Architecture Design

# Amazon Sale Analysis

|  |  |
| --- | --- |
| **Written By** | Yash Shinde |
| **Document Version** | 0.3 |
| **Last Revised Date** |  |

**DOCUMENT CONTROL**

## Change Record:

|  |  |  |  |
| --- | --- | --- | --- |
| **VERSION** | **DATE** | **AUTHOR** | **COMMENTS** |
| 0.1 | 19- Nov -  2023 | Yash shinde | Introduction and architecture defined |
| 0.2 | 20 - Nov -  2023 | Yash Shinde | Architecture & Architecture description appended and  updated. |
|  |  |  |  |
|  |  |  |  |

**Reviews:**

|  |  |  |  |
| --- | --- | --- | --- |
| **VERSION** | **DATE** | **REVIEWER** | **COMMENTS** |
| 0.2 | 21- Nov -  2023 | Yash Shinde | Unit test cases to be added |

**Approval Status:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **VERSION** | **REVIEW**  **DATE** | **REVIEWED BY** |  | **APPROVED BY** | **COMMENTS** |
|  |  |  |  |  |  |

**Contents**

1. [Introduction 04](#_TOC_250005)
   1. [What is Architecture Design Document? 04](#_TOC_250004)
   2. [Scope 04](#_TOC_250003)
2. [Architecture 05](#_TOC_250002)
   1. Tableau Architecture 05
   2. Tableau Server Architecture 05
   3. Gateway/Load Balancer 06
   4. Application Server 06
   5. VIZQL Server 07
   6. Data Engine 07
   7. Backgrounder 07
   8. Data Server 07
   9. Tableau Communication Flow 07
3. Deployment 08
   1. [Deployment Options in Tableau 09](#_TOC_250001)
   2. [Single Node Architecture 10](#_TOC_250000)
   3. Three Node Architecture 11
   4. Five Node Architecture 12

# Introduction

## 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.

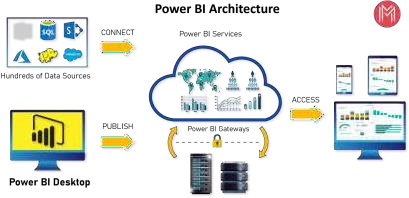
## Scope

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

# Architecture

Power BI is a business suite that includes several technologies that work together. To deliver outstanding business intelligence solutions, Microsoft Power BI technology consists of a group of components such as:

* 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)



### Shape Description automatically generated with low confidenceARCHITECTURE DESIGN

**6**

## 1. Data Sources

An important component of Power BI is its vast range of data sources. You can import data from files in your system, cloud-based online data sources or connect directly to live connections. If you import from data on-premise or online services there is a limit of 1 GB. Some commonly used data sources in Power BI are:

* 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 BigQuery (Beta)
* Azure SQL Database
* Salesforce Reports
* Google Analytics
* Facebook
* GitHub

## 2. Power BI Desktop

*Power BI Desktop is a client-side tool known as a companion development and authoring tool. This desktop-based software is loaded with tools and functionalities to* connect to data sources, transform data, data modeling and creating reports.

You can download and install Power BI Desktop in your system for free. Using Power BI Desktop features, one can do data cleansing, create business metrics and data models, define the relationship between data, define hierarchies, create visuals and publish reports.

## 3. Power BI Service

Power BI Service is a web-based platform from where you can *share reports made on Power BI Desktop, collaborate with other users, and create dashboards.*

It is available in three versions:

* Free version
* Pro version
* Premium version

Power BI Service is also known as, **“Power BI.com”**, **“Power BI Workspace”, “Power BI Site”** and **“Power BI Web Portal”**. This component also offers advanced features like *natural language Q&A* and *alerts*.

## 4. Power BI Report Server

The Power BI Report Server is similar to the Power BI Service. The only difference between these two is that Power BI Report Server is an on-premise platform. It is used by organizations who do not want to publish their reports on the cloud and are concerned about the security of their data.

Power BI Report Server enables you to create dashboards and share your reports with other users following proper security protocols. To use this service, you need to have a Power BI Premium license.

## 5. Power BI Gateway

This component is used to connect and access on-premise data in secured networks. Power BI Gateways are generally used in organizations where data is kept in security and watch. Gateways help to extract out such data through secure channels to Power BI platforms for analysis and reporting.

## 6. Power BI Mobile

Power BI Mobile is a native Power BI application that runs on iOS, Android, and Windows mobile devices. For viewing reports and dashboards, these applications are used.

7. Power BI Embedded

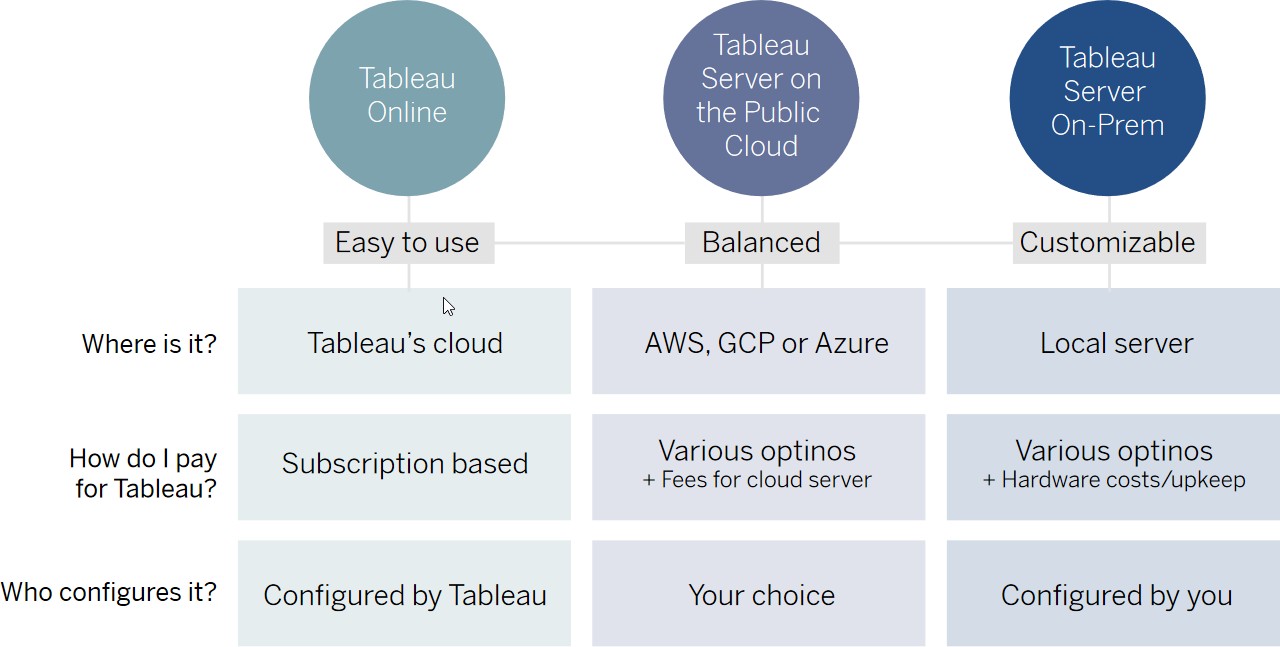
Power BI Embedded offers APIs which are used to embed visuals into custom applications.

1. **Deployment Description**

## Deployment options in Tableau

Tableau’s analytics platform offers three different deployment options depending on your

environment and needs. The below graphic shows each option at a glance:



1. **Tableau Online** Get up and running quickly with no hardware required. Tableau Online is fully hosted by Tableau so all upgrades and maintenance are automatically managed for you.
2. **Tableau Server** deployed on public cloud: Leverage the flexibility and scalability of cloud infrastructure without giving up control. Deploy to Amazon Web Services, Google Cloud Platform, or Microsoft Azure infrastructure to quickly get started with Tableau Server (on your choice of Windows or Linux). Bring your own license or purchase on your preferred marketplace.
3. **Tableau Server deployed on-premises**: Manage and scale your own hardware and software (whether Windows or Linux) as needed. Customize your deployment as you see fit.

## Single Node Architecture



This architecture is a single node architecture. This is the most simple deployment topology.

## 3 Node Architecture



This architecture is a 3 Node Architecture which is more capable to handle concurrent requests.

If we need failover or high availability, or want a second instance of the repository, we must install Tableau Server on a cluster of at least three computers. In a cluster that includes at least three nodes, you can configure two instances of the repository, which gives our cluster failover capability.

## 5 Node Architecture



When we install Tableau Server on a Five-node cluster, we can install server processes on one or both nodes. A five-node cluster can improve the performance of Tableau Server, because the work is spread across multiple machines.

Note the following about five-node clusters:

* + - A five-node cluster does not provide failover or support for high availability.
    - You can't install more than one instance of the repository on a two-node cluster, and the repository must be on the initial node.