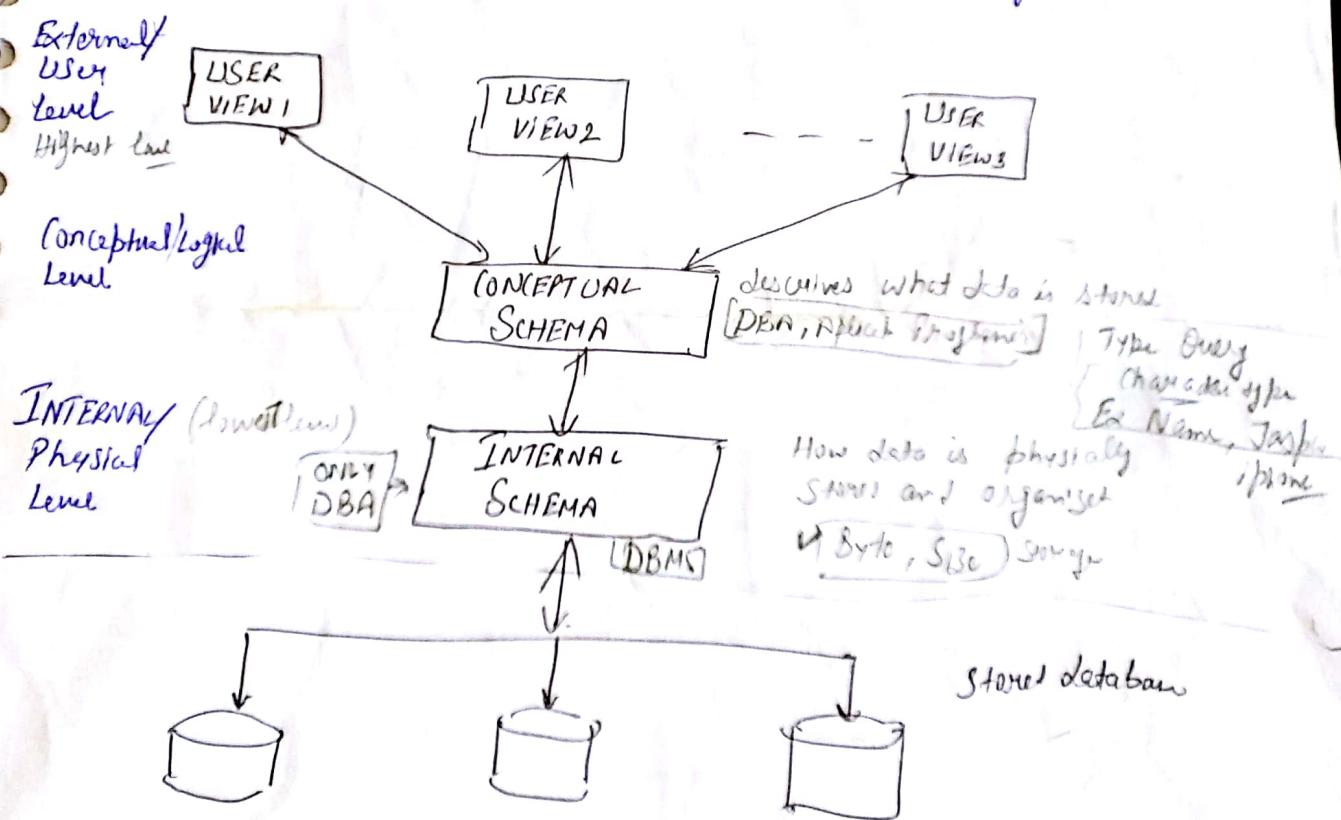


DBMS ARCHITECTURE → describes how data in database viewed by user.

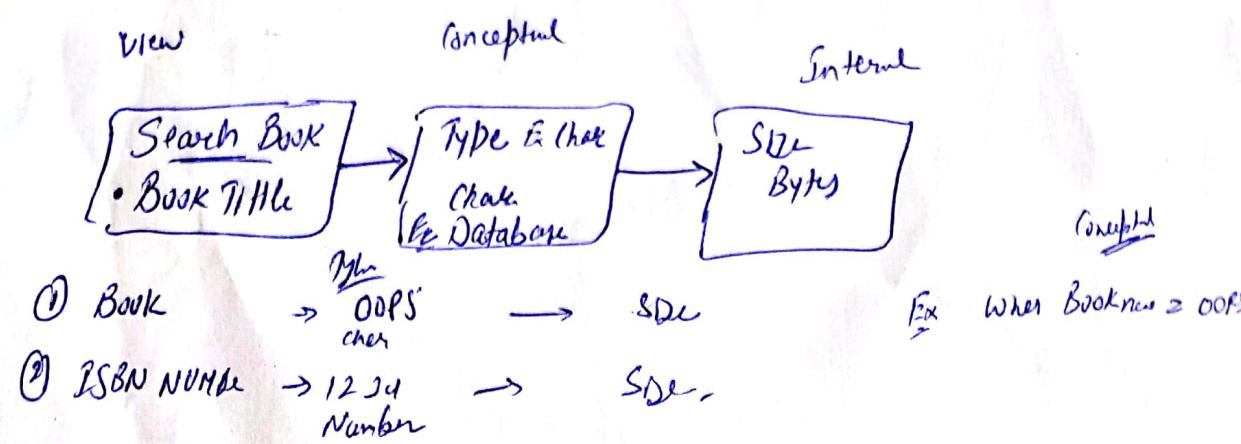
(3)

- DBMS architecture is called three-level architecture.
- The goal of three level architecture is to separate the user application and physical database.
- The major purpose of DBMS is to provide user with an abstract view of data.

Abstraction means to hide certain details of how data is stored and maintained
Complexity



Ex



1 Physical Level

Ex (char [10])

3rd ID [10]

Physical

- It is the lowest level of abstraction.
- It describes how data is actually stored.
- Physical level has an internal schema which describes the physical storage structure of database.

2 Conceptual Level

Table

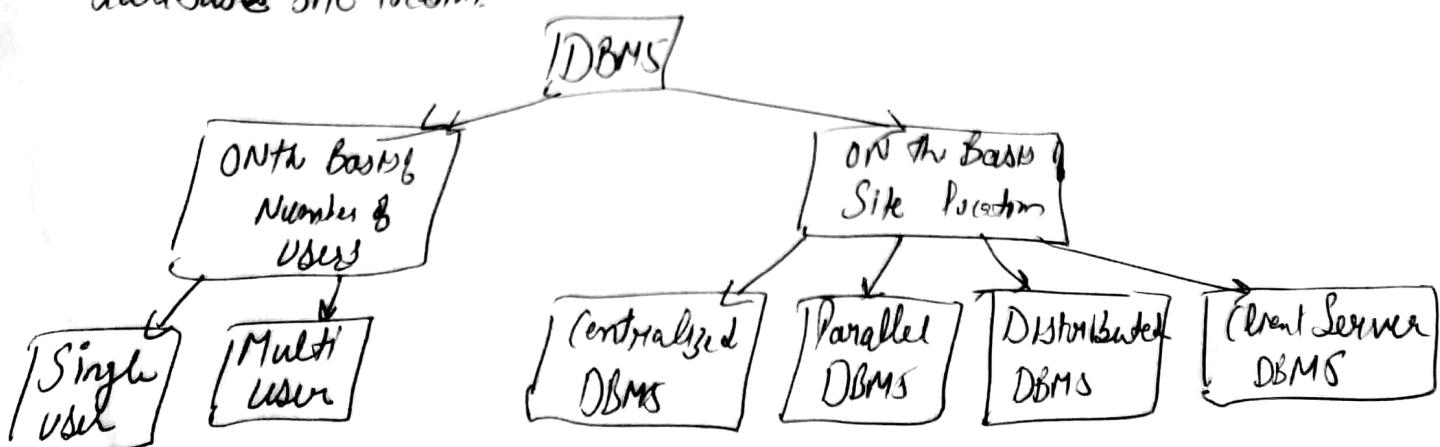
- It is the next higher level of abstraction.
- This schema hides the details of physical storage structure and concentrates on describing entities, data types, relationships, and constraints.

3 View Level

- This is highest level of abstraction.
- The view level includes number of user views.
- Different users may need only a part of the entire database.
- So it describes only required part user

2 Types of Database System

DBMS is classified on the basis of number of users and the database site location.



Classification of DBMS

1) On the Basis of Number of Users

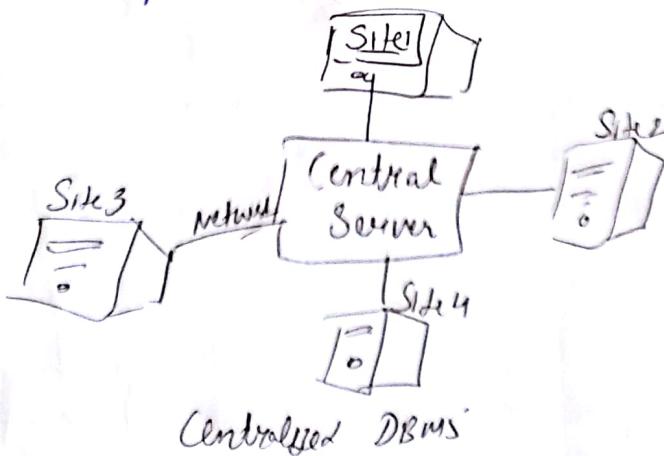
(5)

- ① Single User DBMS + In single user DBMS, database resides on one computer and it is only accessed by one user at a time.
- ② Multiuser DBMS + The multiple users access the data from one central storage area so that the database integrated.
[Integrated means data is not stored in two or more files] \Rightarrow One place only.

2) On the Basis of Site Location

1) Centralized System

- In centralized system, the database resides on some single central location.
- A number of processors can access this central database via some computer network.
- These processors can connect to the central database via some computer network.



Eg Railway Reservation System.

- Where the database is centrally located at New Delhi.

Disadvantage

When central site computer goes down, Then the users are blocked from using system until the system comes back.

2 Parallel Database System

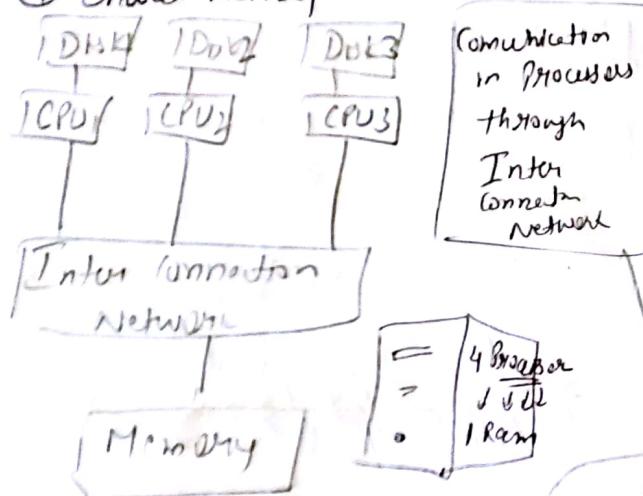
(8)

507

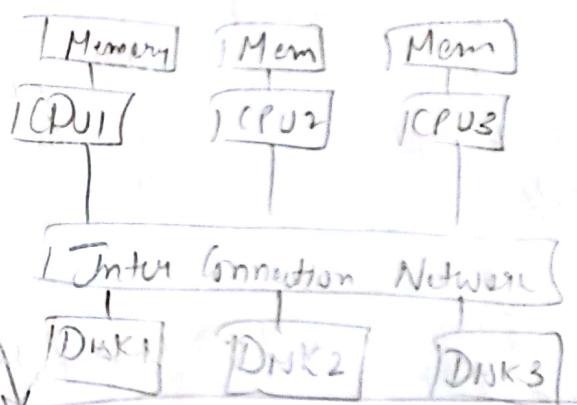
- Parallel database system architecture consists of Central Processing Units and data storage disks in parallel.
- They improve processing and Input/Output (I/O) speed.
- Parallel database systems are used in the applications that have to process an extremely large number of transactions per second or have to query large databases.

There are three architectures in Parallel DBMS

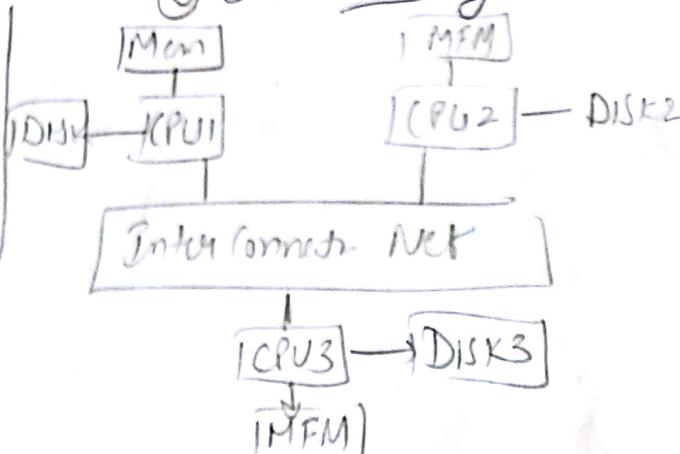
① Shared Memory



② Shares Disk



③ Shared nothing



④ Shared Memory (Ram)

Blocks of Ram can be used by different CPU's (Processors) in Multiprocessor Computer System

Disadvantage

- It uses shared resources.
- For ex - If Disk1 corrupt
Not any system can use Disk1 data,

⑦

1) 507

3) Distributed Database System

- In distributed DBMS, the database is split into number of fragments.

instance

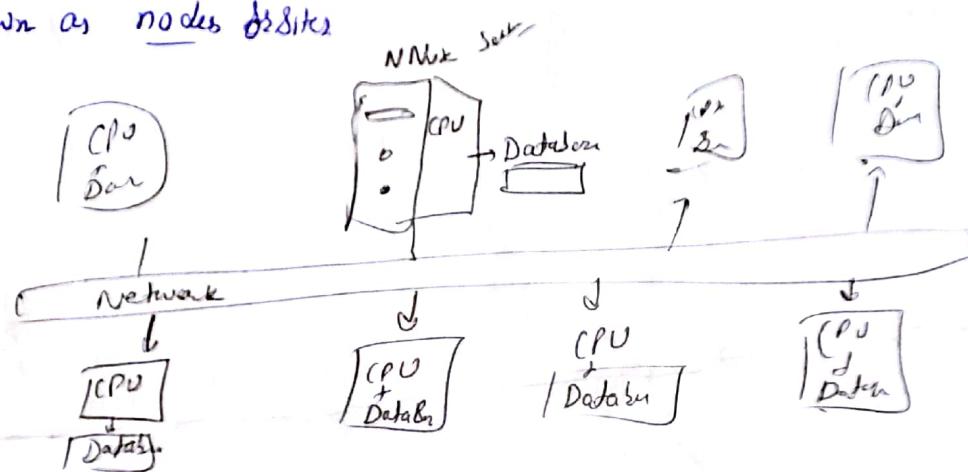
- Each fragment is stored on one or more computers.

- These computers are connected by a communication network.

- The computers in distributed system ~~are~~ may vary in size and function.

techniques

Known as nodes or sites



a time
to

hours

hardly

Advantages

- ① Scalability
- ② No problem if any node fails. (Replication factor)

Disadvantage

- Recovery from failure is complex.

It is

4) Client-Server DBMS

- Client/Server architecture of database management system has two logical components, namely Client and Server.

- Clients are generally personal computers ~~or workstations~~.

- Whereas servers are large workstations.

- The application and tool of DBMS run on one or more client ~~platforms~~ platforms while DBMS software resides on the server.

- Server handles clients requests for each + database access & update.

- Client is front end, Server back end

Data Independence

1) [507]

- In old days, programs stored data in regular files.
- Each program has to maintain its own data
 - ✓ huge overhead.
 - ✓ error prone.
- Development of DBMS helped to achieve data independence (transparency).
- Provide Centralized controller data maintenance and Access.
- Changes made at one level of DB system does not affect other levels.

(1) Logical DI:

Capacity to change conceptual schema without changing external schema/application program.

- ↳ we can change conceptual Schema
- ⇒ expand Database by adding record type, data
- ⇒ To change constraints
- ⇒ To reduce Database (or removing data item).

(2) Physical DI:

Capacity to change internal schema without having to change conceptual Schema.

- ⇒ Changes in Internal Schema needed because
 - Physical files were organized.
 - Improve performance of retrieval or update.

PL/SQL Concepts

- SQL has some disadvantages, PL/SQL came into existence

Disadvantages of SQL

- 1 SQL does not have procedural capabilities:
• SQL does not provide the programmers with the techniques of condition checking, looping and branching.
- 2 SQL statements passed to the Oracle engine one at a time.
This leads to increase in traffic & decrease in the speed of data processing.
- 3 SQL statements are passed to the Oracle engine.
- 4 SQL has no facility for program handling of errors during manipulation of data.
Oracle engg. display their own errors.
But in PL/SQL we have exception handling facility. Ex: exception %

Advantages of PL/SQL1 Block Structure:

- PL/SQL consists of blocks of code which can be nested with in each other
- It send entire block of SQL statement to oracle engine all in one go.
- Reduce network traffic, it processes the code much faster
- All changes made to the data in the table are done or undone in one go.

2 Procedural Language Capabilities:

- It consists of Procedural language constructs such as Conditional statements and loops.

3 Better Performance

- PL/SQL processes multiple SQL statements in a single block of code.
- Reduces network traffic & Improve the performance time of Oracle engine.

4 Error Handling:

- Errors or exceptions are effectively handled during the execution of a program.
- Once an exception is caught specific actions can be taken depending on the type of exception.

5 Portability

- Applications written in PL/SQL are portable to any computer hardware or operating system where Oracle is operational.

PL/SQL What is PL/SQL

- PL/SQL stands for Procedural Language extension of SQL
- PL/SQL is combination of SQL along with procedural features of programming languages.
- It was developed by Oracle Corporation in early 90's to enhance capabilities of SQL

PL/SQL Block Structure

10

PL/SQL Block consists of three sections

- 1 The Declaration Section (optional)
- 2 The Execution Section (mandatory)
- 3 The Exception Section (or Error) Optional.

Sample

DECLARE

Variable Declaration

BEGIN

Program Execution

Exception

Exception handling

END;

• Declaration Section (optional)

Declaration section starts with
DECLARE key word.

• Used to declare placeholder like
variables, constants, records and cursors.
which stores data temporarily -

2 Execution Section

• Starts with keyword BEGIN and ends with END.

• This is mandatory section

• Where the program logic is written to perform any task.

• Like loops & conditional statements.

3 Exception Section

• Starts with EXCEPTION keyword.

• This section is optional

• Any errors in the program can be handled in this section

PL/SQL Block Structure

(13)

(11)

DECLARE	Declaration of memory variable constants, cursors etc.
BEGIN	Executable statements Querries, loops,
Exception	Code to handle errors.
END	END of PL/SQL block.

PL/SQL Execution environment

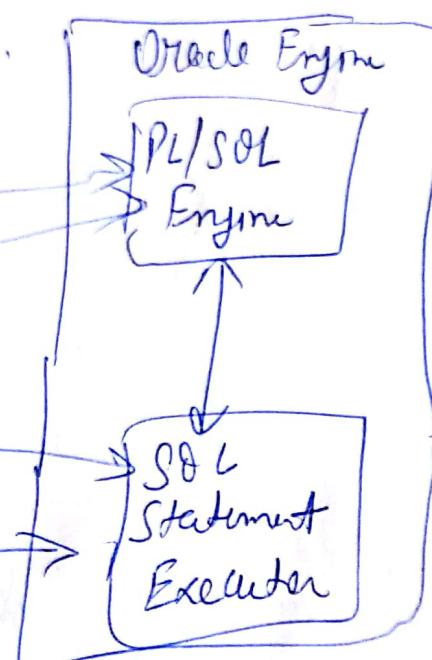
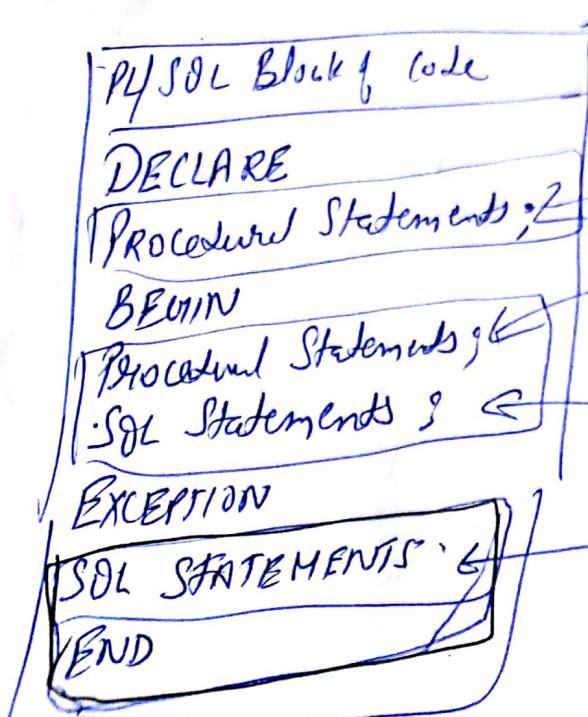
① Oracle Engine consists of

PL/SQL Engine

SQL Statement Executor

Procedural
Statements are executed

SQL Statements are
executed



RDBMS Datatypes

(13)

(12)

- 1 Number → For storing numeric data : 123 number ~~not~~ id (0,0) dev and plan
2 Char → For storing character data ex: a, b, c id (10,3)
3 Date → For storing Date & time data.
4 Boolean → For storing True, False or Null values
5 Varchar → For storing Alpha-Numerical values
- Varchar • Varchar 2
① 2000 bytes 4000 bytes
Length difference • Does not occupy
② Occupies space for null values.

Place holders

3 Literal, Variable, Constant

① Literal = value

It is numeric value, character or a string.

Numeric Literal
↓
Integer Float
25 2.35

String Literal

Ex: 'Tom'

② Single character 't', 'a'

Characters or a set of characters enclosed within single quotes

PL/SQL Placeholder

(13)

- Placeholders are temporary storage area.
- Placeholders can be variable or constant.

Variable

- Variable name must begin with a character & can be followed by maxm 29 characters.
- Reserved words can not be used as variable names unless indosed in double quotes.

Syntax

Variable name datatype (Size);
 or salary number (10);

Assigning Value

- 1) salary number (10) := 10000;
- 2) salary := 20,000;
 Variable := Value;
- 3) Selecting or fetching table (data) value into Variable.
 Select ~~Salary~~ ^{Steps} into ~~Salary~~ ^(variable) from emp
 where id = 2;

- XML
- Stands For extensible self descriptive language to exchange information between systems in any format.
 - Widely used on World Wide Web as a means of communication between different types of web components.
 - It is a ~~old~~ W3C (World Wide Web Consortium) Standard International Web Development Organisation
 - XML is Meta Language. This means that you use it for creating languages. (You can create your own markup language)
 - XML is an extensible, concept. → It is designed to be a self descriptive language.

XML Usage

1 XML is mainly used for Transporting, Storing and Sharing Data.

Ex. • Software Updates.

- Web based feeds, such as Weather reports, stock updates, job etc.
- It is also used in many other aspects of Web Development and Third party Applications.

↳ Every XML document is text based and platform independent.

• XML is a Markup language or XML like HTML.

A Simplified version of SGML Standard Generalized Markup Language

Structure (Hierarchical) • It can describe data in tree or graph structures

Lecture 5

(18)

XML

XML stands for Extensible Markup Language. (9)

Markup :-

① Plain Text
Easy to Edit
Platform Independent

(Importance)

Important Characteristics of XML: (a) Platform Independent

① XML is extensible

It allows you to create your own self-descriptive tags or language, that suits your application.

② XML carries the data, doesn't present it.

It is much more like HTML.
But XML allows you to store data,
irrespective of how it will presented.

③ XML is a public Standard.

XML was developed by organisation called the World Wide Web (W3C).

XML is markup language that defines set of rules for incoming document in a format, that is both human readable & machine readable.

Markup is information added to a document that enhances its meaning in certain ways.

For <message>
<text> Hello (12) </text>
</message>

) Markup symbols or tags. In above

) are <message> - - </message>

- * XML Separates Data From HTML.
 - If you need to display dynamic data in your HTML document. It will take a lot of work to edit HTML each time the data changes.
 - With XML, data can be stored in separate XML files. In this way you can concentrate on using HTML/CSS for display & layout only.
- * XML Simplifies Data Transfer.
 - One of the most time consuming challenges for developer is to exchange data between incompatible systems over the Internet.
 - Exchanging data as XML greatly reduces ~~this~~ this complexity. Because XML is platform independent, either hardware or software.

(16)

XML is not a replacement for HTML.

- It is most important to understand that XML is not a replacement for HTML.

In most applications, XML is used to transport data, while HTML is used to format and display the data.

→ XML Issues / Lacks

XML

Writ

key

① Jaspreet Singh
9592 232401

1 Structured Data

- Relational Databases are highly structured.
- You must define schema before entering any data.
- Every Row conforms to the table schema.
- Changing the schema is hard and may break many things.

2 Unstructured Data

- Data is Free form
- There is no predefined, and it's hard to (like we can insert any data, I Ex restrict strg.)
- Readers need to infer structures and meanings.
(obscure data)
text file, audio, video, images.

3 Semi-Structured Data

- In Semistructured Data, the schema information is mixed with the data values
- Combination of structures and unstructured data.

html code,
Book! - Chapters, sections, titles, Paragraphs, references, index etc.

There are two main structuring concepts to construct an XML.

- ① Element (tag)
- ② Attributes (

(2)

① Elements

Everything from Element start tag to the element's end tag.

RULES

- 1) XML tags are Case Sensitive.

</Adress> This is a wrong </Address>

<Adress> ----- </Adress> ✓

- 2) XML tags must be closed in an appropriate order i.e. an XML tag opened inside another element must be closed before the outer element is closed.

<outer-element>

<Project>

<internal-element>

<Worker>

(Dontent

There are 3 workers.

</internal-element>

</Worker>

</outer-element>

</Project>

- 3) An element name can contain any alphanumeric characters.

The only punctuation marks allowed in names are the hyphen (-), underscore (-) and period (.).

Attributes

- An attribute defines the property of the element. 3
- ~~name its case in the~~
- Attributes of element are separated by white spaces.
- It associates a name with a value.
 - ~~at name = "Value"~~
 - name is followed by an = sign and string value inside " " or ' ' single quotes.

Syntax

```
<element name attribute1 attribute2>  
-- content  
</element name>
```

Empty Element

- Element with no content

```
<Contact_info> </Contact_info>
```

Ex `<recipe name = "Bread" prep_time = "5 mins">`

`<title> Basic Bread </title>`

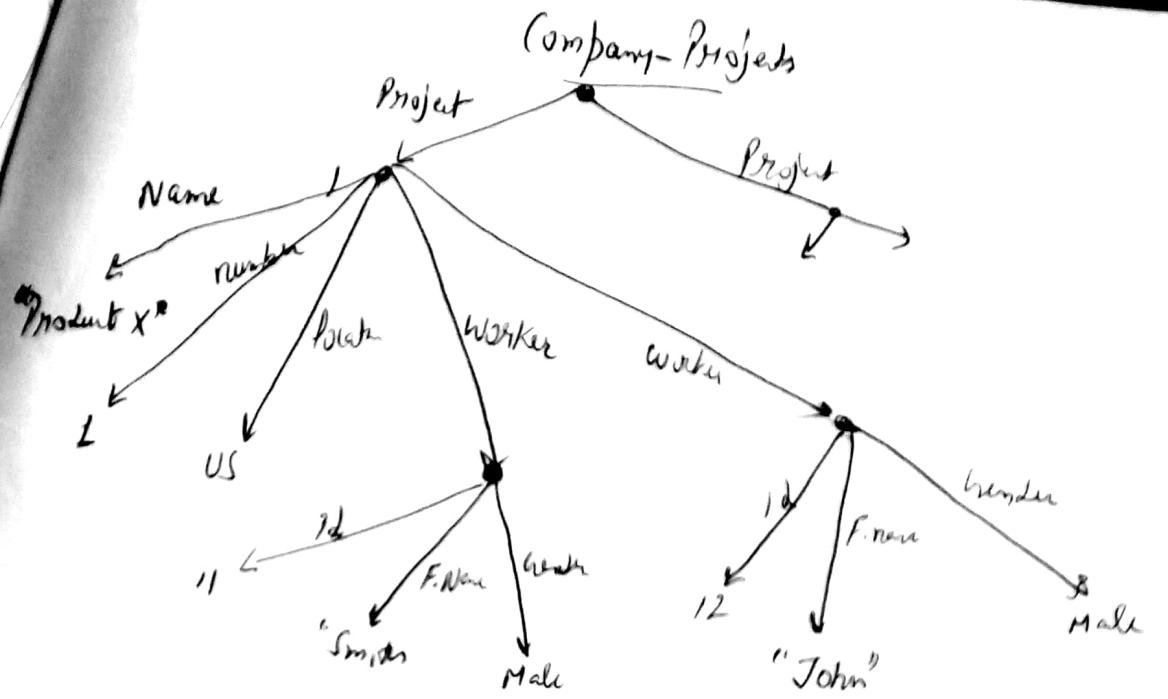
`<ingredient amount = "8" unit = "grams">`

Flour

`</ingredient>`

`<ingredient amount = "10" unit = "teaspoon"> Salt </ingredient>`

`</recipe>`



XML Hierarchical (Tree) Data Model

Representing semistructured data as a graph.

```

<Company-Projects>
  <Project>
    <name> Product. X </name>
    <Number> 1 </Number>
    <Location> US </Location>
    <Worker>
      <Id> 11 </Id>
      <F.Name> Smith </F.name>
      <Gender> Male </Gender>
    </Worker>
    <Worker>
      <Id> 12 </Id>
      <F.Name> John </F.name>
      <Gender> Male </Gender>
    </Worker>
  </Project>
  </Company-Projects>

```

Validation is a process by which an XML document is validated. An XML document is said to be valid if its contents match with the elements, attributes, and associated document type declaration (DTD) and if the document complies with the constraints expressed in it.

Validation is dealt in two ways by the XML parser. They are:

- ✓ • Well-formed XML document
- ✓ • Valid XML document

Well-formed XML Document

An XML document is said to be **well-formed** if it adheres to the following rules:

- Non DTD XML files must use the predefined character entities for **amp(&)**, **apos(single quote)**, **gt(>)**, **lt(<)**, **quot(double quote)**.
- It must follow the ordering of the tag. i.e., the inner tag must be closed before closing the outer tag.
- Each of its opening tags must have a closing tag or it must be a self-ending tag (**<title>....</title>** or **<title/>**).
- It must have only one attribute in a start tag, which needs to be quoted.
- **amp(&)**, **apos(single quote)**, **gt(>)**, **lt(<)**, **quot(double quote)** entities other than these must be declared.

Example

Following is an example of a well-formed XML document:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<!DOCTYPE address
[
    <!ELEMENT address (name,company,phone)>
    <!ELEMENT name (#PCDATA)>
    <!ELEMENT company (#PCDATA)>
    <!ELEMENT phone (#PCDATA)>
]>
<address>
    <name>Tanmay Patil</name>
    <company>TutorialsPoint</company>
    <phone>(011) 123-4567</phone>
```

WELL FORMED AND VALID XML

1 A doc is
Well Formed XML Document

↳ When it is structurally and syntactically correct.
(that means tag are nested properly, case sensitive, elements close properly)

2 Adheres to general XML rules.

- It is not necessary that a Well Formed XML Doc is Valid.
Well Formed XML Doc: may be Valid or Invalid.
↓ either it follows XML rules or it is syntactically correct.

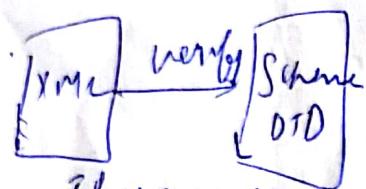
→ Because ~~Valid~~ simply means an XML Doc is Valid.
When it ~~satisfies~~ ~~this document~~ satisfies a schema document.

This mean there is another document, this ^{XML} document points to:
~~schema~~

• This Schema Document is known as DTD [Document Type Declaration]
Valid XML

If an XML document is well formed and
has an associated ~~document~~ DTD,

Then it is said to be a Valid XML doc.



If right then it is valid

```
</address>
```

The above example is said to be well-formed as:

- It defines the type of document. Here, the document type is element type.
- It includes a root element named as address.
- Each of the child elements among name, company, and phone is enclosed in its self-explanatory tag.
- Order of the tags is maintained.

Valid XML Document

If an XML document is well-formed and has an associated Document Type Declaration (DTD), then it is said to be a valid XML document. We will study more about DTD in the chapter XML - DTDs.

15. XML – DTDs



The XML Document Type Declaration, commonly known as DTD, is a way to describe XML language precisely. DTDs check vocabulary and validity of the structure of XML documents against grammatical rules of appropriate XML language.

An XML DTD can be either specified inside the document, or it can be kept in a separate document and then liked separately.)

Syntax

Basic syntax of a DTD is as follows:

```
<!DOCTYPE element DTD identifier  
[ Root elem  
  declaration1  
  declaration2  
  ....  
 ]>
```

In the above syntax,

- ✓ The **DTD** starts with **<!DOCTYPE** delimiter.
- ✓ An **element** tells the parser to parse the document from the specified root element.
- **DTD identifier** is an identifier for the document type definition, which may be the path to a file on the system or URL to a file on the internet. If the DTD is pointing to external path, it is called **External Subset**.
- ✓ The **square brackets []** enclose an optional list of entity declarations called **Internal Subset**.

Internal DTD

A DTD is referred to as an internal DTD if elements are declared within the XML files. To refer it as internal DTD, standalone attribute in XML declaration must be set to yes. This means, the declaration works independent of an external source.)

Syntax

Following is the syntax of internal DTD:

```
<!DOCTYPE root-element [element-declarations]>
```

where *root-element* is the name of root element and *element-declarations* is where you declare the elements.

Example

Following is a simple example of internal DTD:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<!DOCTYPE address [
    <!ELEMENT address (name,company,phone)>
    <!ELEMENT name (#PCDATA)>
    <!ELEMENT company (#PCDATA)>
    <!ELEMENT phone (#PCDATA)>
]>
<address>
    <name>Tanmay Patil</name>
    <company>TutorialsPoint</company>
    <phone>(011) 123-4567</phone>
</address>
```

Let us go through the above code:

Start Declaration - Begin the XML declaration with the following statement.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
```

DTD - Immediately after the XML header, the *document type declaration* follows, commonly referred to as the DOCTYPE:

```
<!DOCTYPE address [
```

The DOCTYPE declaration has an exclamation mark (!) at the start of the element name. The DOCTYPE informs the parser that a DTD is associated with this XML document.

DTD Body - The DOCTYPE declaration is followed by the body of the DTD, where you declare elements, attributes, entities, and notations.

```
<!ELEMENT address (name,company,phone)>
<!ELEMENT name (#PCDATA)>
<!ELEMENT company (#PCDATA)>
<!ELEMENT phone_no (#PCDATA)>
```

Several elements are declared here that make up the vocabulary of the <name> document. <!ELEMENT name (#PCDATA)> defines the element name to be of type "#PCDATA". Here, #PCDATA means parseable text data.

End Declaration - Finally, the declaration section of the DTD is closed using a closing bracket and a closing angle bracket (>). This effectively ends the definition, and thereafter, the XML document follows immediately.

Rules

DTD

- The document type declaration must appear at the start of the document (preceded only by the XML header) — it is not permitted anywhere else within the document.
- Similar to the DOCTYPE declaration, the element declarations must start with an exclamation mark.
- The Name in the document type declaration must match the element type of the root element.

External DTD

In external DTD elements are declared outside the XML file. They are accessed by specifying the system attributes which may be either the legal .dtd file or a valid URL. To refer it as external DTD, standalone attribute in the XML declaration must be set as no. This means, declaration includes information from the external source.

Syntax

Following is the syntax for external DTD:

```
<!DOCTYPE root-element SYSTEM "file-name">
```

where file-name is the file with .dtd extension.

Example

The following example shows external DTD usage:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<!DOCTYPE address SYSTEM "address.dtd">
<address>
    <name>Tanmay Patil</name>
    <company>TutorialsPoint</company>
    <phone>(011) 123-4567</phone>
</address>
```

Q

XMI

Parus

for ~2

The content of the DTD file **address.dtd** is as shown:

```
<!ELEMENT address (name,company,phone)>
<!ELEMENT name (#PCDATA)>
<!ELEMENT company (#PCDATA)>
<!ELEMENT phone (#PCDATA)>
```

ENT

ENT

Types

You can refer to an external DTD by using either **system identifiers** or **public identifiers**.

System Identifiers

A system identifier enables you to specify the location of an external file containing DTD declarations. Syntax is as follows:

```
<!DOCTYPE name SYSTEM "address.dtd" [...]>
```

As you can see, it contains keyword **SYSTEM** and a URI reference pointing to the location of the document.

UNIFORM
RESOURCE Locator

UNIFORM
RESOURCE Locator

Public Identifiers

Public identifiers provide a mechanism to locate DTD resources and is written as follows:

```
<!DOCTYPE name PUBLIC "-//Beginning XML//DTD Address Example//EN">
```

As you can see, it begins with keyword **PUBLIC**, followed by a specialized identifier. Public identifiers are used to identify an entry in a catalog. Public identifiers can follow any format, however, a commonly used format is called **Formal Public Identifiers**, or FPIs.

create → catalog file
store → remote path
remove →

URI = **file:///path/to/book/dtd/addr.dtd** />

// Element PCDATA

<!ELEMENT element-name (#PCDATA)

(1)

[Jasprat Singh]
9592232441]

<!ELEMENT Name (#PCDATA)

Ans <Name> Umesh </Name>

DID

① Element

② Attribute

Attributes

(2)

<!ATTLIST element-name attribute-name attribute-type attribute-value >

<!ATTLIST Employee Employee-Number CDATA #REQUIRED >

<Employee Employee-Number = "105" > XML <Employee ~~105~~ >
 (element) (attribute)
 </Employee>

Wrong

</Employee ?

Attribute Type

CDATA → The value is character Data (Any thing)

ID → Unique
PC DATA → Parser Character Data.

Attribute-Value

① Value → default value of attribute

② # REQUIRED → The attribute is required

③ # IMPLIED → The attribute is optional

④ # FIXED Value → The attribute value is fixed

⑤ <!ATTLIST Employee Employee-Number CDATA #REQUIRED >

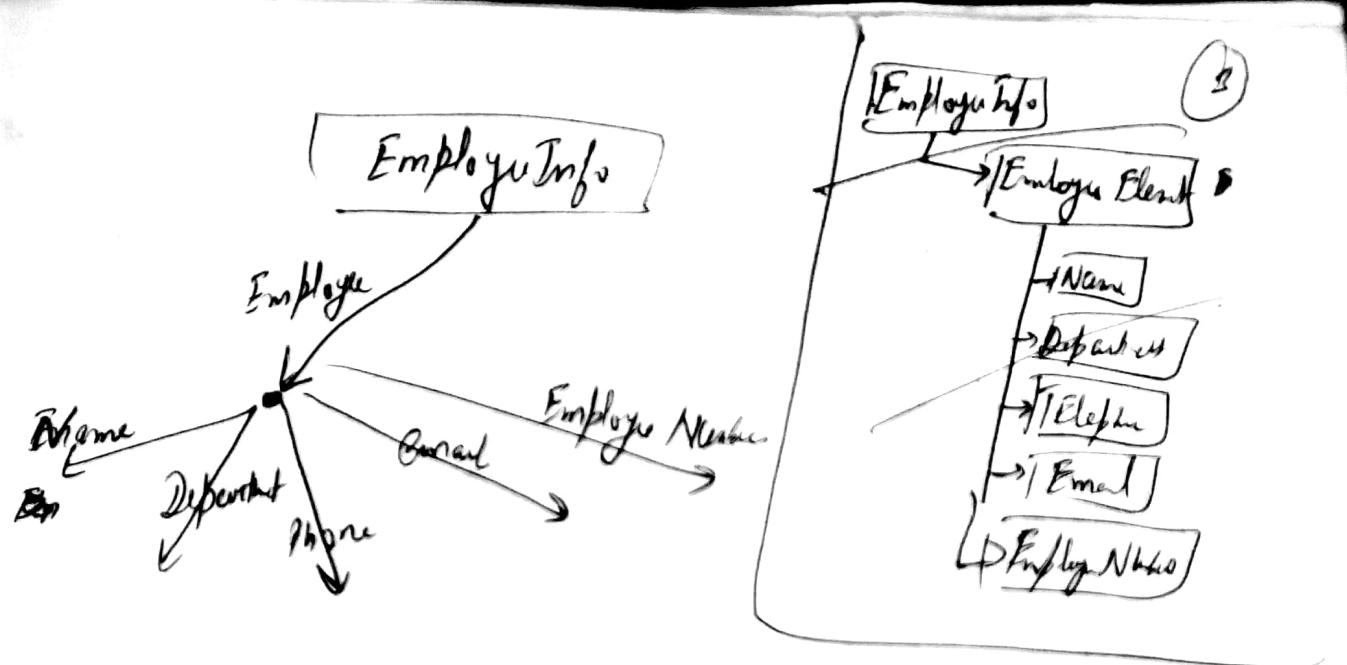
⑥ <!ATTLIST Employee Employee-Number CDATA # IMPLIED >

Implicit (optional)

<Employee Employee-Number = "105" > <Employee >
 </Employee> </Employee>

⑦ <!ATTLIST Employee Employee-Number CDATA # FIXED "105" >

... "105" ...



EmployeeInfo Element : Does not include text data.

Employee Element Element : 0 or more times : Does not include text data.

→ **Name** Element : Once : Text Data

→ **Department** Element : Once : Text Data

→ **Phone** Element : Once : Text Data

→ **Email** Element : Once : Text Data

Employee Number Attribute : Integer Data : Required

External DTD

```
<? XML version = "1.0" encoding = "UTF-8" standalone = "no" ?>
<!DOCTYPE employee-Info SYSTEM "employee-Info.dtd">
```

Note: XML FILE

```
<Employee-Info>
```

```
</Employee-Info>
```

Challenges, Characteristics	Computational Intelligence in Communication	Data Science (ICCIDS 2017)	(IEEE INDEXING)
and Tools			

30-2
1
<?xml version = "1.0" encoding = "UTF-8" stand alone = "yes"?> ④

<!DOCTYPE Employee-Info [

<!ELEMENT Employee-Info (Employee)*>

<!ELEMENT Employee (Name, Department, Phone, Email)>

<!ELEMENTS NAME (#PCDATA)>

<!ELEMENTS DEPARTMENT (#PCDATA)>

<!ELEMENTS PHONE (#PCDATA)>

<!ELEMENT Email (#PCDATA)>

<!ATTLIST Employee Employee-Number CDATA #REQUIRED >

]>

<Employee-Info>

<Employee Employee-Number = "105">

<Name> Umesh </name>

<Department> CSE </Department>

<Phone> 9592000001 </Phone>

<email> ips@gmail.com </email>

</Employee>

</Employee-Info>

Element Other Answer

① + 1 or more Minimum one occurrence of Element
 <!-- Element Employee-Details (Employee) # >

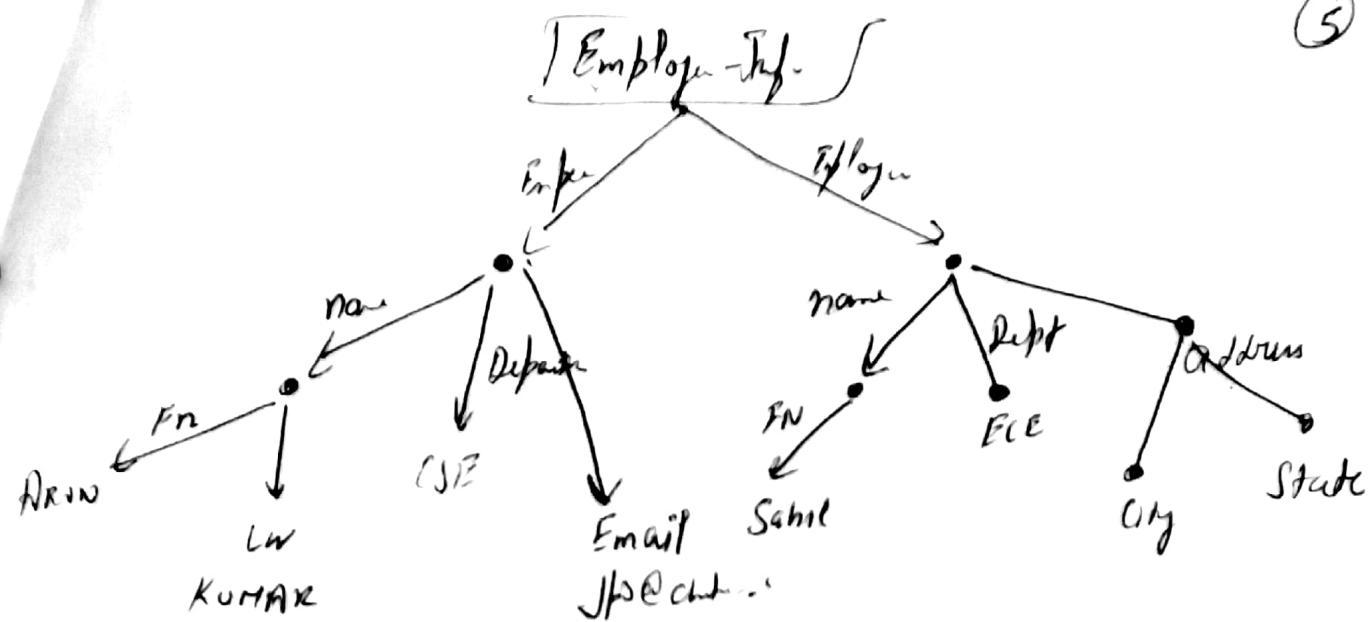
② * 0 or more Zero or more

③ ? 0 or 1 occurrence

④ <! Element Employee-Info (Employee) >

Only one occurrence

5



```

<!DOCTYPE Employee-Info [
  <!ELEMENT Employee-Info (Employee)+>
  <!ELEMENT EMPLOYEE (Name, Department, Email*, Address*)>
  <!ELEMENT Name (FN, LN*)>
  <!ELEMENT FN (#PCDATA)>
  <!ELEMENT LN (#PCDATA)>
  <!ELEMENT Department (#PCDATA)>
  <!ELEMENT Email (#PCDATA)>
  <!ELEMENT Address (City, State)>
  <!ELEMENT City (#PCDATA)>
  <!ELEMENT State (#PCDATA)>
]>
  
```

XML → DTD ~~→ Schema~~

any

<employee id="2" name="Rahul">

[
id unique
name optional]

NB:

<age> 22 </age>

<gender> Male </gender>

<salary> 20,000 </salary>

</employee>

DTD

<?xml version="1.0" encoding="UTF-8"?>

<!DOCTYPE employee [

<!ELEMENT employee (#PCDATA, age, gender, salary) >

<!ELEMENT age (#PCDATA) >

<!ELEMENT gender (#PCDATA) >

<!ELEMENT salary (#PCDATA) >

<!ATTLIST employee id ID #REQUIRED >

<!ATTLIST employee name CDATA #IMPLIED >

]>

~~<! Element Workers (Worker*)>~~
~~<! Element Worker (ssn, lastName?, FirstName?, hours)>~~
~~<! ELEMENT ssn (#PCDATA)>~~
~~<! Element LastName (#PCDATA)>~~
~~<! Element FirstName (#PCDATA)>~~
~~<! Element hours (#PCDATA)>~~
~~]>~~

~~XML DTD Notations~~

- i) A * following the element name means that the element can be repeated zero or more times in the document. This can be called as optional multivalued (repeating) element.
- ii) A + following the element name means that the element can be ~~optional~~ repeated one or more times in the document. This can be called a required multivalued (repeating element)
- iii) A ? following the element name means that the element can be repeated zero or one times. This can be called

an optional single valued (non-repeating) element.

v) An element appearing without any of the preceding three symbols must appear exactly once in the document. This can be called required single-valued (non-repeating) element.

vi) The type of element is specified via parenthesis following the element.

a) If the parenthesis includes names of other elements, these would be the children of the element in the tree structure.

b) If the parenthesis include the Keyword #PCDATA or one of the other types available in XML DTD, the element is a leaf node. #PCDATA stands for parsed Character data, which is roughly similar to a string data type.

vii) Parenthesis can be nested when specifying elements.

viii) A bar symbol symbol ($e_1 | e_2$) specifies that either e_1 or e_2 can appear in the document.

Limitations of XML DTD

- i) The data types in DTD are not very general.
- ii) DTD has its own special syntax & so it requires specialised processors.
→ It would be advantageous to specify XML Schema documents using the syntax rules of XML itself so that the same processors for XML documents can process XML schema descriptions.
- iii) All DTD elements are always forced to follow the specified ordering in the document so unordered elements are not permitted.

Advantages

- i) XML DTD is used to describe XML language precisely.
- ii) To define structure of an XML document.

~~Topic~~] >

(~~Introduction~~)

SQL	PL/SQL
i) one statement is executed at a time.	i) It is executed as a block of one code.
ii) SQL tells the database what to do & not how to do it.	ii) Tells the database has to do things (Procedures)
iii) SQL is used to code Queries, DML & DDL statements.	iii) It is used to code program blocks, triggers, functions, procedures & packages.
iv) we can embed SQL in a PL/SQL program.	iv) We can't embed PL/SQL within a SQL statement.
v) SQL is a data oriented language for selecting & manipulating sets of data.	v) It is a procedural language to create create applications.

XML Schema Definition (XSD)

The XML Schema language is a standard for specifying the structure of XML documents.

XML Schema is based on the tree data model with elements & attributes as the main structuring concepts. It borrows additional concepts from database such as keys, references & identifiers.

DTD

XSD

i) Doesn't support data types. It has only `TYPE` data type for the elements.

i) Supports data types.
 a) primitive / fundamental data types: `string`, `decimal`, `float`, `boolean`.
 b) Custom data types

II) Complex Type: A datatype that contains child elements or attributes & also the mixed contents.

II) Simple Type: A datatype that contains only values.

iii) It doesn't support namespaces. It has its own set of keywords for defining schema.

iv) It uses its own set of namespaces & elements for defining the schema.

- e.g. !DOCTYPE for root tag
 !ELEMENT for an element
- v) There is no restriction of datatype validation.
- vi) It is more suitable for small XML data.
 e.g. bookname ~~at~~

v) doesn't define order for child elements.

vi) e.g. <!DOCTYPE Address[
 <!ELEMENT Address(Named)
 <!ELEMENT Name(#PCDATA)
]>

vii) It allows us to specify restriction on data.
 e.g. → <price>/<price> tag.
 we can write only digits here

viii) It is used in large XML data.
 e.g. → Web Services

vii) order ~~can~~^{is} defined.

e.g. <xsd:element name="Address">
 <xsd:ComplexType>
 <xsd:sequence>
 <xsd:element name="name" type="xsd:string">
 </xsd:sequence>
 <xsd:ComplexType>
 </xsd:element>

✓ Schema

(a)

Schema

- ① Scheme description and XML Name Spaces.
- ② Annotations, documentation, language used (optional).
- ③ Elements and Types. → (root element structure), ex = comp → ~~node~~
- ④ Specify first level elements of root element
- ⑤ ex: dept, emp, project
- ⑥ Specify element type and minimum and maximum occurrences.
 - ① element type =
 - ② Minimum = Null → 0
 - ③ Max = ~~multiple~~ values → unbounded. {
Similar to
? * + of DTD}

Ex: <xsd:element name="employee" type="employee" minOccurs="0" maxOccurs="Unbounded"/>

- ⑥ Specify Keys (constraints)
 - ① Unique Key (xsd:unique)
 - ② Primary Key (xsd:key)
 - ③ Foreign Key (xsd:KeyRef)
- ⑦ Specify structure of complex elements via ComplexType.

→

```
<xsd:complexType name="Department">
  <xsd:sequence>
    <xsd:element name="DeptName" type="xsd:string" ref="DeptNameKey"/>
  </xsd:sequence>
</xsd:complexType>
```

Complex attributes?

Name
In
IN.

(b)

① `<?xml version="1.0" encoding="UTF-8"?>`
`<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">`

~~entwurf doc~~ ~~titel~~
`</xsd:schema>`

② `<xsd:element name="Company">`
`<xsd:complexType>`
`<xsd:sequence>`

③ ④ ⑤ `<xsd:element name="Department" type="string" minOccurs="0" maxOccurs="unbounded"/>`
`</xsd:sequence>`
`</xsd:complexType>`

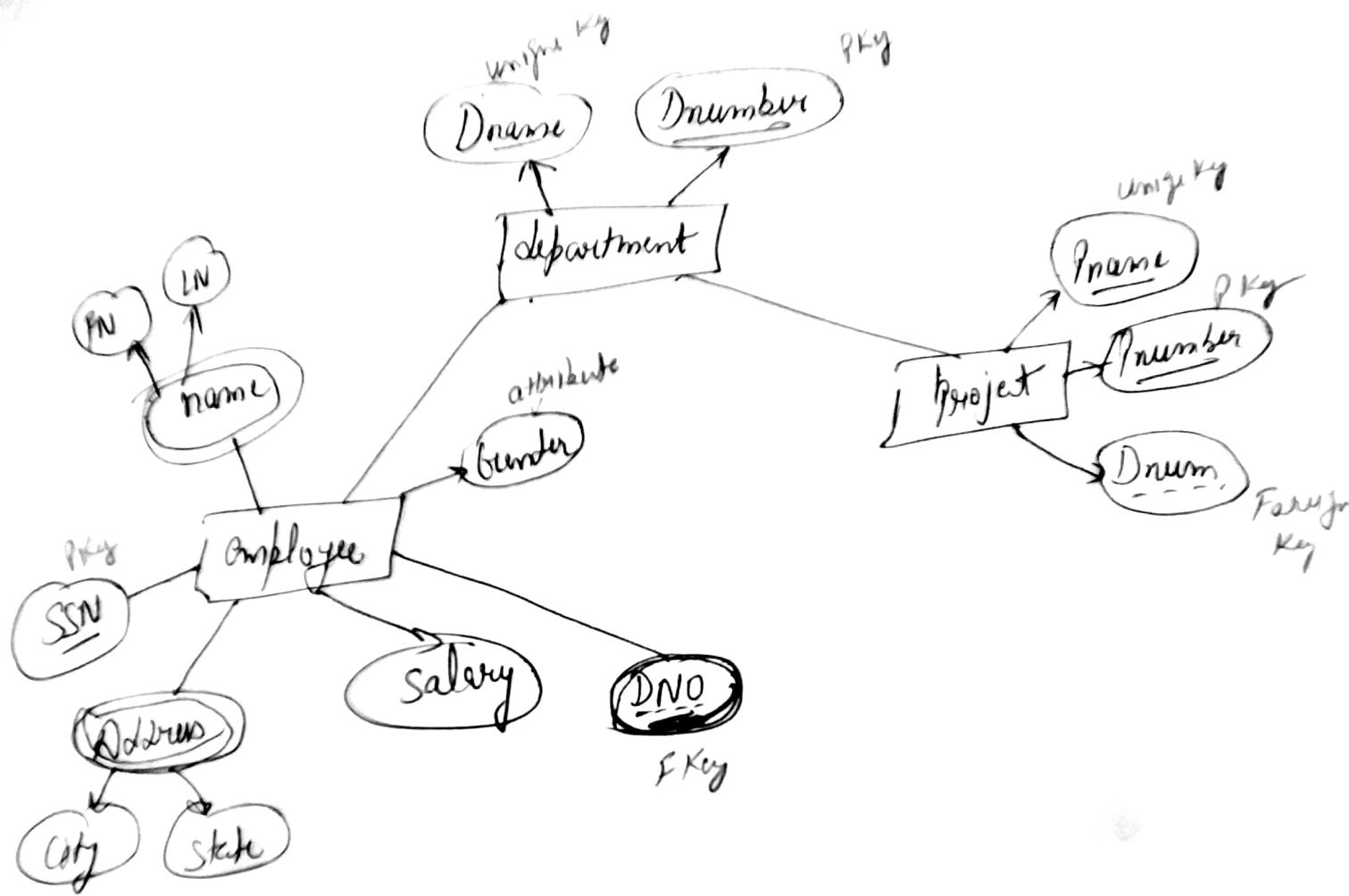
⑥
Keyles

~~Attributes~~
`</xsd:element>`
`<xsd:element>`
tag

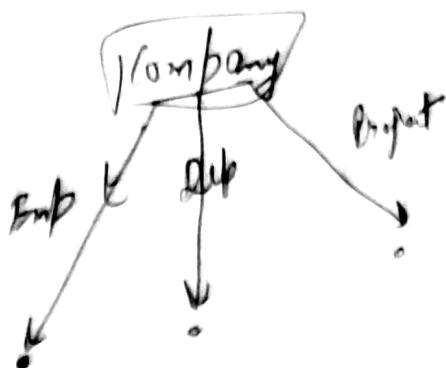
⑦
⑧

①

Company Database



Ex



XML Schema File Company

```
<?xml version = "1.0" encoding = "UTF-8"?>
<xmb: schema xmlns:xsd = "http://www.w3.org/2001/XMLSchema">
```

```
<xsd: element name = "company"> 1st elem
```

```
  <xsd: complexType>
```

```
    <xsd: sequence>
```

```
      <xsd: element name = "department" type = "Department"
```

```
        minOccurs = "0" maxOccurs = "unbounded" />
```

```
      <xsd: element name = "employee" type = "Employee"
```

```
        minOccurs = "0" maxOccurs = "unbounded" />
```

```
      <xsd: element name = "project" type = "Project"
```

```
        minOccurs = "0" maxOccurs = "unbounded" />
```

```
    </xsd: sequence>
```

```
  </xsd: complexType>
```

```
  <xsd: unique name = "departmentNameUnique">
```

```
    <xsd: selector xpath = "department" />
```

```
    <xsd: field xpath = "Dname" />
```

```
  </xsd: unique>
```

```
  <xsd: unique name = "projectNameUnique">
```

```
    <xsd: selector xpath = "project" />
```

```
    <xsd: field xpath = "Pname" />
```

```
  </xsd: unique>
```

```
  <xsd: key name = "departmentNumberKey">
```

```
    <xsd: selector xpath = "department" />
```

```
    <xsd: field xpath = "Dnumber" />
```

```
  </xsd: key>
```

```
  <xsd: key name = "employeeNumberKey">
```

```
    <xsd: selector xpath = "employee" />
```

```
    <xsd: field xpath = "SSN" />
```

```
  </xsd: key>
```

new
space

first level
main
elem

(3) *constraints*

```

<xsd:key name="projectNumberKey">
  <xsd:selector xpath="project"/>
  <xsd:field xpath="Pnumber"/>
</xsd:key>

<xsd:Keyref name="employeeDepartmentNumberKeyRef" refer="DepartmentNumberKey">
  <xsd:selector xpath="employee"/>
  <xsd:field xpath="Dno"/>
</xsd:Keyref>

<xsd:Keyref name="projectDepartmentNumberKeyRef" refer="DepartmentNumberKey">
  <xsd:selector xpath="project"/>
  <xsd:field xpath="Pnum"/>
</xsd:Keyref>

<xsd:Keyref name="WorkerEmployeeNumberKey" refer="EmployeeNumberKey">
  <xsd:selector xpath=""/>
  <xsd:field xpath=""/>
</xsd:Keyref>

<!-- 01 -->
<xsd:complexType name="Department">
  <xsd:sequence>
    <xsd:element name="Dname" type="xsd:string"/>
    <xsd:element name="Dnumber" type="xsd:string"/>
  </xsd:sequence>
</xsd:complexType>

<!-- 02 -->
<xsd:complexType name="Employee">
  <xsd:sequence>
    <xsd:element name="Name" type="xsd:string"/>
    <xsd:element name="SSN" type="xsd:string"/>
  </xsd:sequence>
</xsd:complexType>

```

<xsd:element name="Salary" type="xsd:string"/> />
 <xsd:element name="Address" type="xsd:string"/> />
 <xsd:element name="Dno" type="xsd:string"/> />
 <xsd:attribute name="Grenze" type="xsd:string"/> />

SLT
Link

</xsd:sequence>
 </xsd:complexType>

<xsd:complexType name="Project">

<xsd:sequence>
 <xsd:element name="Pname" type="xsd:string"/> />
 <xsd:element name="Pnumber" type="xsd:string"/> />
 <xsd:element name="Dnum" type="xsd:string"/> />

</xsd:sequence>

</xsd:complexType>

Further
Complex
Elements

<xsd:complexType name="Name">
 </xsd:sequence>

<xsd:sequence>
 <xsd:element name="FN" type="xsd:string"/> />
 <xsd:element name="LN" type="xsd:string"/> />

</xsd:sequence>

</xsd:complexType>

<xsd:complexType name="Address">

<xsd:sequence>
 <xsd:element name="City" type="xsd:string"/> />
 <xsd:element name="State" type="xsd:string"/> />

</xsd:sequence>

</xsd:complexType>

XML Languages

There are several XML query languages

Two standard query languages are:

- ① XPath → (to locate element)
- ② XQuery → (XPath + full features &c)

XSLT
XLink
X

1) XPath:

- XPath provides language constructs for specifying path expressions to identify certain nodes (elements) or attributes within an XML document.

Two main separators are used when specifying paths:

- ① Single slash (/)
- ② Double slash (//)

① Single slash (/): A single slash before a tag specifies that the tag must appear as a direct child of previous (parent) tag.

② Double slash (//): // specifies that the tag can appear as a descendent of previous tag at any level.

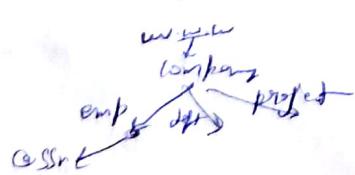
<u>Attribute</u>	/ Wild card
@	*
To specify attribute	any element
@ ISBN	

For example, if a company XML document is stored at the location

• `www.company.com/info.XML`

XPath expression

(1) `doc[www.company.com/info.XML]/company`



(2) `doc[www.company.com/info.XML]/company/department`

(3) `doc[www.company.com/info.XML]//employee`

(4) `doc[www.company.com/info.XML]//employee [Salary > 70000]`

(5) `doc[www.company.com/info.XML]/company/employee [Salary > 70000]`

(6) `doc[www.company.com/info.XML]/company/employee [Salary > 70000]/*`

(7) `doc[www.company.com/info.XML]/company/employee/[Salary > 10000]/*`

Restriction / Limitation of XPath.

pattern

multiple conditions *

• Path expression is that which specifies the path of the item to retrieve.

• Here in XPath path expression specifies the pattern also specifies the item to be retrieved.

• Hence, it is difficult to specify certain conditions on the pattern.

• If we want results differently.

The XQuery language separates these data constraints, and provides more powerful constraints for specifying queries.

- XQuery : Specifying Queries in XML

The typical form of query in XQuery known as FLWR expression.
Four main clauses of XQuery.

FOR LET WHERE RETURN.
 @P @P

1 FOR \$x IN doc(www.company.com/info.xml)/company/employee
[~~employee~~, ~~Salary gt 70000~~] / ~~name~~ Name
RETURN <res> \$x/~~fn~~/FN, \$x/LN </res>

2 FOR \$x IN doc(www.company.com/info.xml)/company/employee
WHERE \$x/Salary gt 70000
RETURN <res> \$x/Name/FN, \$x/Name/LN </res>

3 LET \$d := doc(www.company.com/info.xml)
FOR \$x IN \$d/company/employee
WHERE \$x/Salary gt 70000
RETURN <res> \$x/Name/FN, \$x/Name/LN </res>

• **LET** $\$d := \text{doc}(\text{www.company.com/info.xml})$
FOR $\$x$ IN $\$d/\text{Company}/\text{Project}$ [$\$Pnumber = 5$] / ProjectValue , $\$y$ IN $\$d/\text{Company}/\text{Employee}$
WHERE $\$x/\text{hours} \gt 20$ AND $\$y.\text{ssn} = \$x.\text{ssn}$
RETURN $\text{cres} > \$y/\text{Name}/\text{FN}$, ~~$\$x/\text{hours} < / \text{res} >$~~

68

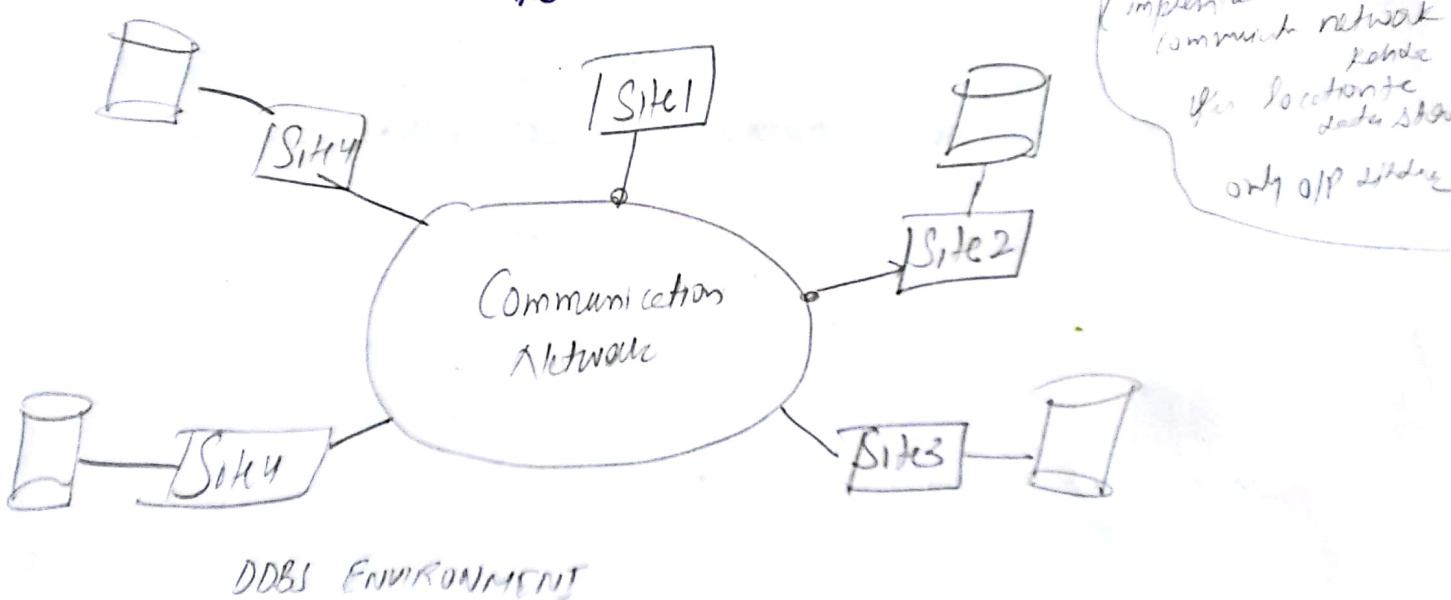


① DISTRIBUTED DATABASE

Distributed Database is a collection of multiple logically interrelated databases, distributed over a computer Network
Relationships (Foreign keys)

2 DISTRIBUTED DATABASE MANAGEMENT SYSTEM

DDBMS is a software System that manages a distributed database while making the distribution transparent to the user.



DDBS ENVIRONMENT

3 Conditions for Distributed Database

For a database to be called distributed the following minimum conditions should be satisfied.

- 1) Connection of Database nodes over a Computer Network: There are multiple computers known as sites or nodes. These sites must be connected by a communication network to transmit

Promises / Features of DBMS

- ① Transparency ② Autonomy ③ Reliability & Availability.

① Transparency

It refers to the idea of hiding implementation details from end users. The advantage of fully transparent DBMS is the high level of support that it provides for the development of complex applications.

Types of Transparency

- ① Data Organisation Transparency.
- ② Replication Transparency
- ③ Fragmentation Transparency.
- ④

① Data organisation Transparency : (also known as ~~but~~ distribution or n/w transparency)
This refers to freedom for the user from the operational details of the network & placement of the data in the distributed system. It is further divided into location & naming transparency.

1.1) Location Transparency

It refers to the fact that the ~~common user do perform~~ query for ~~particular~~ ~~data~~ ~~in particular~~ ~~node~~ a task is independent of the location of data & the location of the nodes in the system on which an operation is carried out.

2) Naming Transparency

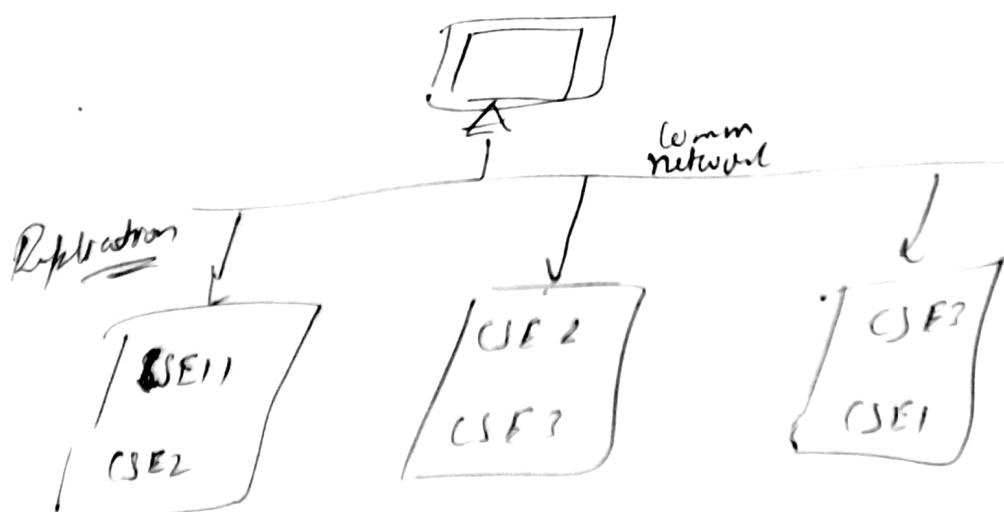
(3)

It means that a unique name is provided for each object in the database. In the absence of naming transparency, users are required to embed the location name as part of the object.

is embedded as a part of object name
Based on (single) location

3) Replication Transparency

The copies of the same data objects may be stored at multiple sites for better availability, performance and reliability. Replication transparency makes the user unaware of the ~~existing~~ existence of these copies.



3) Fragmentation Transparency

One fragment per row

Two types of fragmentation are possible. Horizontal & vertical

• Horizontal Fragmentation distributes a relation (table) into sub relations that are subsets of the tuples (rows) in the original relations.

Vertical Fragmentation

distributes a relation into subrelations where each subrel.
is defined by a subset of the columns of the original relation.

Fragmentation transparency makes the user ~~unknown~~
unaware of the existence of fragments.

Design Transparency & Execution Transparency

User have no idea how the distributed database is designed
and how a transaction executes.

Autonomy

It determines the extent to which individual node or
database in a ~~connected~~ DDB can operate independently.

A high degree of autonomy is desirable for increased
flexibility customised maintenance of an individual node.

Autonomy can be applied to design, communication &
execution.

1) Design Autonomy: Refers to independence of data

model usage & transaction management techniques.

DMS
Network
Tree
Hierarchical
Object oriented

autonomy nodes

2) Communication Autonomy: It determines the extent to
which each node can decide on sharing of information
with other nodes.

3) Execution Autonomy: It refers to the independence
user to act as they ~~please~~ please.

(3)

3) Availability & Reliability

(Availability)

Reliability is defined as the probability that a system is ~~running~~ (not down) at a certain time point.
~~means of cheezheet ka kar shikhai~~
(performing)

Availability is a probability that the system is continuously available during a time interval.
~~Gharas sat kol hei jf koi cheez hoi~~

* Reliability and Availability are related to Database failures, faults & errors.
~~because n k jo system available hai jo release v ho]~~

Advantages of Distributed database

① Improved ease & flexibility of Applications development

Developing and maintaining applications at geographically distributed sites of an organisation is facilitated owing to transparency of data distribution & control.

(Data different sites & distributed nodes has no like transparency maintain need not).

② Increased Reliability & availability

This is achieved by the isolation of faults to their site of origin without affecting the other databases connected to the network.

When the data & DBMS software are distributed over several sites, one site may fail while other sites continue to operate.

Only the data and software that exist at the failed site can not be accessed. This improves data reliability & availability.

~~Replication~~ User can recover all the data of failed site from other sites because In distributed database data is replicated on multiple nodes.

- In Central
- In Parallel

③ Improved Performance

A distributed DBMS fragments the database by keeping the data closer to where it is needed most. (Data localisation).

reduces the contention for CPU or I/O services.

buffer time
traffic

node import time

make process faster
from sp's

Data localisation
closely related
near hence a
response time increase.

3) Eastern Expansion (Scalability)

In distributed environment expansion of the system in term of adding more data increasing database sizes, or add more processors is much easier. It is the main feature of distributed database system which is known as scalability.

Functions of Distributed Database

1) Keeping track of data distribution

imp

The ability to keep track of the data distribution, fragmentation & replication by expanding the DDBMS catalog.

meta data file
inform.

2) Distributed Query Processing:

The ability to access remote sites and transmit queries & data among the various sites via a communication network.

plan

3) Distributed Transaction Management: The ability to devise execution strategy for queries & transactions that access data from more than one site and to synchronize the access to distributed data and maintain the integrity of the overall database.

> just transaction
> evaluate at 2 diff. site for availability
> propagate

(8)

4 Replicated Data management:

The ability to decide which copy of a replicated data item to access and to maintain the consistency of copies of a replicated data item.
[meas per site to same copy
now (chabit hai)]

5 Distributed Database Recovery:

The ability to recover from individual site crashes and from new type of failures.

6 Security:

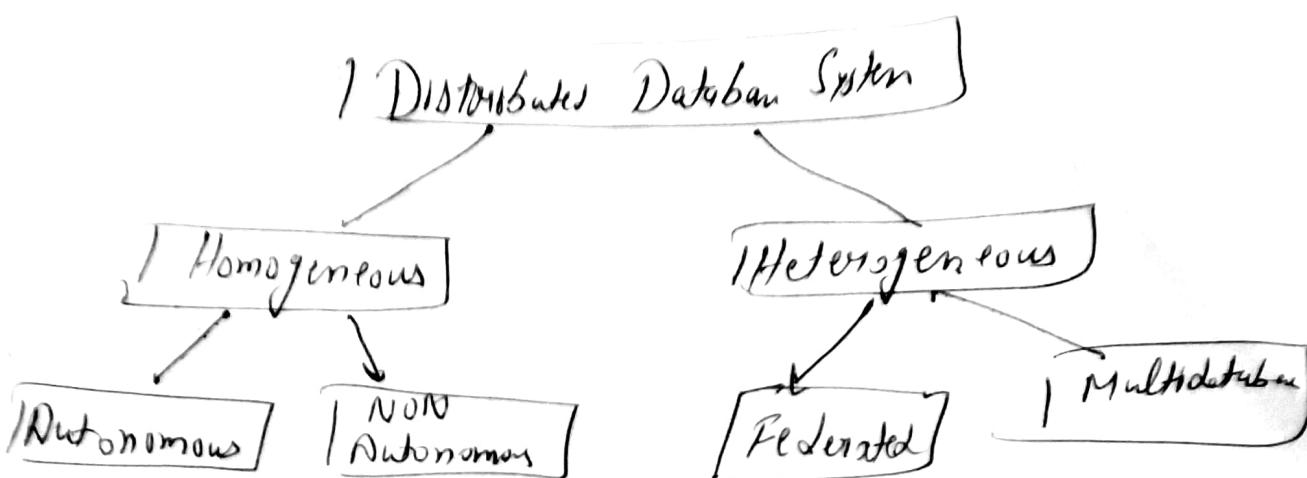
Distributed transaction must be executed with the proper management of the security of the data and the authorization/ access privilege of users.

7 Distributed Directory (Catalog) management:

A directory contain information (metadata) about the data in the database. The directory may be global for the entire DDB or local for each site.

Types of Distributed Database Systems

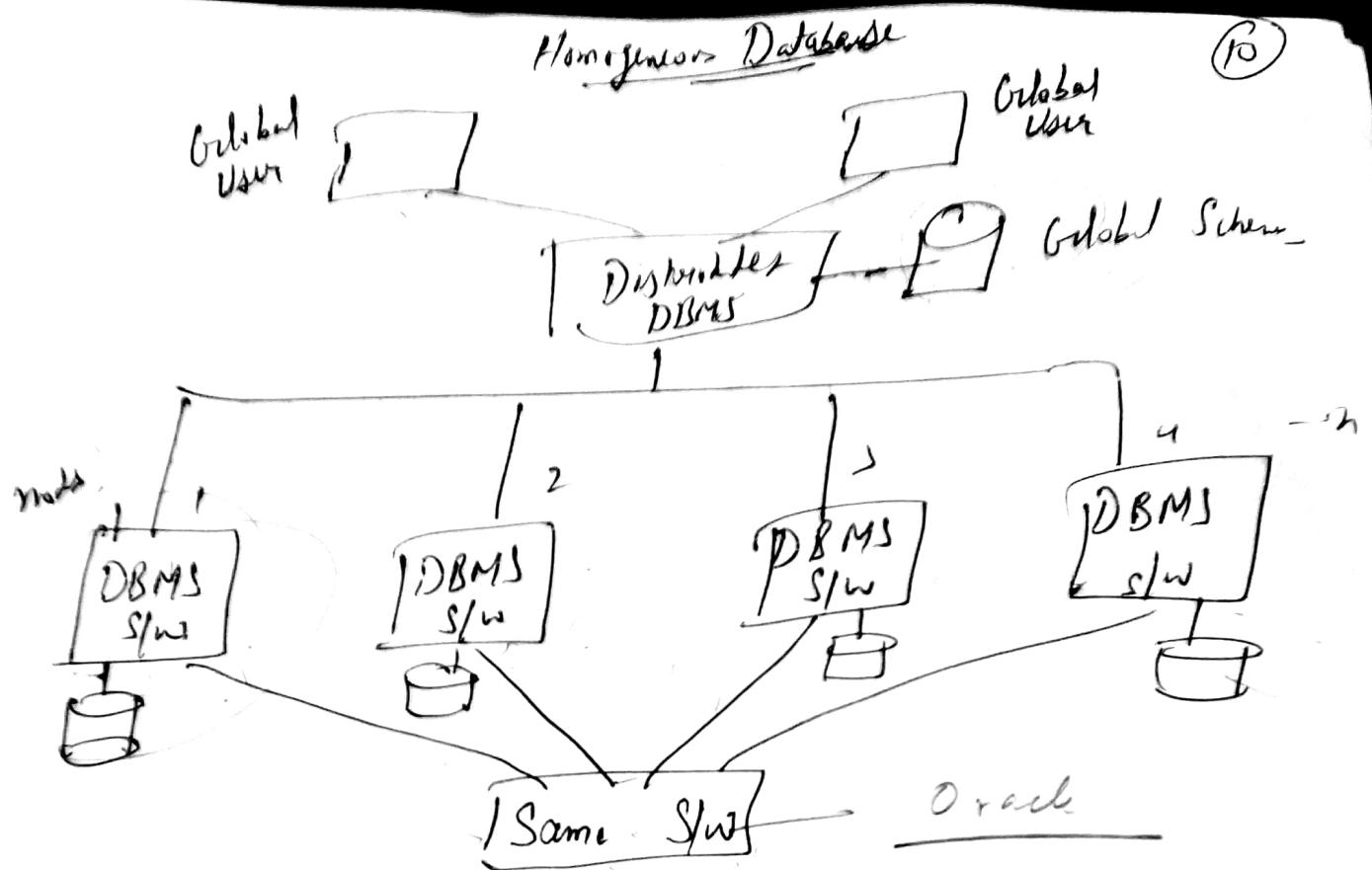
(9)



i) Homogeneous Distributed System:

- i) All the servers and client use identical software and operating system.
- ii) Each site is aware of all other sites & cooperates with other sites to process user requests.
- iii) The database is accessed through a single interface as if it is a single database.
i.e. it appears to user as a single system.
- iv) It is easy to design & manage.
- v) Difficult for most organisations to form a homogenous environment

{ Has many different configurations as
we kisi hai so as bandhan ni ja sake
K iski hi us kina jaiso ch difficult h. i apply kum kar ke bhi



Types of Homogeneous & distributed database

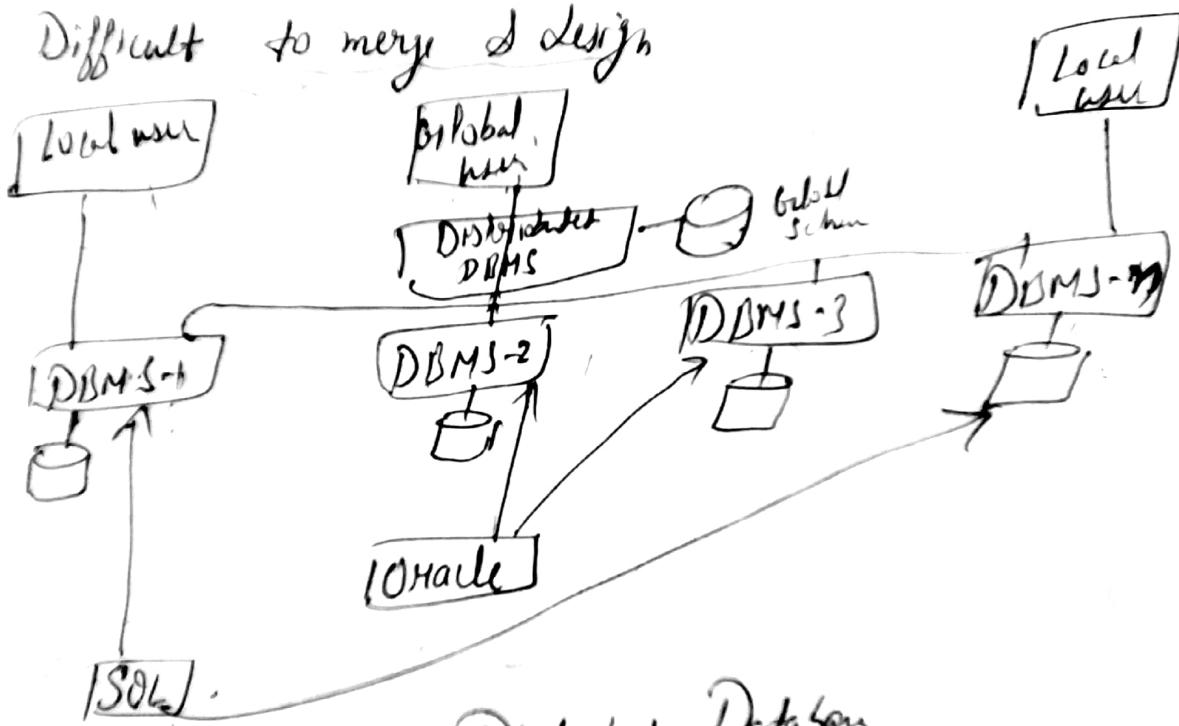
1 Autonomous : (sites are identical)

Each database is independent that functions on its own.
They are integrated by a controlling application & ~~use~~ message passing to share data updates.

2 Non-Autonomous : Data is distributed across the homogenous nodes and a master DBMS co-ordinates. data updates across the sites.

Heterogeneous Distributed Database

- 1) Different sites have different operating systems, data models and schemas & S/W.
- 2) Query processing and transaction processing is complicated due to dissimilar S/W.
- 3) A site may not be aware of other sites & so there is limited co-operation in processing user requests.
- 4) Difficult to merge & design.



Types of Heterogeneous Distributed Database

1) Federated:
The heterogeneous database systems are independent in nature & integrated together so that they function as a single database system. Global schema or view is shared by the applications.

2) Multidatabase System:
There is no global schema, but it ^{may be present} ~~consists of~~ one as needed by the application.

- Homogeneous
- 1) Share common global schema
 - 2) Run identical DBMS s/w.
 - 3) Other sites provide right to change schema s/w.
 - 4) ~~Transactions~~
Transaction process is easy
- Heterogeneous
- 1) Diff sites can have different schema
 - 2) Run different DBMS s/w
 - 3) Every site decide its own schema
 - 4) Transaction has different due to different sites

difference very imp types bus likh
 lo baki types ko explain karne ki
 jaroori nhi sidha age diagram
 bana do

Parallel DBMS

- 1) Machines are physically close to each other,
Ex same server room.
- 2) Machines are connected with high-speed LAN's & switches.
- 3) Communication cost is small
- 4) Follow Shared memory, Shared disk or Shared nothing architecture
- 5) It is a software system where multiple processors are used to execute & run queries in parallel.
- 6) Execution speed is fast
- 7) Nodes should be homogenous.
- 7) Backup at ^{one site} on site only.
- 8) Diagrams Short Disk, Shared memory Shared nothing.
- 9) Replication feature not exist - *(when a notification)* of Replication feature exists.

Distributed DBMS

(12)

- 1) Machines can be far from each other ex in different continent.
- 2) Machines can be connected using public network ex Internet.
- 3) Communication cost is high.
- 4) Follows Shared nothing architecture.
- 5) It is a software system that manages multiple logically inter-related databases distributed over a communication network.
- 6) Nodes ~~can~~ may be homogeneous or heterogeneous.
- 7) Backup at multiple sites.
- 8) Diagram : Distributed Databases
- 9) Replication feature exists. With the help of Replication feature we can retrieve data from any node.

difference very imp

Architectural Models & DBMS

(B)

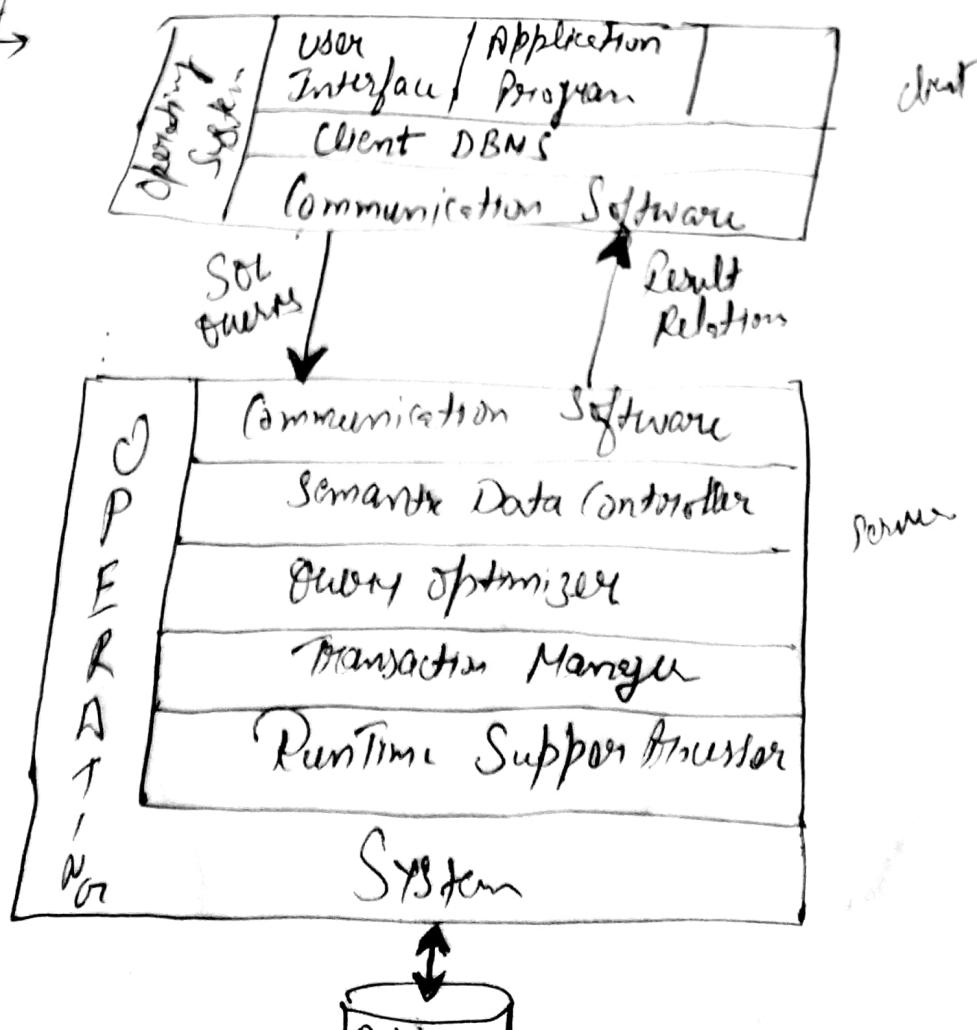
- 1) Client-Server Architecture.
- 2) Peer-to-Peer Architecture.
- 3) Multi DBMS Architecture.

1 Client-Server Architecture

It is a 2 part architecture where functionality is divided into server & clients.

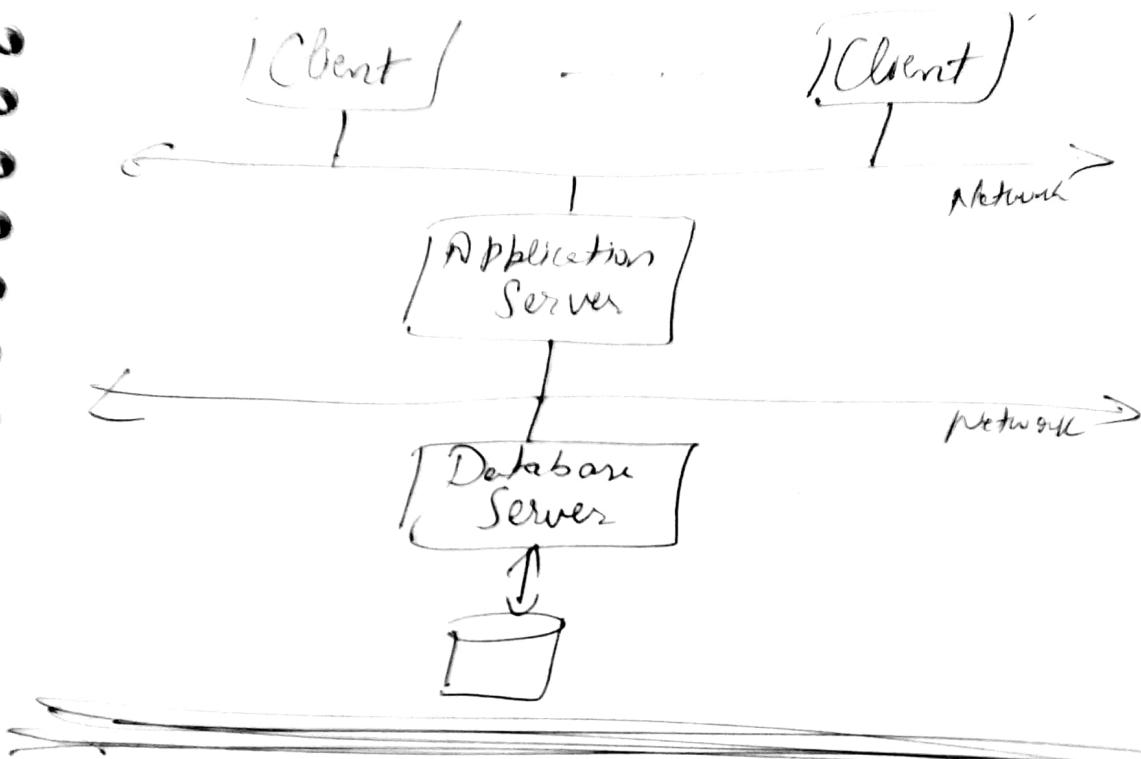
- Server does most of the data management work. This means that all the query processing & optimization, transaction management & storage management is done at server ~~end~~.
- Client functions include applications user interface

Client/Server
Architecture →

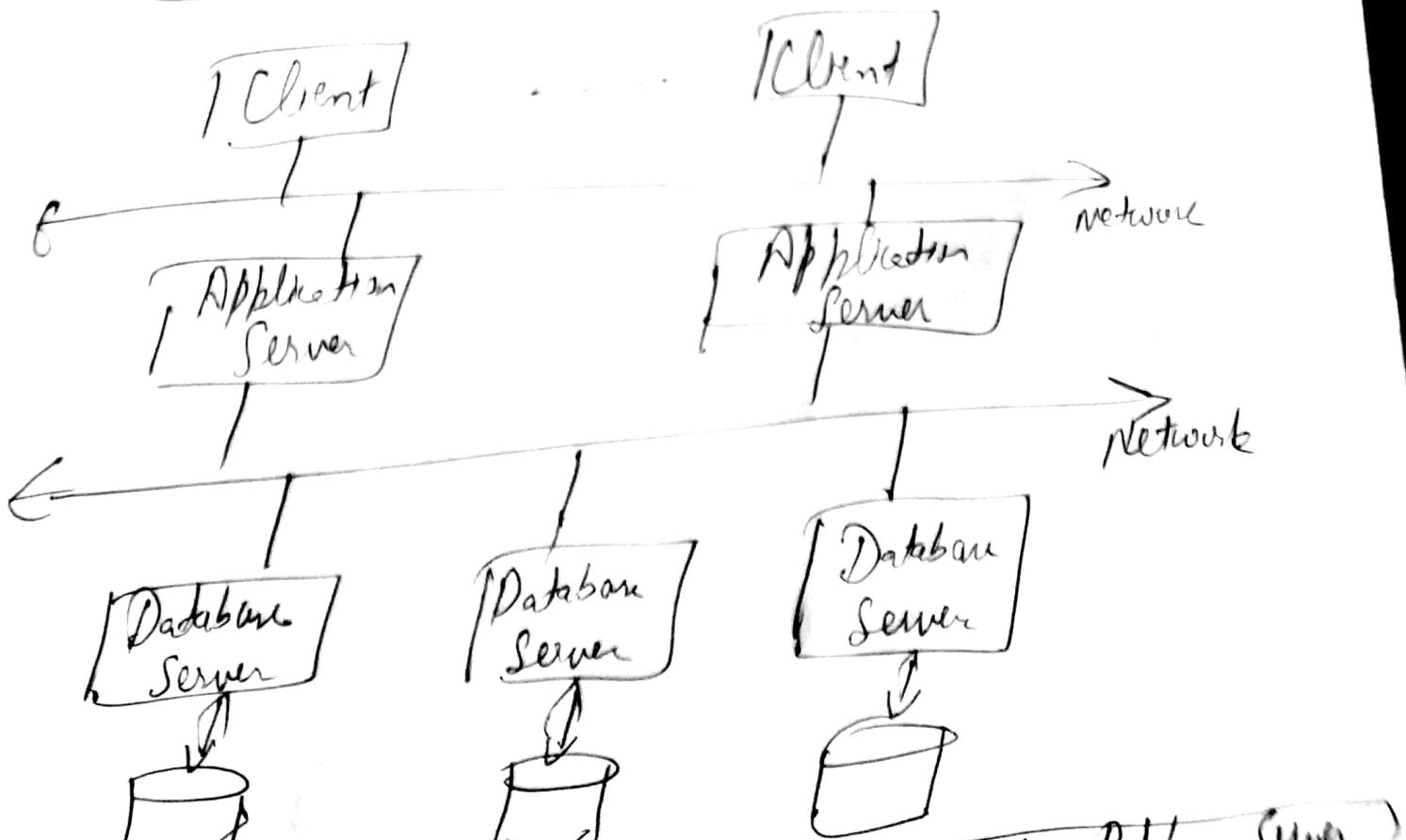


① Single Server / Multiple Clients

(14)



② Multiple Servers / Multiple Clients



① Peer-to-Peer System

(13)

Every system acts as a full DBMS & can communicate with other systems both as client & server to execute queries & transactions.

The peer-to-peer architecture has 4 levels of schema:

1) Global Conceptual Schema (GCS):

- Global logical view of data
- Union of logical conceptual schemes (LCS)

2) LCS Logical Conceptual Scheme:

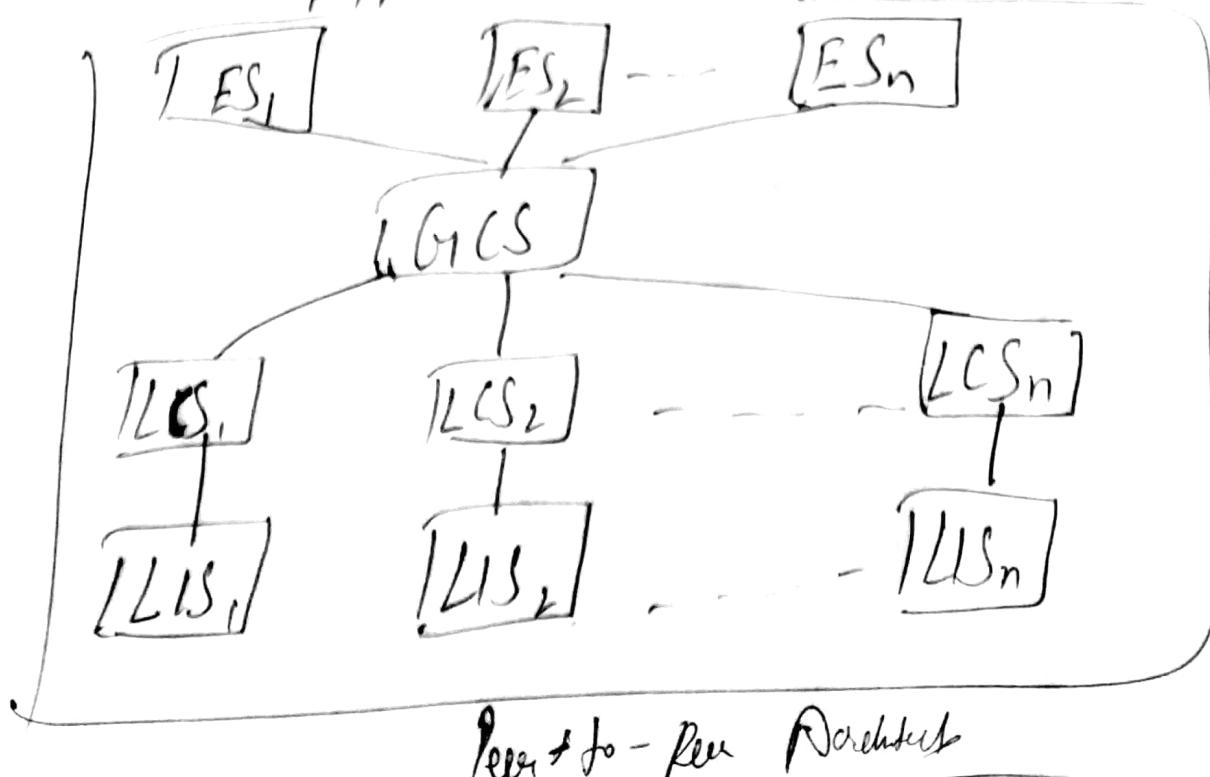
- Describes logical data organisation at each site - enterprise view of entire DBs.

3) Local ~~internal~~ internal Schemas LIS

- Local physical data organisation at each site -

4) External Schema: (ES)

- Describes the user/application view of on the data.



3) Multi-DBMS Architecture

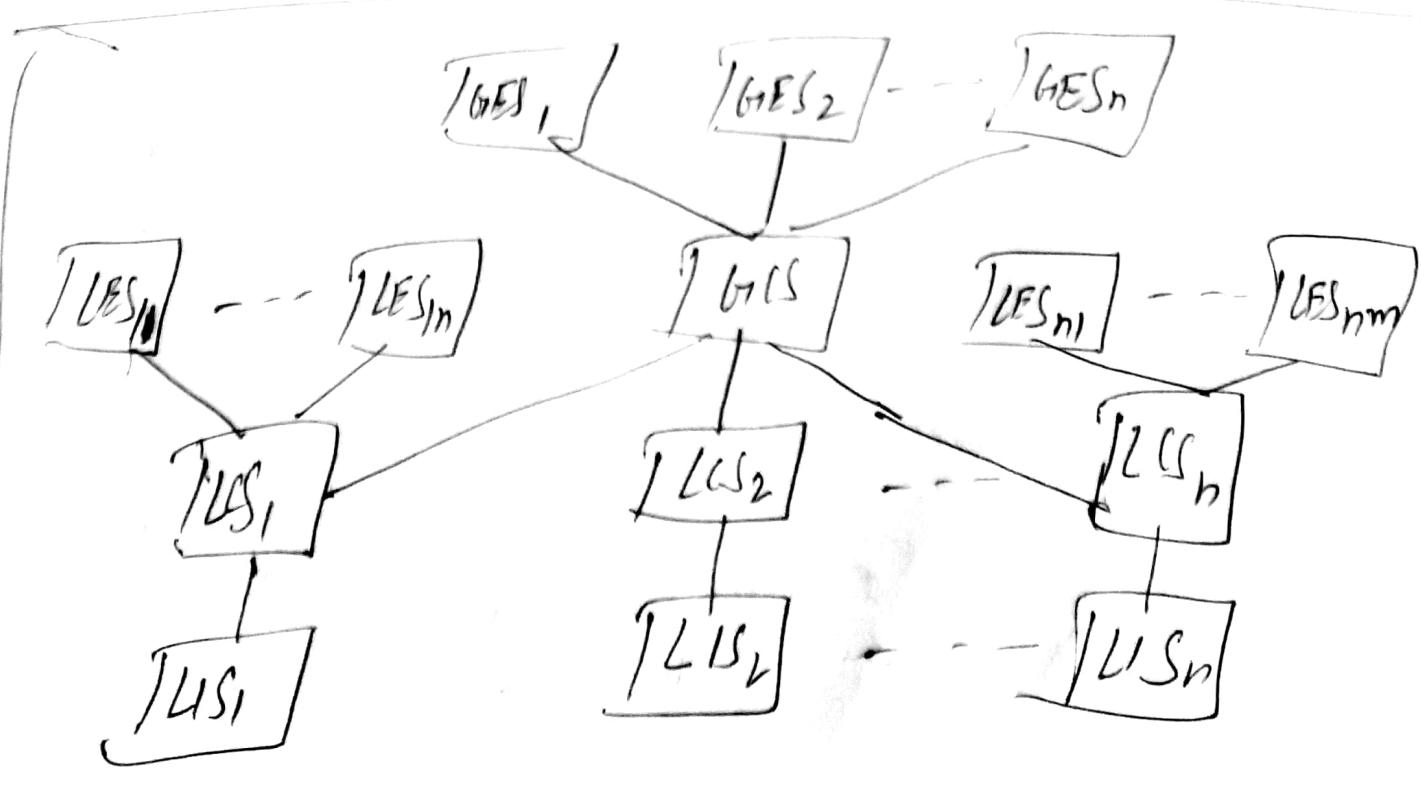
(10)

Fundamental difference to peer-to-peer DBMS is the definition of the Global Consistent Schema (GCS).

- In multi-DBMS, the GCS represents only the collection of some of local databases that each local DBMS want to share.
- Data Distribution across different sites & multi DB to local data mapping.
- Two different architecture models:
 - I Models with a GCS.
 - II Models without GCS.

I) Model with a GCS

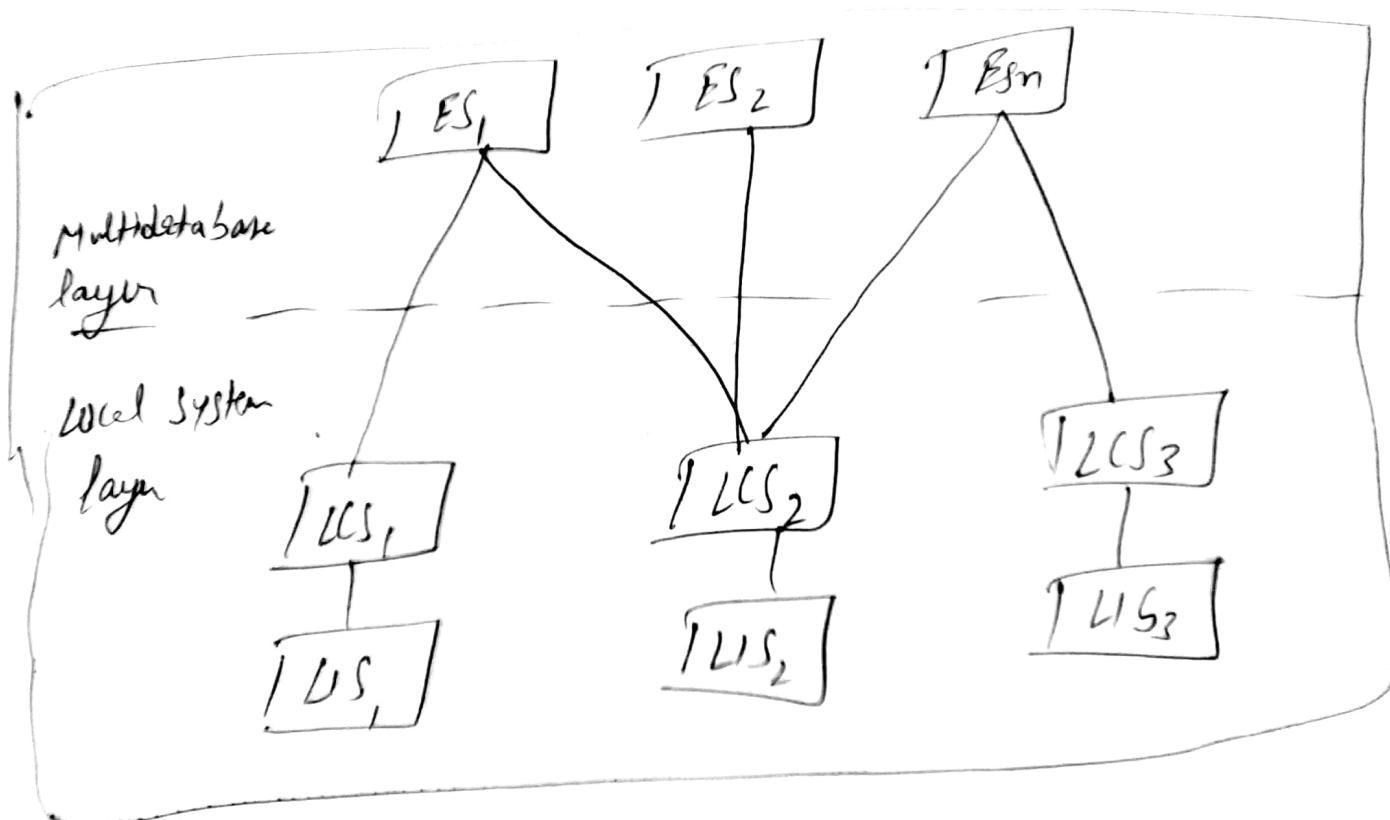
- GCS is a union of parts of the local DBMS - ~~diff~~
- Define their own views in the local DB.



(11)

2. Model without GLCS:

- The local DBMS present to the multi-database layer the part of their local DB they are willing to share.
- External views are defined on top of LCSs.



Data Fragmentation

Types

- 1) Horizontal
- 2) Vertical
- 3) Mixed (Vertical & Horizontal) (Hybrid)

Fragment of a relation is an appropriate unit of distribution.

This means -

- 1) Application views are subsets of relations.
- 2) Locality of access of application is defined on subset of relations.
- 3) Number of transactions are executed concurrently by a transaction manager.

Transaction manager maintains before and after database images, log of transactions & concurrency control.

4. Parallel execution of single query.

~~Segmented data records are stored at different locations.~~

Fragmentation aims to improve

- 1) Reliability
- 2) Performance
- 3) Balance storage capacity & costs.
- 4) Communication costs.
- 5) Security.

Types of Fragmentation

(19)

1) Horizontal Fragmentation:

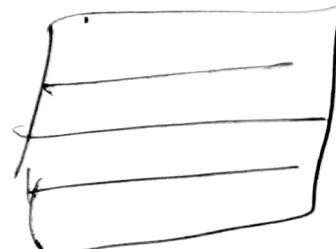
Partitions a relation along its tuples (rows)

2) Vertical:

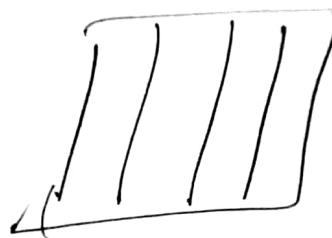
Partitions a relation along its attributes (columns)

3) Mixed (Hybrid)

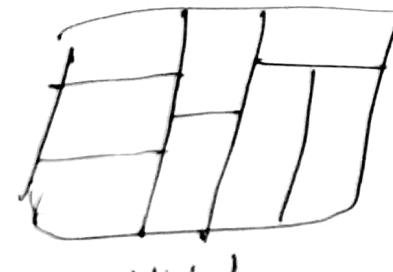
Combination of horizontal & vertical



Horizontal



Vertical



Mixed

Question ~~Ques~~ PROJ

Pno	Pname	Budget	Loc
P ₁	Inst	150000	Montreal
P ₂	OD	135000	New York
P ₃	CAD/CAM	250000	New York
P ₄	Maint	310000	Paris
P ₅	CAD/CAM	500000	Boston

1) Horizontal Fragmentation :

- 1) PROJ1 : projects with budget less than 200000
 2) PROJ2 : projects with budgets greater than or equal to 200000.

PROJ1 σ Budget < 200000 (PROJ)

PNo	PName	Budget	Loc
P ₁	Inst	150000	Montreal
P ₂	DD	135000	New York

PROJ2 σ Budget \geq 200000 (PROJ)

PNo	PName	Budget	Loc
P ₃	CAD/CAM	250000	New York
P ₄	Maint	310000	Paris
P ₅	CAD/CAM	500000	Boston

2) Vertical Fragmentation

- 1) PROJ1 : Information about project budgets.

- 2) PROJ2 : Information about project names & locations

PROJ1 π PNo, Budget (PROJ)

PNo	Budget
P ₁	150000
P ₂	135000
P ₃	250000
P ₄	310000
P ₅	500000

PROJ2 π PNo, PName, Loc (PROJ)

PNo	PName	Loc
P ₁	Inst	Montreal
P ₂	DD	New York
P ₃	CAD/CAM	New York
P ₄	Maint	Paris
P ₅	CAD/CAM	Boston

③ Hybrid Fragmentation

Project		
Pno	Pname	Loc
P1	CAD/CAN	New York
P2	Maint	Paris
P3	CAD/CAN	Boston

$\rightarrow \Pi_{Pno, Pname, Loc} (\sigma_{\text{budget} > 20000})$

$\Downarrow \Pi_{Pno, Pname, Loc} (\sigma_{\text{budget} > 20000})$

~~Math~~ 2 marks

ques aae ga kaa table given
hoga te os ch fragmentation
hogi te us table nu apa dasna

1) Completeness:

Decomposition of relation R into fragments R_1, R_2, \dots, R_n is complete if each data item in R can also be found in some R_i .

2) Reconstruction: If a relation R is decomposed into fragments $R_1, R_2, R_3, \dots, R_n$, then there should exist some relation operator from its fragments.

$$R = R_1 \sqcup \dots \sqcup R_n$$

- Union to combine horizontal fragments.
- Join to combine vertical fragments.

3) Disjointness: (Data item in one fragment should not appear in others)

If a relation R is decomposed into fragments R_1, R_2, \dots, R_n , & data item d_i appears in fragment R_j ,

then d_i should not appear in any other fragment R_K
 $K \neq J$;

- For horizontal fragmentation, data item is tuple.
- For vertical, data item is an attribute.

~~Star and~~

Replication

5 marks

It is useful in improving the availability of data by ~~copying~~ copying data at multiple sites.

- A relation or a fragment can be replicated at one or more sites.
- Fully redundant databases are those in which every site contains a copy of the entire database.

There are 3 types of Replication

- 1) Full Replication
- 2) No Replication
- 3) Partial Replication

(23)

1) Full Replication

Whole database is copied at every site. This ~~matters~~ improves availability because the system can continue to operate as long as atleast one site ^{is} up.

2) Partial Replication

Some fragments of the database may be replicated others may not.

- ⇒ Data may be replicated row by row, table by table, database by database.

mcq on transaction management

+
PLSQL
View
Sequence, Cursor, Index
Exception Handling

data fragmentation nd
its types very very imp
question pakka in
paper
with eg

STI

Slabs

Completed

ek sequence da
ques ya fir cluster
vagara vale topic
cho 5 marks da
question

fragmentation nd replication cho 5 ya 10
marks vala question very very imp