

## PHASE 2

### DATA WAREHOUSING WITH IBM CLOUD DB2 WAREHOUSE

#### INTRODUCTION:

Data Warehouse is a relational database management system (RDBMS) construct to meet the requirement of transaction processing systems. It can be loosely described as any centralized data repository which can be queried for business benefits. It is a group of decision support technologies, targets to enabling the knowledge worker (executive, manager, and analyst) to make superior and higher decisions. So, Data Warehousing support architectures and tool for business executives to systematically organize, understand and use their information to make strategic decisions.

#### Data Warehouse Structure:

A typical data warehouse has four main components: a central database, ETL (extract, transform, load) tools, metadata, and access tools. All of these components are engineered for speed so that you can get results quickly and analyze data on the fly

o **Centralised database:** A database serves as the foundation of your data warehouse. Traditionally, these have been standard relational databases running on premise or in the cloud. But because of Big Data, the need for true, real-time performance, and a drastic reduction in the cost of RAM, in-memory databases are rapidly gaining in popularity



o **Data integration:** Data is pulled from source systems and modified to align the information for rapid analytical consumption using a variety of data integration approaches such as ETL (extract, transform, load) and ELT as well as real-time data replication, bulk-load processing, data transformation, and data quality and enrichment services.



o **Metadata:** Metadata is data about your data. It specifies the source, usage, values, and other features of the data sets in your data warehouse. There is business metadata, which adds context to your data, and technical metadata, which describes how to access data – including where it resides and how it is structured

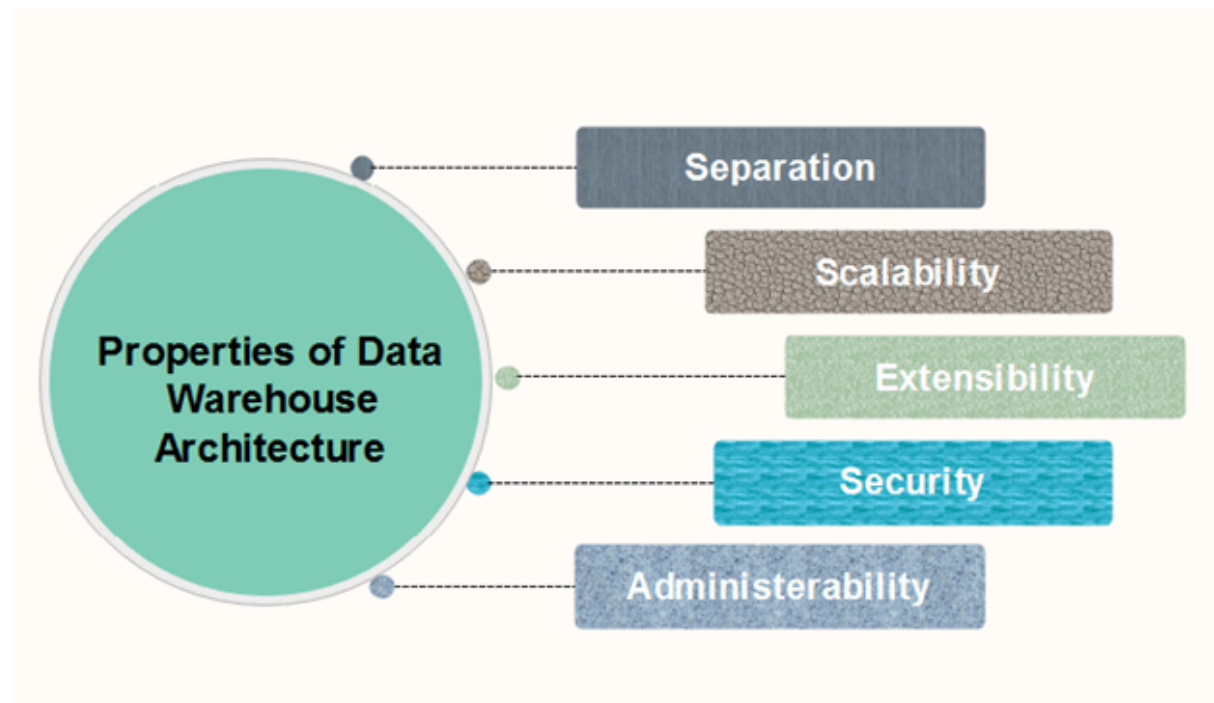


o **Data warehouse access tools:** Access tools allow users to interact with the data in your data warehouse. Examples of access tools include: query and reporting tools, application development tools, data mining tools, and OLAP tools

## Key Characteristics of Data Warehouse

1. **Subject-oriented:** A data warehouse typically provides information on a topic (such as a sales inventory or supply chain) rather than company operations.
2. **Time-variant:** Time variant keys (e.g., for the date, month, time) are typically present.
3. **Integrated:** A data warehouse combines data from various sources. These may include a cloud, relational databases, flat files, structured and semi-structured data, metadata, and master data. The sources are combined in a manner that's consistent, relatable, and ideally certifiable, providing a business with confidence in the data's quality.
4. **Persistent and non-volatile:** Prior data isn't deleted when new data is added. Historical data is preserved for comparisons, trends, and analytics.

## Properties of Data Warehouse Architectures



1. **Separation:** Analytical and transactional processing should be kept apart as much as possible.
2. **Scalability:** Hardware and software architectures should be simple to upgrade the data volume, which has to be managed and processed, and the number of user's requirements, which have to be met, progressively increase.
3. **Extensibility:** The architecture should be able to perform new operations and technologies without redesigning the whole system.
4. **Security:** Monitoring accesses are necessary because of the strategic data stored in the data warehouses.
5. **Administerability:** Data Warehouse management should not be complicated.

## CONCLUSION:

In conclusion, the robust data warehouse project has successfully addressed our organization's data management needs. It has enhanced data accessibility, reliability, and performance, enabling better-informed decision-making. The project's scalability ensures future growth and adaptability to evolving data requirements. With improved data quality and integration, our organization is better positioned to achieve its strategic goals. This data warehouse project represents a vital asset in our quest for data-driven excellence.