Low Level Design

Amazon Sales Data Analysis

Written By	Harshit Nagila Harshit Soni Aniee Sapra Sujal Pandey
Document Version	1.0
Last Revised Date	17th Sept 2022

Contents

1. Introduction 04

1.1 What is Low-Level Design Document? 04

1.2 Scope 04

2. Architecture 05

3. Power BI Architecture 05

4. Power BI Architecture Working 08

5. Working on Power BI Services 11

1. Introduction

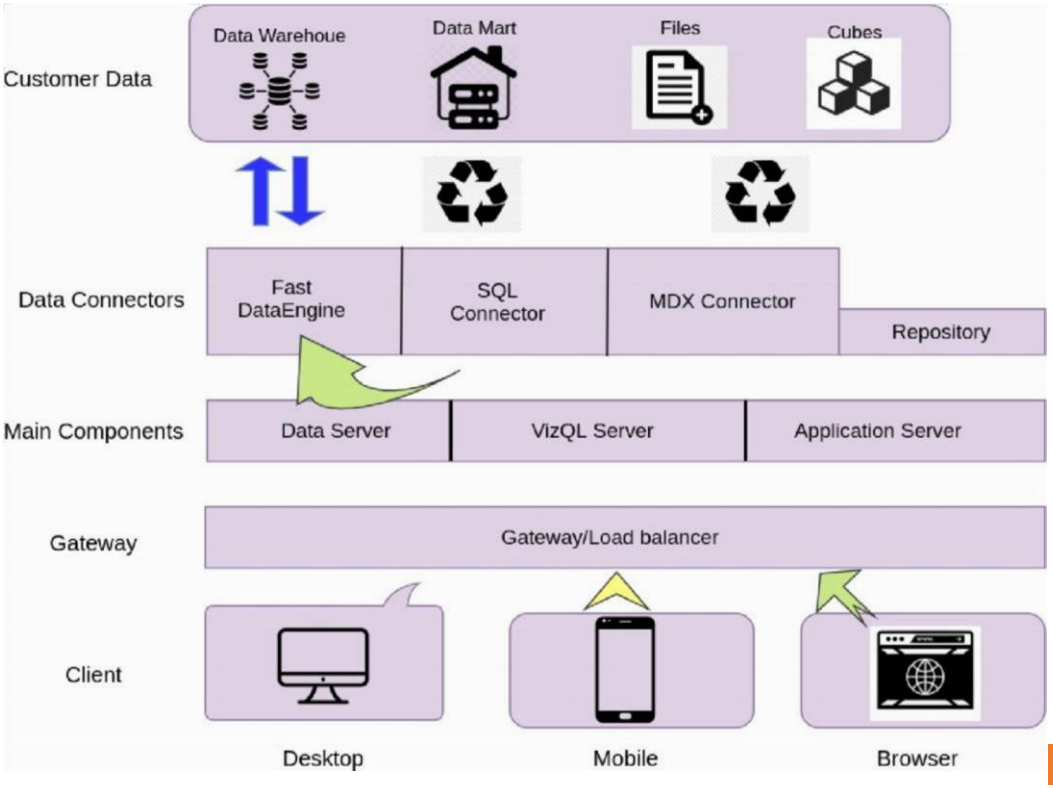
1.1 What is Low-Level design document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the House Price Prediction dashboard. LDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 Scope

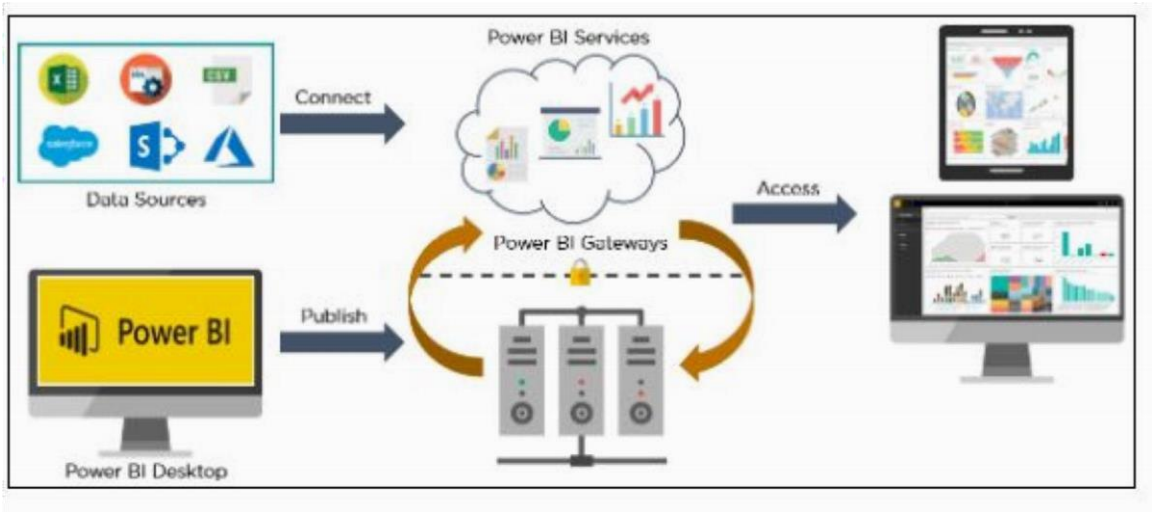
Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

2. Architecture



Power BI Architecture

Power BI architecture is a service built on top of Azure. There are multiple data sources that Power BI can connect to. Power BI Desktop allows you to create reports and data visualizations on the dataset. Power BI gateway is connected to on-premise data s ources to get continuous data for reporting and analytics. Power BI services refer to the cloud services that are used to publish Power BI reports and data visualizations. Using Power BI mobile apps, you can stay connected to their data from anywhere. Powe r BI apps are available for Windows, iOS, and Android platforms.



Let us discuss these four steps giving insightful information about each one of them.

1. Data Integration
2. Data Transforming
3. Report & Publish
4. Creating and Dashboard

1. Data Integration:

Data is extracted from different sources which can be different servers or databases. The data from various sources can be in different types and formats. If you import the file into the Power BI, it compresses the data sets up to 1GB, and it uses a direct query if the compressed data sets exceed more than 1GB. Then the data is integrated into a standard format and stored at a place called a staging area. There are two choices for big data sets. They are as follows. • Azure Analytics Services

• Power BI premium

2. Data Transforming:

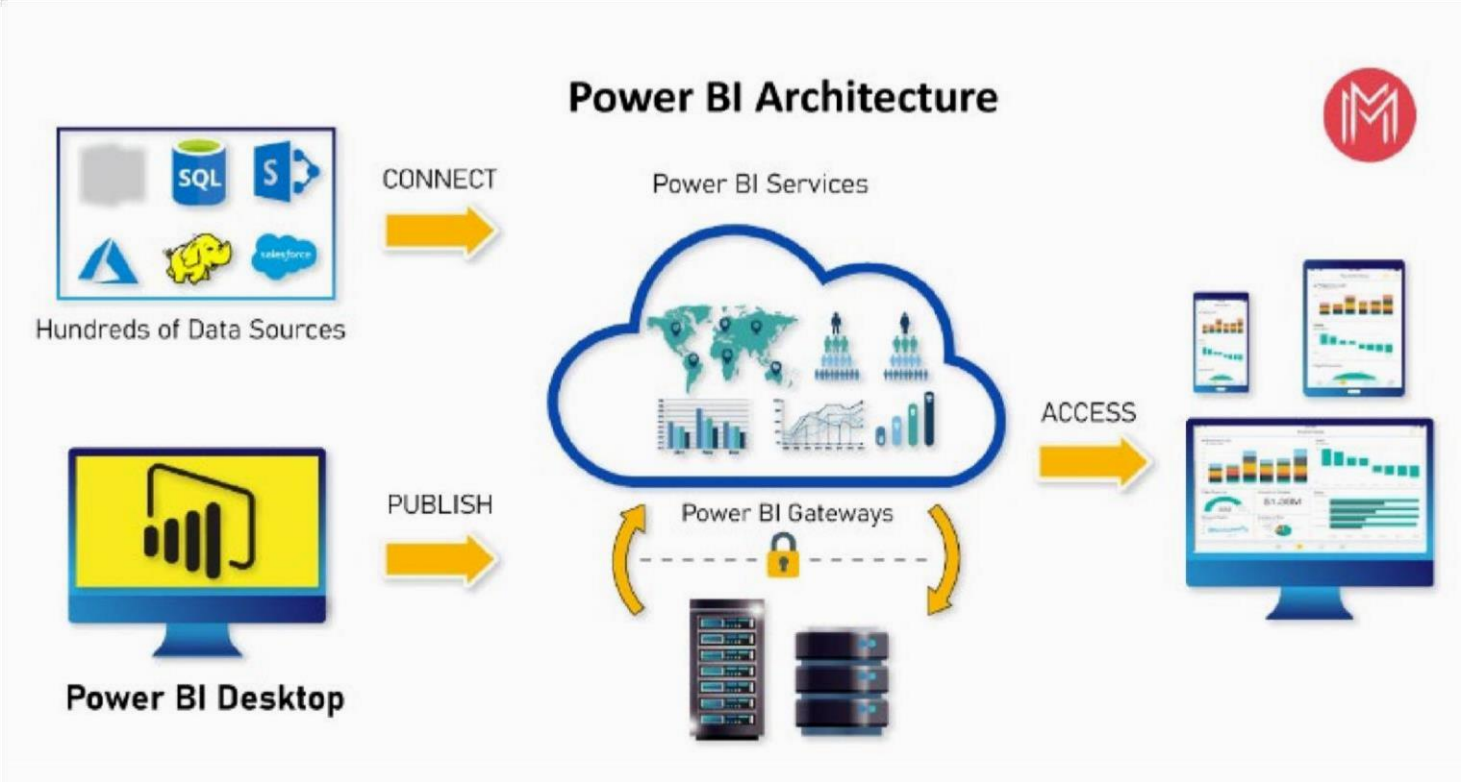
Integrated data is not ready to visualize data because the data should be transformed. To transform the data, it should be cleaned or pre-processed. For example, redundant or missing values are removed from the data sets. After data is pre-processed or cleaned, business rules are applied to transform the data. After processing the data, it is loaded into the data warehouse.

3. Report & Publish:

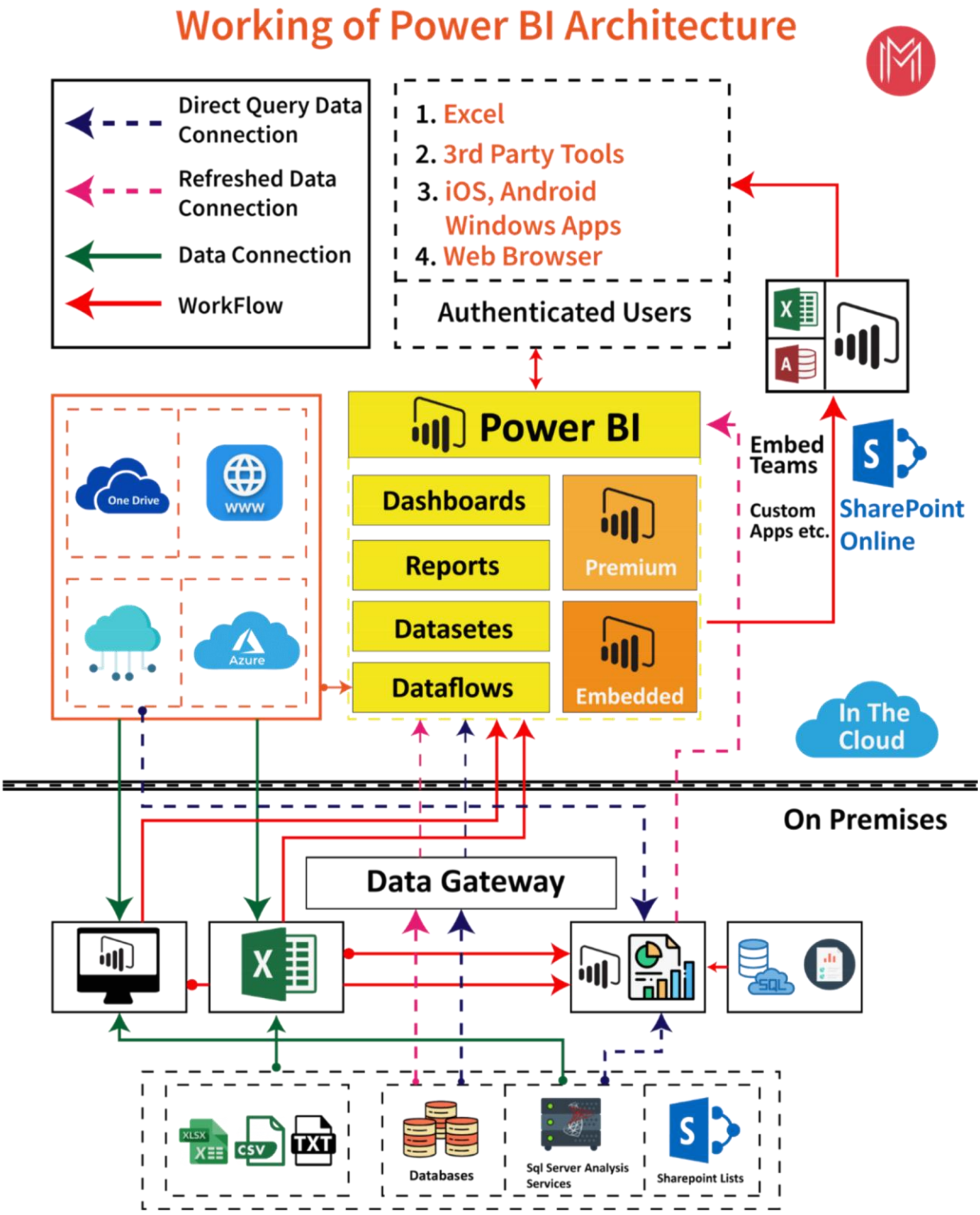
After sourcing and cleaning the data, you can create the reports. Reports are the visualization of the data in the form of slicers, graphs, and charts. Power BI offers a lot of custom visualization to create the reports. After creating reports, you can publish them to power bi services and also publish them to an on-premise power bi server.

4. Creating Dashboards:

You can create dashboards after publishing reports to Power BI services, by holding the individual elements. The visual retains the filter when the report is holding the individual elements to save the report. Pinning the live report page allows the dashboard users to interact with the visual by selecting slicers and filters.


Power BI Architecture
Hundreds of Data Sources
CONNECT
Power BI Services
ACCESS
PUBLISH
Power BI Gateways
Power BI Desktop

Power BI Architecture - Working



If you observe in the top of the image excel, web browsers and other sources are streaming into Power BI components, and they are called data sources. These data sources are authenticated users. Power BI has different data sources like On-Premise, Cloud databases, direct connections, etc.

On-Premise:

Power BI Desktop is accomplished with the authenticating, development and publishing tools. You can transfer the data from data sources to Power BI Desktop. And also, it allows users to create and publish reports on the Power BI Report Server or Power BI Service.

Power BI Publisher allows you to publish the Excel workbooks to the Power BI Report Server. Report Publisher and SQL server Data tools help in creating the KPIs, datasets, paginated reports, mobile reports, etc. All kinds of reports are published at the Power BI Report Server, and from there, reports are distributed to the end-users.

On-Cloud:

Power BI Gateway is the essential component in the Power BI architecture. The Power BI Gateway acts as a bridge or secure channel to transfer the data from On-premise data to On-cloud data sources or apps.

Cloud side architecture consists of a lot of components including Power suite having datasets, dashboards, reports, Power BI Premium, Power BI Embedded, etc. Users can embed the dashboards, reports into applications, SharePoint, Teams, etc. There are Cloud data sources and they are connected to the Power BI tools.

Power BI Service Architecture

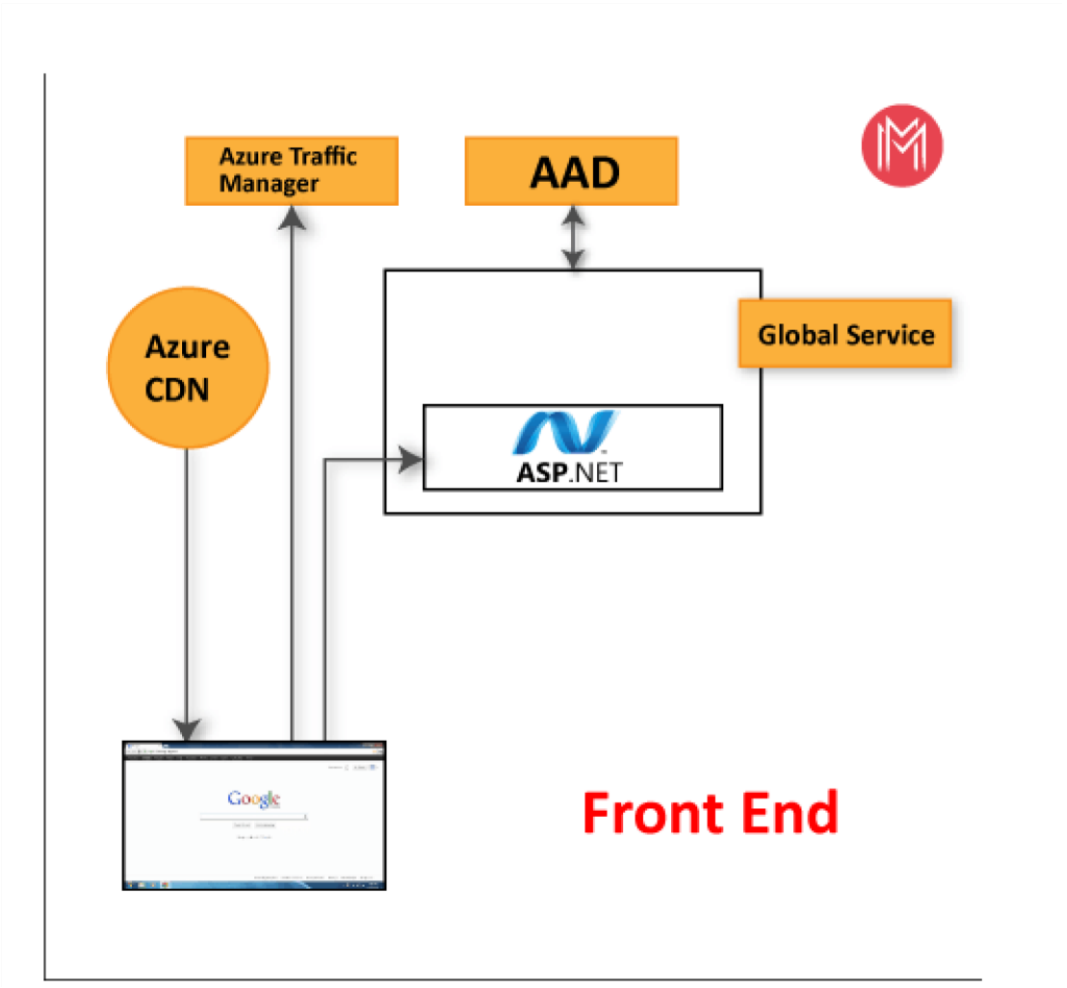
In the previous section, you have learned how to publish the created reports in the Power BI Service.

Power BI Service enables the users to create and access the reports, dashboards from the client platforms like mobile devices, websites, etc. User needs to interact with the Power BI Service whenever they want to access the data that is created on the Power BI. So, now, we will learn how the Power BI Service works.

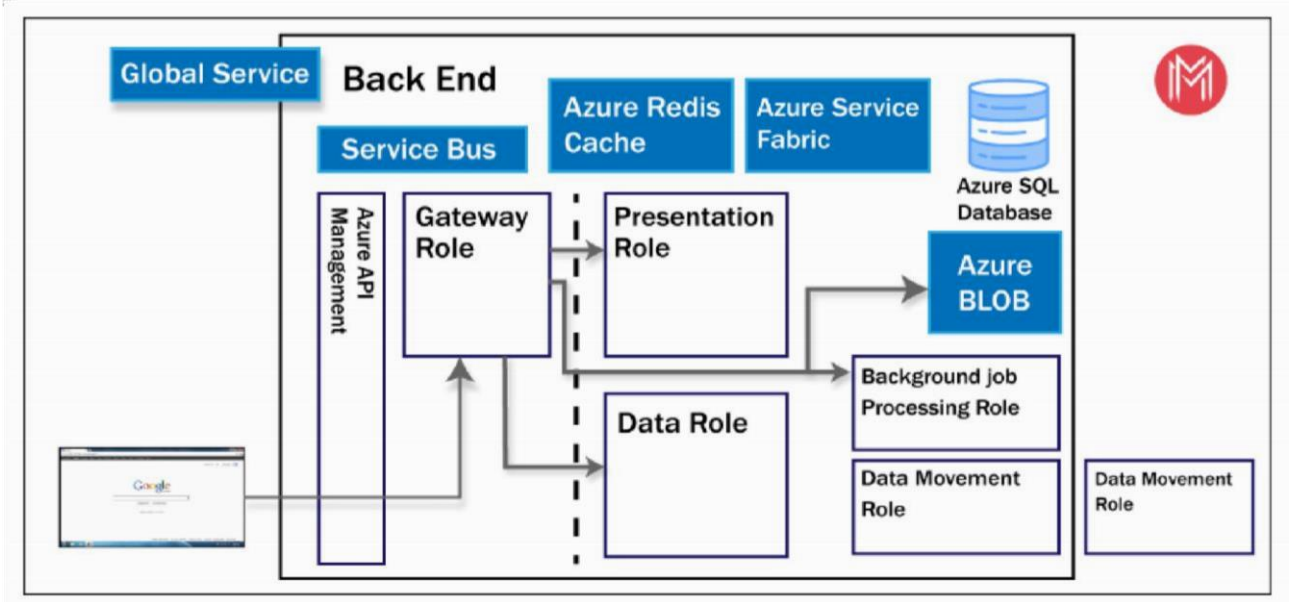
Power BI Service Architecture consists of two clusters. The following are the two clusters.

- Front End Cluster
- Back End Cluster

1. **Front End Cluster:** Front end cluster acts as an intermediate between the back-end cluster and the clients. It is also called a Web Front End Cluster. It establishes the initial connection and authenticates the users or clients using the Azure Active Directory. After user authentication, Azure Traffic Manager directs the user requests to the nearest data centers and Azure Content Delivery Network (CDN) allocates the statice files/content to the users or clients based on the geographical locations.



2. **Back End Cluster:** It manages the datasets, reports, storage, visualizations, data refreshing, data connections, and other services in the Power BI. At the back-end cluster, the web client has only two direct points to interact with the data, i.e., Gateway Role and Azure API Management. These two components are responsible for authorizing, load balancing, routing, authentication, etc.



Working Of Power BI Service

- Power BI stores the data in two leading repositories, i.e., Azure SQL Database and Azure Block Storage. Azure Block Storage enables the users to store the datasets, and all system-related data and metadata are stored in the Azure SQL database.

- It authenticates the user requests and sends them to the Gateway Role. It processes the requests and assigns them to the appropriate components like Background Job Processing Role, Data Movement Role, Presentation Role, and Data Role.

- The presentation role manages all the associated visualization queries like reports and dashboards.

- Presentation Role sends requests to the Gateway Role to the Data Movement Role or Data Role for all relevant datasets.

- Azure Service Bus is used to connect and fetch the data from the On-Premises data sources with the cloud. It sends a request to execute the queries On-Premises data source and retrieve the data from its cloud service.

- The Azure Service Fabric allows all components and microservices which are related to the Power BI Service.
- Azure Cache helps in reporting the data that is stored in the in-memory of the Power BI system.

Next, fill out the title you want your viz to have and click “save”.