ARCHITECTURE DESIGN

ANALYZING AMAZON SALES DATA

Revision Number: 1.0
Last date of revision: 03/08/2023

Chaitanya Nilkanthanawar

iNeuron

DOCUMENT CONTROL

Change Record:

VERSION	DATE	AUTHOR	COMMENTS
1.0	03 Aug 2023	Chaitanya Nilkanthanawar	we present a comprehensive overview that encompasses the introduction and definition of the architecture. Furthermore, we delve into the detailed architecture description, followed by the inclusion of recent updates and enhancements to the architecture.

Reviews:

VERSION	DATE	AUTHOR	COMMENTS
1.0	03 Aug 2023	Chaitanya Nilkanthanawar	Unit test cases need to be incorporated.

Approval Status:

VERSION	REVIEW DATE	REVIEWED BY	APPROVED BY	COMMENTS

iNeuron

CONTENTS

iNeuron

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 (e.g.: 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.

iNeuron

2. Architecture

Power BI is a comprehensive business suite encompassing a synergistic ensemble of technologies, meticulously orchestrated to deliver exemplary business intelligence solutions. At its core, the Microsoft Power BI technology amalgam comprises an array of integral components, each playing a pivotal role in the data journey. These components include:

- POWER QUERY (For Data Mash-Up And Transformation)
- POWER BI DESKTOP (A Companion Development Tool)
- POWER BI MOBILE (For Android, IOS, Windows Phones)
- POWER PIVOT (For In-Memory Tabular Data Modeling)
- POWER VIEW (For Viewing Data Visualizations)
- POWER MAP (For Visualizing 3D Geo-Spatial Data)
- Power Q&A (For Natural Language Q&A)

In essence, Microsoft Power BI orchestrates this symphony of components, harmonizing their capabilities to compose a holistic business intelligence solution. Each element contributes a distinct melody, culminating in a harmonious and potent composition that empowers businesses to unravel the symphony of their data-driven endeavours.
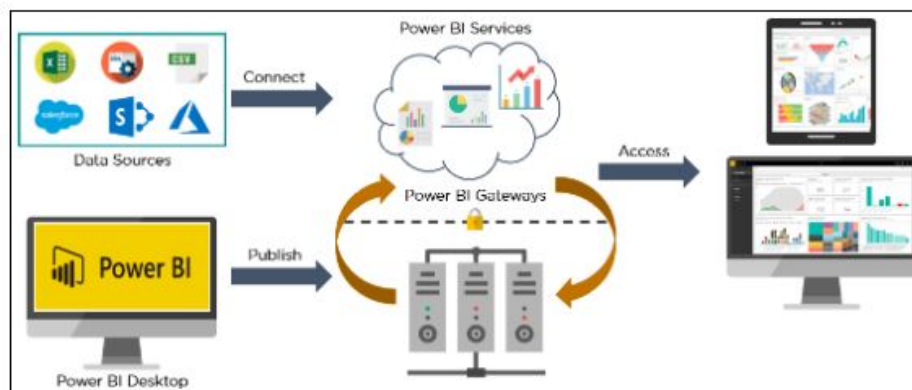
2.1 Data Sources

An integral facet of Power BI lies within its extensive array of data sources. This platform empowers you to seamlessly import data from various origins, including files residing within your local system, data sources hosted on cloud-based platforms, or establish direct connections to real-time data streams. When engaging in data imports, whether from on-premise repositories or online services, it's important to note that there exists a data size threshold of 1 GB. A selection of frequently utilized data sources within Power BI encompasses:

iNeuron

- Excel
- Text/CSV
- XML
- JSON
- Oracle Database
- IBM DB2 Database
- MySQL Database
- PostgreSQL Database
- Sybase Database
- Teradata Database
- SAP HANA Database
- SAP Business Warehouse server
- Amazon Redshift
- Impala
- Google Big Query (Beta)
- Azure SQL Database
- Salesforce Reports
- Google Analytics
- Facebook
- GitHub

2.2 Power BI Desktop

Power BI Desktop stands as a client-side companion development and authoring tool renowned for its comprehensive capabilities. Functioning on the desktop environment, this software is equipped with an extensive array of tools and functionalities, enabling seamless interaction with diverse data sources. Its capabilities span data connectivity, data transformation, robust data modeling, and the creation of insightful reports.

2.3 Power BI Service

The Power BI Service serves as a web-based platform that facilitates seamless sharing of reports crafted using Power BI Desktop, fosters collaborative endeavors among users, and provides tools for dashboard creation. This dynamic platform is accessible through three distinct versions:

- The Free version
- The Pro version
- The Premium version

Recognized under various names such as "Power BI.com," "Power BI Workspace," "Power BI Site," and "Power BI Web Portal," this component extends beyond its basic functions. It encompasses advanced capabilities such as natural language Q&A and alert mechanisms, thereby enhancing its utility and user experience.

iNeuron

2.4 Power BI Report Server

The Power BI Report Server closely parallels the Power BI Service, with the primary distinction lying in its on-premise nature. Unlike the cloud-based Power BI Service, the Power BI Report Server caters to organizations seeking to maintain their reports within their own infrastructure, driven by data security considerations. This localized platform empowers users to develop dashboards and share reports while upholding rigorous security protocols. To leverage the capabilities of the Power BI Report Server, a requisite Power BI Premium license is essential.

2.5 Power BI Gateway

This integral element serves the purpose of establishing connections and accessing on-premise data within highly secure networks. Power BI Gateways find their primary application in organizations that prioritize stringent data security measures. These gateways play a crucial role in extracting such securely held data through protected channels, facilitating its seamless integration into the Power BI platform for comprehensive analysis and reporting functionalities.
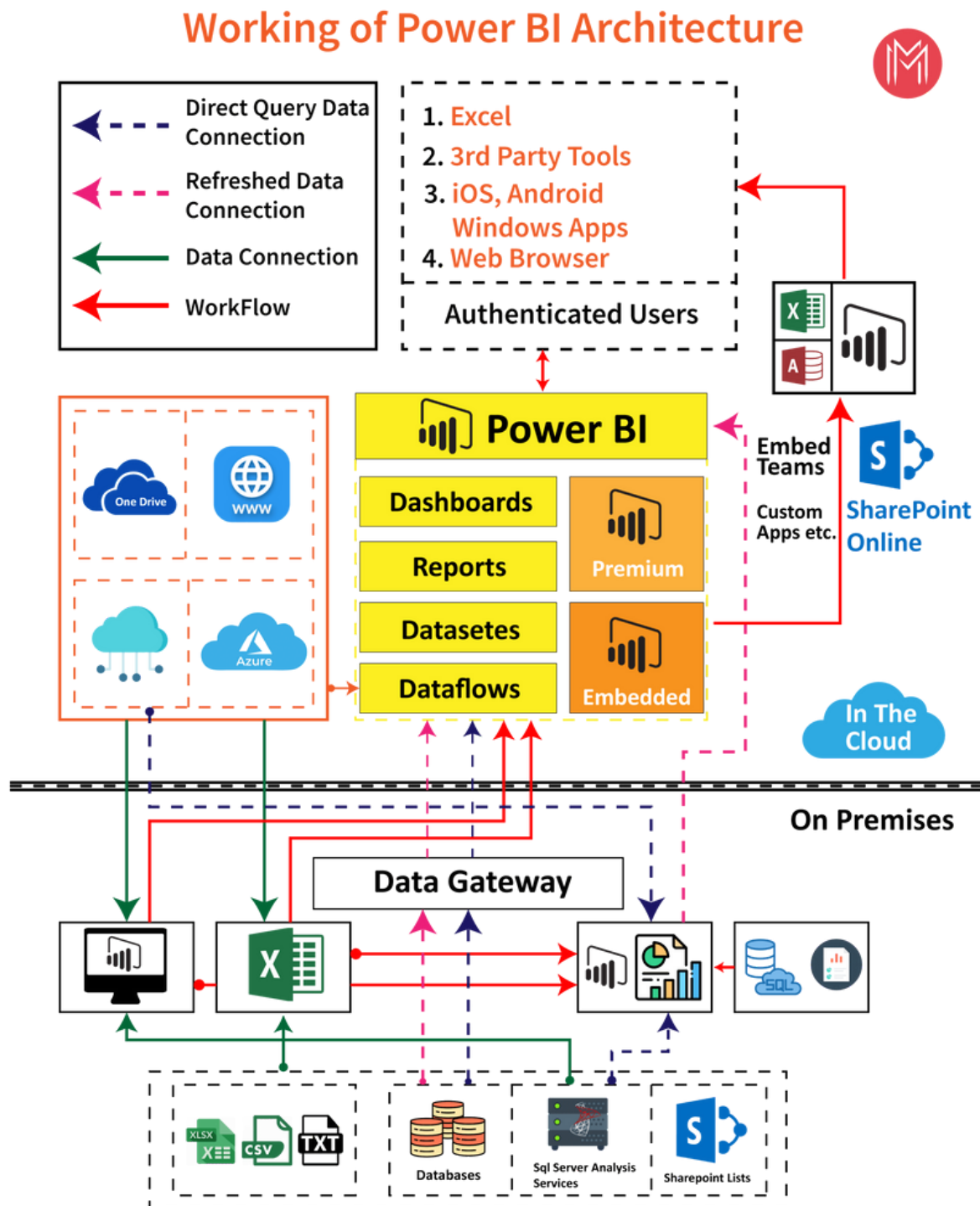
2.6 Power BI Mobile

Power BI Mobile represents a native application designed specifically for iOS, Android, and Windows mobile devices. Its core functionality revolves around the convenient access and visualization of reports and dashboards. This suite of applications serves as the preferred medium for interacting with and gaining insights from Power BI content on the go.

2.7 Power BI Embedded

Power BI Embedded provides a set of Application Programming Interfaces (APIs) that serve the purpose of seamlessly integrating visuals into bespoke and custom applications. This capability empowers developers to incorporate Power BI's interactive data visualizations directly within their own software solutions, enhancing user experiences and enriching applications with data-driven insights.

iNeuron

3. Working of Power BI Architecture

Working of Power BI Architecture

Legend:
- Direct Query Data Connection
- Refreshed Data Connection
- Data Connection
- WorkFlow

Authenticated Users:
1. Excel
2. 3rd Party Tools
3. iOS, Android Windows Apps
4. Web Browser

Power BI:
- Dashboards
- Reports
- Datasetes
- Dataflows
- Premium
- Embedded

Embed Teams, Custom Apps etc.

SharePoint Online

In The Cloud

On Premises

Data Gateway

XLSX, CSV, TXT

Databases | Sql Server Analysis Services | Sharepoint Lists

3.1 On-Premise

Power BI Desktop serves as an essential companion tool for development, authoring, and publishing. It facilitates data import from various sources into Power BI Desktop, enabling the creation of comprehensive reports. Subsequently, these reports can be published on either the Power BI Service or the Power BI Report Server. Additionally, Power BI Publisher for Excel allows for the direct publication of Excel workbooks to the Power BI Report Server. Notably, the SQL Server Data Tools and Report Publisher play pivotal roles in the creation of diverse elements such as datasets, KPIs, mobile reports, and paginated reports. The Power BI Report Server serves as the distribution hub, disseminating reports of varying types to end-users.

3.2 On-Cloud

A cornerstone of the Power BI architecture is the Power BI Gateway, which serves as a secure conduit for transmitting data from on-premise data sources to cloud-based applications or sites. On the cloud side of the architecture, a multitude of components are present. The comprehensive Power BI suite encompasses dataflows, datasets, dashboards, reports, Power BI Embedded, and Power BI Premium, among others. This suite enables versatile embedding of reports and dashboards into platforms such as Teams, SharePoint, and custom applications. Moreover, on-cloud data sources establish connections with Power BI tools through direct connections. At the upper layer, authenticated users play a pivotal role by sharing published reports and dashboards, fostering collaborative decision-making based on insights. Diverse categories of users engage with Power BI reports and dashboards through web browsers, Excel, third-party tools, and mobile devices (iOS, Windows, Android apps).

iNeuron

4. Power BI Service

The availability of reports and dashboards through the Power BI Service extends to a diverse array of client platforms, encompassing websites, mobile devices, and more. Consequently, every individual desiring access to content generated on Power BI is inevitably required to engage with the Power BI Service. Thus, it becomes essential to undertake an exploration of the inner workings of the Power BI Service, unraveling its operational mechanisms. The architecture of the Power BI Service is underpinned by two integral components: a front end and a back end.

4.1 Front End cluster

The front end, referred to as the web front-end cluster, operates as a crucial intermediary bridging the gap between clients and the back end. Within this domain, front end services play a pivotal role by facilitating the establishment of initial connections and client authentication via Azure Active Directory. The repository for user identities lies within Azure Active Directory. Concurrently, the utilization of Azure Traffic Manager proves instrumental in directing user requests towards the closest data center post-authentication. Once successful authentication is achieved for a client or user, the dynamic functioning of the Azure Content Delivery Network (CDN) comes into play, ensuring the distribution of static Power BI content and files to the respective users.

4.2 Back End Cluster

The back-end cluster of Power BI plays a pivotal role in managing user interactions with the Power BI service. This intricate system comprises several essential components. Firstly, Azure API Management serves as the central hub for directing, load balancing, and safeguarding Power BI requests through authentication and authorization mechanisms. The Gateway Role grants access to Power BI's data sources and datasets, enabling seamless connectivity. Power BI Services encompass a suite of functionalities encompassing visualizations, storage, report generation, data connections, and real-time data updates, collectively driving the Power BI experience. The back-end cluster is undeniably integral to the Power BI service, providing swift and dependable access to users' data and reports. Notable attributes of this cluster include its global presence across multiple regions, a composition of multiple servers for handling high user loads, and diligent, ongoing monitoring to guarantee continuous availability.

iNeuron