Project Design Phase-II
Technology Stack (Architecture & Stack)

Date	15 October 2022
Team ID	PNT2022TMID51838
Project Name	Project -GLOBAL SALES DATA ANALYTICS
Maximum Marks	4 Marks

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

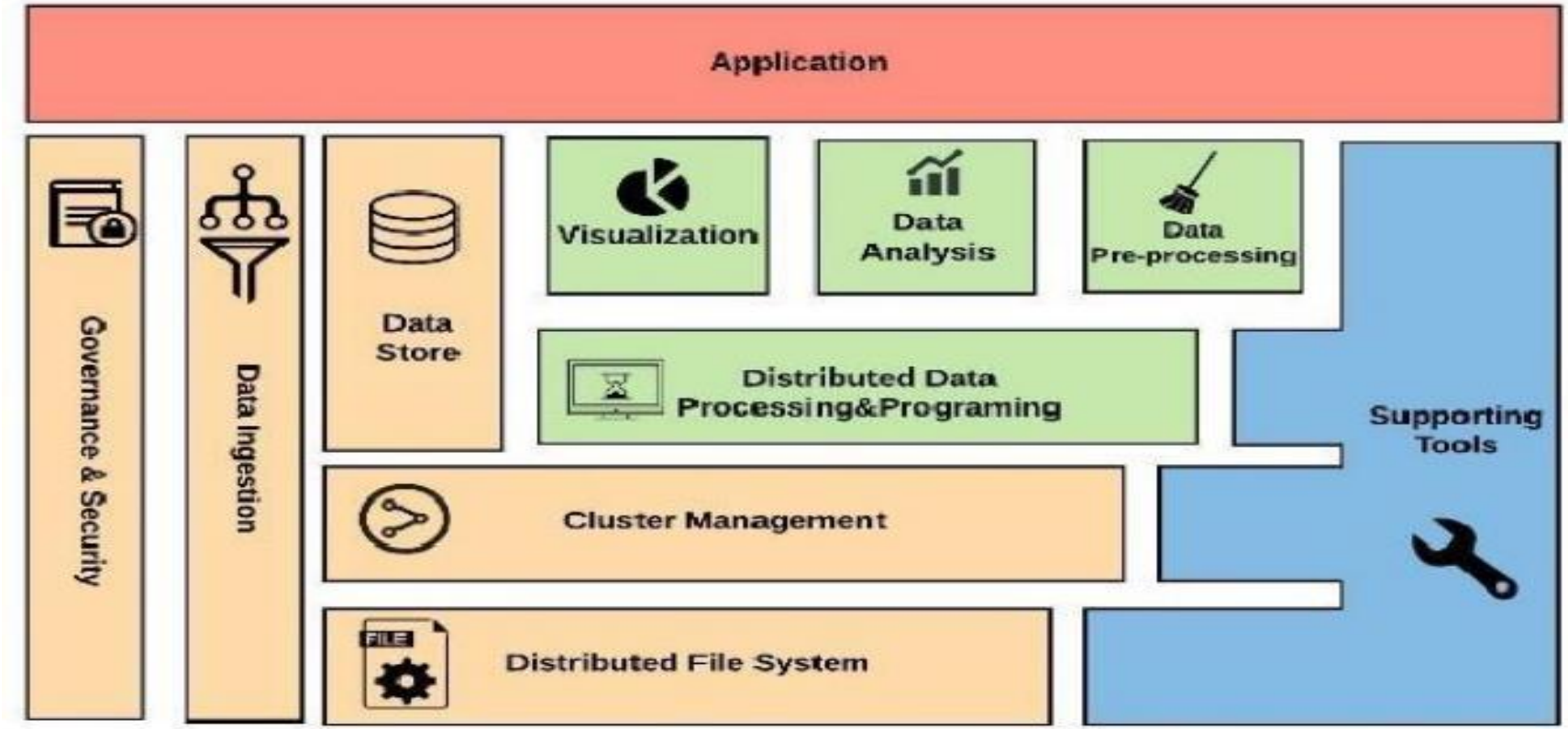

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Application Logic-1	Logic for a process in the application	Java / Python
3.	Application Logic-2	Logic for a process in the application	IBM Watson STT service
4.	Application Logic-3	Logic for a process in the application	IBM Watson Assistant
5.	Database	Data Type, Configurations etc.	MySQL, NoSQL, etc.
6.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
7.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
8.	External API-1	Purpose of External API used in the application	IBM Weather API, etc.
9.	External API-2	Purpose of External API used in the application	Aadhar API, etc.
10.	Machine Learning Model	Purpose of Machine Learning Model	Object Recognition Model, etc.
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	Technology of Opensource framework
2.	Security Implementations	List all the security / access controls implemented, use of firewalls etc.	e.g. SHA-256, Encryptions, IAM Controls, OWASP etc.
3.	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	Technology used
4.	Availability	Justify the availability of application (e.g. use of load balancers, distributed servers etc.)	Technology used
5.	Performance	Design consideration for the performance of the application (number of requests per sec, use of Cache, use of CDN's) etc.	Technology used

Security governance:

Security governance is a process for overseeing the cybersecurity teams who are responsible for mitigating business risks. Security governance leaders make the decisions that allow risks to be prioritized so that security efforts are focused on business priorities rather than their own. Security governance is the means by which you control and direct your organisation's approach to security. When done well, security governance will effectively coordinate the security activities of your organisation. It enables the flow of security information and decisions around your organisation. Included in this definition are three terms that are generally regarded as the high-level security objectives – integrity, availability, and confidentiality.

Data ingestion:

Data ingestion refers to importing data to store in a database for immediate use, and it can be either streaming or batch data. With an increasing number of data sources and types, businesses are challenged with ingesting and processing data fast enough to support business goals. Move data from Salesforce.com to a data warehouse then analyse with Tableau. Capture data from a Twitter feed for real-time sentiment analysis. Acquire data for training machine learning models and experimentation. Data ingestion restructures company data to predetermined formats and makes it easier to utilize, particularly when combined with extract, transform, and load (ETL) operations.

Data storage:

A storage device refers to a computing hardware used to store information permanently or temporarily. The device can be external or internal to a computer, server, and other computing systems. Storage devices are also known as storage medias or storage medium.

Data visualization:

Data visualization is the graphical representation of data using visual elements such as graphs, charts, maps, etc. It transforms your complex raw data into an easily digestible visualization designed to identify trends, patterns, relationships, outliers and more.

Data visualizations tap into the fact that we respond to visual information quicker than textual information. Presenting data in a visual format can empower sales teams to make faster interpretations and deliver timely results.

Data analysis:

Regular sales data analysis provides an understanding of the products that your customers are buying and helps you dissect why they are behaving in a certain way. You can also find patterns in your lead conversions and drop offs. All of these aspects enable you to optimize your sales process.

Data pre-processing:

Data pre-processing is the process of transforming raw data into a useful, understandable format. Real-world or raw data usually has inconsistent formatting, human errors, and can also be incomplete. Data pre-processing resolves such issues and makes datasets more complete and efficient to perform data analysis.

Data Distribution:

Data Distribution are the top four factors that positively impact user satisfaction for Big Data Processing and Distribution products. These factors are determined by an algorithm that selects the attributes that are most likely to predict user satisfaction within this category.

Distributed file system:

A distributed file system (DFS) is a file system that spans across multiple file servers or multiple locations, such as file servers that are situated in different physical places. Files are accessible just as if they were stored locally, from any device and from anywhere on the network. A DFS makes it convenient to share information and files among users on a network in a controlled and authorized way.

DATA STORE TYPE:

4 Types of Computer Data Storage.

- Computer Data Storage #1: Cloud Storage.
- Computer Data Storage #2: Cloud Backup.
- Computer Data Storage #3: USB Flash Drive.
- Computer Data Storage #4: Optical Media Storage.