PROJECT OVERVIEW

Project title : **Data warehousing with IBM Cloud Db2 Warehouse Edit set Access Page Actions**

Domain : **Cloud Application Development – Group 4**

Assignment : **Project submission phase 1**

SUBMITTED BY

Name : k.sanjai

Mail id : srisanjai6@gmail.com

College Name : P.R. ENGINEERING COLLEGE

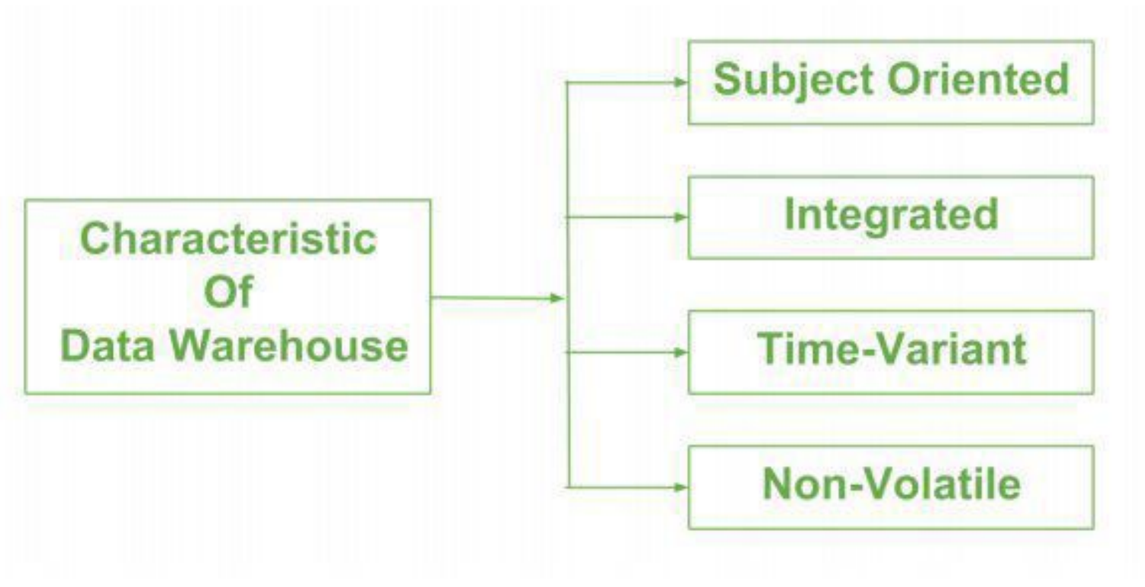
College code : 8212

Group 4 : Zone (13-16)

PROBLEM STATEMENT:

"In today's data-driven business landscape, organizations face a critical challenge: efficiently harnessing and leveraging the vast and diverse data available from various sources for informed decision-making. To address this challenge, our project aims to design and implement a robust data warehouse solution using IBM Cloud Db2 Warehouse. The primary objective is to establish a centralized repository that integrates data from multiple sources, executes efficient ETL processes, and empowers data architects to explore and analyze data effectively. By doing so, we aim to provide actionable insights that enable informed decision-making and enhance the organization's competitive advantage."

DATA WAREHOUSE: **A data warehouse is a central repository of information that can be analyzed to make more informed decisions.**



OBJECTIVE:

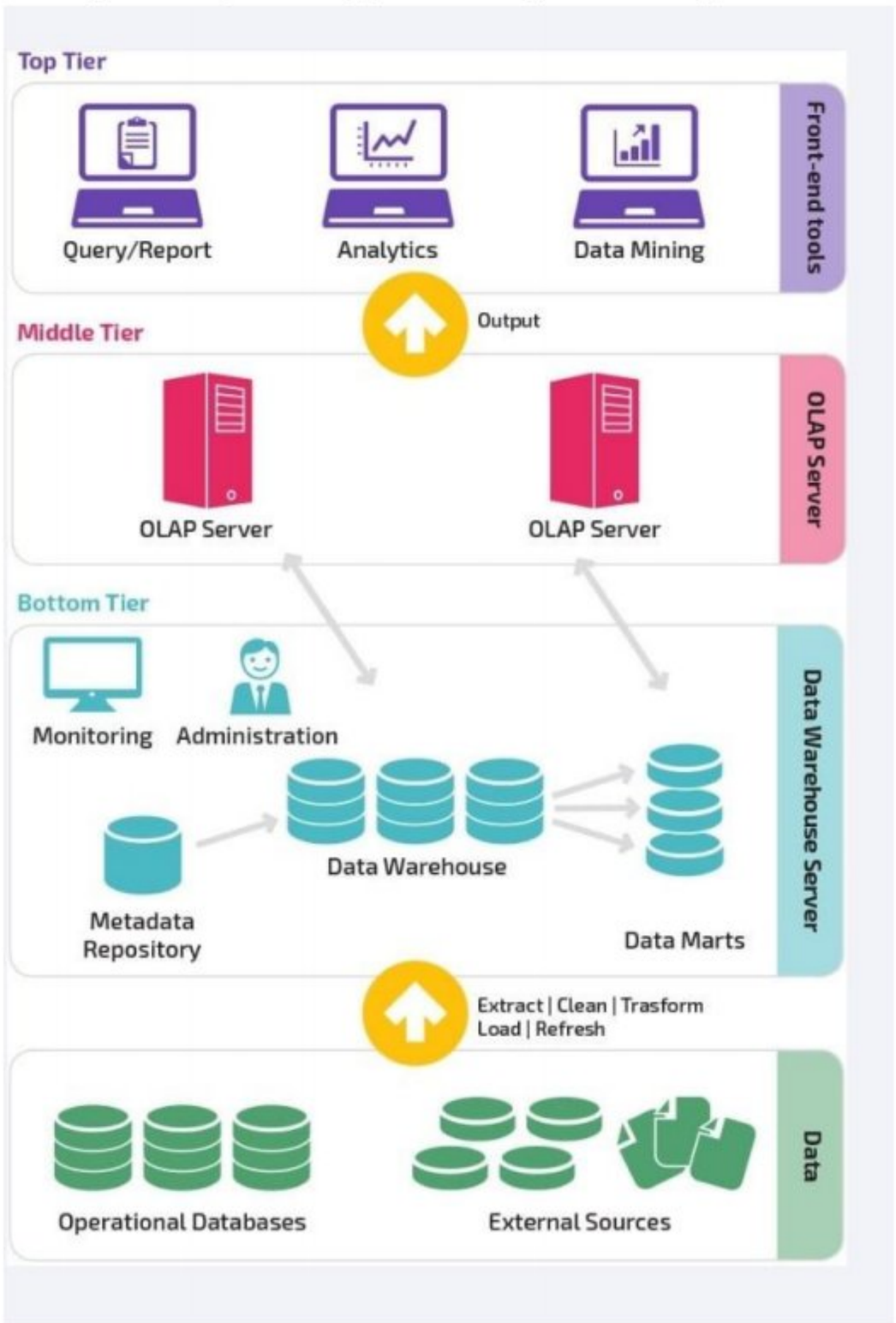
"To create a reliable and efficient data warehouse using IBM Cloud Db2 Warehouse, which consolidates data from various sources, ensures data quality, and empowers users to analyze data for better decision-making."

DESIGN THINKING

1. Data Warehouse Structure:

OBJECTIVE:

- **To design a data warehouse structure that aligns with user needs, facilitates efficient data management, and supports insightful analysis.**



DATA WAREHOUSE THREE TIER ARCHITECTURE

STEPS:

Empathize:

- **Understand the data management challenges faced by users.**
- **Explore their pain points and data access requirements.**

Define:

- **Define the objectives and goals for the data warehouse structure based on user insights.**
- **Create user personas representing the different data warehouse users.**

Ideate:

- **Collaborate with stakeholders to brainstorm schema and structure designs.**
- **Explore innovative ways to accommodate various data sources.**

Prototype:

- **Develop prototype schema designs and data models.**
- **Gather feedback from users to refine these prototypes.**

Test:

- **Conduct usability testing with users to ensure that the schema design aligns with their needs and expectations.**

DELIVERABLES:

- ✓ **User person as**
- ✓ **Defined data warehouse objectives**
- ✓ **Prototype schema designs**
- ✓ **Usability test results and refined schema design**

2. Data Integration:

Objective:

- ➢ **To create seamless data integration processes that simplify data gathering from diverse sources and minimize user effort.**

STEPS

Empathize:

- **Understand the pain points users face when integrating data from different sources.**
- **Identify their frustrations with current integration processes.**

Define:

- **Define data integration requirements based on user insights.**
- **Document data source formats, characteristics, and integration constraints.**

Ideate:

- **Brainstorm creative solutions for data integration, aiming to simplify the process.**
- **Explore user-friendly integration tools and techniques.**

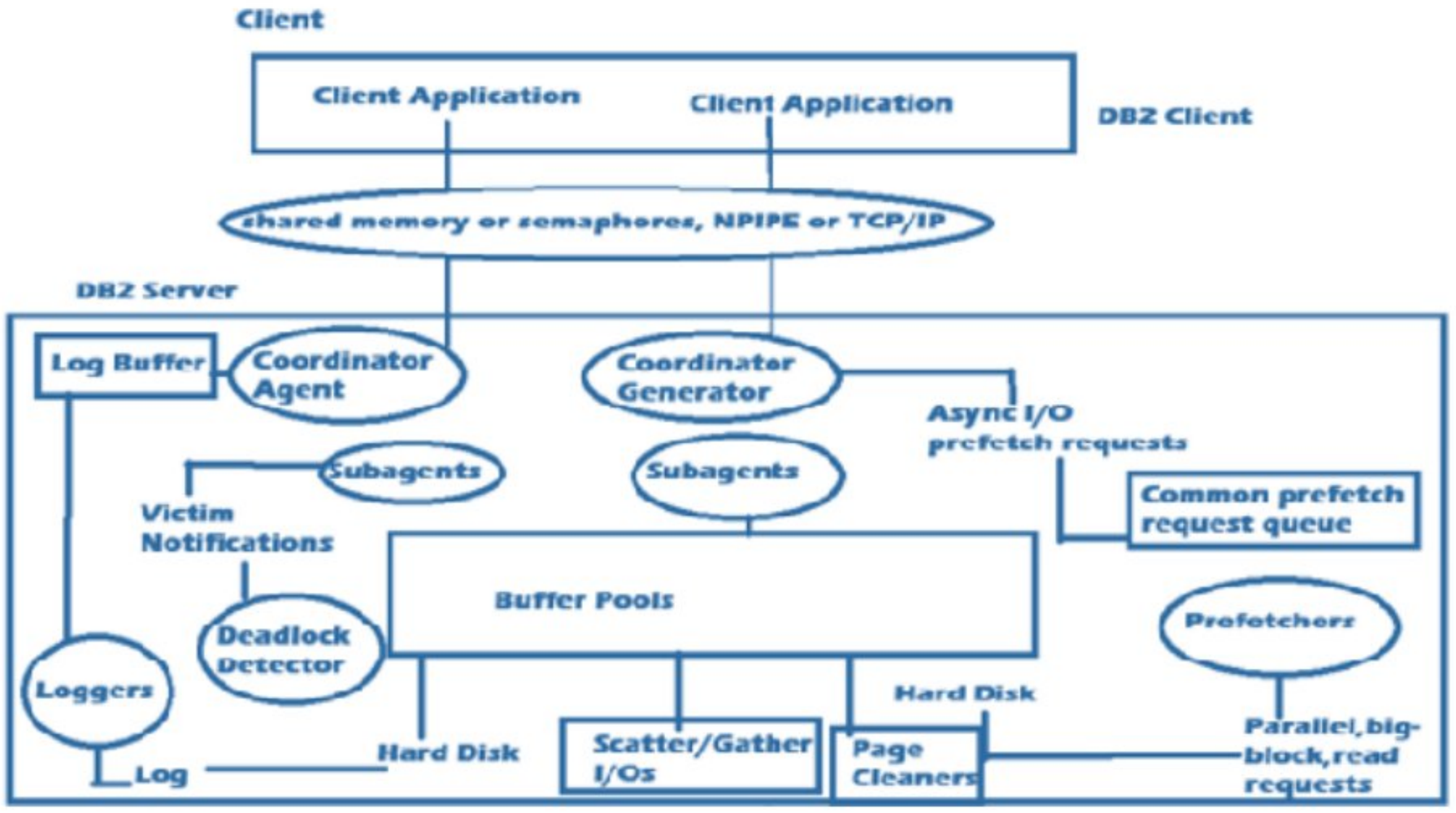
Prototype:

- **Create prototypes of data integration workflows and tools.**
- **Involve users in testing and refining these prototypes.**

Test:

- **Conduct usability testing with data architects and integration teams to refine the integration processes.**

Db 2 Architecture



DELIVERABLES:

- ✓ **Defined data integration requirements**
- ✓ **Prototypes of data integration workflows and tools**
- ✓ **Usability test results and refined integration solutions**

3. ETL Processes:

Objective:

- **To design ETL processes that are user-centric, efficient, and aligned with data engineering needs.**

Steps:

- **Empathize:Understand the challenges data engineers face during ETL processes.Identify pain points related to data extraction, transformation, and loading.**

Define:

- **Define ETL process requirements based on user insights.**
- **Document data transformation rules, quality checks, and automation needs.**

Ideate:

- **Brainstorm innovative ETL solutions that simplify data transformation and ensure data quality.**
- **Explore automation possibilities.**

Prototype:

- **Develop prototypes of ETL workflows and automation scripts.**
- **Involve data engineers in testing and refining these prototypes.**

Test:

- **Collaborate closely with data engineering teams to ensure that the ETL processes are user-friendly, efficient, and aligned with user expectations.**

DELIVERABLES:

- ✓ **Defined ETL process requirements**
- ✓ **Prototypes of ETL workflows and automation scripts**
- ✓ **Usability test results and refined ETL processes**

4. Data Exploration:

Objective:

- **To empower data architects and analysts to intuitively explore and analyze data, fostering a data-driven decision-making culture.**

Steps:

Empathize:

- **Understand the data exploration needs of data architects and analysts.**
- **Identify pain points in accessing and analyzing data.**

Define:

- **Define the requirements for data exploration tools and techniques.**
- **Document user expectations for data visualization and analysis.**

Ideate:

- **Brainstorm creative data exploration solutions that empower users to uncover insights effortlessly.**
- **Explore user-friendly visualization tools.**

Prototype:

- **Develop prototypes of data exploration dashboards and tools.**
- **Collect feedback from data architects to refine these prototypes.**

Test:

- **Conduct usability testing with data analysts to ensure that the data exploration tools facilitate intuitive and insightful data analysis.**

DELIVERABLES:

- ✓ **Defined data exploration requirements**
- ✓ **Prototypes of data exploration dashboards and tools**
- ✓ **Usability test results and refined data exploration solutions**

5. Actionable Insights:

Objective:

- **To deliver actionable insights in a format that is easily digestible and aligned with decision-makers' needs, enabling data-driven decisions.**

STEPS:

Empathize:

- **Understand how decision-makers currently use data to inform their actions.**
- **Identify any barriers to data-driven decision-making.**

Define:

- **Define the requirements for delivering actionable insights.**
- **Document the specific types of insights that decision-makers require.**

Ideate:

- **Brainstorm ways to present insights in a format that is easily digestible and actionable.**
- **Explore interactive dashboards and alerting mechanisms.**

Prototype:

- **Develop prototypes of actionable insights delivery mechanisms.**
- **Involve decision-makers in testing and refining these prototypes.**

Test:

- **Collaborate closely with decision-makers to ensure that the insights provided are truly actionable and align with their decision-making processes.**

DELIVERABLES:

- ✓ **Defined actionable insights requirements**
- ✓ **Prototypes of actionable insights delivery mechanisms**
- ✓ **Usability test results and refined actionable insights solutions**

PYTHON PROGRAM:

```
import ibm_db

# Replace these values with your database credentials

db_hostname = "your_db_hostname"

db_port = "your_db_port"

db_name = "your_db_name"

db_user = "your_db_user"

db_password = "your_db_password"# Establish a connection to the Db2 Warehouse database

conn_str = ( "DATABASE={0};HOSTNAME={1};PORT={2};PROTOCOL=TCPIP;"

   "UID={3};PWD={4};".format(db_name, db_hostname, db_port, db_user, db_password))

try:

 conn = ibm_db.connect(conn_str, "", "")

  if conn:

  print("Connected to the database!")

# Prepare and execute a SELECT query

   sql_query = "SELECT * FROM employees"

  stmt = ibm_db.exec_immediate(conn, sql_query)

    while ibm_db.fetch_row(stmt):

        employee_id = ibm_db.result(stmt, "EMPLOYEE_ID")

        first_name = ibm_db.result(stmt, "FIRST_NAME")

        last_name = ibm_db.result(stmt, "LAST_NAME")

        department = ibm_db.result(stmt, "DEPARTMENT")

print("Employee ID: {0}, First Name: {1}, Last Name: {2}, Department: {3}".format(

            employee_id, first_name, last_name, department))

    ibm_db.close(conn)

    print("Connection closed.")

  else:

    print("Failed to connect to the database.")

except Exception as e:

  print("Error:", e)
```

OUTPUT:

pip install ibm-db

CONCLUSION:

In conclusion, the "Design-Driven Data Warehousing with IBM Cloud Db2" project, approached through the lens of design thinking, aims to create a data infrastructure that not only meets technical requirements but also deeply resonates with the needs and expectations of users. By following the objectives, steps, and deliverables outlined for each project stage, we strive to achieve a user-centric, efficient, and insightful data warehousing solution.Throughout this project, we prioritize empathy for users and stakeholders, engaging them in the design and testing phases to ensure that the data warehouse structure, data integration processes, ETL workflows, data exploration tools, and actionable insights mechanisms align with their goals and challenges.

By fostering a culture of innovation, collaboration, and continuous improvement, this project seeks to empower decision-makers and data professionals with the tools they need to leverage data for informed and agile decision-making. The user-centric approach and design thinking principles embedded in each stage will contribute to the creation of a data warehousing solution that drives value, efficiency, and a data-driven culture within the organization.

Through these efforts, we aim to transform data into a strategic asset, facilitating better decision-making and enabling the organization to stay competitive and responsive in today's data-driven world.

THANK

YOU