

فصل اول: مقدمه



درس پایگاه داده
دانشگاه صنعتی نوشیروانی بابل
مهدی عمادی
m.emadi@nit.ac.ir



فصل اول: مقدمه

- کاربردهای سیستم پایگاه داده
- اهداف سیستم های پایگاه داده
- **View of Data**
- زبان های پایگاه داده
- طراحی پایگاه داده
- موتور پایگاه داده (Database Engine)
- معماری پایگاه داده (Database Architecture)
- کاربران (Users) و راهبران (Administrators) پایگاه داده
- تاریخچه سیستم های پایگاه داده





کاربردهای سیستم پایگاه داده

■ سیستم مدیریت پایگاه داده (Database Management System) یا به طور خلاصه DBMS شامل اطلاعات در مورد یک سازمان (enterprise) خاص است

- مجموعه داده های مرتبط

- مجموعه ای از برنامه ها برای دسترسی به داده ها

- محیطی راحت و کارآمد برای استفاده

■ سیستم های پایگاه داده برای مدیریت مجموعه ای از داده ها استفاده می شود که:

- بسیار ارزشمند (Highly valuable)

- نسبتاً بزرگ (Relatively large)

- و در دسترسی همزمان چند کاربر و برنامه های کاربردی هستند.

■ سیستم پایگاه داده مدرن یک سیستم نرم افزاری پیچیده است که وظیفه مدیریت یک مجموعه بزرگ و پیچیده داده است.

■ پایگاه های داده همه جنبه های زندگی ما را دربر می گیرد.





مثالهایی از کاربردهای پایگاه داده

- **Enterprise Information**
 - **Sales:** customers, products, purchases
 - **Accounting:** payments, receipts, assets
 - **Human Resources:** Information about employees, salaries, payroll taxes.
- **Manufacturing:** management of production, inventory, orders, supply chain.
- **Banking and finance**
 - customer information, accounts, loans, and banking transactions.
 - Credit card transactions
 - Finance: sales and purchases of financial instruments (e.g., stocks and bonds; storing real-time market data
- **Universities:** registration, grade





مثالهایی از کاربردهای پایگاه داده

- **Airlines: reservations, schedules**
- **Telecommunication: records of calls, texts, and data usage, generating monthly bills, maintaining balances on prepaid calling cards**
- **Web-based services**
 - **Online retailers: order tracking, customized recommendations**
 - **Online advertisements**
- **Document databases**
- **Navigation systems: For maintaining the locations of various places of interest along with the exact routes of roads, train systems, buses, etc.**





هدف سیستم های پایگاه داده

در سالهای ابتدایی توسعه کامپیوتر، برنامه های کاربردی پایگاه داده به طور مستقیم بر روی سیستم فایل ساخته می شدند، که مسائل زیر را به دنبال دارد:

- تکرار داده (data redundancy) و ناسازگاری (inconsistency): داده ها در فرمت های مختلف فایل ذخیره می شوند و در نتیجه اطلاعات در فایل های مختلف کپی می شوند
- دشواری در دسترسی به داده ها
 - نیاز به نوشتن یک برنامه جدید برای انجام هر کار جدید
- انزوای داده (Data isolation)
 - استفاده از چندین فایل و فرمت
- مشکلات جامعیتی (Integrity problems)
 - محدودیتهای یکپارچگی (به عنوان مثال، مانده حساب نمی تواند منفی باشد) در کد برنامه می آیند به جای اینکه به طور مشخص در کنار داده ذخیره شوند.
 - در نتیجه اضافه کردن محدودیت های جدید و یا تغییر آنها بسیار مشکل خواهد شد.





هدف سیستم های پایگاه داده

- تجزیه ناپذیری به روز رسانی ها (Atomicity of updates)
 - خرابی ها (Failures) ممکن است پایگاه داده را در یک حالت متناقض (inconsistent) دهد زیرا به روز رسانی جزئی (partial updates) اتفاق افتاده است.
 - به عنوان مثال: انتقال وجوه از یک حساب به حساب دیگر یا باید کامل انجام شود یا اصلا نشود
- دسترسی همزمان (Concurrent) توسط چندین کاربر
 - برای کارآیی (performance) بهتر نیاز به دسترسی همزمان داریم.
 - دسترسی همزمان کنترل نشده می تواند منجر به ناسازگاری شود
- ▶ دو برنامه همزمان موجودی حساب را ۱۰ می خوانند و ۵ تا از آن کم می کنند و هر دو مقدار ۵۰ را به جای آن می نویسند. با برداشت ۱۰ تومان موجودی ۵ می شود.
- مشکلات امنیتی (Security)
 - تعریف سطوح دسترسی برای کاربران بسیار سخت است.

سیستم های پایگاه داده برای تمام این موارد راه حل ارائه می کند





مثال پایگاه داده دانشگاه

■ در این درس ما از پایگاه داده دانشگاه برای نشان دادن تمام مفاهیم استفاده خواهیم کرد
■ شامل:

● دانشجویان (Students)

● مدرسان (Instructors)

● کلاس ها (Classes)

■ مثال برنامه مورد نیاز:

● افزودن دانشجویان، مدرسان و درس های جدید

● ثبت نام دانش آموزان برای درس ها، و ایجاد لیست دانشجویان کلاس

● ثبت نمرات دانشجویان

● محاسبه معدل (GPA) و ساختن کارنامه دانشجو





سیستم مدیریت پایگاه داده (DBMS)

سیستم مدیریت پایگاه داده ها سیستمی است شامل مجموعه ای از داده های به هم مرتبط و همچنین مجموعه ای از برنامه ها برای مدیریت آن داده ها.

DBMS یک محیط راحت و کارا را برای کاربران پایگاه داده ایجاد می کند.

مثال رابطه ای تجاری:

Microsoft SQL Server ●

Oracle ●

IBM DB2 ●

مثال رابطه ای متن باز:

PostgreSQL ●

MySQL ●

NoSQL ●

Hbase ●

Cassandra ●

mongoDB ●

redis ●

Neo4j ●





دیدگاه داده (View of Data)

■ یک سیستم پایگاه داده مجموعه ای از داده های مرتبط و مجموعه ای از برنامه هایی است که کاربران را قادر می سازد تا این اطلاعات را به دست آورند و اصلاح کنند.

■ از اهداف اصلی سیستم پایگاه داده، ارائه دیدگاه انتزاعی (abstract view) از داده ها است.

● مدل های داده (Data models)

▶ مجموعه ای از ابزارهای مفهومی برای توصیف داده ها، روابط داده ها (data relationships)، معانی داده ها (data semantics) و محدودیت های سازگاری (consistency constraints).

● انتزاع داده (Data abstraction)

▶ پیچیدگی های ساختار داده ها مورد استفاده برای نمایش داده ها در پایگاه داده از نظر کاربران در سطوح مختلف انتزاعی داده مخفی است.





مدل های داده

■ مجموعه ای از ابزارها برای توصیف

● داده

● ارتباط داده ها

● معانی داده

● محدودیت های داده

■ مدل رابطه ای (Relational model)

■ مدل داده موجودیت-رابطه (Entity-Relationship) – مرسوم ترین مدل مورد استفاده در طراحی پایگاه داده

■ مدل داده شیء گرا (Object-based)

● (Object-oriented and Object-relational)

■ داده های نیمه ساختیافته (Semi-structured) – (XML , JSON)

■ دیگر مدل های داده:

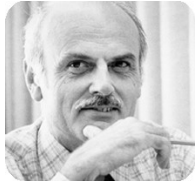
● مدل شبکه (Network model)

● مدل سلسله مراتبی (Hierarchical model)





مدل رابطه ای



Ted Codd

Turing Award 1981

■ همه داده ها در چند جدول مجزا ذخیره می شوند.

■ نمونه یک جدول:

ستون ها (Columns)

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
22222	Einstein	Physics	95000
12121	Wu	Finance	90000
32343	El Said	History	60000
45565	Katz	Comp. Sci.	75000
98345	Kim	Elec. Eng.	80000
76766	Crick	Biology	72000
10101	Srinivasan	Comp. Sci.	65000
58583	Califieri	History	62000
83821	Brandt	Comp. Sci.	92000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
76543	Singh	Finance	80000

سطرها (Rows)





یک نمونه پایگاه داده رابطه ای

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
22222	Einstein	Physics	95000
12121	Wu	Finance	90000
32343	El Said	History	60000
45565	Katz	Comp. Sci.	75000
98345	Kim	Elec. Eng.	80000
76766	Crick	Biology	72000
10101	Srinivasan	Comp. Sci.	65000
58583	Califieri	History	62000
83821	Brandt	Comp. Sci.	92000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
76543	Singh	Finance	80000

(a) The *instructor* table

<i>dept_name</i>	<i>building</i>	<i>budget</i>
Comp. Sci.	Taylor	100000
Biology	Watson	90000
Elec. Eng.	Taylor	85000
Music	Packard	80000
Finance	Painter	120000
History	Painter	50000
Physics	Watson	70000

(b) The *department* table





سطوح انتزاعی (Levels of Abstraction)

■ سطح فیزیکی (Physical level) :

- در مورد نحوه ذخیره سازی رکوردهای داده ای در فایلها صحبت می کند .

■ سطح منطقی (Logical level) :

- در مورد داده های ذخیره شده در پایگاه داده و رابطه های بین آنها صحبت می کند.

```
type instructor = record
```

```
  ID : string;
```

```
  name : string;
```

```
  dept_name : string;
```

```
  salary : integer;
```

```
end;
```

■ سطح دیدگاه (View level) :

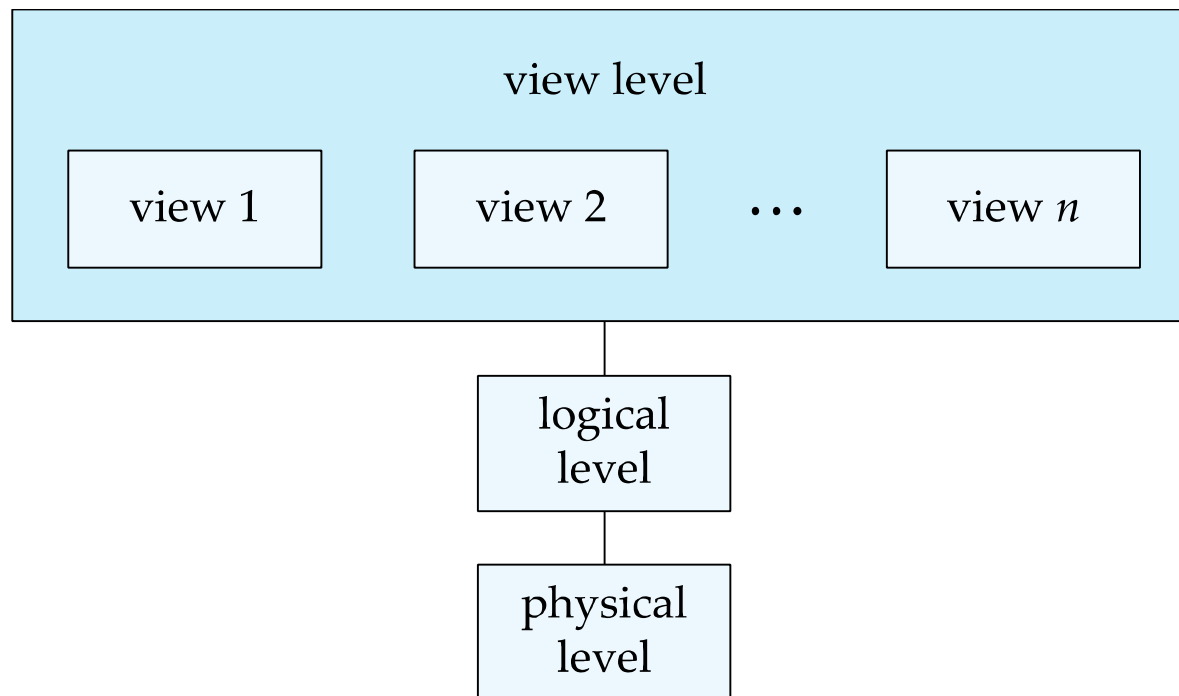
- این سطح در واقع همان برنامه هایی هستند که تمام جزئیات را از دید کاربر مخفی می کنند. (برای سادگی، امنیت و ...)





View of Data

■ معماری یک سیستم پایگاه داده بر اساس سطوح انتزاعی





نمونه ها و شما ها (Instances and Schemas)

- شبیه به نوع داده و متغیر از آن نوع داده در زبان های برنامه نویسی
- **شمای منطقی (Logical Schema) – ساختار منطقی عمومی پایگاه داده**
 - به عنوان مثال: یک پایگاه داده بانکی شامل اطلاعات مشتریان، حساب ها و ارتباط بین آنها
 - در حقیقت شبیه به نوع داده متغیرهای موجود در یک برنامه می باشد.
- **شمای فیزیکی (Physical schema) – ساختار فیزیکی عمومی پایگاه داده**
- **نمونه (Instance) – محتویات واقعی یک پایگاه داده در یک لحظه از زمان**
 - در حقیقت شبیه به مقادیر متغیرهای موجود در یک برنامه می باشد.





زبان های پایگاه داده

■ SQL زبانی است برای توصیف داده ها و دستکاری آنها در مدل رابطه ای

■ زبان توصیف داده ها (DDL)

- `create table account (
 account-number char(10),
 balance integer)`

■ زبان دستکاری داده ها (DML)

- `select customer.customer-name
 from customer
 where customer.customer-id = '192-83-7465'`

■ برای دیگر مدل های داده نیز زبان های خاص آنها تعریف شده است.

- برای مثال `Xquery` و `Xpath` زبان هایی هستند برای دستکاری داده در مدل `XML`

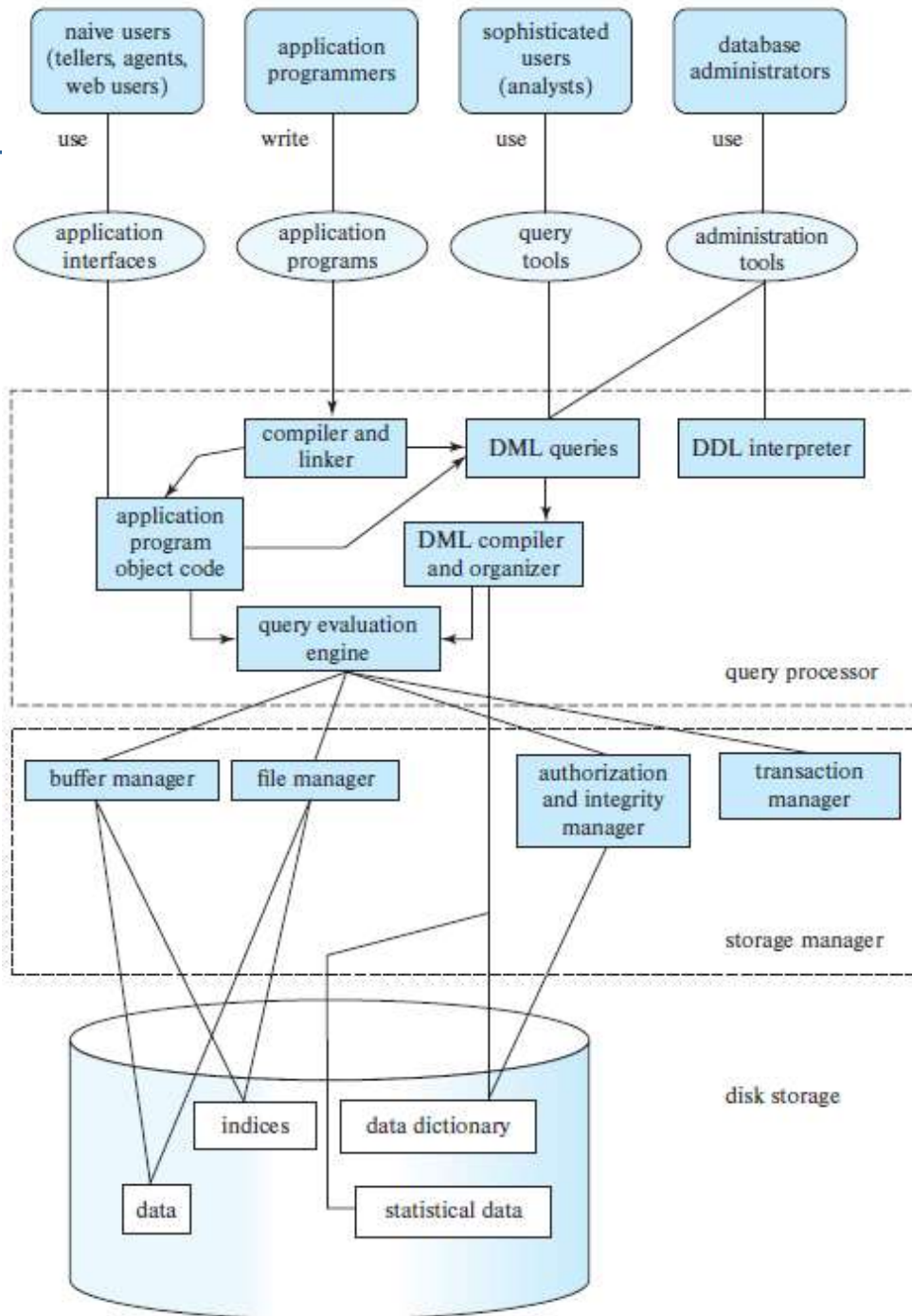




موتور پایگاه داده (Database Engine)

- A database system is partitioned into modules that deal with each of the responsibilities of the overall system.
- The functional components of a database system can be divided into
 - The storage manager,
 - The query processor component,
 - The transaction management component.





موتور پایگاه داده





معماری پایگاه داده (Database Architecture)

- **Centralized databases**
 - One to a few cores, shared memory
- **Client-server,**
 - One server machine executes work on behalf of multiple client machines.
- **Parallel databases**
 - Many core shared memory
 - Shared disk
 - Shared nothing
- **Distributed databases**
 - Geographical distribution
 - Schema/data heterogeneity





برنامه های مبتنی بر پایگاه داده (Database Applications)

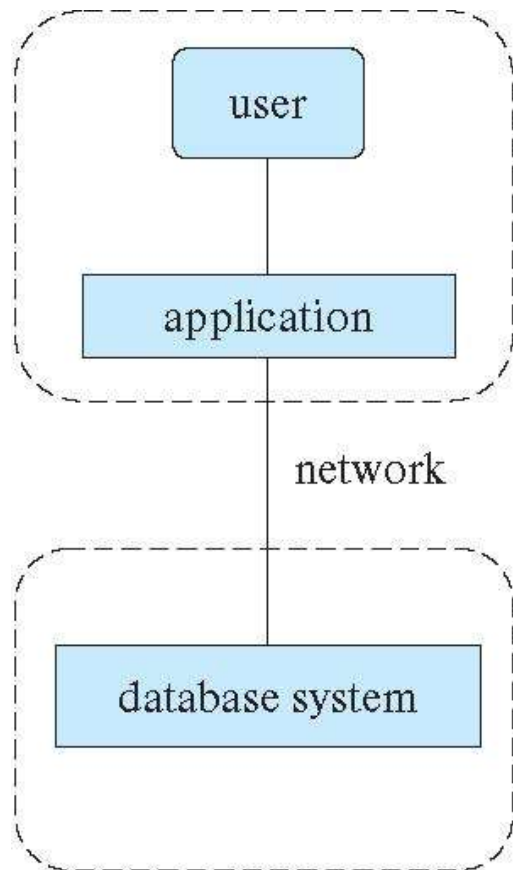
Database applications are usually partitioned into two or three parts

- **Two-tier architecture** -- the application resides at the client machine, where it invokes database system functionality at the server machine
- **Three-tier architecture** -- the client machine acts as a front end and does not contain any direct database calls.
 - The client end communicates with an application server, usually through a forms interface.
 - The application server in turn communicates with a database system to access data.



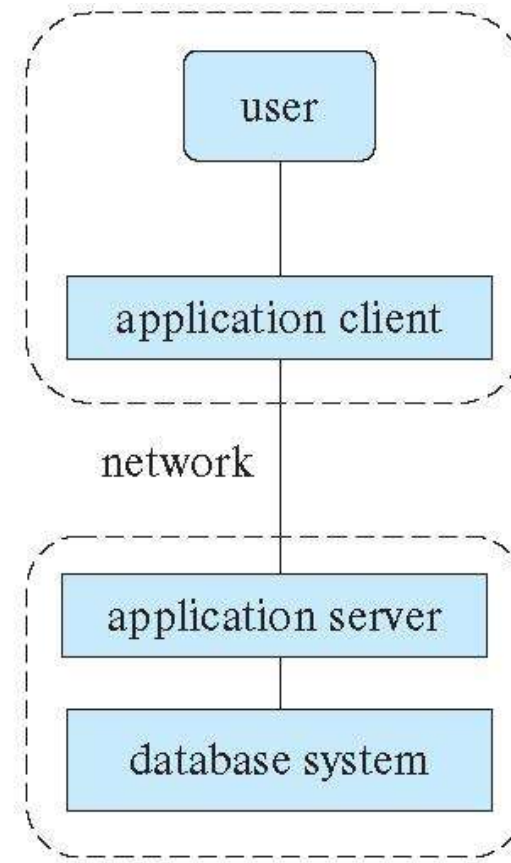


Two-tier and three-tier architectures



(a) Two-tier architecture

client



server

(b) Three-tier architecture





کاربران پایگاه داده

چهار نوع مختلف از کاربران سیستم پایگاه داده وجود دارد

- کاربران بی اطلاع (Naive users): کاربرانی که با استفاده از یکی از برنامه های کاربردی که قبلا نوشته شده است با سیستم ارتباط برقرار می کنند.
- برنامه نویسان (Application programmers): متخصصان کامپیوتر هستند که برنامه های کاربردی را می نویسند.
- کاربران پیشرفته (Sophisticated users): ارتباط برقرار کردن با سیستم بدون نوشتن برنامه ها
 - کار با پایگاه داده به کمک زبان های پرس و جو (SQL, ...)
 - ارتباط به کمک ابزارهای دیگر مانند نرم افزارهای تحلیل داده
- کاربران متخصص (Specialized users): نوشتن برنامه های پایگاه داده تخصصی که در چارچوب پردازش داده های سنتی متناسب نیستند. این افراد به نوعی خود توسعه دهند یک سیستم مدیریت پایگاه داده هستند. به عنوان مثال، CAD، داده های گرافیکی، صوتی، تصویری.
 - این سیستم های در بیشتر موارد به صورت یک افزونه (Extension) به سیستم های مدیریت پایگاه داده موجود اضافه می شوند.





راهنمای پایگاه داده (Database Administrator)

فردی که کنترل مرکزی بر سیستم دارد، یک راهنمای پایگاه داده (DBA) نامیده می شود که مسئولیت های آن عبارتند از:

- تعریف شما (Schema definition)
- تعریف ساختارهای ذخیره سازی و متد دسترسی به داده ها
- بهبود مستمر شما و سازمان فیزیکی داده ها
- اعطای مجوز برای دسترسی (Granting of authorization) به داده ها
- پشتیبانی های معمول
- پشتیبان گیری (backing up) دوره ای از پایگاه داده
- کنترل فضای دیسک و بررسی وجود حافظه آزاد کافی برای عملیات عادی و ارتقاء فضای دیسک در صورت نیاز
- نظارت (Monitoring) بر کارهای (jobs) در حال اجرا در پایگاه داده و پایش پیوسته کارایی (performance) آن.





تاریخچه سیستم های پایگاه داده

- **1950s and early 1960s:**
 - **Data processing using magnetic tapes for storage**
 - ▶ **Tapes provided only sequential access**
 - **Punched cards for input**
- **Late 1960s and 1970s:**
 - **Hard disks allowed direct access to data**
 - **Network and hierarchical data models in widespread use**
 - **Ted Codd defines the relational data model**
 - ▶ **Would win the ACM Turing Award for this work**
 - ▶ **IBM Research begins System R prototype**
 - ▶ **UC Berkeley (Michael Stonebraker) begins Ingres prototype**
 - ▶ **Oracle releases first commercial relational database**
 - **High-performance (for the era) transaction processing**





تاریخچه سیستم های پایگاه داده (ادامه...)

- **1980s:**
 - **Research relational prototypes evolve into commercial systems**
 - ▶ **SQL becomes industrial standard**
 - **Parallel and distributed database systems**
 - ▶ **Wisconsin, IBM, Teradata**
 - **Object-oriented database systems**
- **1990s:**
 - **Large decision support and data-mining applications**
 - **Large multi-terabyte data warehouses**
 - **Emergence of Web commerce**





تاریخچه سیستم های پایگاه داده (ادامه...)

- **2000s**
 - **Big data storage systems**
 - ▶ Google BigTable, Yahoo PNuts, Amazon,
 - ▶ “NoSQL” systems.
 - **Big data analysis: beyond SQL**
 - ▶ Map reduce and friends
- **2010s**
 - **SQL reloaded**
 - ▶ SQL front end to Map Reduce systems
 - ▶ Massively parallel database systems
 - ▶ Multi-core main-memory databases



پایان فصل اول

