Normalization

(*) What is Normalization?

→ is a database design technique that reduce data redundancy and eliminates undesinable characteristic like (isertion, update delete Anomalies)

1NF (First Normal form)

① a table cell should contain a single value

cant have multiple value in a column attribute

✗
101, 102

✓
101
102

② Each Records need to be unique

Column name unique

✗	
ID	ID

✓	
ID	Roll

ID	Name	course
1	M	NET
2	G	JAVA
3	S	C, JAVA

→

ID	Name	course
1	M	NET
2	G	JAVA
3	S	C
3	S	JAVA

2NF → second Normal form

① Table should be ~~be~~ in 1NF form

② Pantial dependency is not allowed

cannot identify uniquely

PK

ID	Name	address
1	G	Dhaka
2	G	Barishal
3	B	Dhaka

Non key

key

↓

Uniquely Identifiyable

→ Functional dependency

Dependency

all non key column dependent on key column (Primary)

... Here Name, address dependent on Id to uniquely Identify them

Dependendent on ID not on course ID

Dependent on course ID Not on ID

PK PK

ID	Name	Addno	course ID	course Name	credit	Grade
101	A	M	1	C	3	A+
101	A	M	2	JAVA	3	A
102	B	S	1	C	3	B
102	B	S	2	JAVA	3	B+
Key	Non key	Non	Key	Non	Non	Non

Here table need 2 PK for Identifying a record uniquely

① Candidate key

(multiple primary key in a table to Uniquely Identify)

Partial Dependency

if a table has candidate key (multiple primary key)

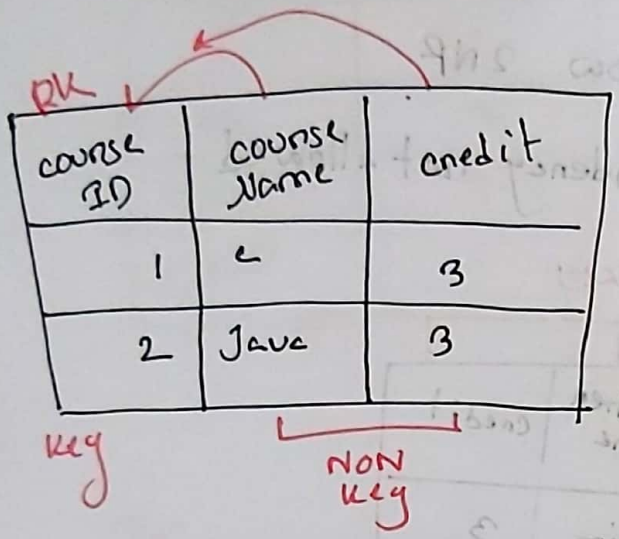
and any non key value is dependent only on one of the any primary key

∴ Here

course_name and credit is non key and dependent only on course ID

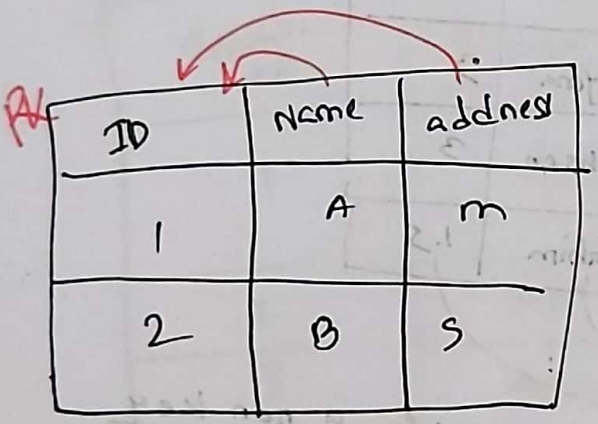
To apply, 2NF Remove pantial Dependency

(can have fully on functional depency)



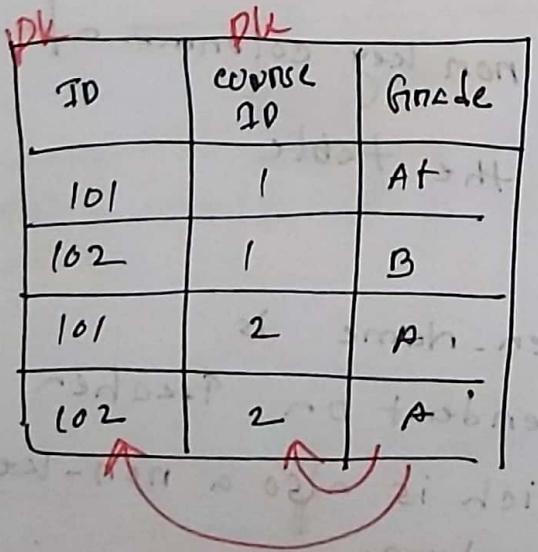
Functional depency

2NF → ✓



Functional dependency

2NF → ✓



Non key / Grade dependent on both primary key

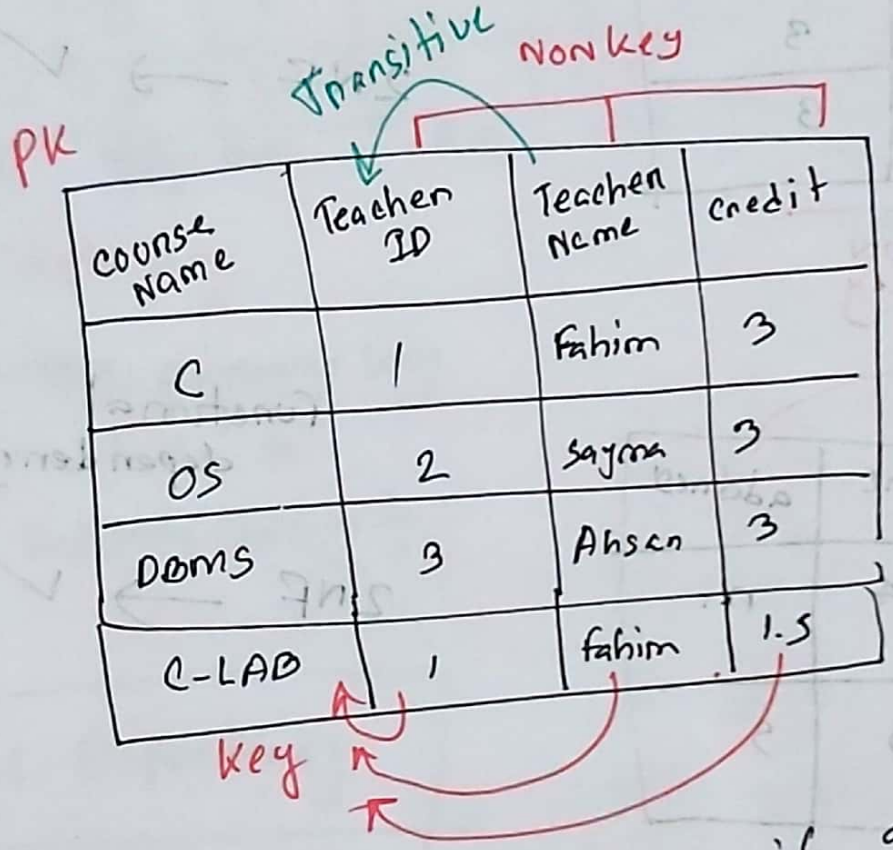
pantial dependency → X

2NF → ✓

3NF

(1) Table must follow 2NF

(2) Transitive Dependency not allowed



Transitive dependency → if a non key column of table is dependent on another non key column of the table

∴ Here Teacher_Name is dependent on Teacher i.d which is also a non-key column

PK

course name	Teacher ID	credit
C	1	3
OS	2	3
DBMS	3	3
C-LAB	1	1.5

Transitive dependency → X

3NF → ✓

PK

Teacher id	Teacher name
1	Fahim
2	Sajma
3	Ahsan

Transitive dependency → X

3NF → ✓

Boyce-codd

3.5NF → BNCF → Key column, nonkey কলাম dependent হবে না

4NF → No multi valued dependency

Q Fix 3NF

transitive dependency

member (membershipID, fullname, physical address, saturation)

NON key

NON key

movienented (membership ID , movies rented)

(1)

membershipID	Fullname	physical address	Saturation ID

(2)

Saturation ID	Saturation

Q Fix 2NF

memben (Fullname, Physical address, Movie nented, saturation)

(1)

membership ID	Full name	Physical address	saturation

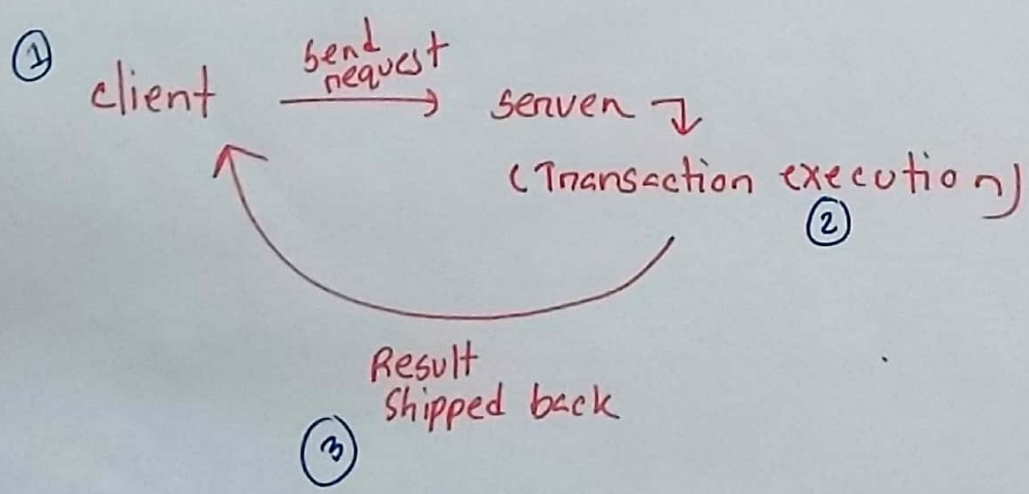
(2)

membership ID	~~title~~ Movie Rented

Database System Architecture

Centralized Data system

- Single-user system
- multi-User System (Server System)
  - Transaction Server (Query Server / SQL server system)
  - data Server



A typical transaction server consists of

multiple process accessing data in shared memory

transaction process

1. Database writer process
2. Log writer process
3. checkpoint process
4. Process Monitor Process
5. Lock Manager Process

shared data

1. Buffer pool
2. Lock table
3. Log buffer
4. Cached Query plan

Two main performance Monitor?

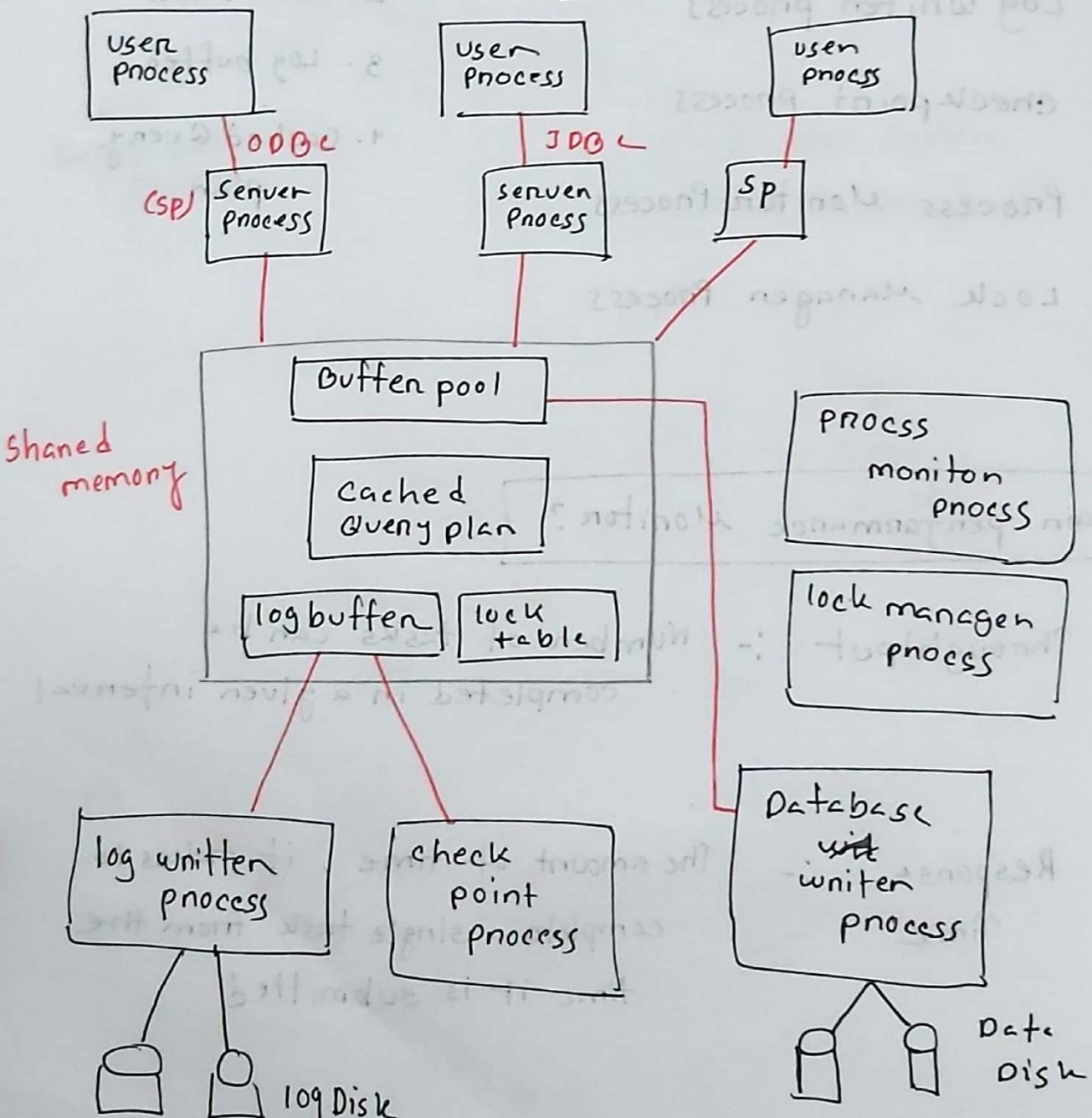
(1) Throughtput :- Number of tasks can be completed in a given interval

(2) Response Time :- The amount of time, it takes to complete a single task from the time it is submitted

What is transaction?

→ Transaction is a logical unit of work that consists of one or more database operations (reading, writing or modifying data) stored in database.

State Diagram



4 properties of transaction

1. Atomicity → Test and set
   → compare and swap
2. Consistency
3. Isolation
4. Durability

Data Server System

→ is a software application on hardware infrastructure designed to manage and store data efficiently

Data caching

→ is a technique to improve performance of data retrival operations by storing frequently accessed data in cache

Lock caching

→ is a technique to improve efficiency of Lock operations

(*) what is parallel system ? (PS)

→ ps refers to database architecture that utilizes multiple processor / computing Resources to process data operations co-currently

Motivation → improve performance by dividing the workload into smaller tasks than can be executed simultaneously on multiple processor

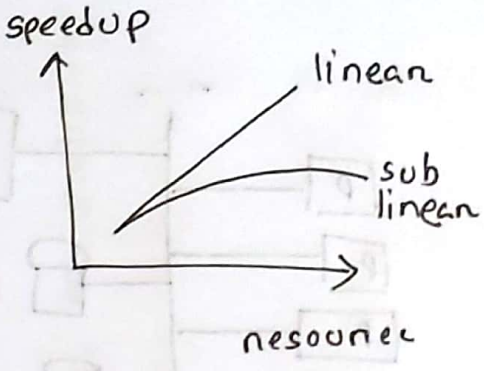
Benifits

1. High performance
2. Decision support

Types

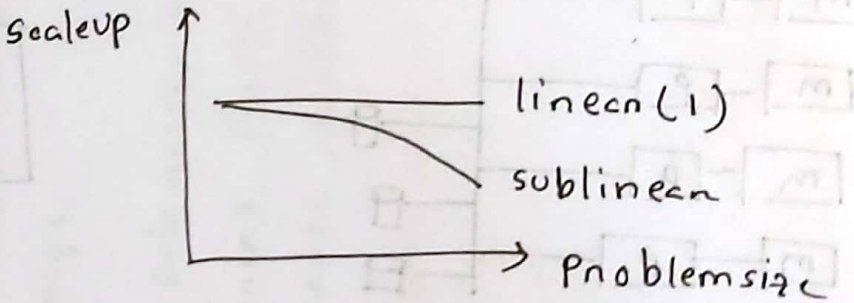
1. Coarse gain parallel
2. ~~fine~~ Fine gain parallel

$$\text{Speedup} = \frac{\text{small system elapsed time}}{\text{large system elapsed time}}$$



$$\text{Scale Up} = \frac{\text{small system small problem elapsed time}}{\text{large system large problem elapsed time}}$$

(T_S) / (T_L)



Factor Limiting Speedup and Scale Up

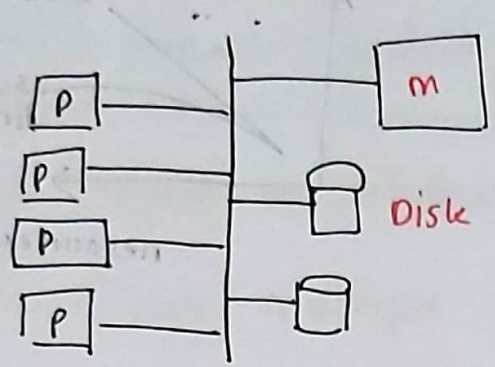
① Startup / Sequential costs

Amdahl's law :- Speed limited $= \frac{1}{(1-P) + (P/N)}$

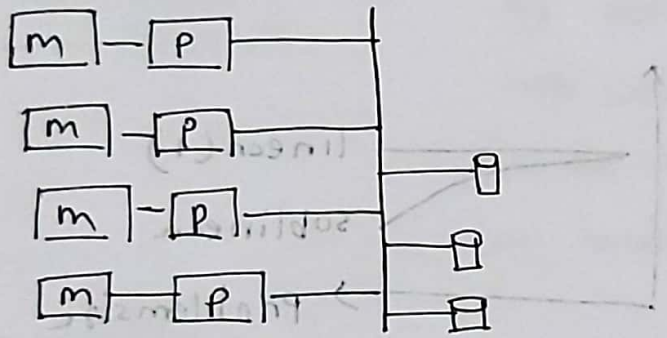
Gustafor's law $= \frac{1}{n(1-P)+P}$

② Interference

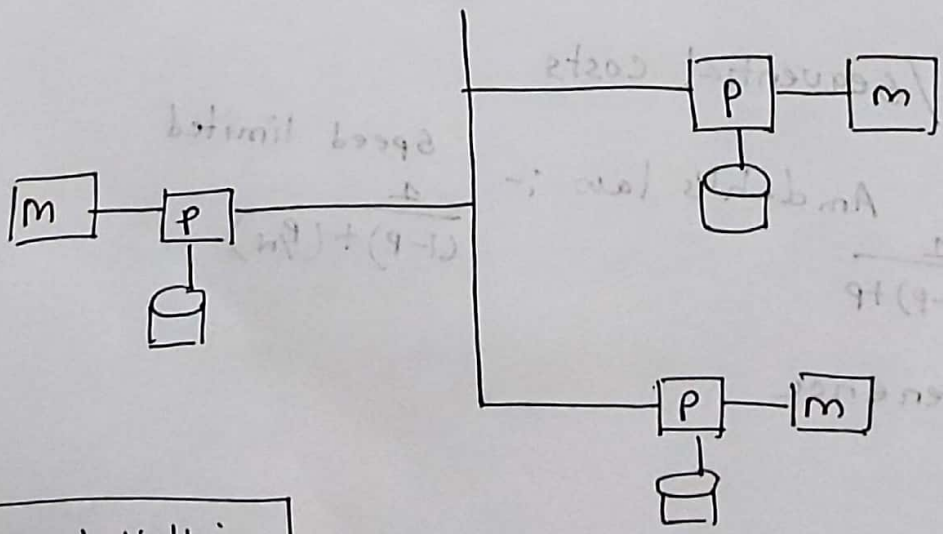
③ Skew



Shaned memony



Shaned Disk



Shaned Nothing

shared memory	shared Disk	shared Nothing
① multiple processon onNodes accecss a common physical memony spacc	① multiple processon on node have thein own pnivate memony but shane acess to common disk stonage	① each pnoeesson on node has its own pnivate disk and memony
② connected by bus on cross ban stwtich	② connected via stonage anea Netwonk on Netwonk attached stonage	② Donot shane memony on stonage dinectly
③ key chanactenistic :- communicaton and Data consistency - (not scaleable) - (Bottle neck issue)	③ Data access and shanning, fault tolenaance, scalibility (Bottle neck)	③ Data pantitioning, Distnibution and panallelism fault tolenance, scalibility (costing high)