CSE 4125: Distributed Database Systems
Chapter - 2

Review of databases and computer networks.

Outline

- Review of Database.
- Review of Computer Networks.

Review of Databases

The Relational Model

- *Relations*: Data stored in tables.
- *Attributes*: Fixed number of columns.
- *Tuples*: Dynamic number of rows.
- *Grade*: Number of attributes.
- *Cardinality*: Number of tuples.

Relational Schema

- Representation of a relation.
- Name of the relation and the attributes appearing in it.
- Example:

$$EMP\ (EMPNUM, NAME, AGE, DEPTNUM)$$

KEYS

- Subset of the attributes whose values are unique.
- Example:

$$EMP\ (\underline{EMPNUM}, NAME, AGE, DEPTNUM)$$

Relational Algebra

- A collection of operations.
- Takes relation(s) as input.
- Produces one relation as result.
- Two types –
  i. *Unary*: 1 input, 1 result
  ii. *Binary*: 2 inputs, 1 result

Unary: Selection

- Example: $\boldsymbol{SL}_{A = a}\, R$

R

A	B	C
a	1	a
b	1	b
a	1	d
b	2	f

S

A	B	C
a	1	a
a	3	f

T

B	C	D
1	a	1
3	b	1
3	c	2
1	d	4
2	a	3

result

Unary: Projection

- Example: $\boldsymbol{PJ}_{A,B}\, R$

R

A	B	C
a	1	a
b	1	b
a	1	d
b	2	f

S

A	B	C
a	1	a
a	3	f

T

B	C	D
1	a	1
3	b	1
3	c	2
1	d	4
2	a	3

result

Binary: Union

- Example: *R* ***UN*** *S*

R

A	B	C
a	1	a
b	1	b
a	1	d
b	2	f

S

A	B	C
a	1	a
a	3	f

T

B	C	D
1	a	1
3	b	1
3	c	2
1	d	4
2	a	3

result

Binary: Cartesian Product

- Example: *R* ***CP*** *S*

R

A	B	C
a	1	a
b	1	b
a	1	d
b	2	f

S

A	B	C
a	1	a
a	3	f

T

B	C	D
1	a	1
3	b	1
3	c	2
1	d	4
2	a	3

result

Binary: Join

- Example: R ***JN*** $_{R.C = T.C}$ T

R

A	B	C
a	1	a
b	1	b
a	1	d
b	2	f

S

A	B	C
a	1	a
a	3	f

T

B	C	D
1	a	1
3	b	1
3	c	2
1	d	4
2	a	3

result

Binary: Semi-join

- Example: $R \; \boldsymbol{SJ}_{R.C = T.C} \; T$

R

A	B	C
a	1	a
b	1	b
a	1	d
b	2	f

S

A	B	C
a	1	a
a	3	f

T

B	C	D
1	a	1
3	b	1
3	c	2
1	d	4
2	a	3

result

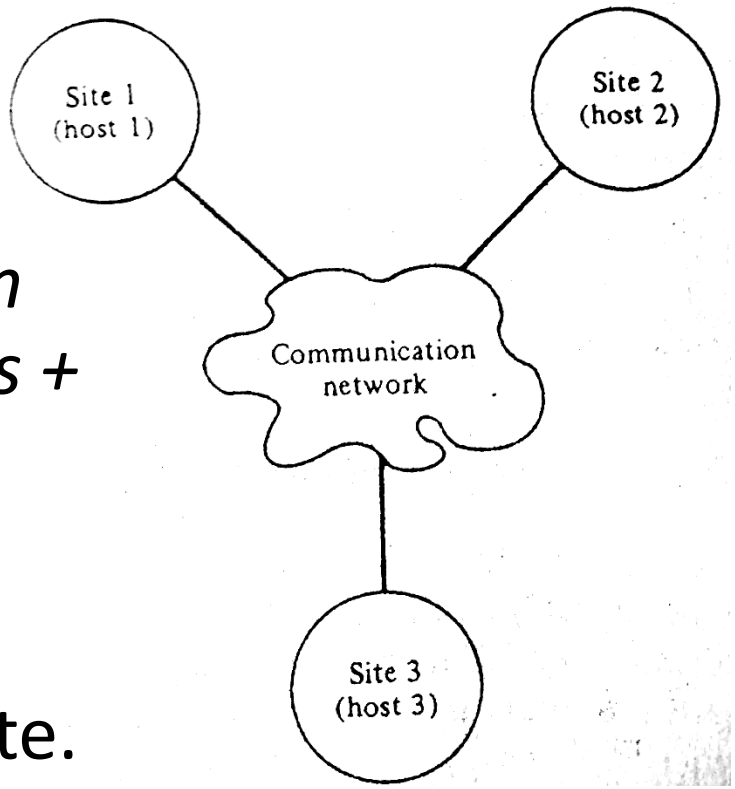
Application, Program and Query

- Database Applications:
  - Sequence of operations requested by end users (not a programmer).
  - Examples: `read (fileName, var).`
- Database Programs:
  - Implementation of the application.
- Query:
  - An expression in a suitable language.
  - Defines a portion of data contained in DB.

Review of Computer Networks

Computer Network

- *Computers (hosts/ sites)* capable of performing autonomous work.
- Connected by *Communication network (communication links + computers).*
- A process running at any site can send a message to a process running at another site.

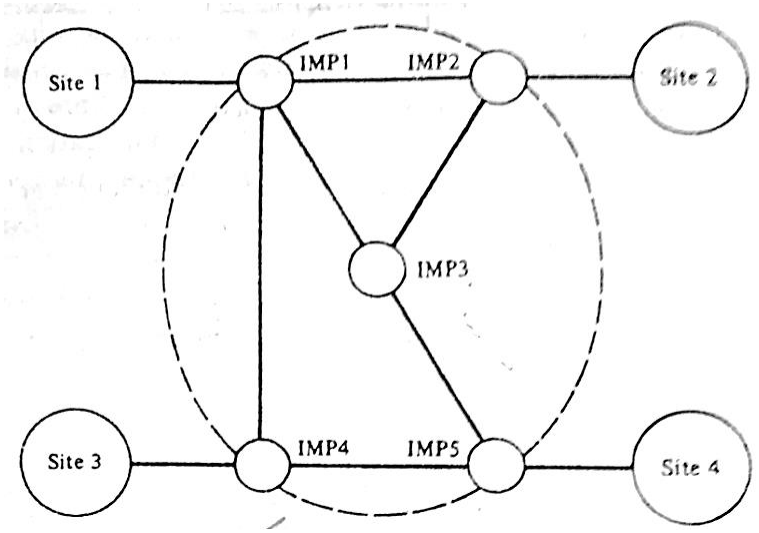


Parameters of CN

- *Delay:* Time with which the message is delivered.
- *Costs:* Fixed costs + costs proportional to message length.
- *Reliability:* Probability of correct delivery of the message.

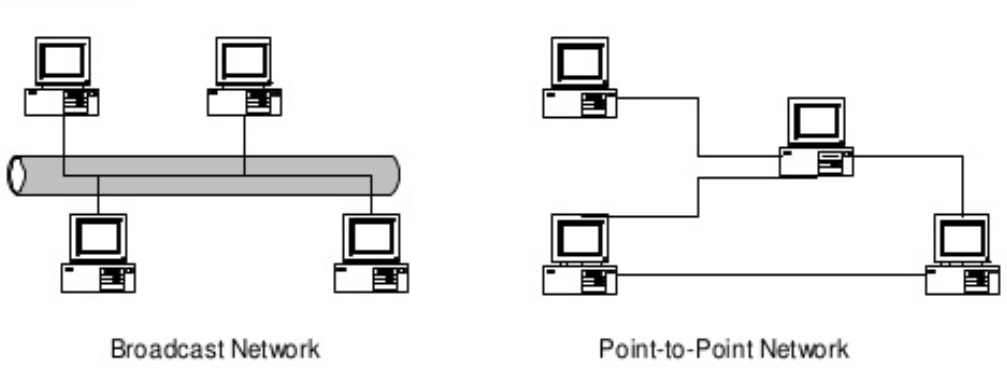
IMPs

- Interface Message Processors:
  - Dedicated processors pairwise connected by communication links.
  - *Routing:* function of choosing path from source to destination.



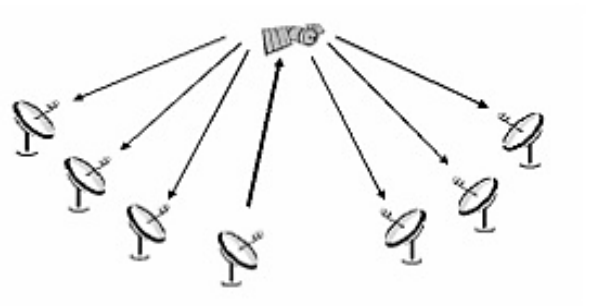
Types of CNs

- Point-to-point:
  - Pairs of sites/ hosts are directly connected.
- Broadcast:
  - All sites/ hosts shares a single communication channel.

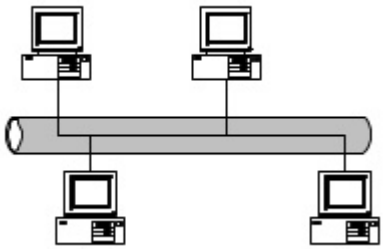


Types of Broadcast

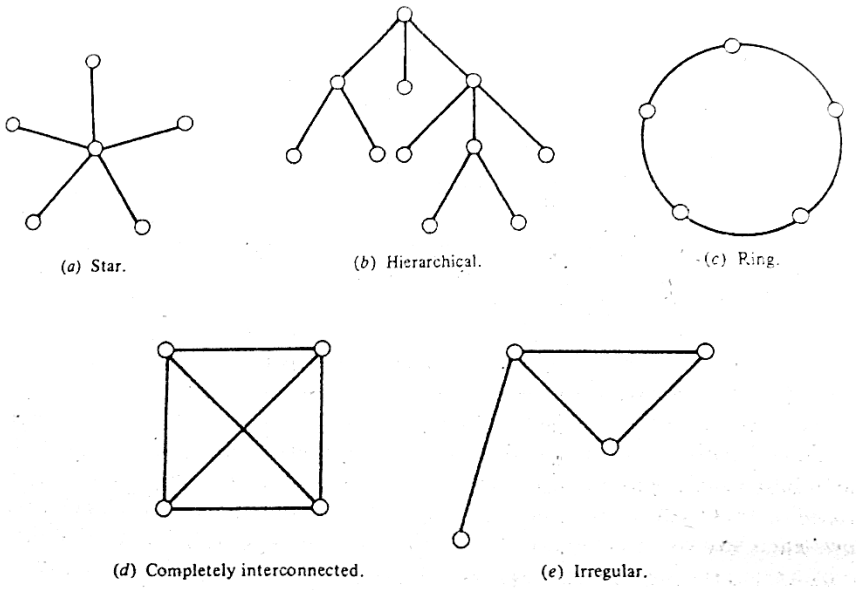
- Satellite



- Common bus



Network Topologies



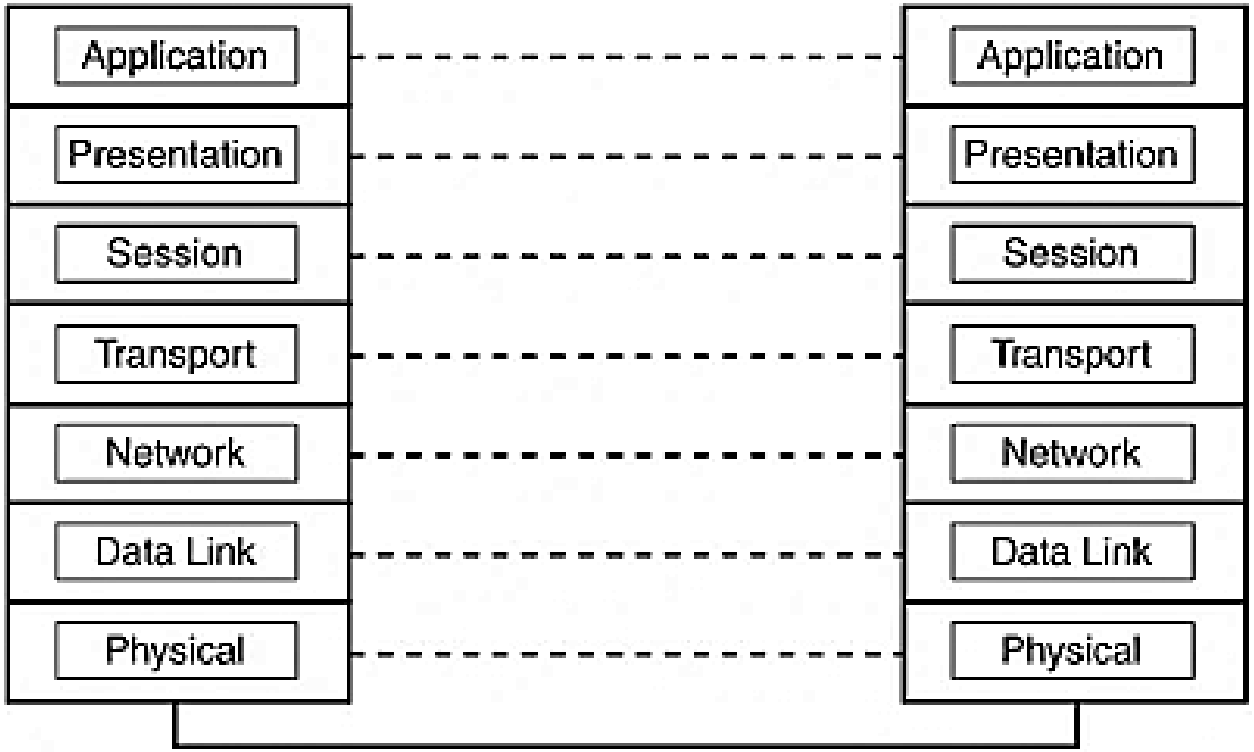
1. Star
2. Hierarchical
3. Ring
4. Completely connected
5. Irregular

Protocols and Sessions

- Protocols:
  - Rules followed by two or more processes for communicating.
  - Example: making a phone call.
- Sessions:
  - A session is established between two processes and is held until all necessary messages have been exchanged.
  - Example: conversation over a phone.

The ISO/ OSI Reference Architecture

- ISO : International Standards Organization
- OSI: Open System Interconnection



The ISO/ OSI Reference Architecture (cont.)

- Application layer:
  - Algorithms and protocols.
- Presentation layer:
  - Conversion of information.
  - Example: character code conversion from sender end to receiver end.
- Session layer:
  - Establishing and maintaining sessions.

The ISO/ OSI Reference Architecture (cont.)

- Transportation layer:
  - True source-to-destination layer.
  - Implements point-to-point channel between source and destination.
  - To perform efficiently, it uses the services of –
    - Network layer.
    - Data-link layer.
    - Physical layer.

Additional Reading

- Aspects of the relational model
- Relational algebra:
  - Difference
  - Natural join
  - Natural semi-join
- Different types of database applications
- Datagram

Sample Questions

a) If *R* and *S* are the input relations, and *T* is the output relation, for which relational algebraic operation(s) the following statements are true?
    i. grade(R) = grade(S) = grade(T)
    ii. grade(R) + grade(S) = grade(T)

b) Give examples of protocol and session from the context of DDB.