__/__/____

Data Warehousing

→ constructed from multiple heterogeneous sources.

→ support analytical reporting, structured queries and decision making.

→ provides us generalized and consolidated data in multidimensional views.

→ It also provide us OLAP tools
↓
Online Analytical Processing

→ Analysis results in data generalization and data mining.
↓
Association, Clustering, Prediction.

Data Warehouse: Database which is kept separate from the organisation's operational database.
It has consolidated historical data.

→ OLTP (Online Transactional Processing)

Operational Database	Data warehouse
→ concurrent processing of multiple transactions	
Concurrent control & recovery mechanism req. to ensure robustness and consistency of data base.	→ time variant: It tells about the time stamp of that transaction
→ Read & modify operations	→ read only access.
→ maintains current data (Day to Day processing)	- historical data (Historical processing of information)
→ High performance	- high flexibility
→ ER model	→ Star schema, Snowflake schema fact constellation ..
100 MB – 1 GB	→ 100 GB – 1 TB.

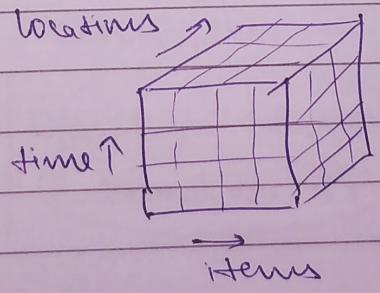
__/__/____

→ DWH is constructed by integrating data from multiple heterogeneous sources that supports analytical reporting, structured queries and decision making.

DWH
- → comparing quarterly & yearly sales
- → Analyse customer buying preferences, buying time, budget cycles etc.

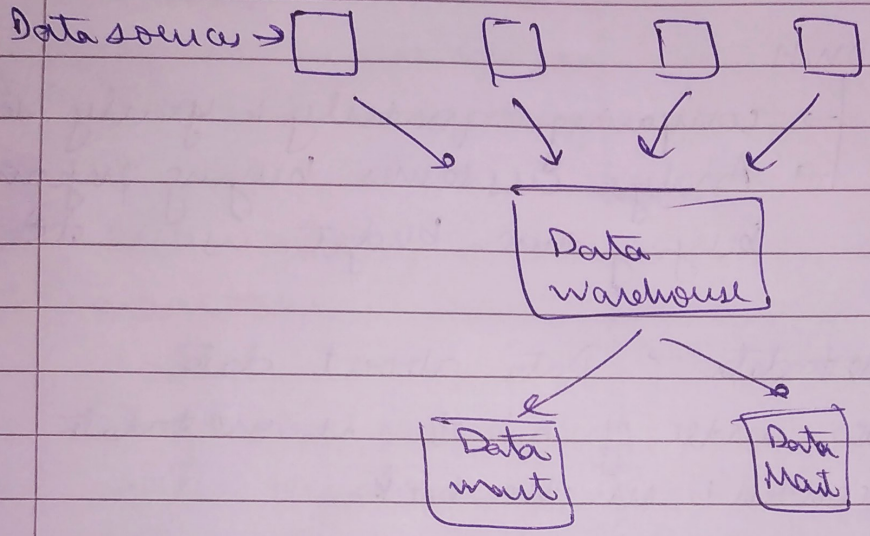
Metadata → Data about data
eg. index of a book is metadata for contents in the book
- → summarised data that leads to detailed data
- → Road map to data warehouse.

Data Cube:



__/__/____

Data Mart : Subset of organisation wide data that is valuable to specific group of people in organisation

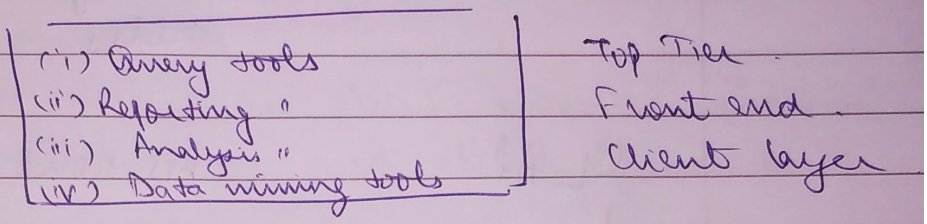


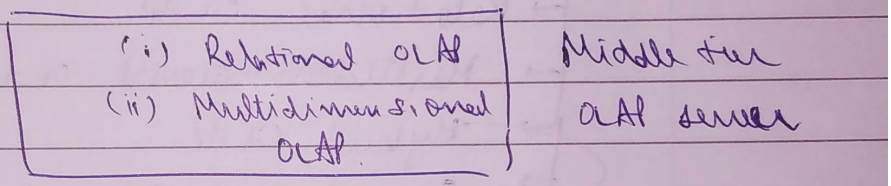
Virtual warehouse : View over an operational data warehouse

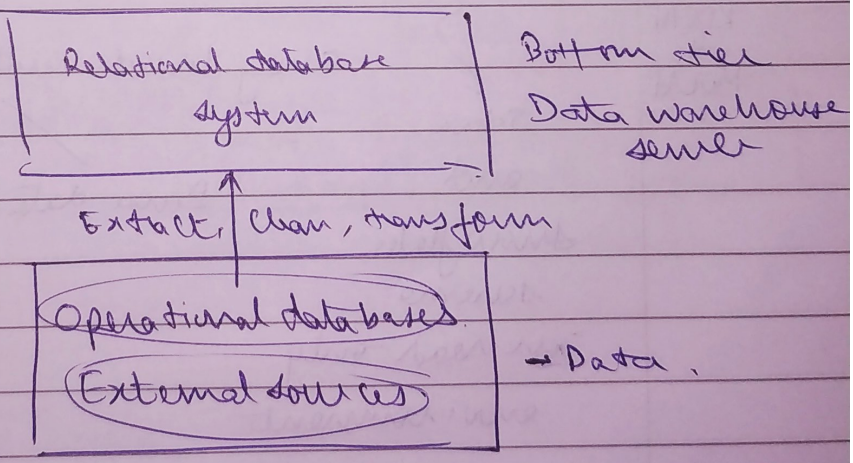
__/__/____

DWH – Architecture

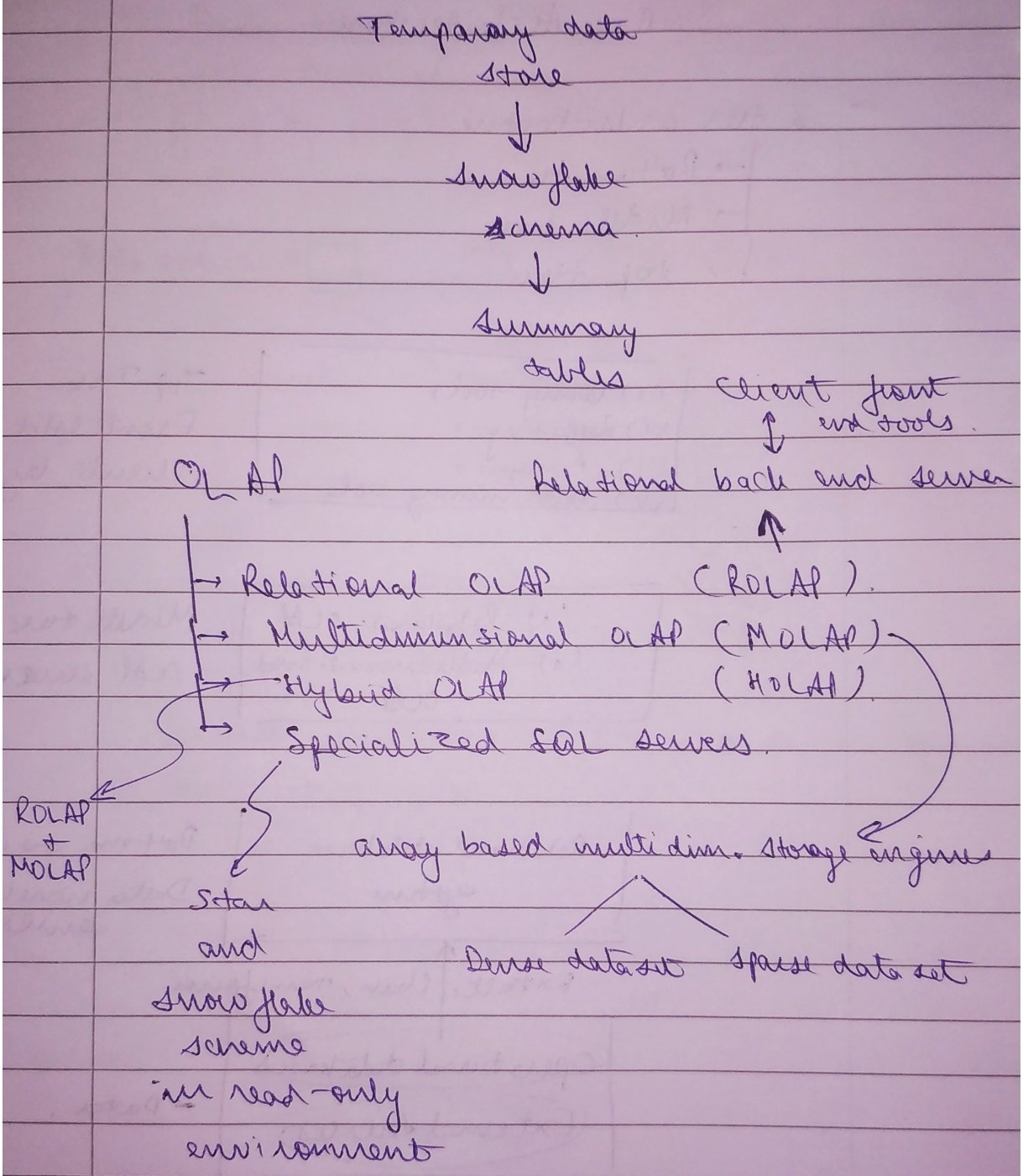
→ 3 tier architecture
- → Bottom tier
- → Middle tier
- → top tier







__/__/____


Temporary data store
Snow flake schema
Summary tables
Client front end tools
Relational back end server
OLAP
Relational OLAP (ROLAP)
Multidimensional OLAP (MOLAP)
Hybrid OLAP (HOLAP)
Specialized SQL servers
ROLAP + MOLAP
array based multidim. storage engines
Dense dataset
Sparse dataset
Star and snow flake schema in read-only environment

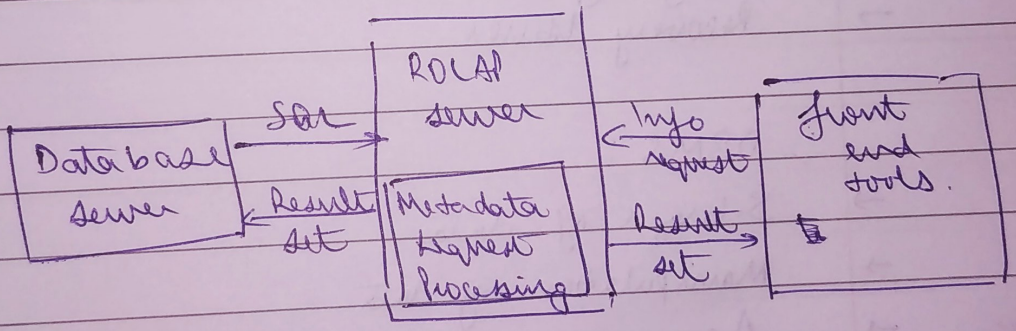
__/__/____

OLAP operations:

- Roll up
- Drill down (Reverse of Roll up)
- Slice & Dice.
- ~~Dice~~ Pivot (rotate)

Relational OLAP:
ROLAP servers are placed between relational back end server and client front end tools.

ROLAP architecture:



__/__/____

Schema → Logical description of entire database.

Star schema:
Snowflake "
Fact constellation schema:

File Based System Drawbacks:
→ Data Redundancy & Inconsistency
→ Data Isolation.
→ Unanticipated queries
→ Concurrent access anomalies
→ Security Issues
→ Integrity Issues
→ Recovery Issues

DBMS:
→ Storage of data
→ Manipulation of data
→ Access restrictions for unauthorized users

__/__/__

RDBMS

- multiple user access.
- access control feature.
- Support SQL.
- Reduce data redundancy