01- Database Introduction & Architectures

School of Computer Science University of Windsor

Dr. Shafaq Khan

Agenda

- Course Introduction
- Team Introduction
- Lecture
- Project hand-out

Deadline

Forming Groups: 20th May 2023 [11:59 PM]

- Section 4 (MO 08:30 am): Section 4 Project Groups.docx
- Section 1 (MO 11:30 am): <u>Section1 Project Groups.docx</u>
- Section 2 (TU 08:30 am): <u>Section2 Project Groups.docx</u>
- Section 3 (WE 11:30 am): Section 3 Project Groups.docx



Introductory Questions

What is a database

What is the history of databases?

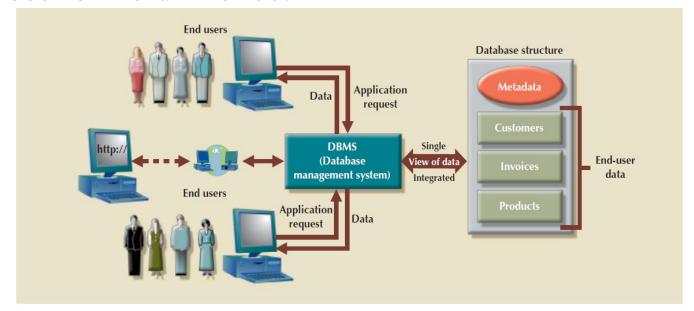
What are the different database system architectures?

What is a data warehouse?

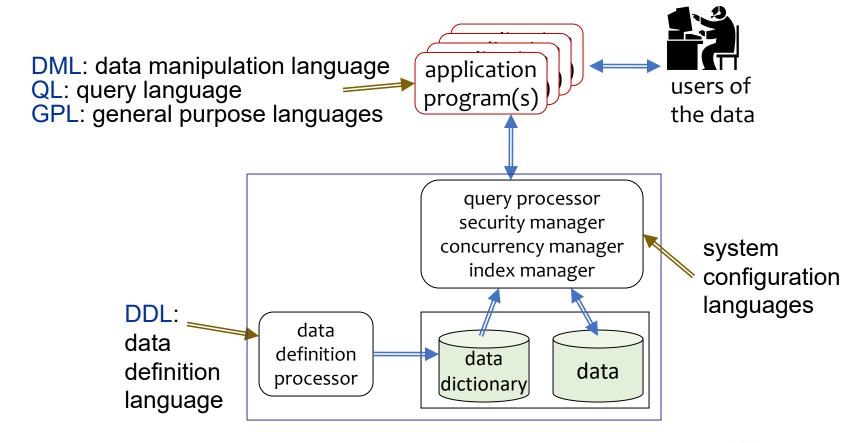


What is a Database?

- A database is any collection of data.
- A DBMS is a software that manages and controls access to the database.



DBMS Languages





Advantages of Using a DBMS

Anything you can do with a DBMS, you can do with a file system, a network and a heap of C code. So why spend the money to buy a DBMS?

- Integrity: A DBMS maintains the consistency of stored data
- Concurrency: A DBMS supports access by concurrent users
- Access Control: A DBMS can restrict access to authorized users
- Redundancy Control: A DBMS can assist in controlling redundancy
- Backup and Recovery: A DBMS can provide backup and recovery.
 - *backup* = snapshots of the data particular times
 - recovery = restoring the data to a <u>consistent</u> state after a system crash



History of Database Systems

IDS

Honeywell - The first network DBMS.
Charles Bachman

OBJECT-ORIENTED DATABASES

In 1983

DB₂

IBM- "SEQUEL" becomes the standard (SQL).- by Donald D. Chamberlin and Raymond F. Boyce

NoSQL

MongoDB Cassandra HBase Oracle NoSQL

Data Warehouses

Distributed / Shared-Nothing
Netezza, ParAccel

MonetDb, Greenplum Datallegro, Vertica

1960

1970

1980

1990

2000

2010

2020

IDMS

Cullinet – by Charles
Bachman- primarily a network
model (CODASYL) database
management system for
mainframes.

IMS

IBM - to keep track of the supplies and parts inventory for the Saturn V and Apollo space exploration projects.

Relational Model

IBM— by **Ted Codd-** high level of abstraction of databases to avoid all this sort of maintenance burden on humans.

System R, INGRES, Oracle

SQL SERVER MySQL PostgreSQL SQLite

NewSQL

a relational database with the scalable properties of NoSQL VoltDB, Clustrix, MemSQL, Spanner

Hybrid Systems

Execute fast OLTP like a NewSQL system while also executing complex OLAP queries like a data warehouse system.

Hyper, nappy, JustOne

Cloud System

Snowflake, Xeround

First database-as-a-service (DBaaS) offerings were "containerized" versions of existing DBMSs.

RedShift &Aurora from Amazon,

A Aurgra snowflak

amazon

REDSHIFT

Amazon

415 systems in ranking, May 2023

Rank					Score		
May 2023	Apr 2023	May 2022	DBMS	Database Model	May 2023	Apr 2023	May 2022
1.	1.	1.	Oracle 😷	Relational, Multi-model 👔	1232.64	+4.36	-30.18
2.	2.	2.	MySQL	Relational, Multi-model 👔	1172.46	+14.68	-29.64
3.	3.	3.	Microsoft SQL Server 🚹	Relational, Multi-model 📵	920.09	+1.57	-21.11
4.	4.	4.	PostgreSQL 🚹	Relational, Multi-model 🔞	617.90	+9.49	+2.61
5.	5.	5.	MongoDB ☐	Document, Multi-model 🚺	436.61	-5.29	-41.63
6.	6.	6.	Redis 🛨	Key-value, Multi-model 👔	168.13	-5.42	-10.89
7.	7.	7.	IBM Db2	Relational, Multi-model 👔	143.02	-2.48	-17.31
8.	8.	8.	Elasticsearch	Search engine, Multi-model 👔	141.63	+0.56	-16.06
9.	9.	1 0.	SQLite [Relational	133.86	-0.68	-0.87
10.	10.	4 9.	Microsoft Access	Relational	131.17	-0.20	-12.27
11.	1 2.	1 4.	Snowflake 😷	Relational	111.73	+0.60	+18.22
12.	4 11.	4 11.	Cassandra 😷	Wide column	111.14	-0.67	-6.88
13.	13.	4 12.	MariaDB 🚹	Relational, Multi-model 🔞	96.87	+0.93	-14.26
14.	14.	4 13.	Splunk	Search engine	86.64	+1.20	-9.71
15.	1 6.	1 6.	Amazon DynamoDB 😷	Multi-model 👔	81.11	+3.66	-3.35
16.	4 15.	4 15.	Microsoft Azure SQL Database	Relational, Multi-model 🔞	79.19	+0.13	-6.14
17.	17.	17.	Hive	Relational	73.61	+1.96	-8.00
18.	1 9.	1 24.	Databricks	Multi-model 🔞	63.94	+2.98	+16.09
19.	4 18.	4 18.	Teradata	Relational, Multi-model 👔	62.71	+1.12	-5.67
20.	20.	1 23.	Google BigQuery 🖽	Relational	54.87	+1.55	+6.26

Source: https://db-engines.com/en/ranking



University of Windsor

Multi-user DBMS Architectures

Two-Tier Client-Server

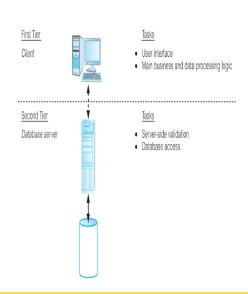
Client (tier 1) manages user interface and runs applications.
Server (tier 2) holds database and DBMS.

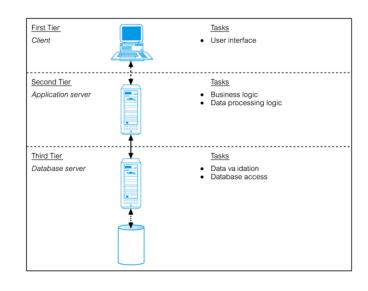
Three-Tier Client-Server

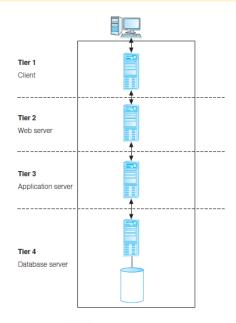
The need for enterprise scalability challenged the traditional two-tier client—server model

n-Tier Client-Server (e.g. 4-Tier)

The three-tier architecture can be expanded to n tiers, with additional tiers providing more flexibility and scalability.



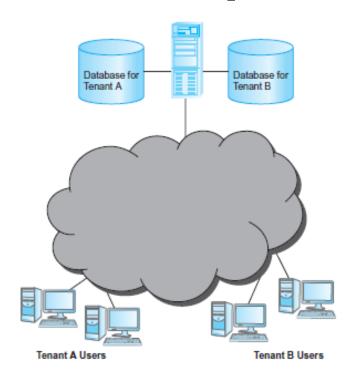


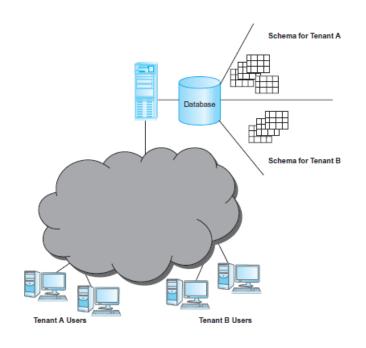




Cloud-based database solutions

- Multi-tenant cloud database
 - shared DBMS server, separate databases.
 - shared database, separate schema architecture



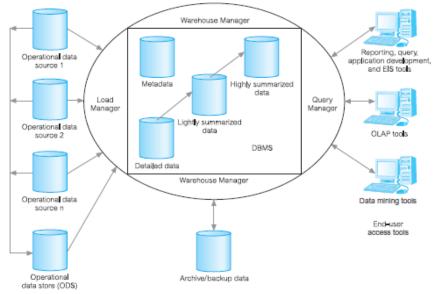




Data Warehousing

• A consolidated/integrated view of corporate data drawn from disparate operational data sources and a range of end-user access tools capable of supporting simple to highly complex queries to support decision making.

Data Warehouse Video



Data Warehouse Video



Inside The World's Largest Data Center - YouTube



Recap and Conclusion

- Introduction
- Database Structure
- Advantages of the Database Approach
- History of Database Systems
- Multi-User DBMS Architecture
- Cloud-based Database Architecture
- Data Warehousing



Test your understanding

- Which of the following is **not** a valid NoSQL database?
 - A. Cassandra
 - B. HBase
 - C. MongoDB
 - D. PostgreSQL

Project Intro



Any Questions

