

Cloud Application Development

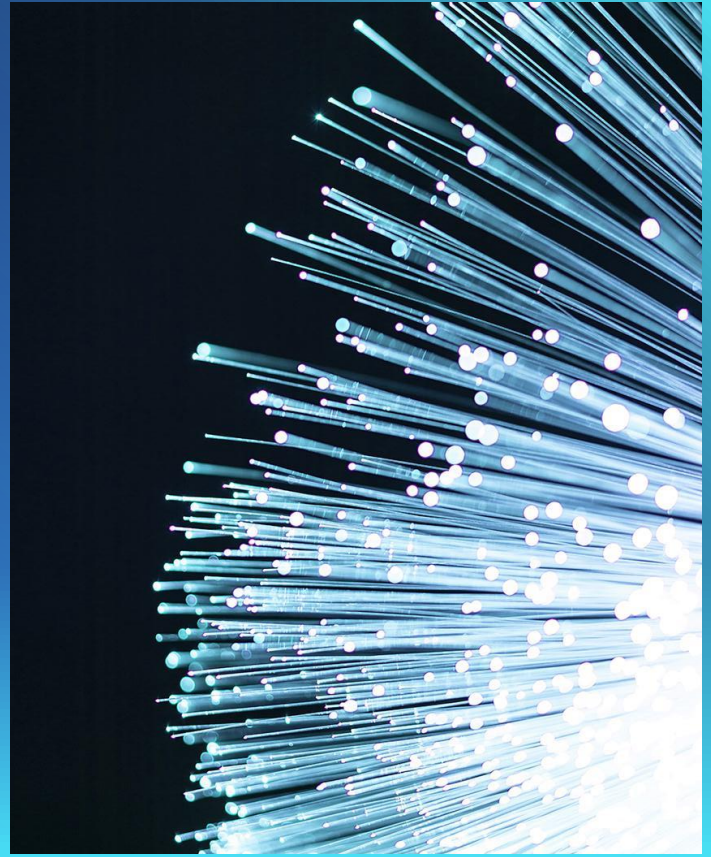
**Data Warehousing with IBM
Cloud Db2 Warehouse**

IBM: CAD101



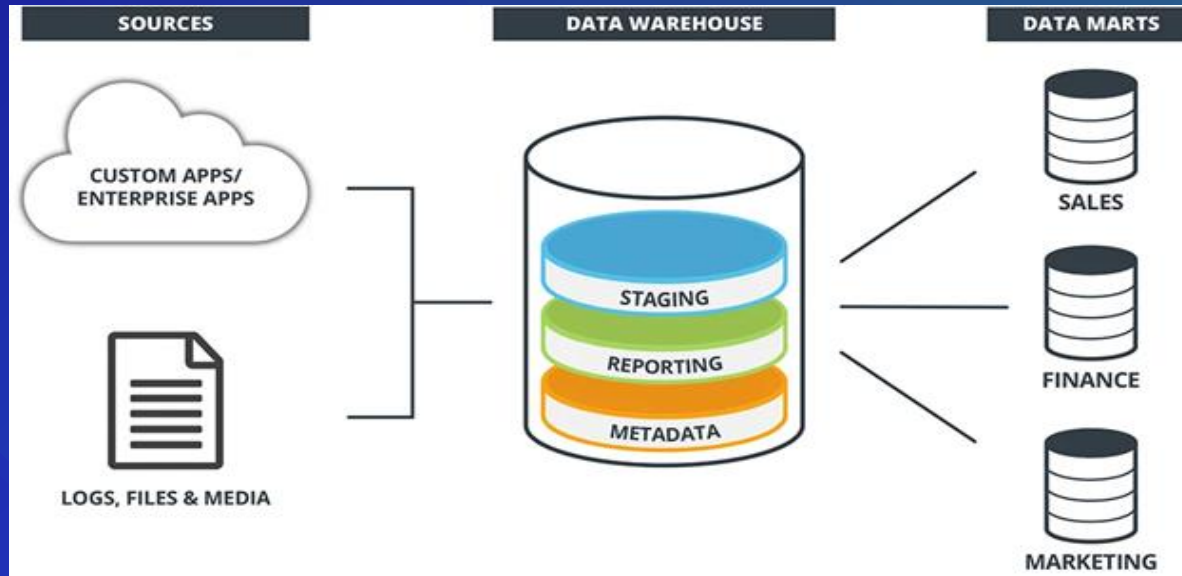
01

DATA WAREHOUSE INNOVATION



DATA WAREHOUSING WITH IBM CLOUD DB2 WAREHOUSE

BLOCK DIAGRAM OF DATA WAREHOUSING:





IBM CLOUD DB2 WAREHOUSE

IBM Db2 Warehouse is a data warehousing solution provided by IBM as part of its cloud services. It is designed to help organizations efficiently manage and analyze large volumes of data from various sources, enabling data-driven decision-making

DATA WAREHOUSING:

- A data warehouse will help you to build accurate forecasting models and identify impactful trends. When building a data warehouse it's important to recognise the following steps and thoroughly address each one.



PREPROCESSING:

Defining Business Requirements:

Since a data warehouse encompasses all areas of your business it's vital that every department is involved with the design process. The process of requirements gathering involves all stakeholders and helps every department understand the purpose of the data warehouse, how they will benefit, and what results they can expect.



RESOURCES

- * Requirement gathering can happen as one-to-one or collective meetings. This phase often turns out to be one of the hardest parts of data warehousing implementation. Because a data warehouse includes data from so many sources, spanning multiple departments, there can often be negotiations over information sharing and prioritisation. A skilled business analyst can act as an external mediator to ensure all stakeholders are happy with the defined project goals.
- * The requirements gathering phase is so important in ensuring that department goals are aligned with the overall project and that buy-in from all the relevant stakeholders is achieved. It can also help to highlight current and future needs from taking a deep dive into the data which will be used for analysis, which will likely uncover where your data is and isn't being used effectively.

SETTING UP PHYSICAL ENVIRONMENTS:

- * Your data warehouse will typically have three environments which mimics software development best practice. The three environments are development, testing, and production and these are used in tandem to ensure changes are tested for integrity and security before they are pushed to live in the production environment. They allow for development and Quality Assurance to occur without affecting the productive environment.
- * Three separate environments are also needed to run test data, identify breakpoints that need to be rectified, and to reduce stress on server workloads. It is not enough to simply have different physical environments set up. The different processes (such as ETL, OLAP cube and reporting) also need to be set up properly for each environment.

INTRODUCING DATA MODELLING:

- * Data modelling is the blueprint from which the data warehouse is built. It can help you visualise data relationships, standardise naming conventions and establishing security process compliance. This is known as the most complex phase of data warehouse design. A good data model will allow the data warehousing system to grow easily and perform well!
- * Data modelling normally takes place at the data mart level and branches out into the data warehouse. The three most popular data models for warehouses are the Snowflake, Star, and Galaxy Schemas. The chosen model will impact the structure of your data warehouse and data marts, and it will help to guide the overall architecture within the warehouse.

CHOOSING YOUR ETL SOLUTION

- * ETL stands for Extract, Transform and Load and represents the collection and processing of data from various sources into one central data store where it can be later analysed. Your business has access to many data sources but often it's presented in a way that is hard or impossible to consume.
- * A good ETL process can be the difference between a slow and hard-to-use data warehouse and a sleek warehouse that adds value to every part of your organisation. For this reason, it's vital that the right ETL solution is selected.

ONLINE ANALYTICS PROCESSING (OLAP) CUBE

An OLAP cube helps you to analyse the data in your data warehouse or data mart. Since your warehouse will be sorting data from multiple sources, the OLAP cube helps you to organise all of that data in a multi-dimensional format that makes it easier to analyse.

FEATURE EXTRACTION

Creating A Front End:

This stage refers to the front-end visualisation, where users can understand and apply the results of data queries. If users cannot visualise the reports, the data warehouse is likely to provide little value to them, making front end development an important part of a data warehouse initiative. Another area of importance is the complexity of the reporting tool.



Optimising Queries:

The more data returned from a query, the more resources the database needs to expand to process and store this data. This is why it's important to only minimise data retrieval, especially if you're paying for query power separately.

This stage is very specific to each organisation's needs.



Rolling Out The End Product:

:

The hard work has now been done and you're close to getting value from your shiny, new data warehouse. It's at this point that team members will need to be trained in using it. Throughout the process Quality Assurance and Testing have been ensuring there are no bugs or usability issues.



Although these are the standard steps in creating a data warehouse, it's important to remember that every scenario is different. There may be additional steps that your business needs to take based on the requirements or complexity of your organisation's needs. Ultimately, a successfully implemented data warehouse will deliver value at every level of your organization

IBM Cloud Db2 Warehouse is a robust and versatile cloud-based data warehousing

