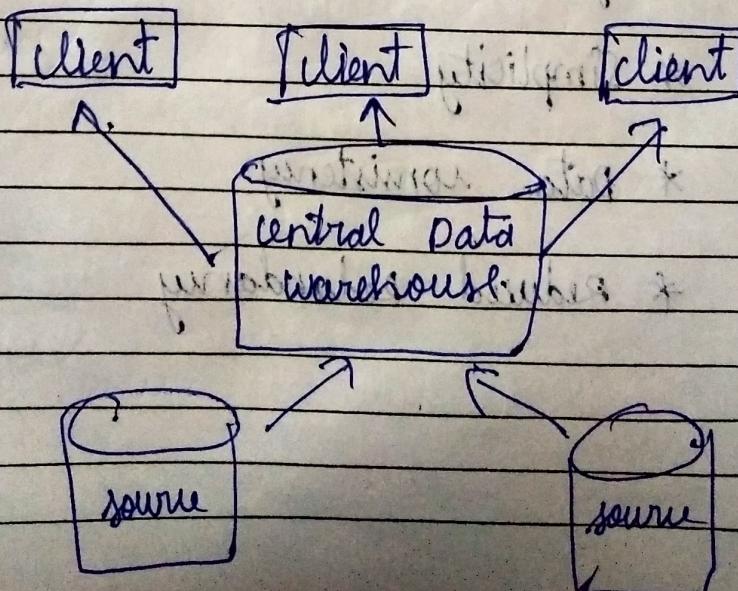


Data Warehouse process Architecture:

- > The process architecture in a data warehouse defines how data is collected, processed, and made available for analysis.
- > It defines the methods and procedures used to extract, transform, and load data from various sources into a centralized repository.
- > There are two fundamental types of process architectures in data warehouse.
 - i) centralized process architecture
 - ii) distributed architecture

i) centralized process architecture:



- > In a centralized process architecture, all data management processes, including data storage, processing, and analysis, are performed in a single, centralized system.
- > All data is stored in one central data warehouse, making it easier to manage and maintain.
- > ETL processes occur centrally, which can simplify data integration and data consistency.
- > It is easier to monitor, manage, and secure since all data and its processing occur in one place.

Advantages:

- * Simplicity
- * Data consistency
- * Reduced redundancy

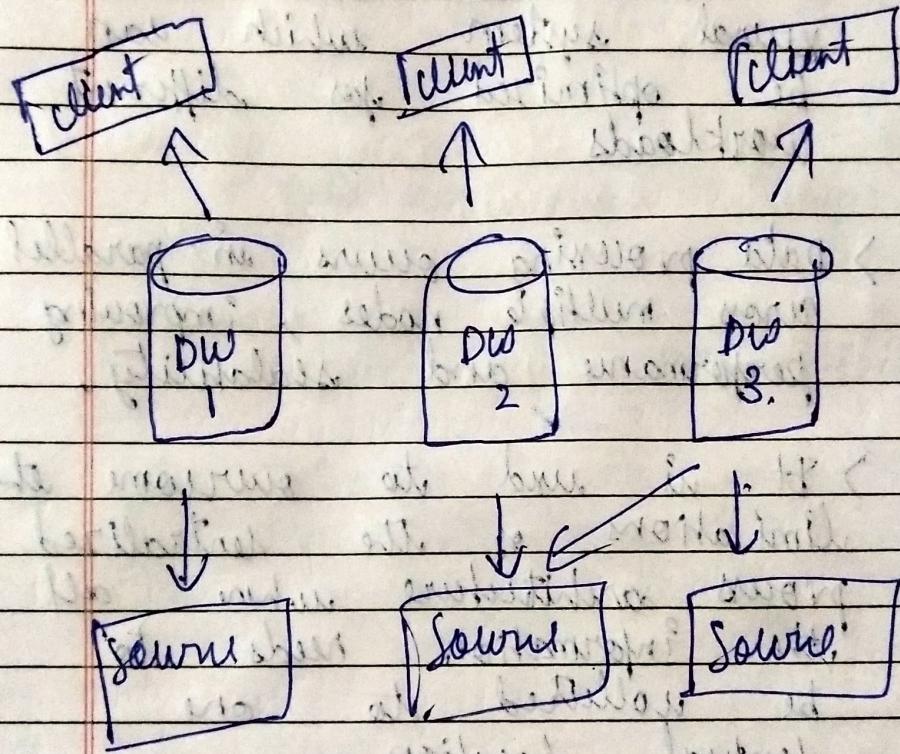
Disadvantages:

- * Scalability

- * Latency

- * Limited flexibility.

ii) Distributed process architecture



- > In a distributed processing architecture, data management processes are distributed across multiple systems.
- > This approach leverages the capabilities of various nodes to handle data processing, storage, and analysis.
- > Multiple data repositories are used to store data across several systems which can be optimized for different workloads.
- > Data processing occurs in parallel across multiple nodes, improving performance and scalability.
- > It is used to overcome the limitations of the centralized processing architecture where all the information needs to be collected to one central location.
- > Eg: Client - server, three-tier architecture, n-tier architecture, cluster architecture, peer - to - peer architecture.

Adv:

* Performance Increased

* Scalability

* Flexibility.

Dis:

* Complexity

* Latency

* Data consistency challenges.

Key components of process Architecture

1. Data Extraction : The process of retrieving data from various source systems, such as DB, flat files or other data repositories.

2. Data Transformation : Once data is extracted, it must be transformed to fit in the Data warehouse.

> It includes data cleaning, standardization, formatting.

3. Data Loading : The final step where the transformed data is loaded into the data warehouse.

Benefits :

- * Improve the efficiency of data handling & Management
- * Increase the amounts of data to scaled
- * Maintain the data consistency and Accuracy
- * Optimize the performance of data processing.