ARCHITECTURE
ANALYZING OF GOOGLE APP STORE

Revision Number: 1.0

Last date of Revision: 13/09/2021

Submitted by:

Dattatreya Thunuguntla

Contents:

S. No	Topic	Page No
1	Introduction	3
1.1	What Is Architecture Design Document	3
1.2	Scope	3
2	Architecture	4
2.1	Data Models	5
2.2	Data Warehousing	6
2.3	Data Sources	7
2.4	Data Ingestion	7
2.5	Data Consumption	8
3	Deployment	8

1. INTRODUCTION:

1.1 WHAT IS ARCHITECTURE DESIGN DOCUMENT

Any software needs the architectural design to represents the design of software. IEEE defines architectural design as “the process of defining a collection of hardware and software components and their interfaces to establish the framework for the development of a computer system.” The software that is built for computer-based systems can exhibit one of these many architectures.

Each style will describe a system category that consists of:

• A set of components (Ex: a database, computational modules) that will perform a function required by the system.

• The set of connectors will help in coordination, communication, and cooperation between the components.

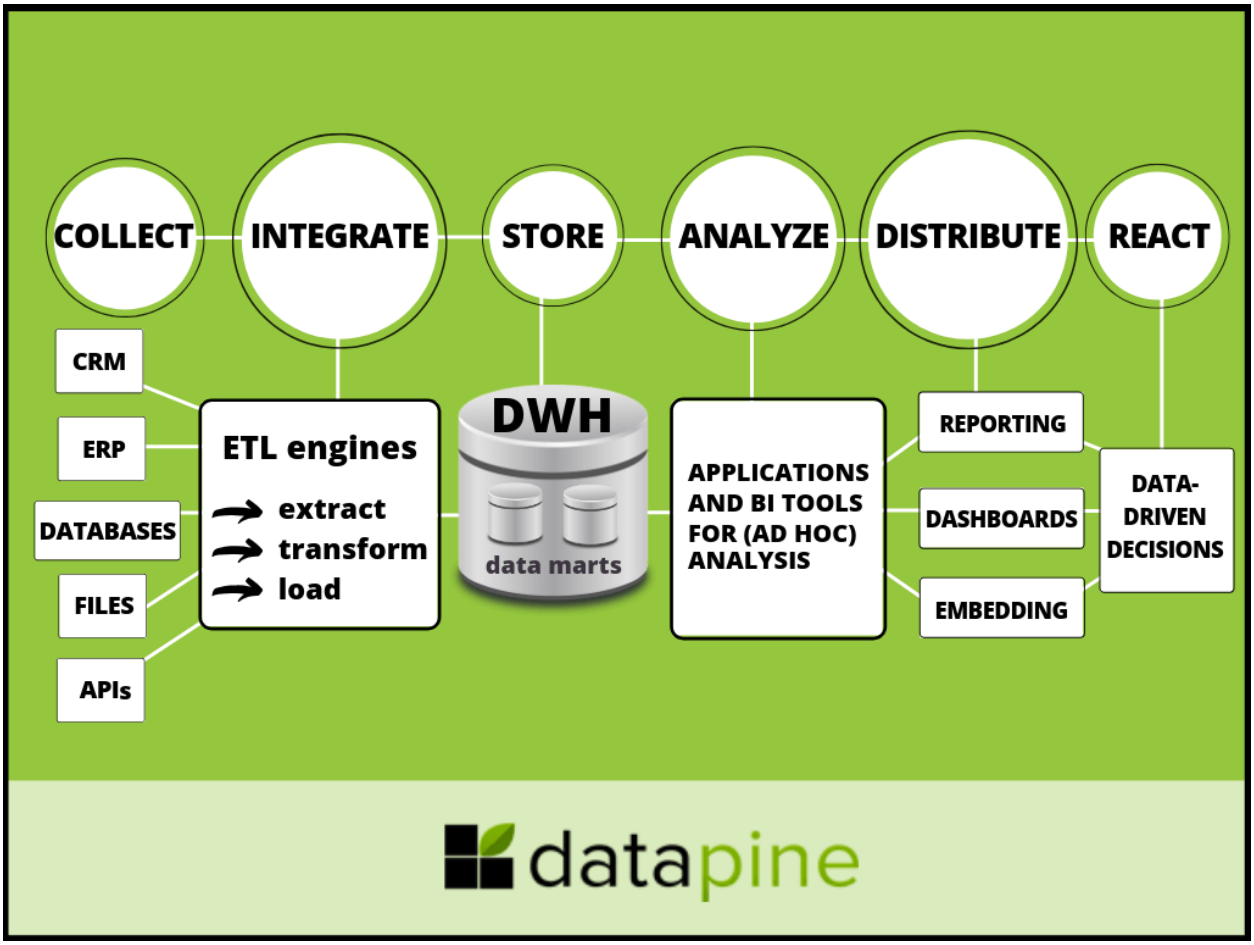
• Conditions that how components can be integrated to form the system.

• Semantic models that help the designer to understand the overall properties of the system.

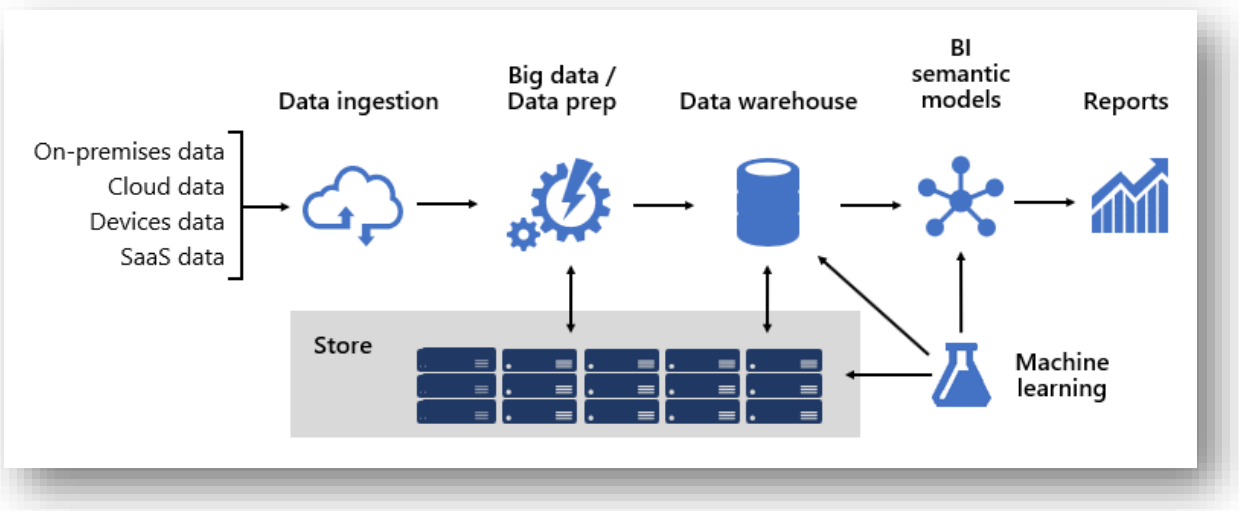
1.2 SCOPE

Architecture Design Document (ADD) is an architecture design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the design principles may be defined during requirement analysis and then refined during architectural design work.

2. ARCHITECTURE:



POWER BI SOLUTION ARCHITECTURE:



Designing a robust BI platform is somewhat like building a bridge; a bridge that connects transformed and enriched source data to data consumers. The design of such a complex structure requires an engineering mindset, though it can be one of the most creative and rewarding IT architectures you could design. In a large organization, a BI solution architecture can consist of:

- Data source
- Data ingestion
- Big data / data preparation
- Data warehouse
- BI semantic models
- Reports

The platform must support specific demands. Specifically, it must scale and perform to meet the expectations of business services and data consumers. At the same time, it must be secure from the ground up. And, it must be sufficiently resilient to adapt to change—because it's a certainty that in time new data and subject areas must be brought online.

2.1 Data Models:

Data models provide you with control over how data is structured and accessed. To business services and data consumers, data models are their interface with the BI platform.

A BI platform can deliver three different types of models:

- Enterprise models
- BI semantic models
- Machine Learning (ML) models

1. Enterprise models:

Enterprise models are built and maintained by IT architects. They're sometimes referred to as dimensional models or data marts. Typically, data is stored in relational format as dimension and fact tables. These tables store cleansed and enriched data consolidated from many systems and they represent an authoritative source for reporting and analytics. In a cloud BI platform, enterprise models can be deployed to a Synapse SQL pool in Azure Synapse. The Synapse SQL pool then becomes the single version of truth the organization can count on for fast and robust insights.

2. BI Semantic Models:

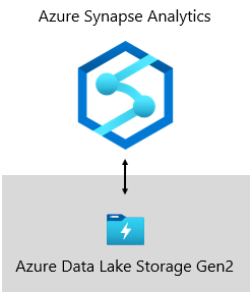
BI semantic models represent a semantic layer over enterprise models. They're built and maintained by BI developers and business users. BI developers create core BI semantic models that source data from enterprise models. Business users can create smaller-scale, independent models—or, they can extend core BI semantic models with departmental or external sources. BI semantic models commonly focus on a single subject area, and are often widely shared. In a cloud BI platform, BI developers can deploy BI semantic models to Azure Analysis Services or Power BI Premium capacities.

3. Machine Learning Models:

Machine Learning (ML) models are built and maintained by data scientists. They're mostly developed from raw sources in the data lake. Trained ML models can reveal patterns within your data. In many circumstances, those patterns can be used to make predictions that can be used to enrich data. For example, purchasing behavior can be used to predict customer churn or segment customers. Prediction results can be added to enterprise models to allow analysis by customer segment. In a cloud BI platform, you can use Azure Machine Learning to train, deploy, automate, manage, and track ML models.

2.2 Data Warehousing:

Sitting at the heart of a BI platform is the data warehouse, which hosts your enterprise models. It's a source of sanctioned data—as a system of record and as a hub—serving enterprise models for reporting, BI, and data science. Many business services, including line-of-business (LOB) applications, can rely upon the data warehouse as an authoritative and governed source of enterprise knowledge. At Microsoft, our data warehouse is hosted on Azure Data Lake Storage Gen2 (ADLS Gen2) and Azure Synapse Analytics.



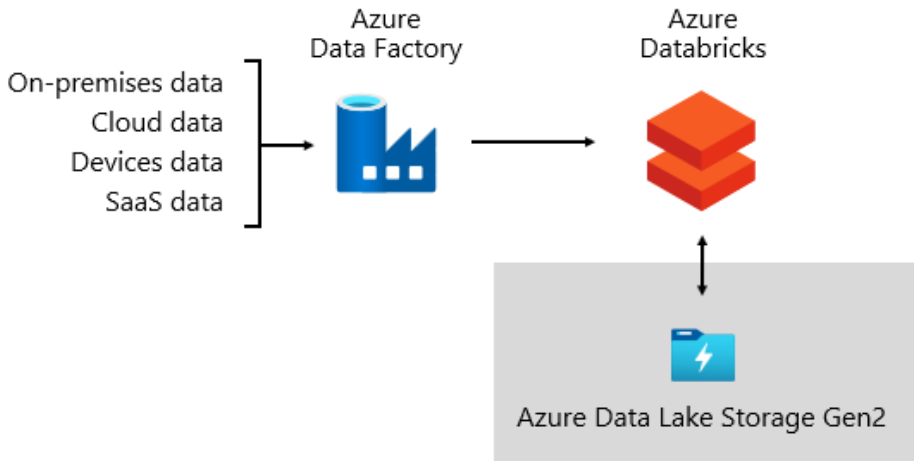
2.3 Data Sources:

A data warehouse can consolidate data from practically any data source. It's mostly built over LOB data sources, which are commonly relational databases storing subject-specific data for sales, marketing, finance, etc. These databases can be cloud-hosted or they can reside on-premises. Other data sources can be file-based, especially web logs or IOT data sourced from devices. What's more, data can be sourced from Software-as-a-Service (SaaS) vendors.

At Microsoft, some of our internal systems output operational data direct to ADLS Gen2 using raw file formats. In addition to our data lake, other source systems comprise relational LOB applications, Excel workbooks, other file-based sources, and Master Data Management (MDM) and custom data repositories. MDM repositories allow us to manage our master data to ensure authoritative, standardized, and validated versions of data.

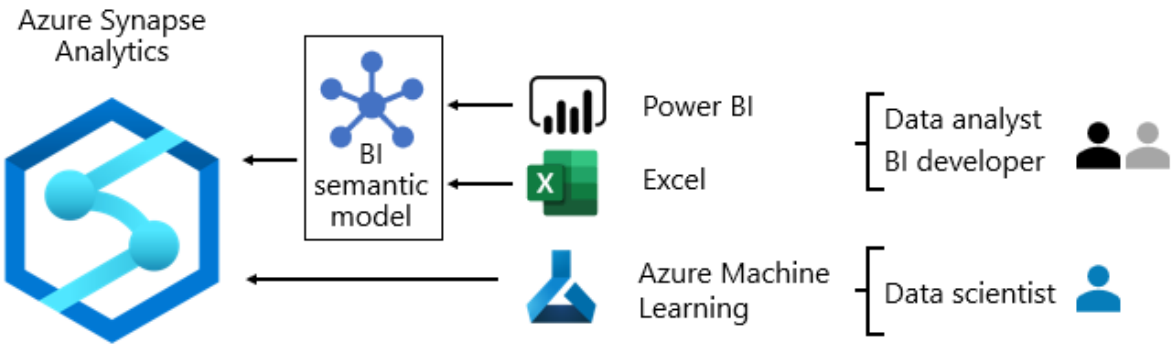
2.4 Data Ingestion:

On a periodic basis, and according to the rhythms of the business, data is ingested from source systems and loaded into the data warehouse. It could be once a day or at more frequent intervals. Data ingestion is concerned with extracting, transforming, and loading data. Or, perhaps the other way round: extracting, loading, and then transforming data. The difference comes down to where the transformation takes place. Transformations are applied to cleanse, conform, integrate, and standardize data. For more information, see Extract, transform, and load (ETL). Ultimately, the goal is to load the right data into your enterprise model as quickly and efficiently as possible.



2.5 Data Consumption:

At the reporting layer, business services consume enterprise data sourced from the data warehouse. They also access data directly in the data lake for ad hoc analysis or data science tasks. Fine-grained permissions are enforced at all layers: in the data lake, enterprise models, and BI semantic models. The permissions ensure data consumers can only see the data they have rights to access. At Microsoft, we use Power BI reports and dashboards, and Power BI paginated reports. Some reporting and ad hoc analysis is done in Excel—particularly for financial reporting. We publish data dictionaries, which provide reference information about our data models. They're made available to our users so they can discover information about our BI platform. Dictionaries document model designs, providing descriptions about entities, formats, structure, data lineage, relationships, and calculations.



3. Deployment:

After generating the reports, we can deploy it in Power BI Service or else other Web Applications, Web Sites via APIs