DBMS NOTES CS302

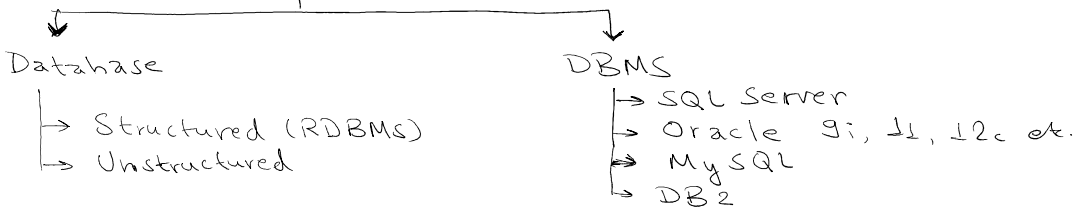
09 September 2021 16:08

– Subhojit Ghimire

✱ Syllabus:

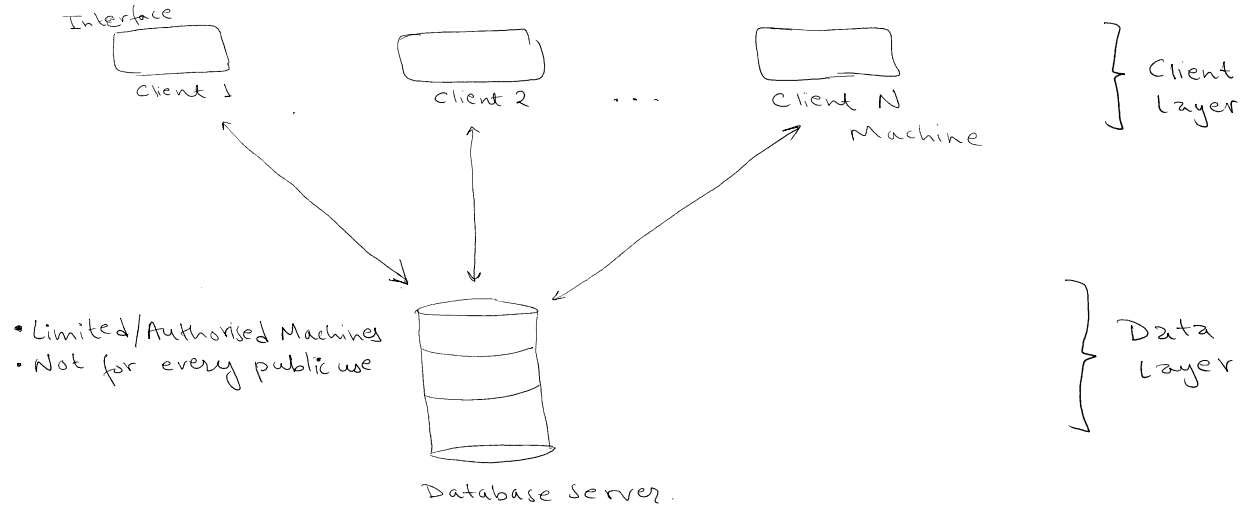
1. Data Independence: 2 tier, 3 tier, 3 level of abstraction (aka 3 schema)
2. Data Models: Network, Heirarchical, Relational, ER, Object Oriented
3. ER Models: Attributes, Relationships and Types.
4. Basics of Keys: Primary, Candidate, Super, Foreign.
5. Normalisation: Closure Method, Functional Dependicies, 1st - 2nd - 3rd Normal forms, BCNF
6. Transaction Control and Concurrency: ACID, R-W, W-R, W-W. Problems, Conflict Serialisability, Recoverability, 2-Phase Locking, Timestamp Ordering Protocol.
7. SQL and Relational Algebra: DDL, DML, DCL, Constraints, Aggregrate functions, Joins, Nested Query
8. Indexing → Primary, Cluster, Secondary, B tree, B^+ tree.

✱ Database System:



✱ File System vs DBMS:

✱ 2-tier and 3-tier Architecture:



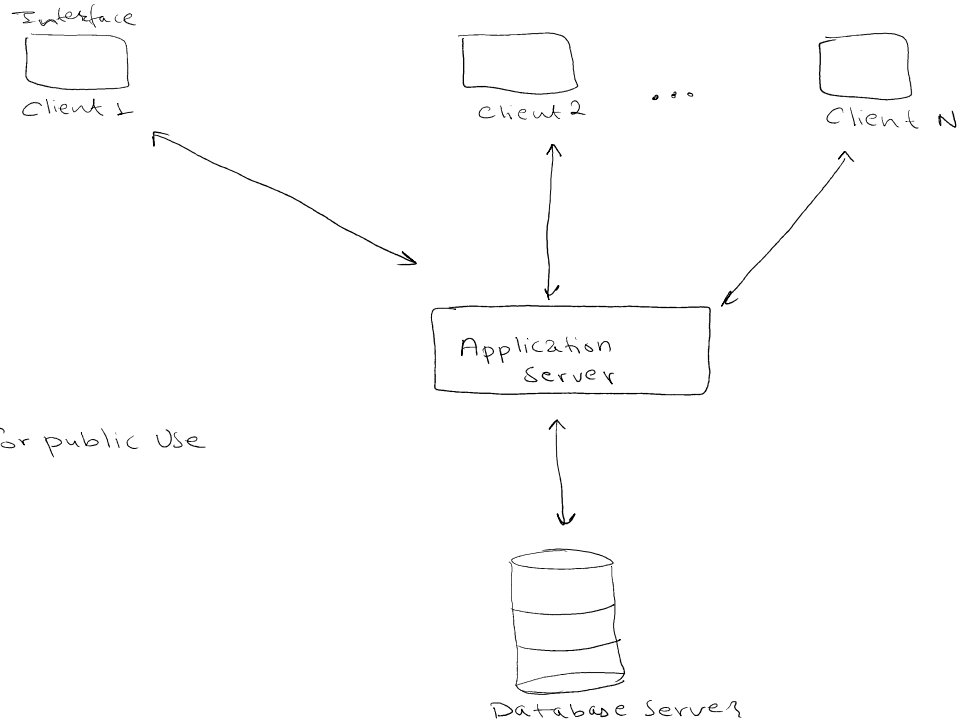
2-tier Architecture

Pros ⟷ Cons

2-tier Architecture

Pros
- Maintenance

Cons
- Scalability
- Security



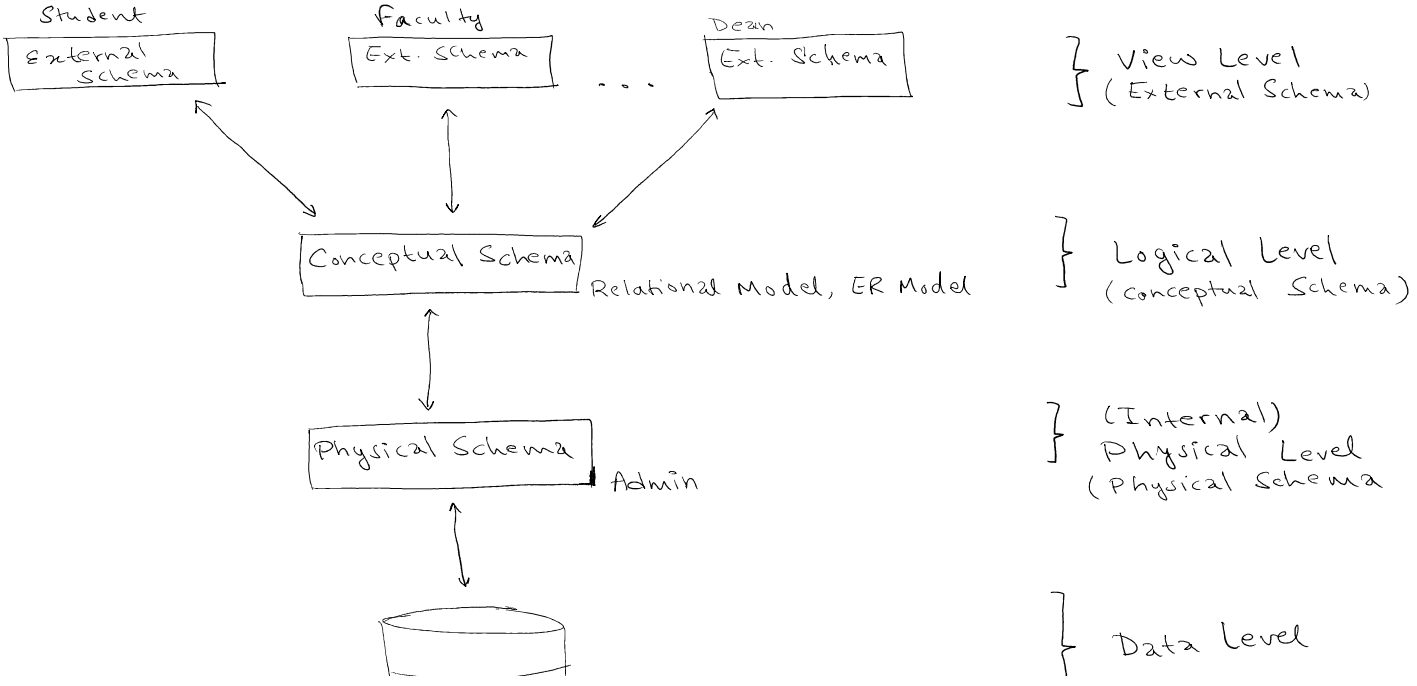
- For public Use

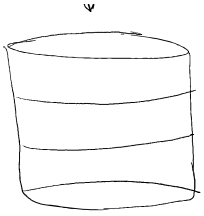
3-tier Architecture

Pros
- Security
- Scability

Cons
- Maintenance

* Schema: [Logical Representation]
  - 3-Schema Architecture: [Three level of Abstraction]



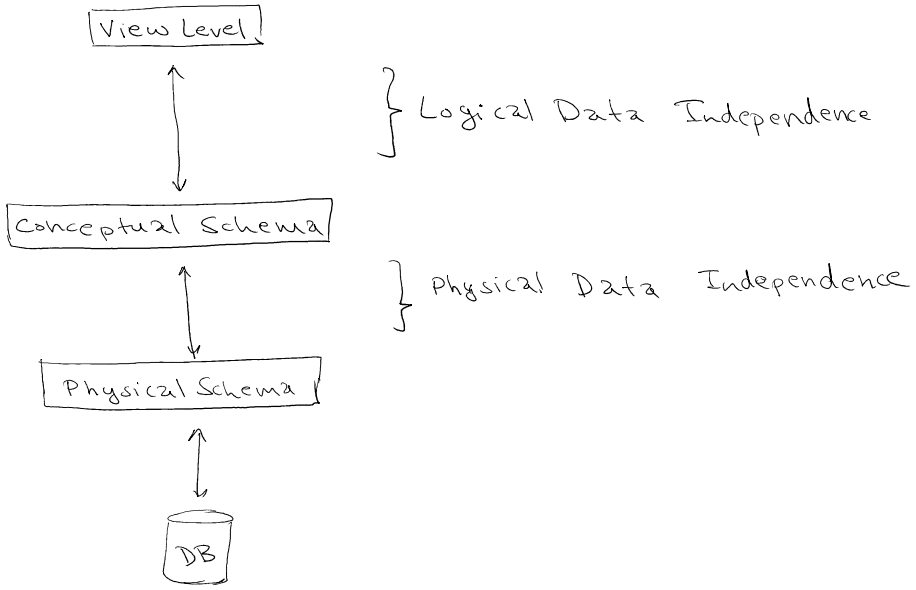


Database (Disk)

} Data Level

NOTE: In database (disk), the data/information is stored as file.
But, data/information is viewed/modified by user in the form of table.
DBMS is all about table, everything is done in tabular form.

* Data Independence.



* Key: Candidate and Primary

→ what is key?
Key is an attribute that uniquely identifies any two tuples in the table.

- Candidate key:
  Set of valid attributes that can be used to uniquely identify and distinguish two or more tuples in the table. For example: Roll No, License No, Registration number, Voter id etc.

- Primary key:
  of all the candidate keys, only one key is selected which is the most appropriate and is made into the primary key.

- Alternative Key:
  After the primary key is chosen, the rest of the candidate keys are called the alternative keys.

NOTE: There can only be ONE primary key per table.
There can be multiple candidate keys, but only one primary key.

PRIMARY KEY = Unique + Not NULL

* Foreign Key:
It is an attribute or set of attributes that references to primary key of same table or another table (relation)

→ It maintains referential integrity.

- Referenced (Base) Table
  - → Insert : No violation
  - → Delete : May Cause Violation → Soln: (On delete Cascade, On delete Set NULL, On delete No Action)
  - → Update: May Cause Violation → Soln:

- Referencing Table:
  - → Insert : May cause Violation
  - → Delete : No Violation
  - → Update : May cause Violation

EXAMPLE Q: Let $R_1(a,b,c)$ and $R_2(x,y,z)$ be two relations in which 'a' is foreign key in R_1 that refers to Primary key of R_2. Consider four options
(a) Insert into R_1 (c) Delete from R_1
(b) Insert into R_2 (d) Delete from R_2

Which is correct regarding referential integrity?
1) Option a and b cause violation
2) Option b and c will cause violation
3) Option c and d will cause violation
4) Option d and a will cause violation

* Super Key:
It is a combination of all possible attributes which can uniquely Identify two tuples in a table.

Super key of any Candidate Key is Super key

for example, A table has attributes Roll, Name, Age.
and, Roll is a candidate key.
Then, Roll alone can be a super key.
Combination of Roll and Name can be super key.
Combo. of Roll and Age can be super key
Combo. of Roll, Name and Age can be super key
However, combo. of Name and Age can NOT be super key.
because neither Name nor age is candidate key.

- Given, $R(A_1, A_2, A_3 \ldots A_n)$, then how many super keys are possible?

i) If, only A_1 is candidate key?
→ 2^{n-1} possible super keys.

ii) If, A_1 and A_2 are candidate keys?
→ $2^{n-1} + 2^{n-1} - 2^{n-2}$ possible super keys.
$= 2(2^{n-1}) - 2^{n-2}$

$= 2(2^{n-1}) - 2^{n-2}$

Similarly,
if three candidate keys, then $3(2^{n-1}) - 2^{n-3}$ super keys.
- And, if k candidate keys, then $k(2^{n-1}) - 2^{n-k}$ super keys.

iii) If, A_1A_2 together makes for one candidate key.
→ 2^{n-2} super keys possible.

iv) If, A_1A_2 and A_3A_4 make for two candidate keys,
→ $2^{n-2} + 2^{n-2} - 2^{n-4}$ possible super keys.
$2(2^{n-2}) - 2^{n-(2\times2)}$

Similarly, if K such candidate keys of combo 2.
- Then, $k(2^{n-2}) - 2^{n-2k}$

That means, if K such candidate keys of combo m.
Then, super keys = $k(2^{n-m}) - 2^{n-mk}$

Final formula,

* super keys = $\begin{cases} k(2^{n-m}) & ; \text{if } k = 1 \\ k(2^{n-m}) - 2^{n-mk} & ; \text{if } k > 1 \end{cases}$

where, k is number of candidate keys.
m is combination that makes for each candidate key.

* E-R Model (Entity-Relationship Model)

Entity → 

Attributes → 

Relationship → 

Types of Attributes:

1. Single vs Multivalued Attributes

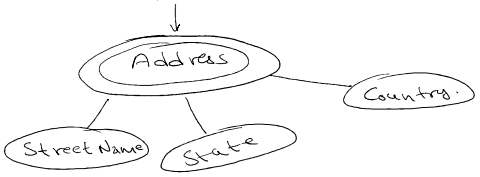
2. Simple vs Composite Attributes.

3. Stored vs Derived Attributes

4. Key vs Non-key Attributes

5. Required vs Optional Attributes

6. Complex Attribute = Composite + Multivalued.



* Cardinality (Degree of Relationship)
  1. One to one (1 - 1)
  2. one to many (1 - M) (1 - ∞)
  3. many to one (M - 1)
  4. many to many (M - N)

E.g. :- Given

Then Tables can be formed such.

Employee

Eid	Ename	Age
E1	abcd	20
E2	efgh	20
E3	abcd	21
E4	xyza	25

Primark key. Eid

Works

Eid	Did
E1	D1
E3	D2
E2	D3

Here, we took Eid as the P.K, However, we can take Did as PK as well.

Department

Did	Dname	Loc
D1	IT	Bang
D2	HR	Delhi
D3	Product	Delhi

Primary key Did

∴ It is 1 - 1 relationship, so both Eid and Did will be present only once, hence can be taken as P.k.

Note If: Given M - N relationship, we need to make new Primary key attribute for relationship table. In case of 1 - M or M - 1, we can take the unique attribute (ie, M) as Primary key.

Also, the ids taken in relationship tables are foreign keys.

The above table can be further simplified by merging as: (or reduced)

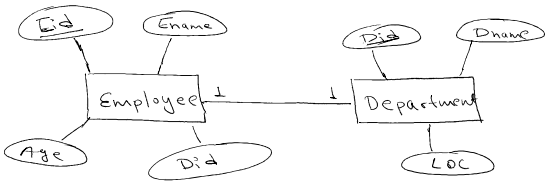
Employee

Eid	Ename	Age	Did
E1	abcd	20	D1
E2	efgh	20	D3
E3	abcd	21	D2
E4	xyza	25	

Department

Did	Dname	Loc
D1	IT	Bang
D2	HR	Delhi
D3	Product	Delhi

After Merging, ER obtained,



Had we taken Did as the primary key in relationship table instead of Eid, we could have merged Work table with Department table like

Eid	Ename	Age

Did	Dname	Loc	Eid

Such merging is only possible with 1 - 1, 1 - M and M - 1 Because we take one of the foreign keys as Primary key

But in case of M - N, we need to create a new attribute as Primary key which will represent the relationship table and hence cannot be merged.

Example: Given

Roll	Name	Age
1	A	20
2	B	20
3	C	21

Roll	Cid	RollCid
1	C1	1C1
2	C2	2C2
1	C2	1C2
2	C1	2C1
3	C1	3C1

Cid	Name	Credit
C1	Math	4
C2	Phy	4
C3	chem	2

Here, New primark key created is the composite key of Roll and Cid. as, RollCid.

NO. M-N cannot be reduced or merged.