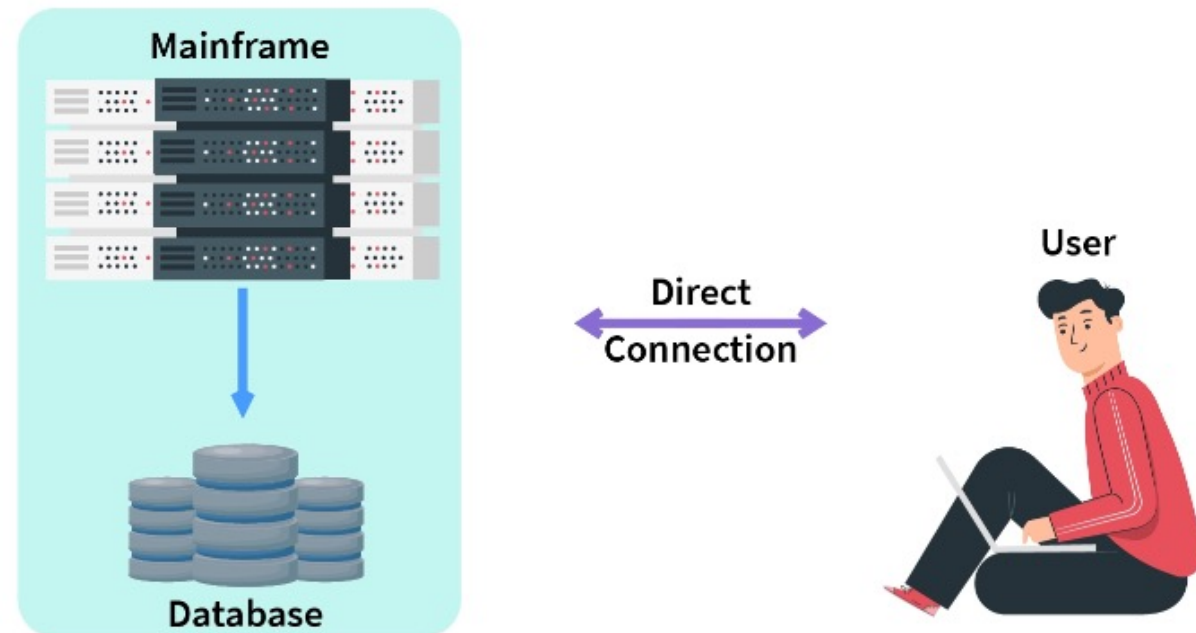
1-рівнева архітектура

Однорівнева архітектура в СУБД — це найпростіша архітектура бази даних, у якій клієнт, сервер і база даних знаходяться на одній машині.

Простий приклад однорівневої архітектури – це будь-який час, коли ви встановлюєте базу даних у своїй системі та отримуєте до неї доступ для виконання запитів SQL. Але на виробництві така архітектура використовується рідко.

Основні моменти:

- 1 - Найпростіша архітектура СУБД.
2. Усі компоненти СУБД, тобто сервер, база даних і клієнт, знаходяться в одній системі.
- 3 - Користувач може отримати прямий доступ до бази даних.
- 4 – використовується, коли дані не змінюються часто.
- 5 - Підходить для програмістів, розробників баз даних і для однокористувацького доступу.



Типи архітектури СУБД

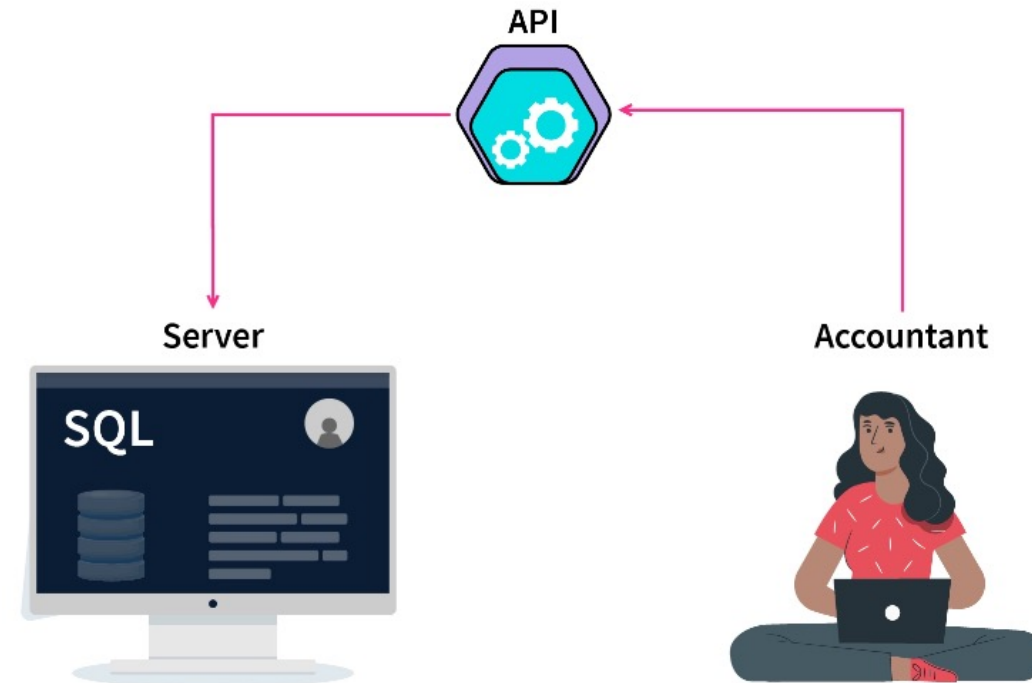
2-рівнева архітектура

Дворівнева **архітектура** в СУБД — це архітектура бази даних, де рівень презентації працює на клієнті (ПК, мобільний пристрій, планшет тощо), а дані зберігаються на сервері, який називається другим рівнем.

Дворівнева архітектура забезпечує додаткову безпеку СУБД, оскільки вона не доступна безпосередньо кінцевому користувачеві. Він також забезпечує пряме та швидше спілкування.

Основні моменти:

- 1 – Подібно до архітектури клієнт-сервер.
- 2 – Швидший доступ, простіше в обслуговуванні та може працювати з кількома користувачами одночасно.
- 3 - Використовується, коли ми хочемо отримати доступ до СУБД через програми та API.
- 4 - Має проблеми з масштабованістю та безпекою через пряме підключення клієнт-сервер.



Типи архітектури СУБД

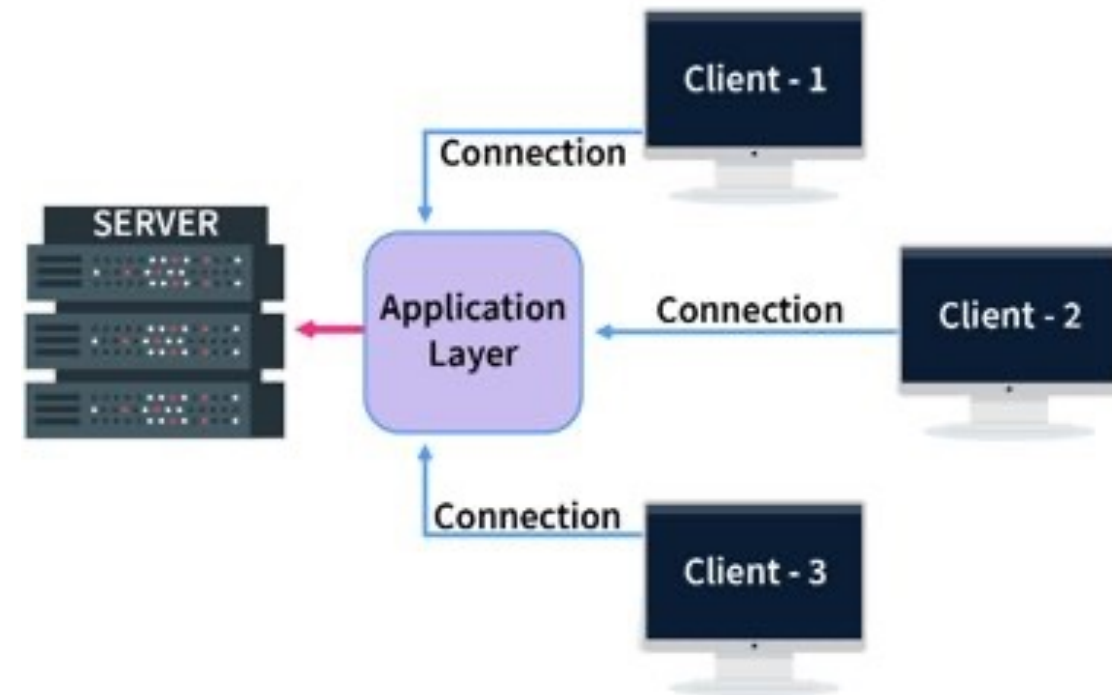
3-рівнева архітектура

3 - **рівнева архітектура** в СУБД — це найпопулярніша клієнт-серверна архітектура в СУБД, у якій розробка та підтримка функціональних процесів, логіки, доступу до даних, зберігання даних та інтерфейсу користувача здійснюється незалежно як окремі модулі. Трирівнева архітектура містить презентаційний рівень, прикладний рівень і сервер бази даних.

Дизайн 3-рівневої архітектури бази даних є розширенням 2-рівневої архітектури клієнт-сервер.

3-рівнева архітектура має такі рівні:

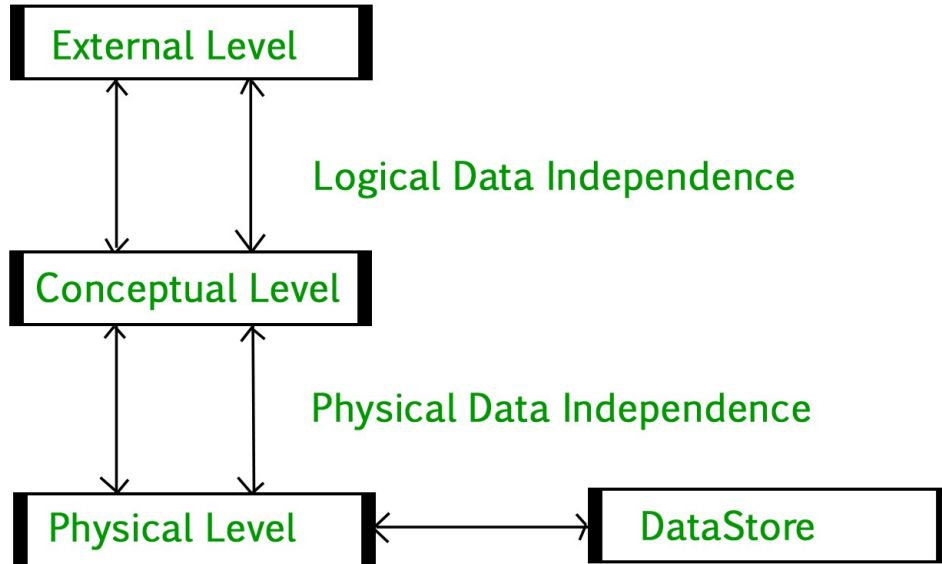
- 1.Рівень презентації (ваш ПК, планшет, мобільний пристрій тощо)
- 2.Прикладний рівень (сервер)
- 3.Сервер бази даних



Типи архітектури СУБД

Метою трирівневої клієнт-серверної архітектури є:

- Щоб відокремити програми користувача та фізичну базу даних
- Для підтримки характеристик СУБД
- Незалежність програми від даних
- Підтримка кількох переглядів даних



Прикладний рівень знаходиться між користувачем і СУБД, який відповідає за передачу запиту користувача системі СУБД і надсилання відповіді від СУБД користувачеві.

Прикладний рівень (рівень бізнес-логіки) також обробляє функціональну логіку, обмеження та правила перед передачею даних користувачеві або вниз до СУБД.

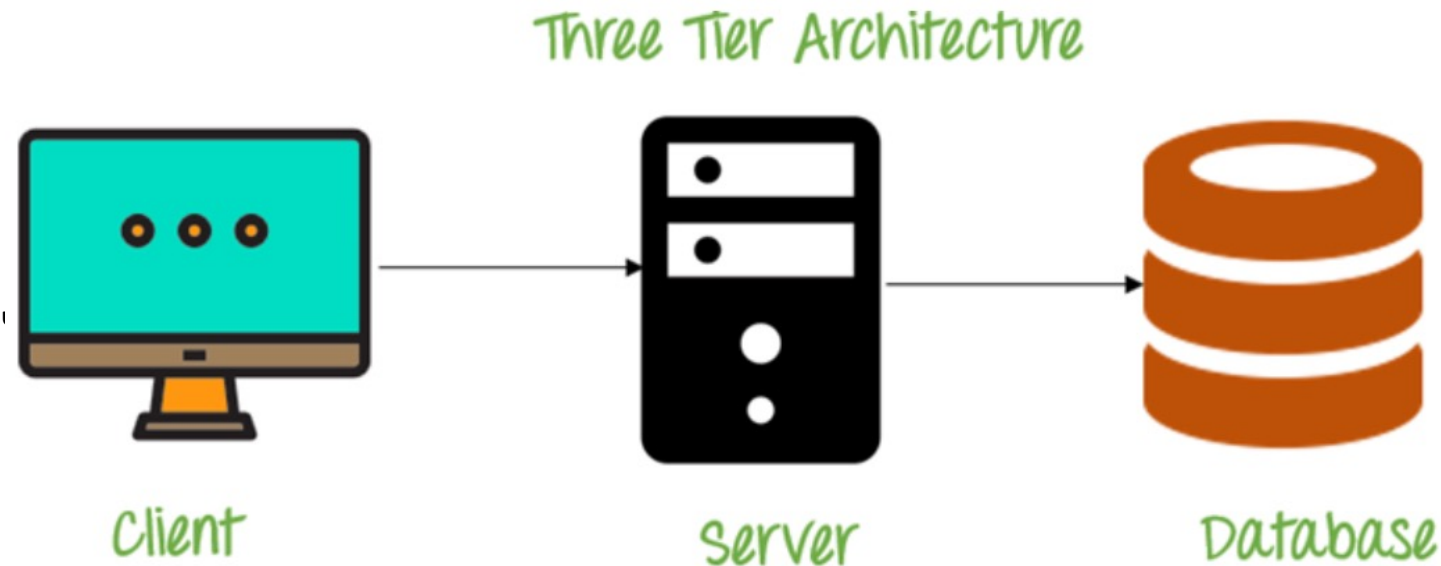
Типи архітектури СУБД

Основними перевагами трирівневої архітектури СУБД є:

- **Масштабованість.** Оскільки сервер бази даних не знає про користувачів за межами прикладного рівня, а прикладний рівень реалізує балансування навантаження, клієнтів може бути скільки завгодно.
- **Цілісність даних** – можна уникнути пошкодження даних і неправильних запитів завдяки перевіркам, які виконуються на прикладному рівні для кожного клієнтського запиту.
- **Безпека.** Видалення прямого зв'язку між клієнтською та серверною системами через абстракцію зменшує неавторизований доступ до бази даних.

Основні моменти:

- 1 - Найпоширеніша архітектура СУБД.
- 2 – відповідає архітектурі клієнт-додаток-сервер.
- 3 - Покращена безпека, цілісність даних і масштабованість.
- 4 - Має складність і проблеми з обслуговуванням додаткового рівня.



Questions

Questions

