

Multi-tier architecture of Data Warehouse

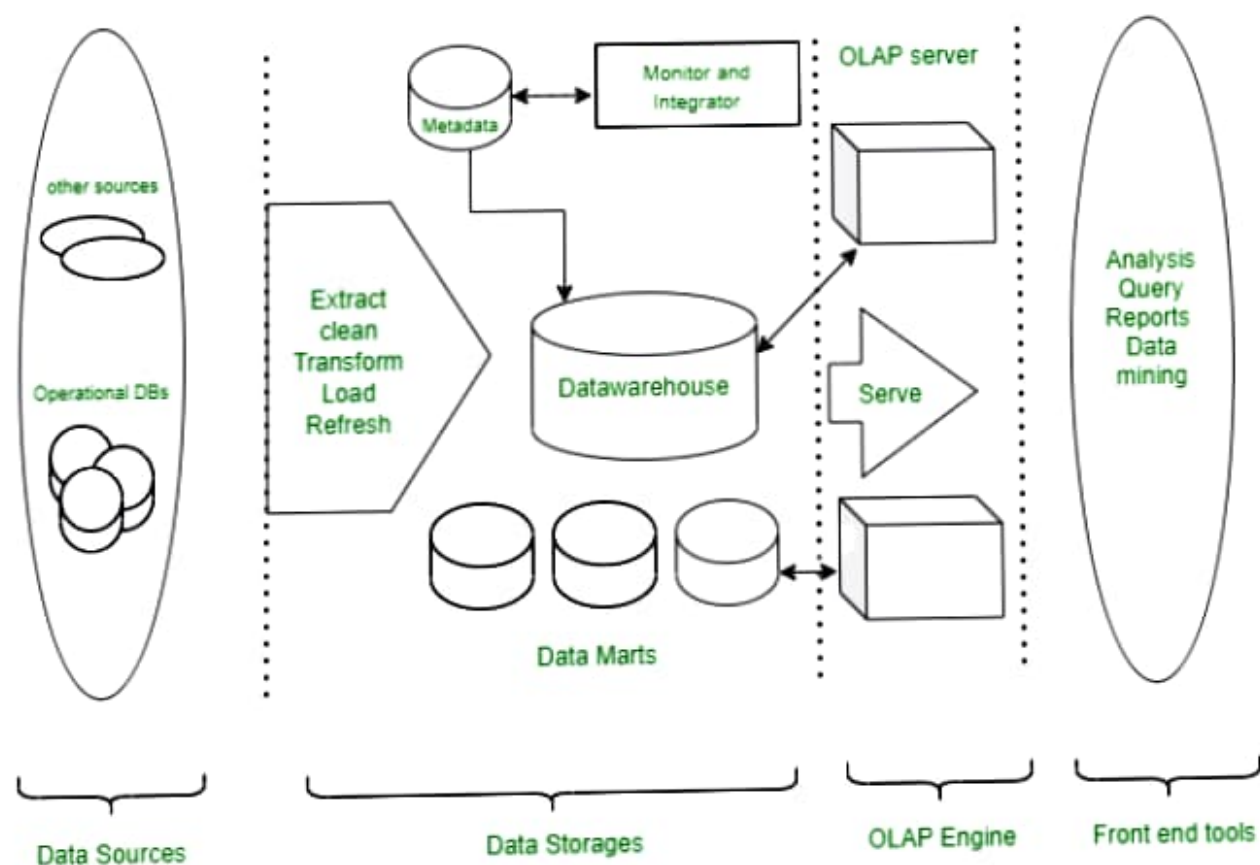
A [data warehouse](#) is Representable by data integration from multiple heterogeneous sources. It was defined by **Bill Inmon** in 1990.

The data warehouse is an integrated, subject-oriented, time-variant, and non-volatile collection of data. A Data Warehouse is structured by data integration from multiple heterogeneous sources. It is a system used for data analysis and reporting. A data warehouse is deliberate a core factor of business intelligence. BI technology provides a historical, current, and predictive view of business operations without data mining many businesses may not be able to perform effective market analysis, the strength and weakness of their competitors, profitable decisions, etc.

Data Warehouse is referred to the data repository that is maintained separately from the organization's operational data. **Multi-Tier**

Data Warehouse is referred to the data repository that is maintained separately from the organization's operational data. **Multi-Tier Data Warehouse Architecture** consists of the following components:

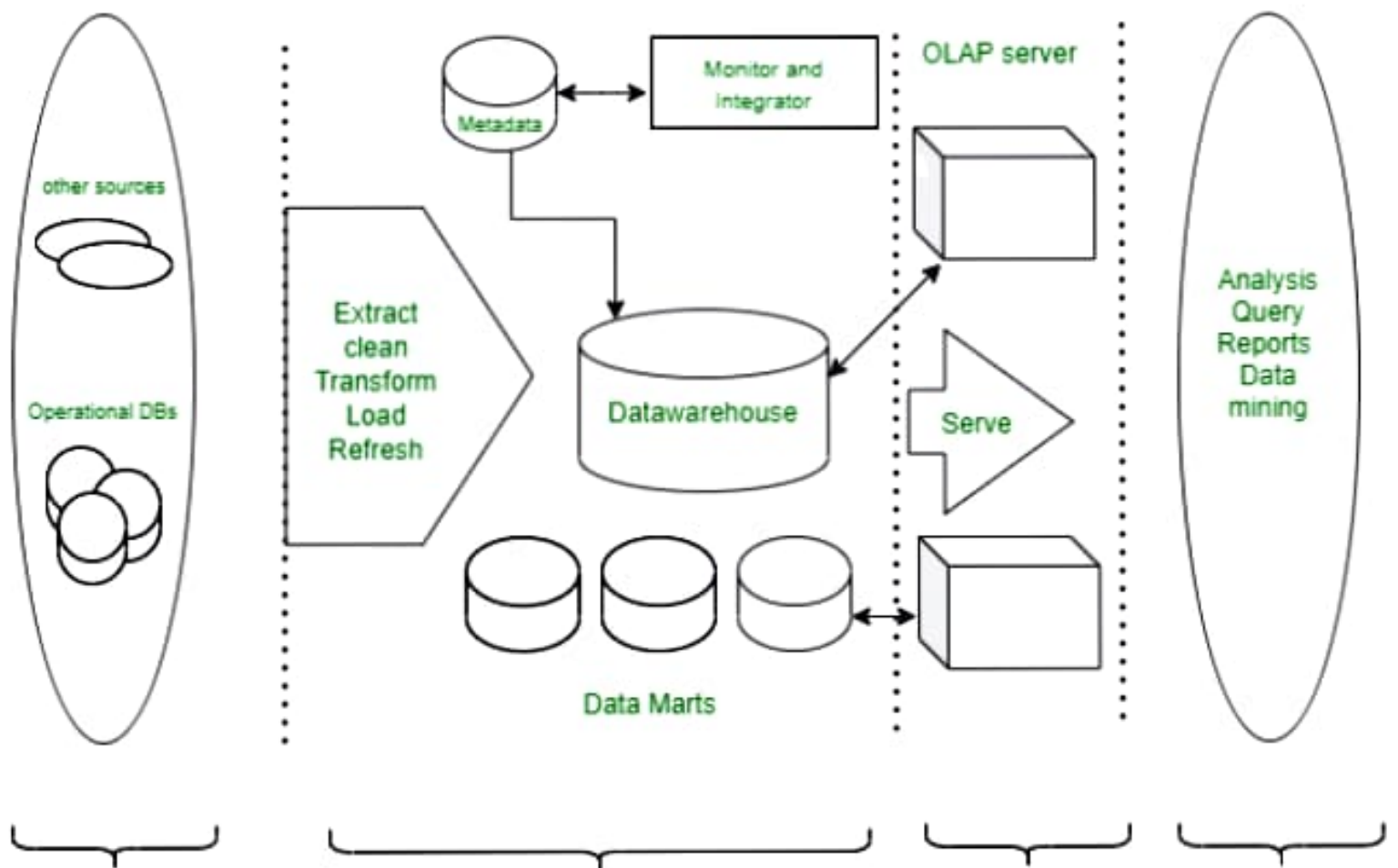
1. Bottom Tier
2. Middle Tier
3. Top Tier



Three/Multi-tier Architecture of Data Warehouse

Data Warehouse Architecture consists of the following components:

1. Bottom Tier
2. Middle Tier
3. Top Tier



Bottom Tier(Data sources and data storage) :

1. The bottom Tier usually consists of Data Sources and Data Storage.
2. It is a warehouse database server. For Example RDBMS.
3. In Bottom Tier, using the application program interface(called gateways), data is extracted from operational and external sources.
4. Application Program Interface likes ODBC(Open Database Connection), OLE-DB(Open-Linking and Embedding for Database), JDBC(Java Database Connection) is supported.
5. ETL stands for Extract, Transform, and Load.

Several popular ETL tools include:

- I.** IBM Infosphere
- II.** Informatica
- III.** Confluent
- IV.** Microsoft SSIS
- V.** Snaplogic
- VI.** Alooka

Middle Tier :

The middle tier is an OLAP server that is typically implemented using either :

A relational OLAP (ROLAP) model (i.e., an extended relational DBMS that maps operations from standard data to standard data); **or** A multidimensional OLAP (MOLAP) model (ie, a special purpose server that directly implements multidimensional data and operations).

OLAP server models come in three different categories, including:

1. **ROLAP:** A relational database is not converted into a multidimensional database; rather, a relational database is actively broken down into several dimensions as part of relational online analytical processing(ROLAP). This is used when everything that is contained in the repository is a relational database system.
2. **MOLAP:** A different type of online analytical processing called multidimensional online analytical processing(MOLAP) includes directories and catalogs that are immediately

1. **ROLAP:** A relational database is not converted into a multidimensional database; rather, a relational database is actively broken down into several dimensions as part of relational online analytical processing(ROLAP). This is used when everything that is contained in the repository is a relational database system.
2. **MOLAP:** A different type of online analytical processing called multidimensional online analytical processing(MOLAP) includes directories and catalogs that are immediately integrated into its multidimensional database system. This is used when all that is contained in the repository is the multidimensional database system.
3. **HOLAP:** A combination of relational and multidimensional online analytical processing paradigms is hybrid online analytical processing(HOLAP). HOLAP is the ideal option for a seamless functional flow across the database systems when the repository houses both the relational database

Top Tier :

The top tier is a front-end client layer, which includes query and reporting tools, analysis tools, and/or data mining tools (eg, trend analysis, prediction, etc.).

Here are a few Top Tier tools that are often used:

- I. SAP BW**
- II. SAS Business Intelligence**
- III. IBM Cognos**
- IV. Crystal Reports**
- V. Microsoft BI Platform**

Data Warehouse Models :

From the architecture point of view, there are three warehouse models-

Enterprise Warehouse:-

- An enterprise warehouse collects all information topics spread throughout the organization.
- It provides corporate-wide data integration, typically from one or several operational systems or external information providers.

Data Warehouse Models :

From the architecture point of view, there are three warehouse models-

Enterprise Warehouse:-

- An enterprise warehouse collects all information topics spread throughout the organization.
- It provides corporate-wide data integration, typically from one or several operational systems or external information providers, and is cross-functional in scope.
- It usually contains detailed data as well as summarized data and can range in size from a few gigabytes to hundreds of gigabytes, terabytes, or beyond. Can be an enterprise data warehouse.
- The traditional mainframe, computer super server, or parallel architecture has been implemented on platforms. This requires extensive commercial modeling and may take years to design and manufacture.

Data Mart:-

- A data mart contains a subset of corporate-wide data that is important to a specific group of users.
- The scope is limited to specific selected subjects.
- For example, a marketing data mart may limit its topics to customers, goods, and sales.
- The data contained in the data marts are summarized. Data marts are typically applied to low-cost departmental servers that are Unix/Linux or Windows-based.
- The implementation cycle of a data mart is more likely to be measured in weeks rather than months or years. However, in the long run, complex integration is involved in its design and planning were not enterprise-wide.

Virtual Warehouse:-

- A virtual warehouse is a group of views on an operational database.
- For efficient query processing, only a few possible summary views can be physical.

Virtual Warehouse:-

- A virtual warehouse is a group of views on an operational database.
- For efficient query processing, only a few possible summary views can be physical.
- Creating a virtual warehouse is easy, but requires additional capacity on operational database servers.

Advantages of Multi-Tier Architecture of Data warehouse

1. **Scalability:** Various components can be added, deleted, or updated in accordance with the data warehouse's shifting needs and specifications.
2. **Better Performance:** The several layers enable parallel and efficient processing, which enhances performance and reaction times.
3. **Modularity:** The architecture supports modular design, which facilitates the creation, testing, and deployment of separate components.
4. **Security:** The data warehouse's overall

Advantages of Multi-Tier Architecture of Data warehouse

1. **Scalability:** Various components can be added, deleted, or updated in accordance with the data warehouse's shifting needs and specifications.
2. **Better Performance:** The several layers enable parallel and efficient processing, which enhances performance and reaction times.
3. **Modularity:** The architecture supports modular design, which facilitates the creation, testing, and deployment of separate components.
4. **Security:** The data warehouse's overall security can be improved by applying various security measures to various layers.
5. **Improved Resource Management:** Different tiers can be tuned to use the proper hardware resources, cutting expenses overall and increasing effectiveness.
6. **Easier Maintenance:** Maintenance is simpler because individual components can be

4. **Security:** The data warehouse's overall security can be improved by applying various security measures to various layers.
5. **Improved Resource Management:** Different tiers can be tuned to use the proper hardware resources, cutting expenses overall and increasing effectiveness.
6. **Easier Maintenance:** Maintenance is simpler because individual components can be updated or maintained without affecting the data warehouse as a whole.
7. **Improved Reliability:** Using many tiers can offer redundancy and failover capabilities, enhancing the data warehouse's overall reliability.