CSE 4125: Distributed Database Systems
Chapter – 4
(Part – A)

Distributed Database Design

Framework for DDB Design

- Designing ***conceptual schema*** (description of all the data which are used by the database applications).

- Designing ***physical database*** (mapping conceptual schema to storage area and defining access methods).

Framework for DDB Design

The distribution of the database adds two following problems to the previous problems of the design of the data distribution –

- Designing fragmentation (how global relations are sub divided into fragments).
- Designing allocation of fragments (how fragments are mapped to physical images).

Objectives of the Design of Data Distribution

- **Processing locality.**
  - –Maximize processing locality, Placing data as close as possible to the application using them.

- **Availability and reliability.**
  - –Multiple copies of data.
  - –Recovery.

- **Workload distribution.**
  - –Taking advantage of the powers and computer resources at each site.
  - –Maximize parallel execution of application.

- **Storage costs and availability.**
  - –CPU, I/O and transmission costs.
  - –Considering the storage limitation.

Approaches to Design the Data Distribution

- **Top-Down approach.**
  - Most attractive for systems which are developed from scratch.

Design global schema → Design Fragmentation schema → Allocate fragments to sites → Perform physical design of the data

Problem: When the distributed database is developed as the aggregation of existing databases, is not easy to follow top down approach.

Approaches to Design the Data Distribution

- **Bottom-Up approach.**
  - –Have existing databases at different sites.
  - –How to integrate them
  - –How to deal with heterogeneity and autonomy (i.e. independence).

Selection of a common database model → Translation of local schema → Integration of local schema into global schema

The Design of Fragmentation

1. Design of Horizontal Fragmentation
   I. Primary
   II. Derived

2. Design of Vertical Fragmentation

3. Design of Mixed Fragmentation

The Design of Primary Horizontal Fragmentation

Simple Predicate

Given a global relation $\boldsymbol{R\ (A_1, A_2, \ldots, A_n)}$ where attribute $\boldsymbol{A_i}$ has domain $\boldsymbol{D_i}$,

A simple predicate $\boldsymbol{p_j}$ ***defined on R has the form***

$$\boldsymbol{p_j : A_i\, \theta\ Value}$$

Where $\boldsymbol{\theta \in \{=, <, \neq, \leq, >, \geq\}}$ and $\boldsymbol{Value \in D_i}$

Example: Given global relation ***J.***

J

JNO	JNAME	BUDGET	LOC
J1	Instrumental	150,000	Montreal
J2	Database Dev.	135,000	New York
J3	CAD/CAM	250,000	New York
J4	Maintenance	350,000	Orlando

Simple predicates: $p_j : A_i \theta \ Value$

p_1: *JNAME = "Maintenance"*

p_2: *BUDGET <= 200,000*

Minterm Predicate

Given a set of simple predicates for relation ***R:***

$$\boldsymbol{P = \{ p_1, p_2, \ldots, p_m \},}$$

The set of minterm predicates: $\boldsymbol{M = \{ m_1, m_2, \ldots, m_n \}}$ is defined as,

$$M = \{ m_i \mid m_i = \bigwedge_{p_j \in P} p_j^* \}$$

where $p_j^* = p_j$ or $p_j^* = \neg(p_j)$.

Provided that, $\boldsymbol{m_i \neq false.}$

Example:

TITLE	SAL
Elect. Eng.	40,000
Syst. Analy.	54,000
Mech. Eng.	32,000
Programmer	42,000

Possible simple predicates:

P_1: TITLE="Elect. Eng."
P_2: TITLE="Syst. Analy"
P_3: TITLE="Mech. Eng."
P_4: TITLE="Programmer"
P_5: SAL<=35,000
P_6: SAL > 35,000

Some corresponding minterm predicates:

$$m_1 : TITLE = "Elect.Eng." \wedge SAL \leq 35{,}000$$

$$m_2 : TITLE \neq "Elect.Eng" \wedge SAL > 35{,}000$$

Exercise

b) Consider the following relation R:

Title	*Experience (months)*	*Salary (thousands)*
Electrical Engg.	12	50
System Analyst	10	60
Mechanical Engg.	5	30
Programmer	7	80
Sales Manager	9	42

Given, the following minterm predicates for R:

$$m_1 : Title =' Programmer' \wedge Experience \leq 12 \wedge Salary \geq 35$$

$$m_2 : Title \neq' Programmer' \wedge Experience > 12 \wedge Salary \geq 35$$

$$m_3 : Title \neq' Programmer' \wedge Experience > 12 \wedge Salary \geq 35$$

Write the set P of all possible simple predicates for m_1, m_2 and m_3.