

Assignment 02

Making Sense of the Market

Submitted by:

Jaya Varshini Prabakar

SUID: 463948991

NetID: jprabaka

Part 03:

For each of the offerings, answer the following questions:

1. Website?
2. What does it do?
3. Is it a product (hardware/software), service (we do the work for you), or both?
4. Is it installed onsite, or is it in the cloud?
5. Is it systems or technical architecture?
 - a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce?
 - b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on

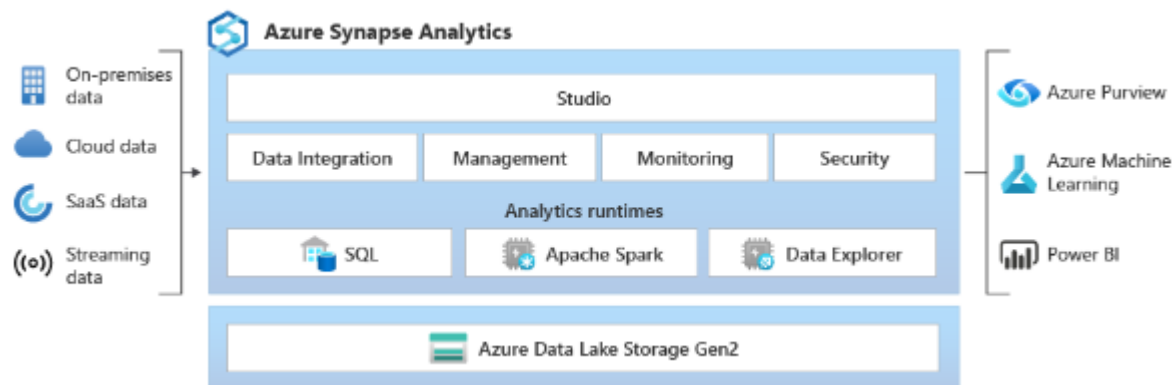
Offerings

1. Azure Synapse
2. Birst
3. dbt Labs
4. Fivetran
5. MemSQL
6. Oracle Essbase
7. Pentaho Data Integration
8. PostgreSQL
9. Reltio
10. Snowflake.com

1. Azure Synapse:

Brief Definition:

Azure Synapse is a cloud-based analytics service provided by Microsoft Azure. It provides services to ingest, prepare, manage, and serve data for business intelligence and advanced analytics purpose. This platform combines data warehouse, big data processing and data integration into one.



Azure Synapse Architecture

Source for architecture: <https://learn.microsoft.com/en-us/azure/synapse-analytics/overview-what-is>

Questions:

1. Website? <https://web.azuresynapse.net/>
2. What does it do? From the architecture diagram we can see that azure synapse analytics is used for integration of data from different sources systems and implements analytics which is later visualized.
3. Is it a product (hardware/software), service (we do the work for you), or both? It is a **cloud-based service** offered by Microsoft Azure.
4. Is it installed onsite, or is it in the cloud? It is offered in **Cloud**.
5. Is it systems or technical architecture? It is a **System Architecture**
- 5 a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **MMP**
- 5 b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so, on

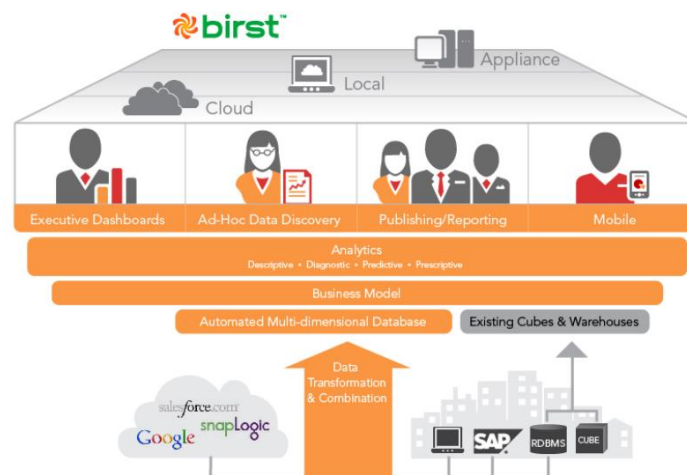
The Azure Synapse Architecture also address the following parts of CIF:

Data Mart Creation, Data integration using **ETL**, **Cross Media Storage**.

2. Birst:

Brief Definition:

Birst is a cloud-based business intelligence tool used for reporting analysis purpose that will connect the entire enterprise through virtualized BI instances in a common base of analysis. It also serves as an automation tool which automates process of preparing data and adds an adaptive layer that works across any device.



Birst Architecture

Source of Birst Framework: https://www.informatica.com/content/informatica-marketplace/en_us/listings/on-premise/solutions/birst_agile_business_analytics.html

Questions:

1. Website? <https://www.infor.com/solutions/advanced-analytics/business-intelligence/birst>

2. What does it do? **Analytical reporting service** on cloud

3. Is it a product (hardware/software), service (we do the work for you), or both?

It is a **product** and also a **SaaS** (Software as Service)

4. Is it installed onsite, or is it in the cloud? It is in **Cloud** and can be also **installed** in company's internal cloud architecture.

5. Is it systems or technical architecture? It is a **Technical Architecture**.

a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **N/A**

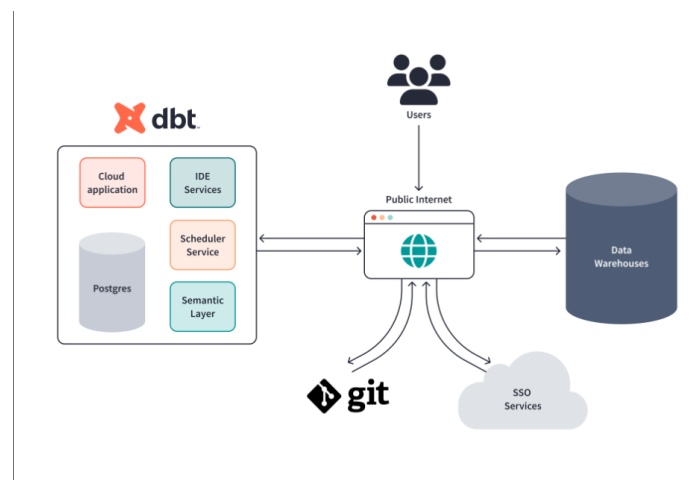
b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so, on

The birst address some parts of CIF which are:

1. **ODS**(since it stores reports in a centralized repo and also it has a part of internal access and also user friendly)
2. **DDS**(supports ad hoc query from end user)
3. **ETL** (since it extracts or prepares data automatically from various sources)

3. dbt Labs:

Brief Definition: **dbt Labs** (data build tools) Labs is used for deploying and maintaining dbt. **dbt** is used to create a clear picture about data pipelines, how the data is produced and also maps business logic by creating a lineage graph.



dbt Cloud Architecture

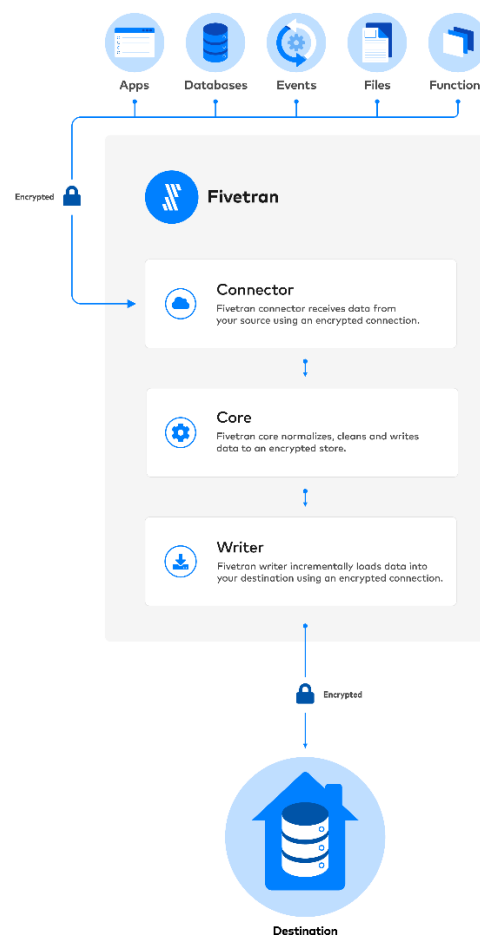
Source for dbt Cloud Architecture: <https://docs.getdbt.com/docs/cloud/about-cloud/architecture>

Questions:

1. Website? <https://www.getdbt.com/>
2. What does it do? dbt Labs maintains and develops dbt and that dbt is helping for visualizing data pipelines and for transforming data.
3. Is it a product (hardware/software), service (we do the work for you), or both?
dbt Labs is a **cloud-based service and can be also used as software product.**
4. Is it installed onsite, or is it in the cloud? dbt Labs offer two services 1. dbt Core which is onsite infrastructure. 2. dbt cloud is a cloud based platform.
5. Is it systems or technical architecture? It is a **Technical Architecture.**
 - a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **N/A**
 - b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on
 1. ODS/ DDS
 2. External world application

4. Fivetran:

Brief Definition: Fivetran is an automated data movement platform used for moving data in, out and across cloud locations. Fivetran has automated ETL process thus providing more space for pipeline construction.



Fivetran Architecture

Source of Fivetran Architecture: <https://fivetran.com/docs/getting-started/architecture>

Questions:

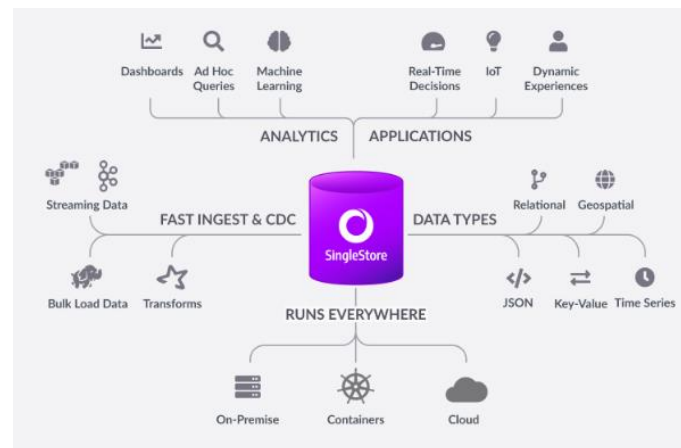
1. Website? <https://www.fivetran.com/>
2. What does it do? **Acts as a connector between cloud platforms for data movements.**
3. Is it a product (hardware/software), service (we do the work for you), or both? It is a **service offering**.
4. Is it installed onsite, or is it in the cloud? It is in **Cloud**
5. Is it systems or technical architecture? Fivetran is a **technical architecture**.
 - a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **N/a**

b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on

The main area of CIF that Fivetran address is **IM & T/ ETL** , it integrates data from various source and acts as a extraction, transformation and load through schema changes and incremental updates.

5. MemSQL:

Brief Definition: MemSQL is currently known as SingleStoreDB. It is a **distributed, relational database** which handles both transaction and real time analytics at scale. It scales **horizontally** on cloud instance or in enterprise hardware thus providing throughput across wide range of platforms.



SingleStoreDB Architecture

Source of SingleStore Architecture: <https://docs.singlestore.com/db/v8.1/introduction/how-singlestoredb-works/>

Questions:

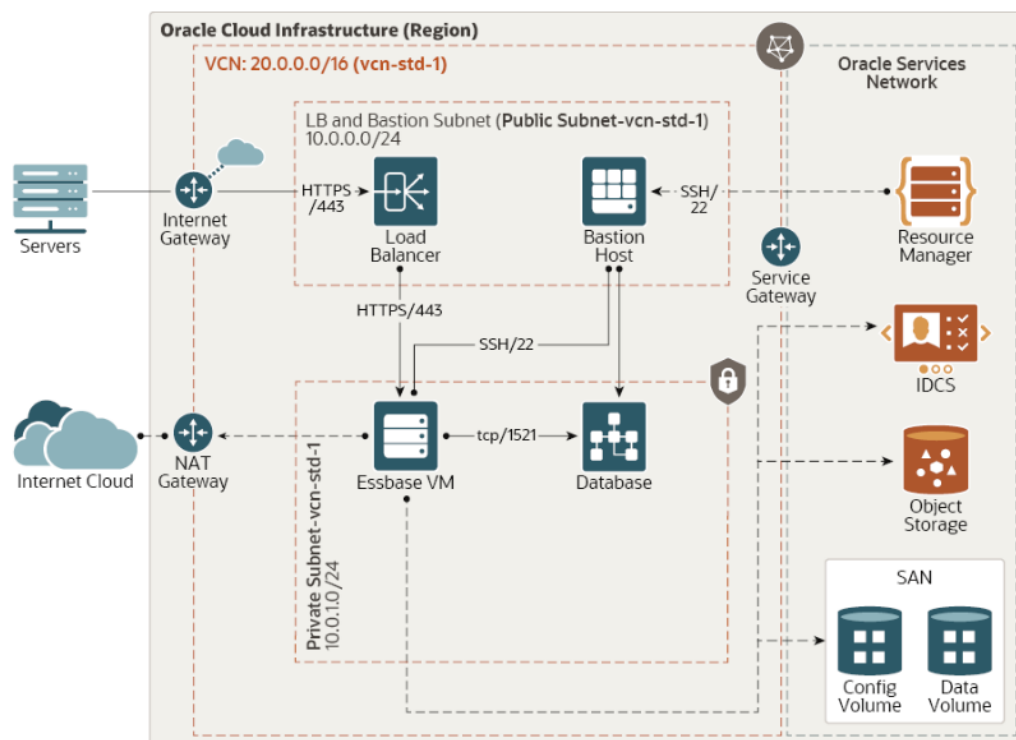
1. Website? <https://www.singlestore.com/>
2. What does it do? It is a distributed relational Database that do **Real time analytics at scale**.
3. Is it a product (hardware/software), service (we do the work for you), or both? It's a **Product**.
4. Is it installed onsite, or is it in the cloud? **It can be install on premise and in the cloud as well**.
5. Is it systems or technical architecture? It is a **Technical Architecture**.
 - a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **N/A**
 - b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on

From the architecture SingleStoreDB(MemSQL) address the following concepts of CIF:

1. **Data Mart**
2. **External World Application**(supports OLTP and real time analytics)
3. **DSS Application**

6. Oracle Essbase:

Brief Definition: Oracle Essbase is a **Business Analytical solution** that uses flexible, best-in-class architecture for analysis, reporting and calculation. It is a multi threaded OLAP database software which uses symmetric multiprocessing.



Essbase Architecture

Source for diagram: <https://docs.oracle.com/en/solutions/essbase-deployment-in-cloud/index.html>

Questions:

1. Website? <https://www.oracle.com/business-analytics/essbase.html>

2. What does it do?

Gives flexibility to work with a wide range of data and also helps to design, manage and scale in cloud.

3. Is it a product (hardware/software), service (we do the work for you), or both?

It is a **Software Product**.

4. Is it installed onsite, or is it in the cloud? It can be **installed on premise** and also can be used in **cloud environment**.

5. Is it systems or technical architecture? It is a **technical architecture**.

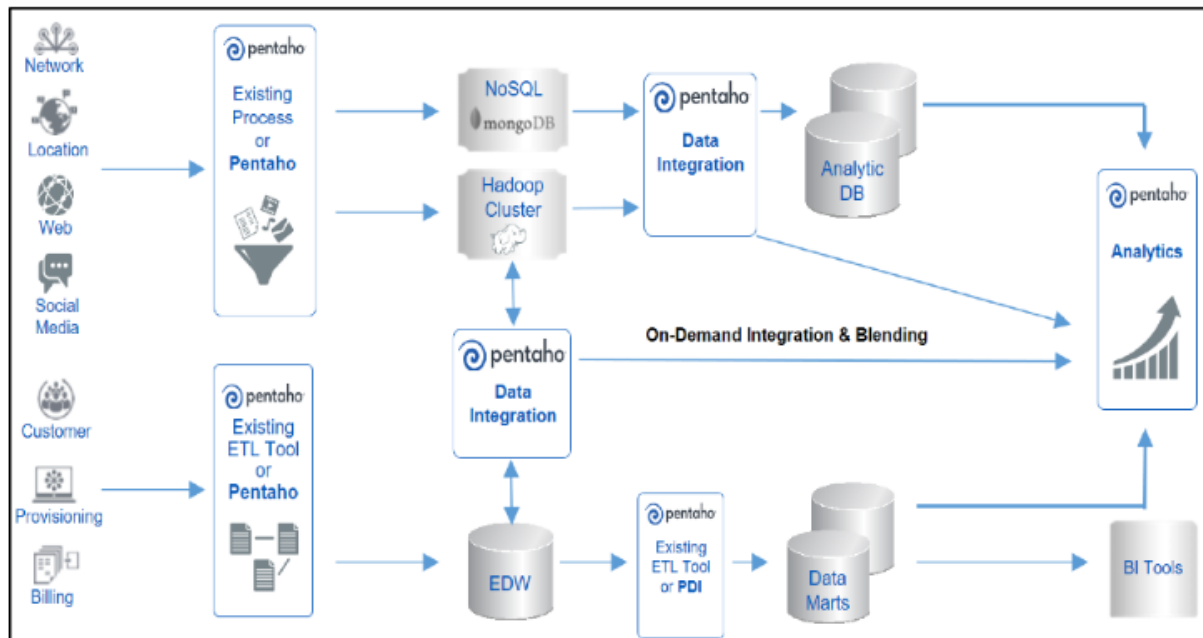
a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **N/A**

b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on

The main component of CIF, Oracle Essbase address is: **ETL process and Cross media storage**.

7. Pentaho Data Integration:

Brief Definition: Pentaho Data Integration provides Extract, Transform and Load (ETL) capabilities that facilitate the process of capturing, cleansing and storing data using uniform and consistent format that is used by end users and IoT technologies.



Pentaho Data Integration Use Case

Source of Use Case: <https://dwbi.org/pages/173>

Questions:

1. Website? <https://www.hitachivantara.com/en-us/products/pentaho-platform/data-integration-analytics.html>

2. What does it do?

Blending multiple data source into one platform. Delivers real world analytic capabilities by consuming integration job as a single data source.

3. Is it a product (hardware/software), service (we do the work for you), or both?

It is a **Software product**.

4. Is it installed onsite, or is it in the cloud?

Pentaho can be installed in both on-premise and also in cloud.

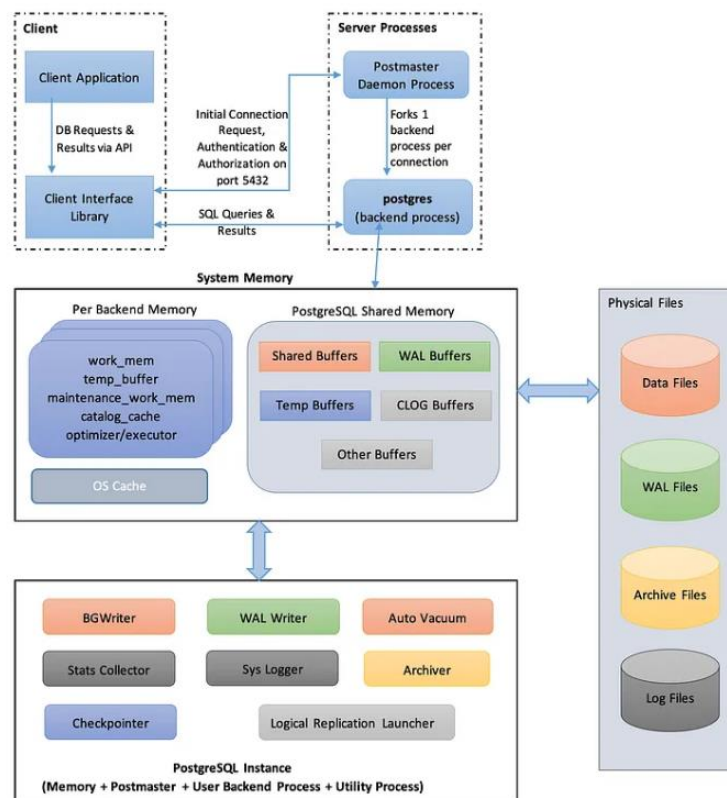
5. Is it systems or technical architecture? It is a **System Architecture**.

a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **Supports both SMP and MMP**

b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on **N/A**

8. PostgreSQL:

Brief Definition: It is a relational database management system which is designed to handle a range of workloads, from single machines to data warehouses or Web Services. It is free and open-source RDBMS.



PostgreSQL Architecture

Source of Architecture: <https://medium.com/@sumeet.k.shukla/postgresql-architecture-6df259dc1145>

Questions:

1. Website? <https://www.postgresql.org/>
2. What does it do? It is a **Relational Database Management System**
3. Is it a product (hardware/software), service (we do the work for you), or both?

It is a **Product**

4. Is it installed onsite, or is it in the cloud?

It can be installed **on-premise and also be used in cloud**

5. Is it systems or technical architecture? It is a **Technical Architecture**.

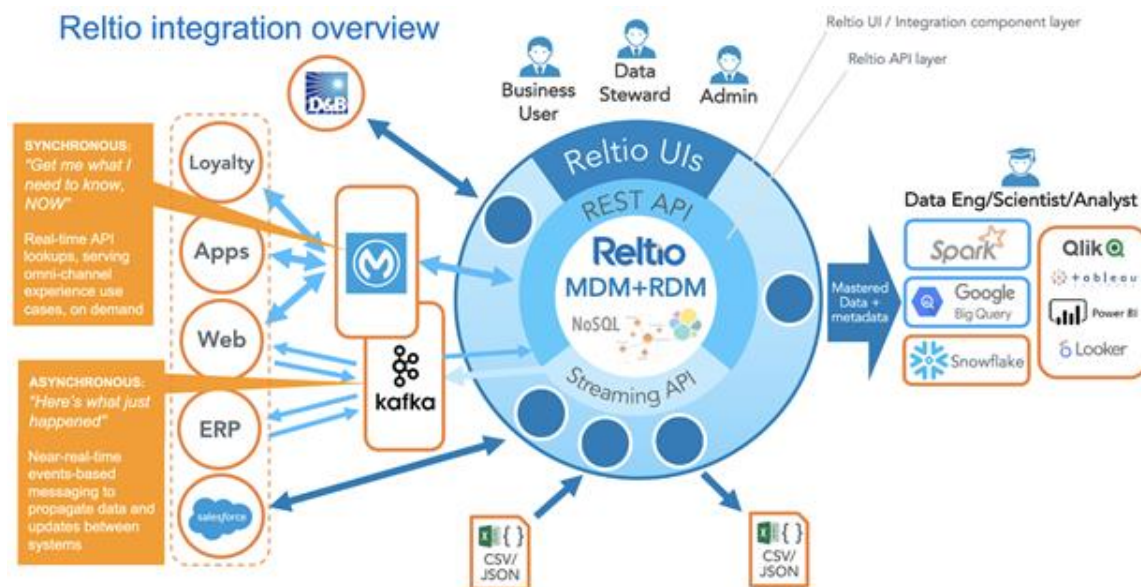
a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **N/A**

b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on

The main components of CIF, which PostgreSQL address are **Data mart, ODS,NDS,DDS. ETL and Cross media storage.**

9. Reltio:

Brief Definition: Reltio is a cloud based Data Management platform which helps to create a unified view of original master data. It acts as a centralized hub for managing the master data.



Reltio Overview

Source: <https://community.reltio.com/blogs/matthew-gagan/2021/07/17/api-first-integration-in-master-data-management-an>

Question:

1. Website? <https://www.reltio.com/>

2. What does it do? It is a cloud native Master Data Management platform that consolidates millions of profile from hundreds of data sources.

3. Is it a product (hardware/software), service (we do the work for you), or both?

It is a **Software Product**

4. Is it installed onsite, or is it in the cloud?

It is a cloud based service, which operates fully on cloud.

5. Is it systems or technical architecture?

It is a **Technical Architecture**

a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? **N/A**

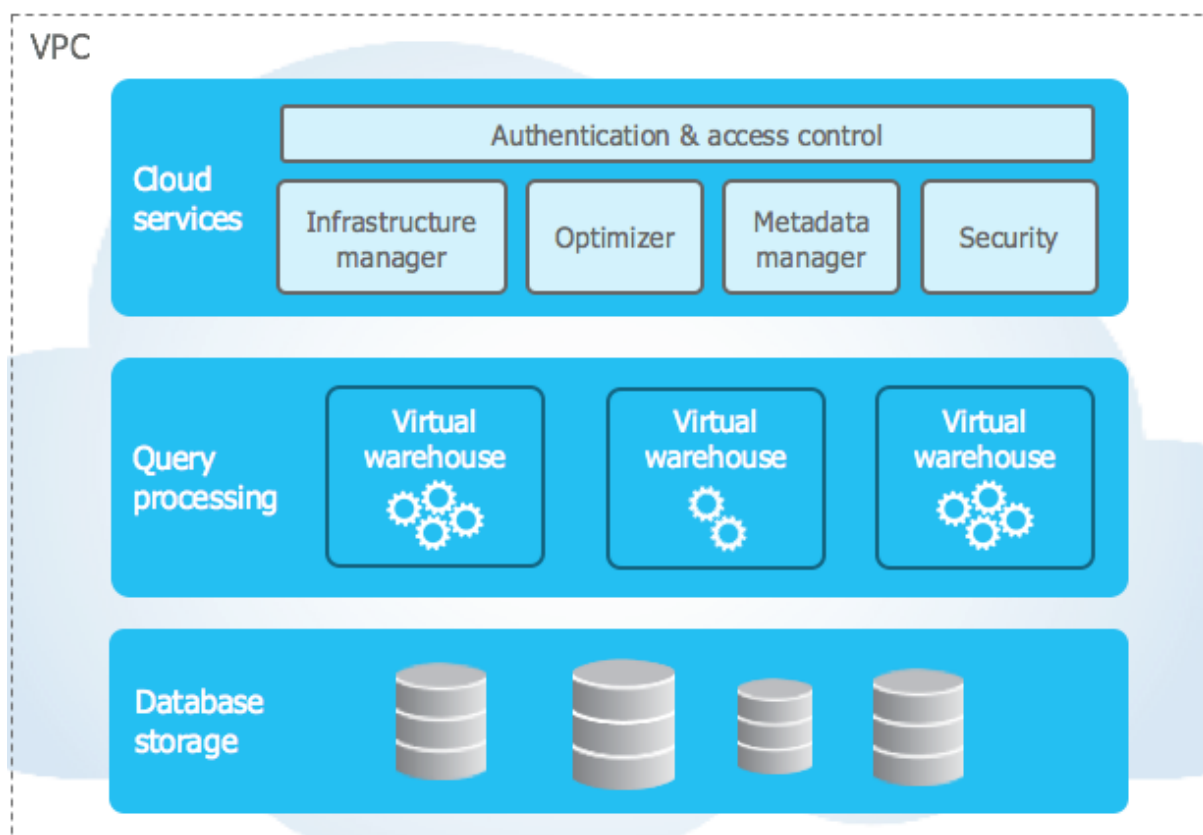
b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart? IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on

It address the following in CIF, it serves as a centralized technical architecture

1. Data mart
2. DDS

10. Snowflake.com

Brief Definition: Snowflake enables data processing and analytic solutions that are faster than traditional offerings. It acts as enterprise analytic database along with additional special features.



Snowflake Architecture

Source: <https://docs.snowflake.com/en/user-guide/intro-key-concepts>

Questions:

1. Website? <https://www.snowflake.com/en/>
2. What does it do? It provides organisations with fully managed, scalable and flexible solution for data storing, processing and analysing.
3. Is it a product (hardware/software), service (we do the work for you), or both?

It is a cloud based service.

4. Is it installed onsite, or is it in the cloud?

It operates entirely on cloud.

5. Is it systems or technical architecture? It is a systems architecture

a. If systems architecture, is it SMP, MPP, or Hadoop/MapReduce? It uses MMP

b. If technical architecture, which part of the corporate information factory does it aim to address? Data mart?
IM&T/ETL? ODS/NDS/DDS? External world application? Crossmedia storage? And so on **N/A**