

BACS3183

Advanced Database Management

Chapter 9

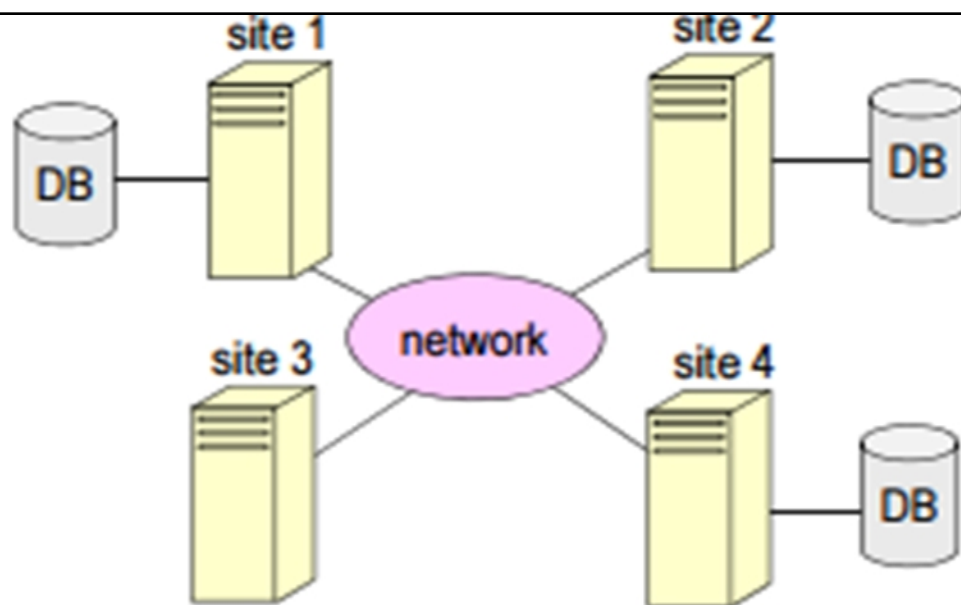
Distributed Database

Learning Outcomes

- Describe **characteristics** of distributed database environments
- Explain **strategies and options** for distributed database design
- Discuss **data replication** and **partitioning**
- Discuss **query processing** in distributed databases
- Describe **distributed concurrency control**

1. Distributed Database

A logically interrelated collection of shared data (and a description of this data), **physically distributed** over a computer **network**.



- A given site may have a local copy of all, part, or none of a particular database.

Centralized Database

Centralized Database is a database that is located, stored, and maintained in a **single location**.

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
-------	-------	-------	------------	-------	---------	-----	--------	-----------	-----

DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
-------	----------------	---------	----------------

DEPT_LOCATIONS

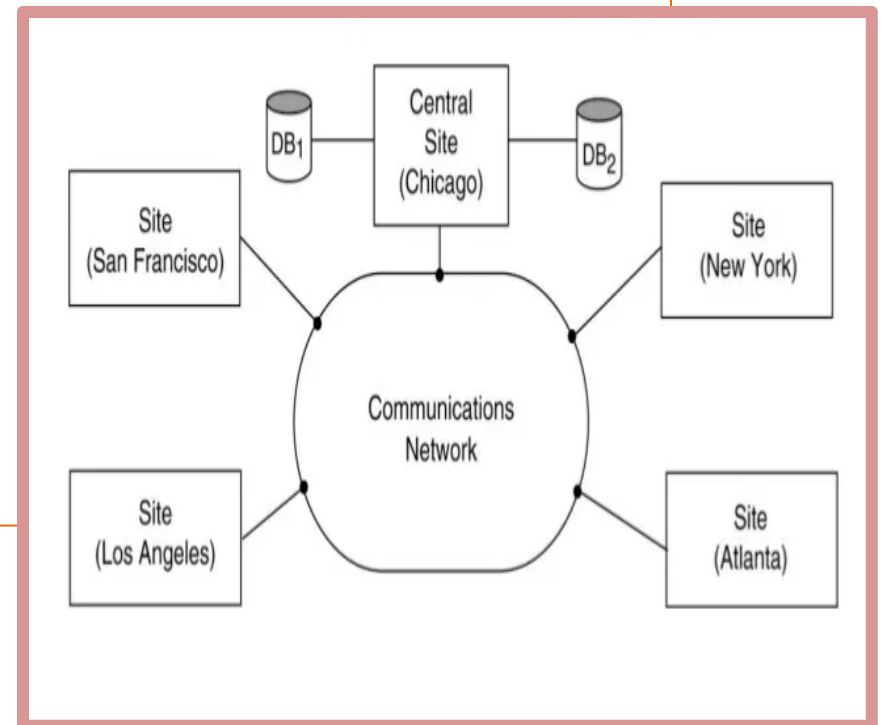
<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
-------	----------------	-----------	------

WORKS_ON

<u>Essn</u>	<u>Pno</u>	Hours
-------------	------------	-------



Distributed Database

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
-------	-------	-------	------------	-------	---------	-----	--------	-----------	-----

DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
-------	----------------	---------	----------------

DEPT_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
-------	----------------	-----------	------

WORKS_ON

<u>Essn</u>	<u>Pno</u>	Hours
-------------	------------	-------

distributed databases.

EMPLOYEES San Francisco
and Los Angeles

PROJECTS San Francisco

WORKS_ON San Francisco
employees

San Francisco

Los Angeles

EMPLOYEES Los Angeles

PROJECTS Los Angeles and
San Francisco

WORKS_ON Los Angeles
employees

EMPLOYEES All

PROJECTS All

WORKS_ON All

Chicago
(Headquarters)

EMPLOYEES New York

PROJECTS All

WORKS_ON New York
employees

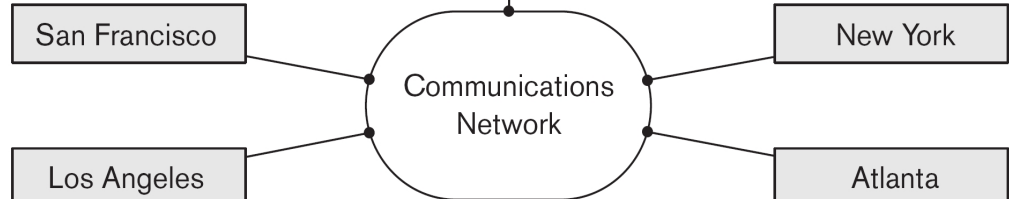
New York

Atlanta

EMPLOYEES Atlanta

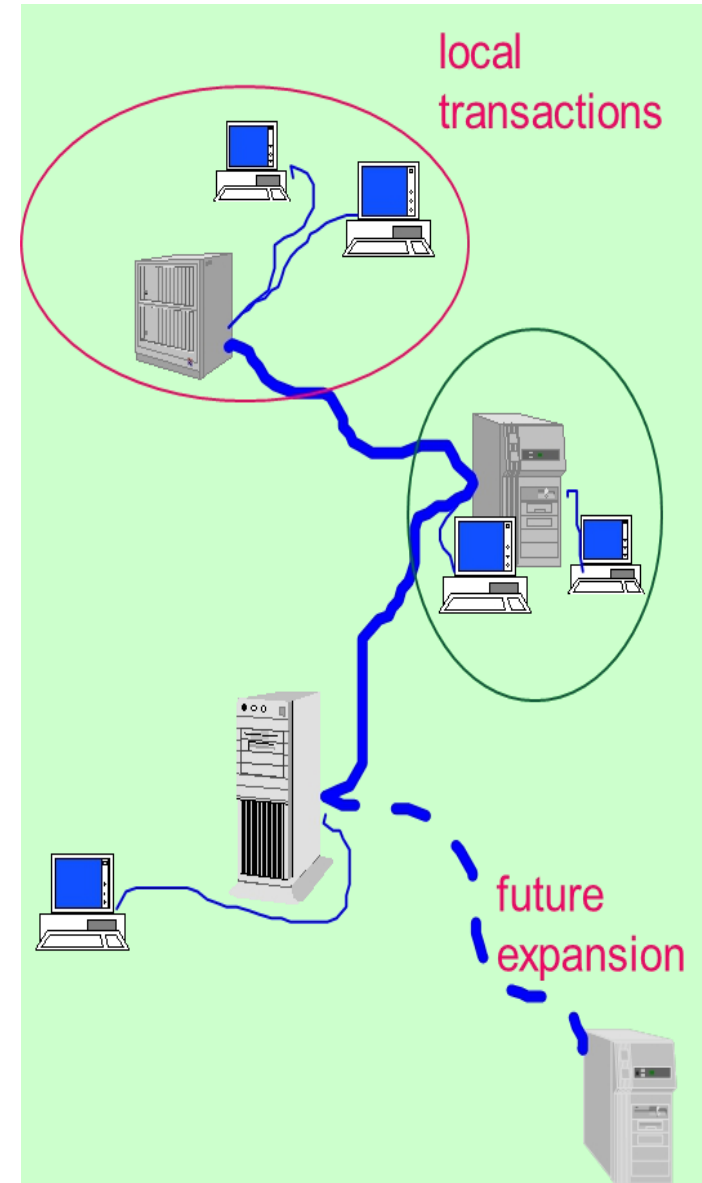
PROJECTS Atlanta

WORKS_ON Atlanta
employees



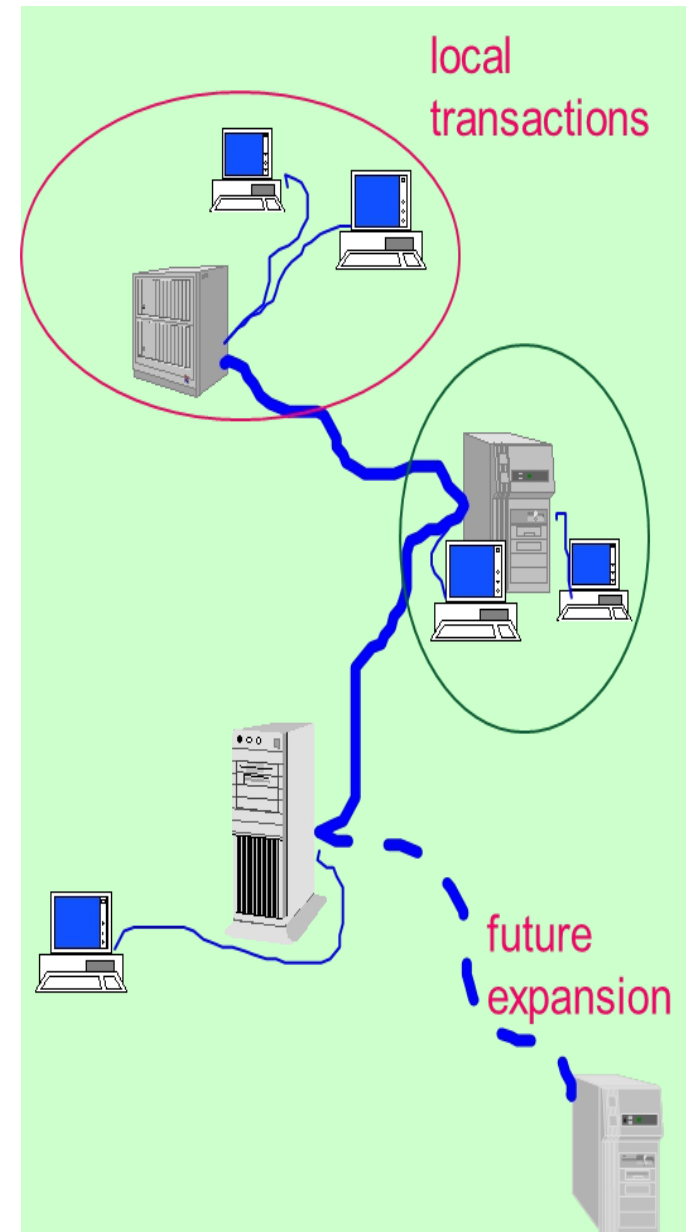
Advantages of DDBMSs

- **Increased reliability/availability** – Even if a component fail, a distributed system will continue to function at some reduced level.
- **Local control over data**
 - Each site controls its own data, security, logging, recovery.
 - Users can access nonlocal data when needed
- **Modular growth** – Easier and more economical to add another computer and data to distributed network
- **Lower communication costs** – Data located closer to their point of use
- **Faster response for certain queries** - Data located closer to their point of use



Disadvantages of DDBMSs

- **Software cost and complexity**
- **Processing overhead** – To ensure **proper coordination** among data at the different sites
- **Data integrity exposure** – Harder to ensure integrity because of increased complexity and need for coordination among data at different sites
- **Slower response for certain queries** – If data are not distributed properly, not located close to their point of use

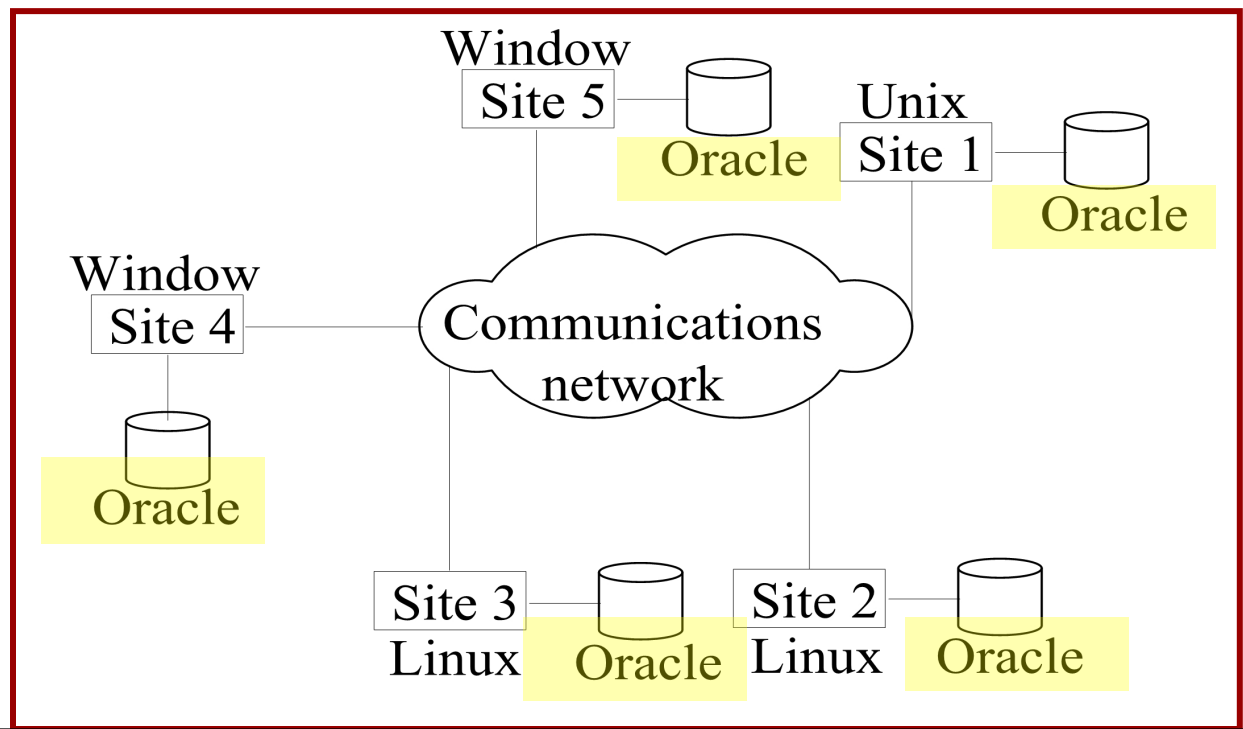


Types of DDBMS

- **Homogeneous DDBMS**
- **Heterogeneous DDBMS**

Homogeneous DDBMS

- All sites use **same DBMS product**.
- The underlying operating systems can be a mixture of Linux, Window, Unix, etc.
- Much **easier to design and manage**.
- Approach provides **incremental growth** and allows **increased performance**.



Heterogeneous DDBMS

- Sites may run **different DBMS products** with **different data models**.
- Occurs when sites have implemented their own databases and integration is considered later.
- **Translations** required to allow for different hardware and/or different DBMS products.
- Typical solution - use **gateways**
 - convert the language and model of each different DBMS into the language and model of the relational system

