



دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

طراحی پایگاه داده‌ها

(فصل اول : مقدمات)

مهدی دادبخش

mahdi.dadbakhsh@sharif.edu

شماره درس : ۴۰۳۸۴

یکشنبه - سه‌شنبه (۱۶:۳۰ الی ۱۸:۰۰)

۱۴۰۱ - ۱۴۰۲

مفاهیم اولیه پایگاه داده

مراحل تولید سیستم پایگاه داده

امکانات مورد نیاز در ایجاد پایگاه داده

عناصر محیط عملیاتی

انواع معماری سیستم پایگاه داده

سیستم مدیریت پایگاه داده

سیستم مدیریت پایگاه داده رابطه‌ای

مزایای سیستم پایگاه داده

پایان

داده (Data) :

داده عبارت است از یک کمیتی (عددی/حرفی) که معنا و مفهوم مشخصی داشته باشد.

اطلاعات (Information) :

تعریف دقیق و جامعی برای مفهوم اطلاع وجود ندارد. برخی، اطلاع را داده پردازش شده می‌دانند و برخی دیگر، آن را معنایی می‌دانند که انسان به داده نسبت می‌دهد.

پایگاه داده (Database) :

مجموعه‌ای است از داده‌های ذخیره شده، پایا (Persistent)، مجتمع (Integrated) و به هم مرتبط (Interconnected)، تا حد امکان فاقد افزونگی (Redundancy)، دارای معماری خاص خود، مبتنی بر یک مدل داده‌ای (Data Model) مشخص، مورد استفاده یک یا چند کاربر (Multi User) در یک سازمان (در یک محیط) به صورت اشتراکی (Shared) و همروند (Concurrent).

مثال : سیستم یا محیط عملیاتی دانشگاه : این محیط از تعدادی زیرمحیط یا زیرسیستم تشکیل شده است : آموزش، امور مالی، امور دانشجویی و زیرمحیط‌های یک محیط عملیاتی ممکن است در نوع موجودیت‌ها با یکدیگر اشتراک داشته باشند.

در هر محیط عملیاتی مجموعه‌ای از موجودیت‌ها وجود دارند که کاربران به اطلاعات آنها نیاز داده‌ای و یا پردازشی دارد.

در طراحی سیستم برای یک محیط عملیاتی باید برای هر زیر محیطی یک سیستم کاربردی ایجاد کنیم. برای این منظور دو راهکار وجود دارد :

۱ - سیستم فایلینگ یا سنتی

۲ - سیستم پایگاه داده

سیستم پایگاه داده

سیستم فایلینگ یا سنتی

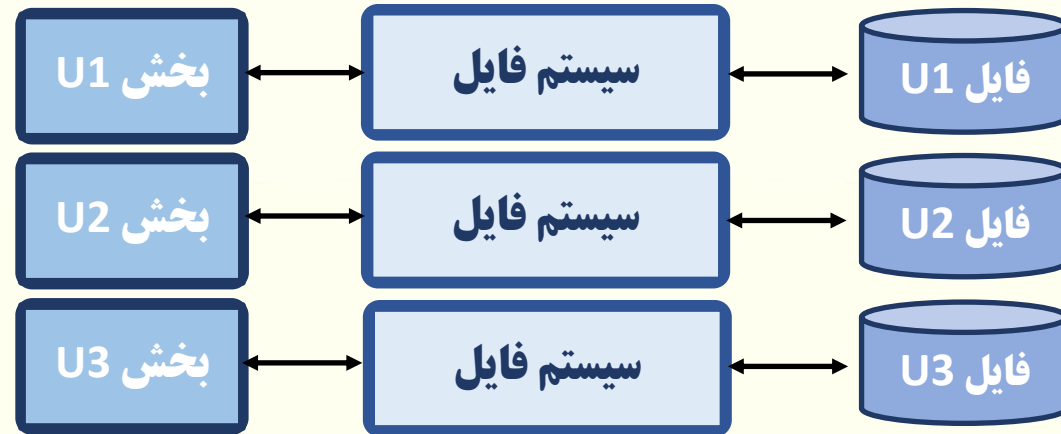


سیستم فایلینگ یا سنتی

فرض کنید می خواهیم سیستمی را طراحی کنیم که از سه بخش U1 ، U2 و U3 تشکیل شده است. در روش سنتی یا فایلینگ، هر یک از بخشهای U1 ، U2 و U3 بصورت مجزا سیستم خاص خود را دارند و با استفاده از امکانات سیستم فایل موجود در سیستم عامل به فایلهای مربوط به خود دسترسی دارند.

ویژگیهای این روش عبارتند از :

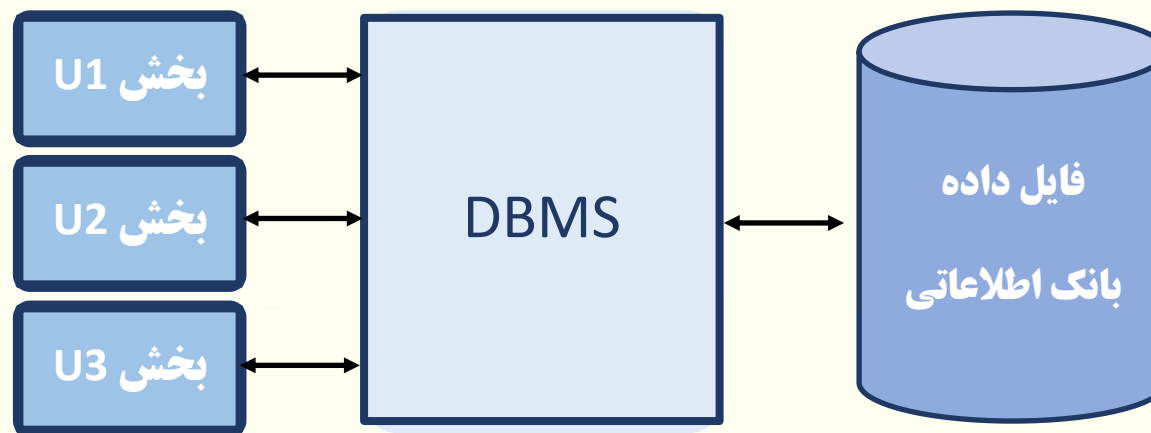
- ۱ - پراکندگی داده ها : اطلاعات هر بخش در فایلهای مجزا ذخیره می شود.
- ۲ - افزونگی یا تکرار : ممکن است چندین بخش به داده های مشترک نیاز داشته باشند، چون فایلهای هر بخش مجزا می باشد باید داده های مشترک در فایل مربوط به بخشهای مختلف تکرار شود.
- ۳ - امنیت کم : دسترسی کاربران به داده ها از طریق سیستم فایل انجام می شود و سیستم فایل کنترل چندانی روی دسترسی ها ندارد ، بنابراین امنیت در درجه پایینی قرار دارد.
- ۴ - وجود سیستمهای نامجتمع و نامرتبط به هم
- ۵ - ناسازگاری دادهها



سیستم پایگاه داده

در این روش، تمامی داده های مورد نیاز سیستم در فایلی به نام فایل بانک اطلاعاتی ذخیره می شود، لذا پراکندگی داده ها وجود ندارد و داده ها بطور متمرکز قرار دارند. هر داده ای فقط یکبار ذخیره می شود و بخشهای مختلف در صورت نیاز و داشتن مجوز می توانند بطور اشتراکی از آنها استفاده کنند. کاربر بطور مستقیم به داده ها دسترسی ندارد و دسترسی از طریق سیستم مدیریت بانک اطلاعاتی (DBMS) انجام می شود. بنابراین امنیت در سطح بسیار بالایی اعمال می شود.

* DBMS : Data Base Management System



مراحل تولید سیستم پایگاه داده



امکانات مورد نیاز در ایجاد پایگاه داده



سخت افزار

سخت افزار انواع مختلفی دارد : سخت افزار ذخیره سازی (دیسک) ، سخت افزار پردازشگر (کامپیوتر) و سخت افزار ارتباطی (محلی و شبکه ای)

نرم افزار

نرم افزار انواع مختلفی دارد : سیستم عامل (OS) و سیستم فایل (FS) ، سیستم مدیریت پایگاه داده (DBMS) ، ابزارها (Tools) و برنامه های کاربردی (Applications)

کاربر

کاربر انواع مختلفی دارد : مدیر بانک اطلاعاتی (DBA) ، برنامه نویس بانک اطلاعاتی (DBP) و کاربر نهایی (End User)

داده

داده ها به دو دسته تقسیم می شوند : داده های کاربران و داده های سیستمی



انواع معماری سیستم پایگاه داده

معماری سیستم پایگاه داده چیست؟

- هر پایگاه داده‌ای برای مدیریت سیستم خود از یک معماری تبعیت می‌کند. هدف از معماری پایگاه داده، به وجود آوردن یک راه ساده است که کاربران با هر سطحی بتوانند با یک زبان با سیستم تعامل داشته باشند.
- در توصیف معماری سیستم باید موارد زیر را مشخص کنیم :
 - از چه مولفه‌هایی، از هر مولفه چه تعداد و با چه کیفیتی تشکیل شده است.
 - مولفه‌ها چگونه با هم ترکیب شده‌اند (جنبه ساختاری سیستم).
 - مولفه‌ها چگونه با هم در تعامل هستند (جنبه رفتاری سیستم).

انواع معماری پایگاه داده :

- سیستم‌های متمرکز
- سیستم‌های مشتری / خدمتگزار
- سیستم‌های موازی
- سیستم‌های توزیع شده

سیستم‌های مشتری / خدمتگزار

سیستم‌های متمرکز

سیستم‌های توزیع شده

سیستم‌های موازی



سیستم‌های پایگاه داده متمرکز (Centralized Database System)

- در سیستم‌های متمرکز کلیه داده‌ها در یک پایگاه داده ذخیره و بازیابی می‌شوند. از این رو کاربران برای دسترسی به اطلاعات باید از دستگاهی که پایگاه داده در آن ذخیره شده است استفاده کنند. در این معماری یک حافظه مشترک بین پایگاه داده، سیستم و چندین پردازنده وجود دارد.
- در این معماری یک پایگاه داده (متمرکز و مجتمع) روی یک سیستم کامپیوتری و بدون ارتباط با سیستم کامپیوتری دیگر ایجاد می‌شود.
- معمولاً به صورت تک کاربری و برای کاربردهای کوچک و با امکانات محدود از این معماری استفاده می‌شود.
- امروزه با توجه به افزایش استفاده از پایگاه داده‌ها و حجم اطلاعات و تعداد کاربران این سیستم کمتر مورد استفاده قرار می‌گیرد و سیستم‌هایی با قابلیت‌های بیشتر جایگزین سیستم‌های متمرکز شده‌اند.

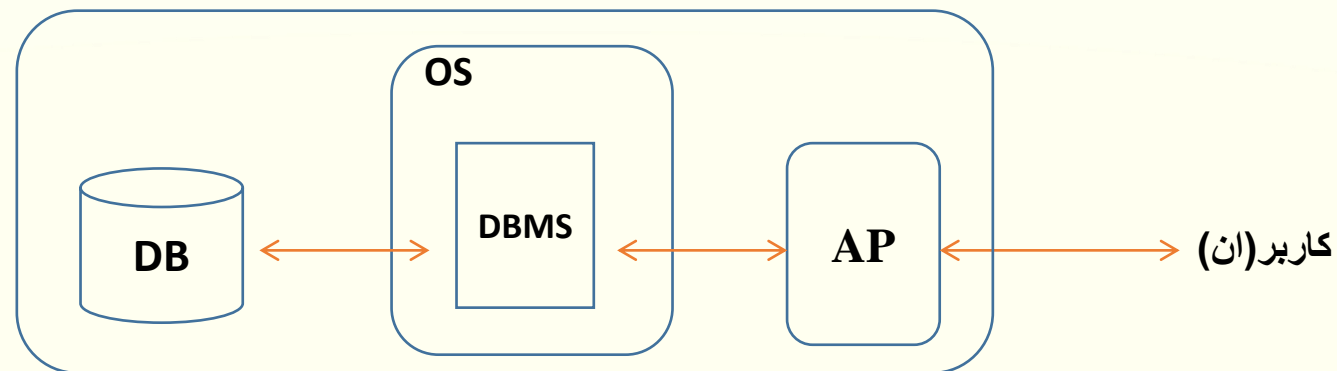
مزایا :

- سادگی در طراحی
- سادگی در استفاده
- عدم نیاز به امکانات سخت افزاری یا نرم افزاری خاص

معایب :

- تک کاربره بودن
- مشکل بودن استفاده در سازمانهای بزرگ

ماشین

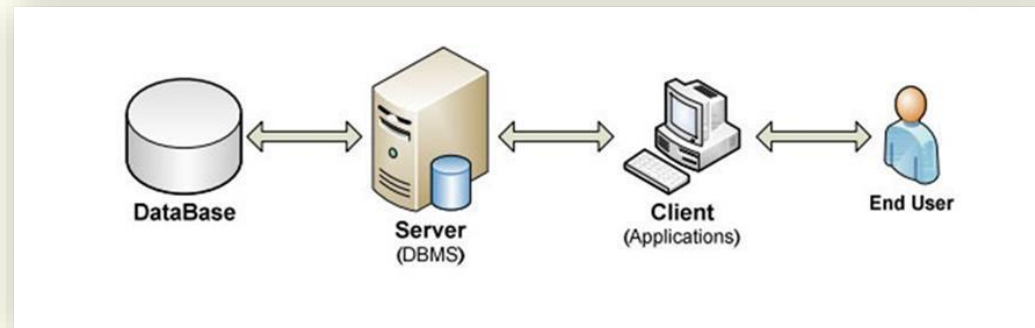


سیستم‌های پایگاه داده مشتری / خدمتگزار (Client/Server Database System)

- این معماری بعد از معماری متمرکز و در جهت کارایی بیشتر طراحی شد. در این معماری یک کامپیوتر به عنوان سرور عمل می‌کند و هر کاربر با کامپیوتر خود به سرور و در نتیجه به پایگاه داده متصل می‌شوند.
- دلیل اصلی استفاده از معماری مشتری-خدمتگزار (Client-Server)، تقسیم وظایف سیستم می‌باشد.
- هر ماشینی (فیزیکی یا منطقی) که خدمتی را به ماشین دیگر بدهد، خدمتگزار (Server) و هر ماشینی که خدمتی دریافت می‌کند، مشتری (Client) نامیده می‌شود.
- نمونه‌هایی از انواع خدمتگزارها عبارتند از : DB Server, Message Server, Print Server, File Server.

مزایای سیستم های مشتری / خدمتگزار:

- افزایش راندمان سیستم و صرفه‌ی اقتصادی
- سادگی بیشتر در گسترش و توزیع منابع
- تولید واسطه‌های کاربر بهتر
- راحتی در نگهداری سیستم



چند مشتری - چند خدمتگزار

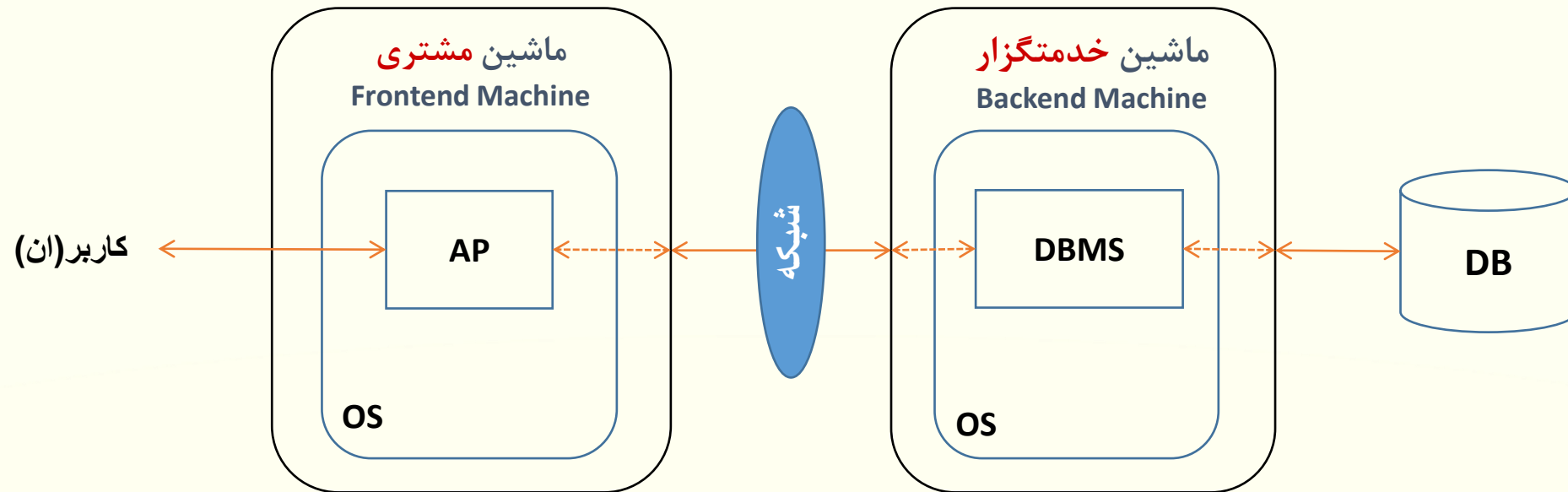
معماری سه لایه

معماری دو لایه

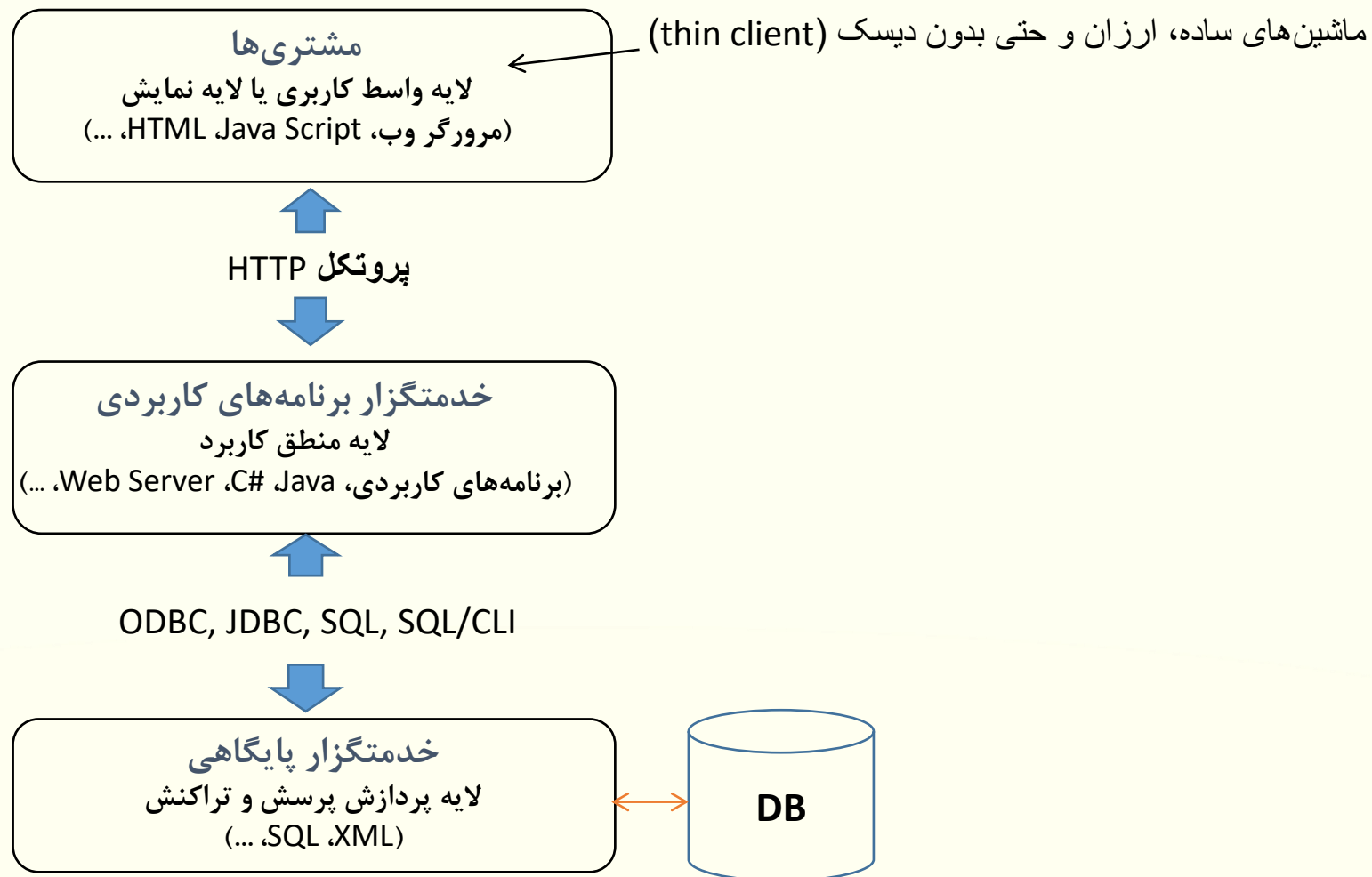


معماری مشتری / خدمتگذار دو لایه

- معمولا شامل دو سایت:
 - سایت مشتری: تمام برنامه‌های کاربردی در آن اجرا می‌شوند.
 - سایت خدمتگذار: تمام داده‌ها در آن ذخیره می‌شوند

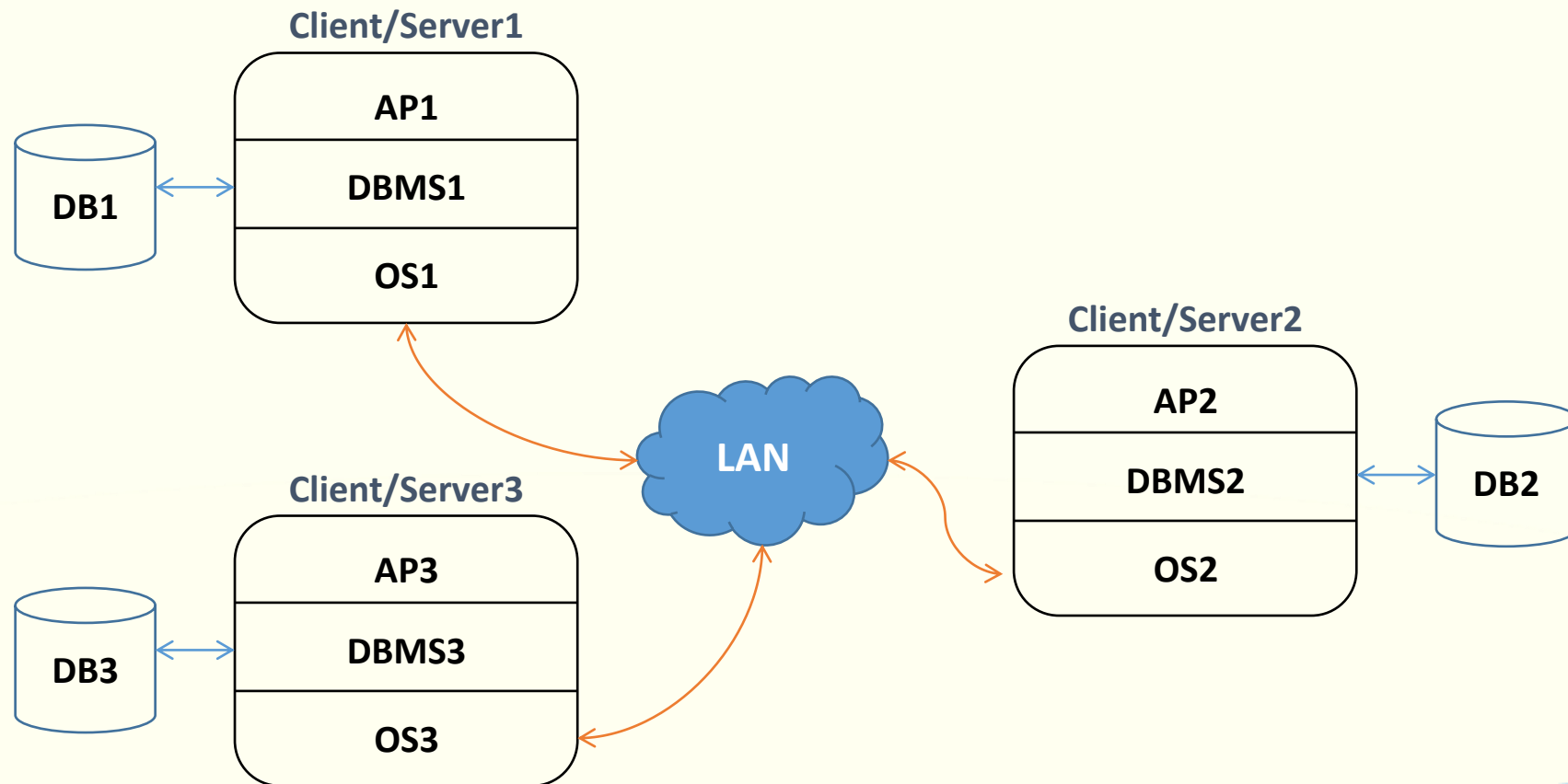


معماری مشتری / خدمتگذار سه لایه



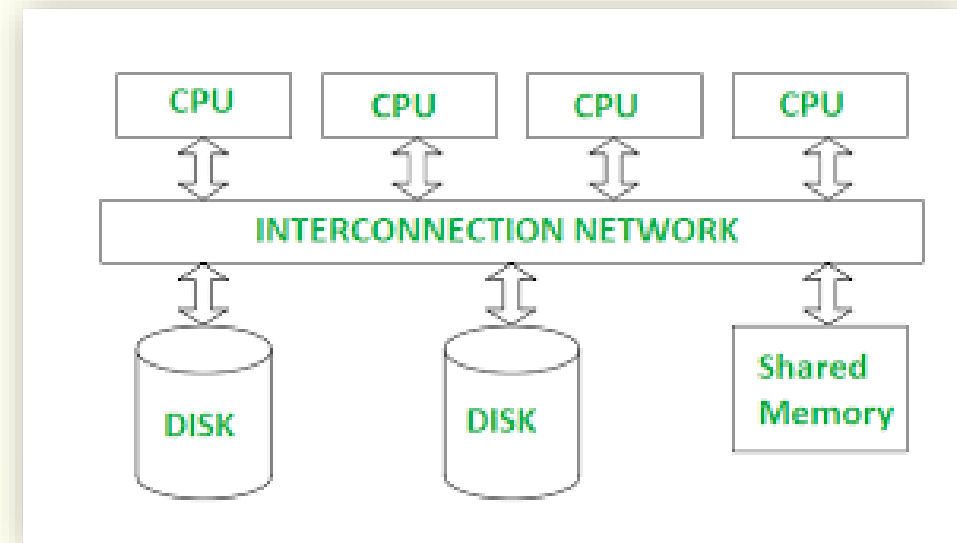
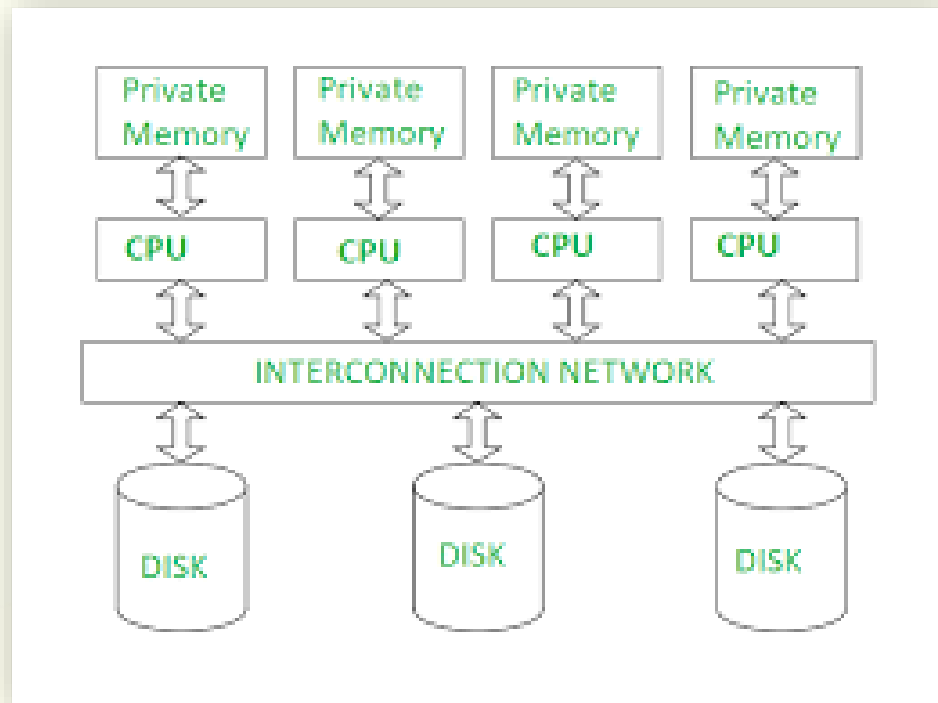
معماری چند مشتری - چند خدمتگزار

- سیستم‌های پایگاه داده همزمان یا ناهمزمان ایجاد می‌شوند.
- اجزای تشکیل‌دهنده سیستم‌ها (OS ها و DBMS ها) معمولاً همگن هستند.
- برخی سایت‌ها ممکن است فقط مشتری و یا خدمتگزار باشند.
- مسئولیت اینکه داده مورد نیاز کاربر هر سیستم مشتری نزد کدام خدمتگزار است برعهده خود مشتری (یا همان برنامه کاربردی) است.



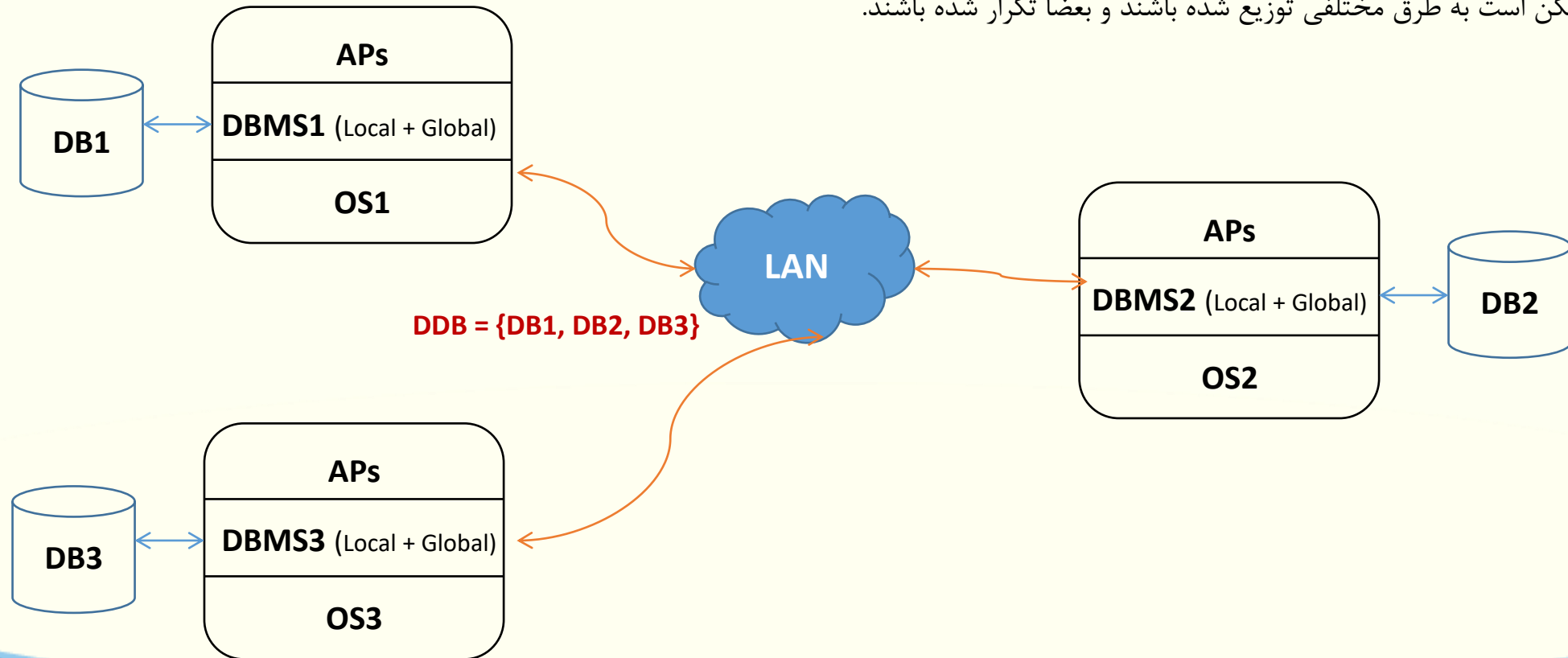
سیستم‌های پایگاه داده موازی (Parallel Database System)

- در معماری موازی با انجام کارها به صورت موازی تلاش بر این است که راندمان سیستم افزایش پیدا کند. عملیات‌هایی که به صورت موازی پردازش می‌شوند می‌تواند خواندن داده‌ها و پردازش درخواست‌ها باشد.



سیستم‌های پایگاه داده توزیع شده (Distributed Database System)

- در این نوع معماری پایگاه داده، داده‌ها در محل‌های مختلفی نگهداری می‌شود. ویژگی مهم این نوع معماری مستقل بودن داده‌های توزیع شده است به این معنا که کاربران بدون نیاز به دانستن محل داده‌ها می‌توانند درخواست‌هایشان را به پایگاه داده ارسال کنند.
- سیستم پایگاه داده توزیع شده، مجموعه‌ای است از چند پایگاه داده منطقاً یکپارچه (مجتمع)، ولی به طور فیزیکی توزیع شده روی یک شبکه کامپیوتری.
- توزیع شدگی از دید برنامه‌ها و کاربران پایگاه داده پنهان است.
- هر سایت دارای یک سیستم مدیریت داده محلی و یک سیستم مدیریت داده توزیع شده است (و می‌تواند هر سایت خود معماری چند مشتری-تک خدمت‌گزار داشته باشد).
- داده‌ها ممکن است به طرق مختلفی توزیع شده باشند و بعضاً تکرار شده باشند.



سیستم مدیریت پایگاه داده (Database Management System)

- سیستم مدیریت پایگاه داده (DBMS) نرم افزاری است که از مجموعه ای از ابزارها و بخش های مرتبط با هم به منظور فراهم آوردن امکان مدیریت کامل اطلاعات ذخیره شده در پایگاه داده تشکیل شده است.
- DBMS پل ارتباطی میان پایگاه داده، کاربر و اپلیکیشن است که وظیفه دارد ارتباط و تعامل میان این اجزا را آسان سازد.

انواع سیستم های مدیریت پایگاه داده :

- انواع مختلفی از سیستم های مدیریت پایگاه داده از نظر ساختار و قوائد ذخیره اطلاعات توسعه داده شده است. هر یک از آنها برای هدف خاصی ایجاد شده اند که نسبت به یکدیگر برتری ها و معایبی نسبی دارند. هر یک از انواع DBMS ها باید با توجه به هدف و نیازی که وجود دارد، به کار گرفته شوند.
- Hierarchical DBMS : ساختار این سیستم به صورت درختی است.
- Network DBMS : این سیستم از ساختار شبکه ای برای ذخیره اطلاعات استفاده می کند.
- Relational DBMS : این سیستم، یکی از پرکاربردترین سیستم های مدیریت پایگاه داده است که ساختار داده ها را به صورت رابطه ای در نظر می گیرد.
- Object Oriented DBMS : در این سیستم، داده ها به صورت شی در نظر گرفته می شوند.
- NoSQL : این سیستم جدیدترین سیستم مدیریت پایگاه داده است که برای زمانی کاربرد دارد که داده ها به صورت جدولی نباشند.

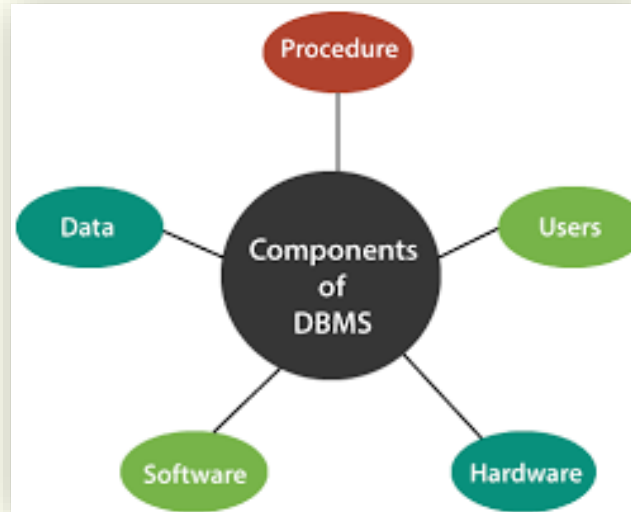
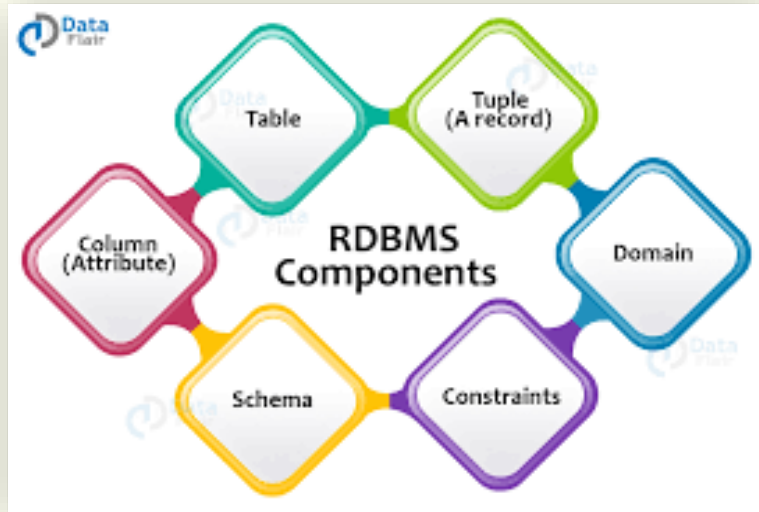
مثالهایی از DBMS :

MySQL , PostgreSQL , SQLite , SQL Server , Oracle , Access , ...



سیستم مدیریت پایگاه داده رابطه‌ای (Relational Database Management System)

سیستم مدیریت پایگاه داده رابطه‌ای (RDBMS) یک نرم افزار سیستم مدیریت پایگاه داده است که داده ها را در قالب جدولی مدیریت و نگهداری می کند.



مزایای سیستم مدیریت پایگاه داده رابطه‌ای :

- امکان ذخیره بسته های بزرگ داده
- حداقل افزونگی داده
- حفظ یکپارچه سازی داده ها
- ابزاری بهتر برای ساختار و سازماندهی داده ها
- اجازه دسترسی چند کاربر

تفاوت DBMS و RDBMS :

- عملکرد : DBMS یک نرم افزار سیستمی برای ایجاد، ذخیره، مدیریت، به روز رسانی و بازیابی داده‌ها از پایگاه‌های داده است ولی RDBMS نرم افزاری است که به منظور بازیابی، به روز رسانی و ذخیره کارآمد داده‌ها، امکان ایجاد و مدیریت پایگاه‌های داده را در قالب جدول فراهم می‌کند.
- ذخیره سازی : در DBMS ذخیره داده‌ها در فایل‌ها است ولی در RDBMS داده‌ها در جداول ذخیره می‌شوند و هیچ سلسله مراتبی وجود ندارد. در عوض از یک مدل رابطه‌ای پیروی می‌کند.
- تعداد کاربران : DBMS تنها می‌تواند از یک کاربر پشتیبانی کند ولی در RDBMS از چندین کاربر پشتیبانی می‌شود.
- نرمال سازی (سازماندهی داده‌ها در پایگاه داده) : DBMS از نرمال سازی پشتیبانی نمی‌کند ولی در RDBMS نرمال سازی فعال است.
- نوع داده : DBMS نمی‌تواند مقدار زیادی داده را ذخیره کند ولی در RDBMS این چنین نیست.
- سایر تفاوت ها : روابط داده‌ها، توزیع پایگاه‌های داده، افزونگی داده‌ها، سخت افزار و نرم افزار مورد نیاز، یکپارچگی داده‌ها، دسترسی به داده‌ها امنیت داده‌ها و ...

- به اشتراک گذاشتن داده‌ها
- کاهش میزان افزونگی
- سازگاری داده‌ها
- تضمین جامعیت و دقت داده‌ها
- امکان اعمال قوانین دقیق امنیتی
- سهولت در دریافت گزارشات متنوع
- تعدد زبان‌های میزبان
- تعدد انواع کاربران
- در دسترس بودن داده‌ها





پایان فصل اول

مهدی دادبخش

mahdi.dadbakhsh@sharif.edu

۱۴۰۱ – ۱۴۰۲