Architecture Design

Amazon Sales Analysis

Written By	Ankita Gupta, Sanjeev Kumar
Document Version	0.3
Last Revised Date	

DOCUMENT CONTROL

Change Record:

VERSION	DATE	AUTHOR	COMMENTS
0.1	30- July - 2021	Ankita Gupta, Sanjeev Kumar	Introduction and architecture defined
0.2	20 - Aug - 2023	Ankita Gupta, Sanjeev Kumar	Architecture & Architecture description appended and updated.

Reviews:

VERSION	DATE	REVIEWER	COMMENTS
0.2	30-Sept - 2023	Ankita Gupta, Sanjeev Kumar	Unit test cases to be added

Approval Status:

VERSION	REVIEW DATE	REVIEWED BY		APPROVED BY	COMMENTS

Contents

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 (eg: 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.
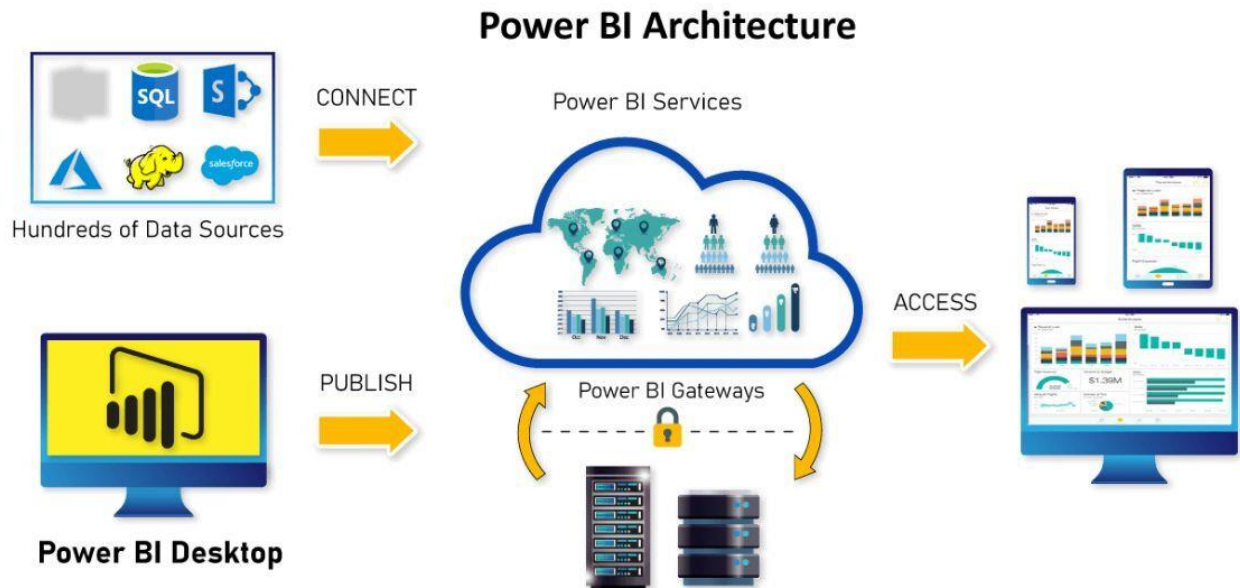
2. Architecture
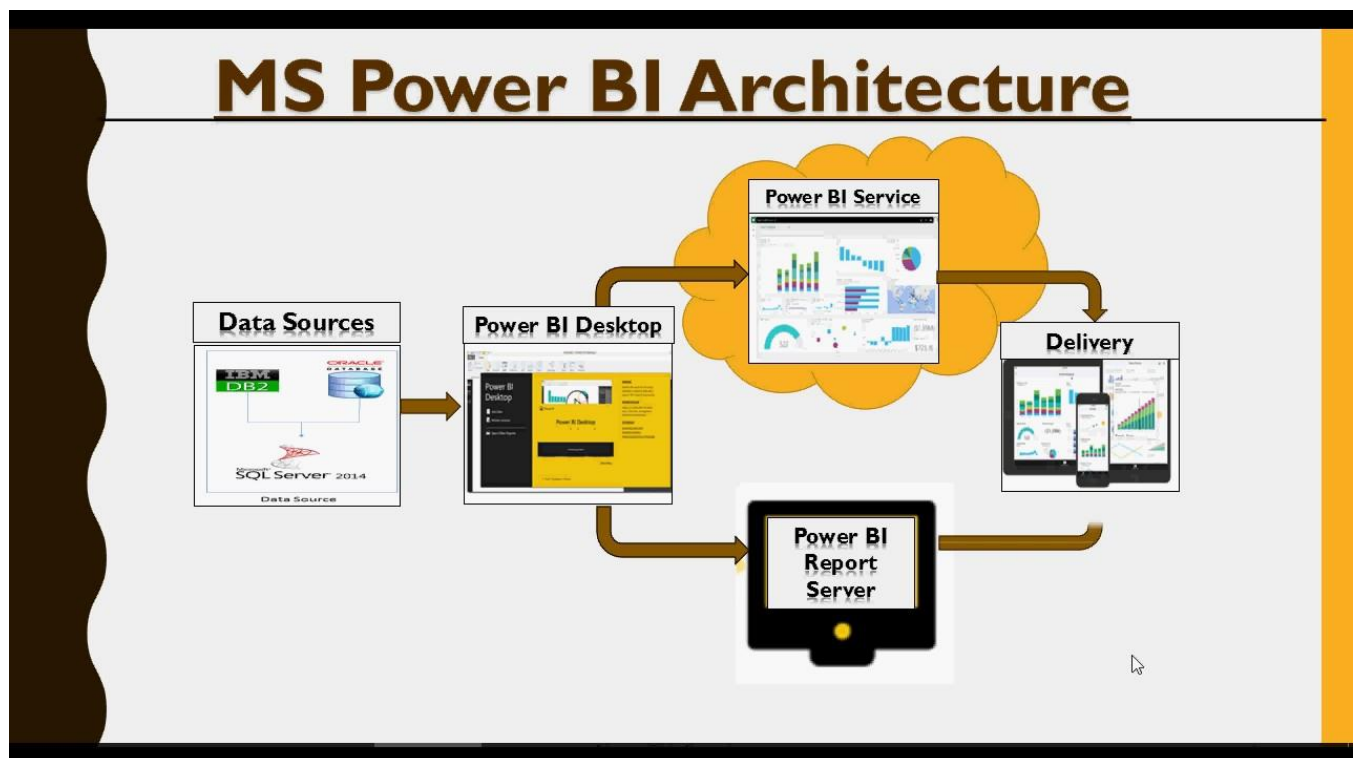
Power BI Server Architecture

In Power BI, the platform also utilizes a scalable n-tier client-server architecture that serves mobile clients, web clients, and desktop-installed software. This architecture ensures flexibility and efficiency in deployments.

The following diagram shows Power BI Server's

Power BI Architecture

PowerBI Server is internally managed by the multiple server processes.

1. Gateway/Load Balancer

It serves as an entry point to the Power BI Server, ensuring efficient access. Additionally, it balances

the load to the server when multiple processes are configured

2) Application Server:-

Application Server processes (e.g., wgserver.exe) manage browsing and permissions for the Power BI Server web and mobile interfaces. When a user initiates a session on Power BI Server by accessing a view from a client device, an Application Server thread is triggered. The thread is responsible for verifying the permissions of the user and the requested view

3) Repository:-

The Power BI Server Repository is a PostgreSQL database designed to store essential server-related data. This data encompasses details concerning Power BI Server users, groups, group memberships, permissions, projects, data sources, as well as extract metadata and refresh information.

4) VIZQL Server:-

When a view is accessed, the client initiates a request to the VizQL process (vizqlserver.exe). The VizQL process subsequently sends queries directly to the data source and retrieves a result set. This result set is then transformed into images and displayed to the user. Each VizQL Server maintains its own cache, which can be utilized by multiple users for improved performance and efficiency.

5) Data Engine:-

The Data Engine serves as a repository for data extracts and handles query responses

6) Backgrounder:-

The Backgrounder executes server tasks, including scheduled extract refreshes and tasks initiated from external tools like tabcmd. It also manages various background tasks within the server environment.

7) Data Server:-

The Data Server is responsible for overseeing connections to Power BI Server data sources. It also preserves metadata originating from Power BI Desktop, which includes calculations, definitions, and groups.
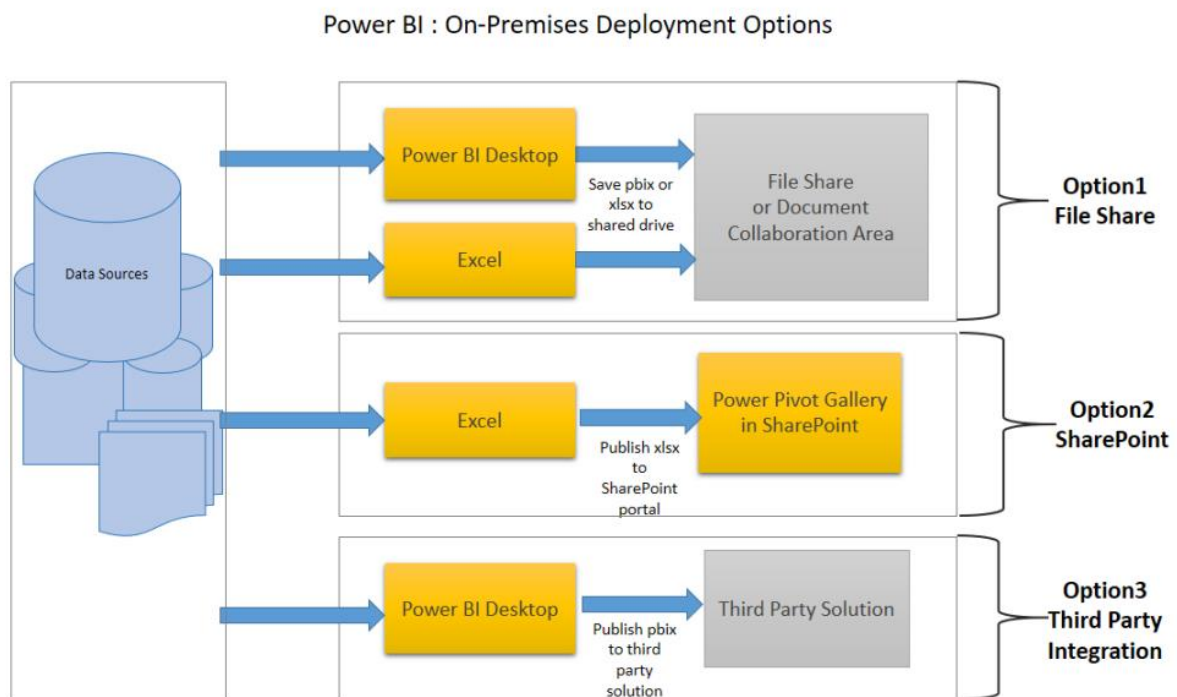
8) Power BI Communication Flow
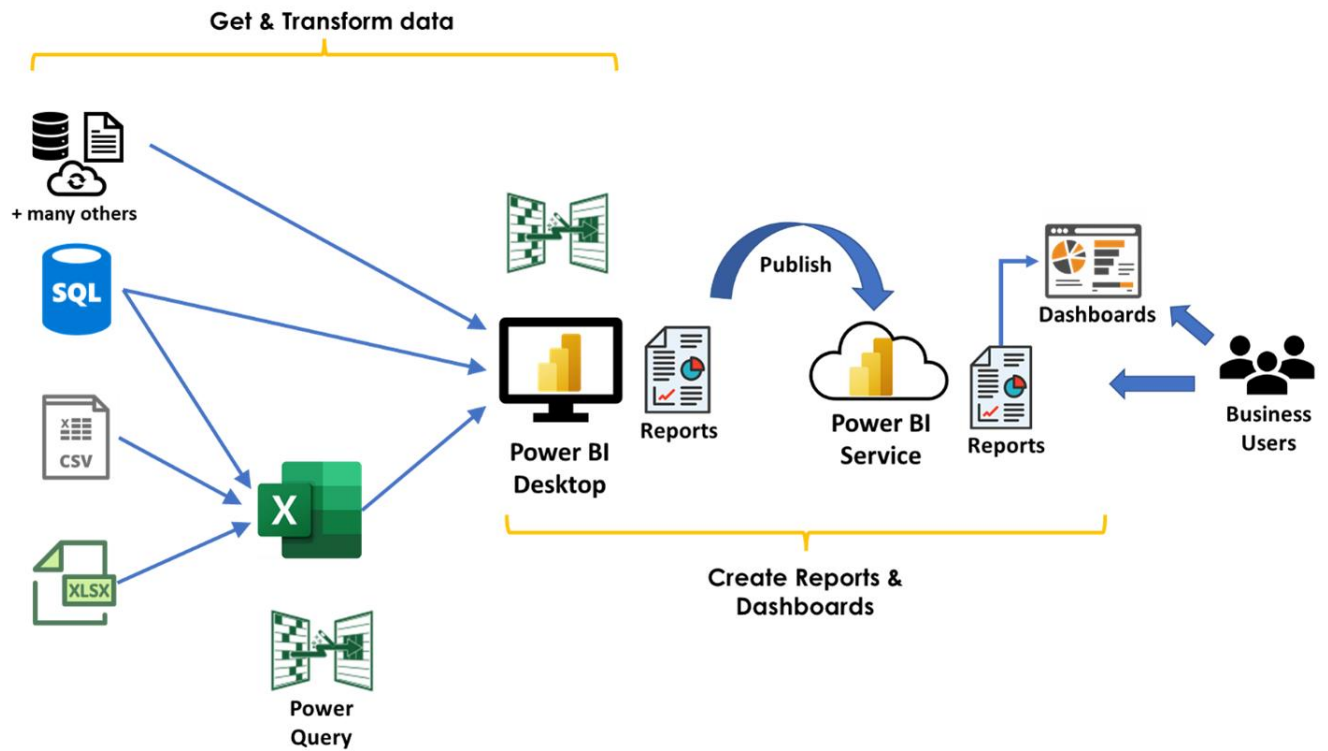
3. Deployment Description

3.1 Deployment options in Power BI

Power BI analytics platform offers three different deployment options depending on your environment and needs. The below graphic shows each option at a glance:

Power BI : On-Premises Deployment Options

1. **Power BI Online** In Power BI, you can easily collaborate with team members by sharing and collaborating on reports and dashboards in real-time, facilitating effective data-driven decision-making

2. **Power BI**, you can also take advantage of the cloud by deploying your reports and dashboards on public cloud platforms like Microsoft Azure, ensuring flexibility and scalability while retaining control. This allows you to use your existing licenses or acquire new ones through your preferred marketplace for seamless integration and deployment.

3. Power BI, you have the option to deploy your solution on-premises, giving you full control over your hardware and software resources, whether you prefer Windows or Linux environments. This flexibility allows you to tailor your deployment to your specific requirements and scale it as needed to meet your organization's unique needs..

4 Power BI Service

Users can access the reports and dashboards from Power BI Service using client platforms like websites, mobile devices, etc. This means that every client who wants to access content created on Power BI needs to interact with Power BI Service. And so, we must take a look under the hood and learn how Power BI Service works.

Power BI Service's architecture consists of two parts:

A front end

A back end

4.1 Front End cluster

The front end also called the web front-end cluster acts as an intermediary between clients and the back end. The front end services are used for establishing an initial connection and authenticating clients using Azure Active Directory. The Azure Active Directory stores user identities.

Along with this, Azure Traffic Manager is used to direct user requests to the nearest data center after authentication. Once a client/user is authenticated, the **Azure Content Delivery Network (CDN**) distributes static Power BI content/files to users.

4.2 Back End Cluster

The front end also called the web front-end cluster acts as an intermediary between clients and the back end. The front end services are used for establishing an initial connection and authenticating clients using Azure Active Directory. The Azure Active Directory stores user identities.

Along with this, Azure Traffic Manager is used to direct user requests to the nearest data center after authentication. Once a client/user is authenticated, the Azure Content Delivery Network (CDN) distributes static Power BI content/files to users.

4.3 Working of Power BI Service

- Power BI stores its data in two main repositories; **Azure block storage** and **Azure SQL database**. Azure block storage stores the datasets uploaded by users and all the metadata and system-related data is stored in the Azure SQL database.
- After Azure API Management authenticates a user request, it is sent to the Gateway Role. The Gateway Role processes the requests and directs them to suitable components like *Presentation Role, Background Job Processing Role, Data Role, and Data Movement Role.*
- For instance, the Presentation Role handles all the visualization related queries like for dashboards and reports.
- For all the data related queries, the request is sent by the Gateway Role to the Data Role or Data Movement Role.
- Power BI Service back end uses Azure Service Bus to connect on-premise datasources with the cloud. Azure Service Bus receives all the requests to fetch data from the on-premise data source. Then it processes the request and executes the query on the on-premise data source to retrieve data from it to the cloud service.
- The Azure Service Fabric manages all the microservices and components associated with running Power BI.
- Azure AD Cache helps in real-time reporting using the data stored in the in-memory of the Power BI system