

Architecture Design Document

Big Game Census Data Visualization

-Aishwarya Srivastava

| | |
|--------------------------|-----------------------------------|
| Written By | Author 1- Aishwarya Srivastava |
| Document Version | 0.1 |
| Last Revised Date | |

DOCUMENT CONTROL

Change Record:

| VERSION | DATE | AUTHOR | COMMENTS |
|---------|---------------|----------|----------|
| 0.1 | 7- Sep - 2021 | Author 1 | |
| | | | |
| | | | |
| | | | |

Reviews:

| VERSION | DATE | REVIEWER | COMMENTS |
|---------|----------|----------|-------------------------------|
| 0.1 | 8-9-2021 | Author 1 | Node architecture to be added |

Contents

| | | |
|-----------|---------------------------------------------|-----------|
| 1. | Introduction | 04 |
| 1.1 | What is Architecture Design Document? | 04 |
| 1.2 | Scope..... | 04 |
| 2. | Architecture | 05 |
| 2.1 | Tableau Server Architecture | 05 |
| 2.2 | Components of Tableau Server..... | 07 |
| 2.3 | Summary..... | 08 |
| 3. | Deployment | 09 |
| 3.1 | Deployment Options in Tableau | 09 |
| 3.2 | Single Node Architecture | 10 |
| 3.3 | Three Node Architecture..... | 11 |
| 3.4 | Five Node Architecture..... | 12 |

1. Introduction

1.1 What is Architecture design document?

Any software needs the architectural design to represent 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.

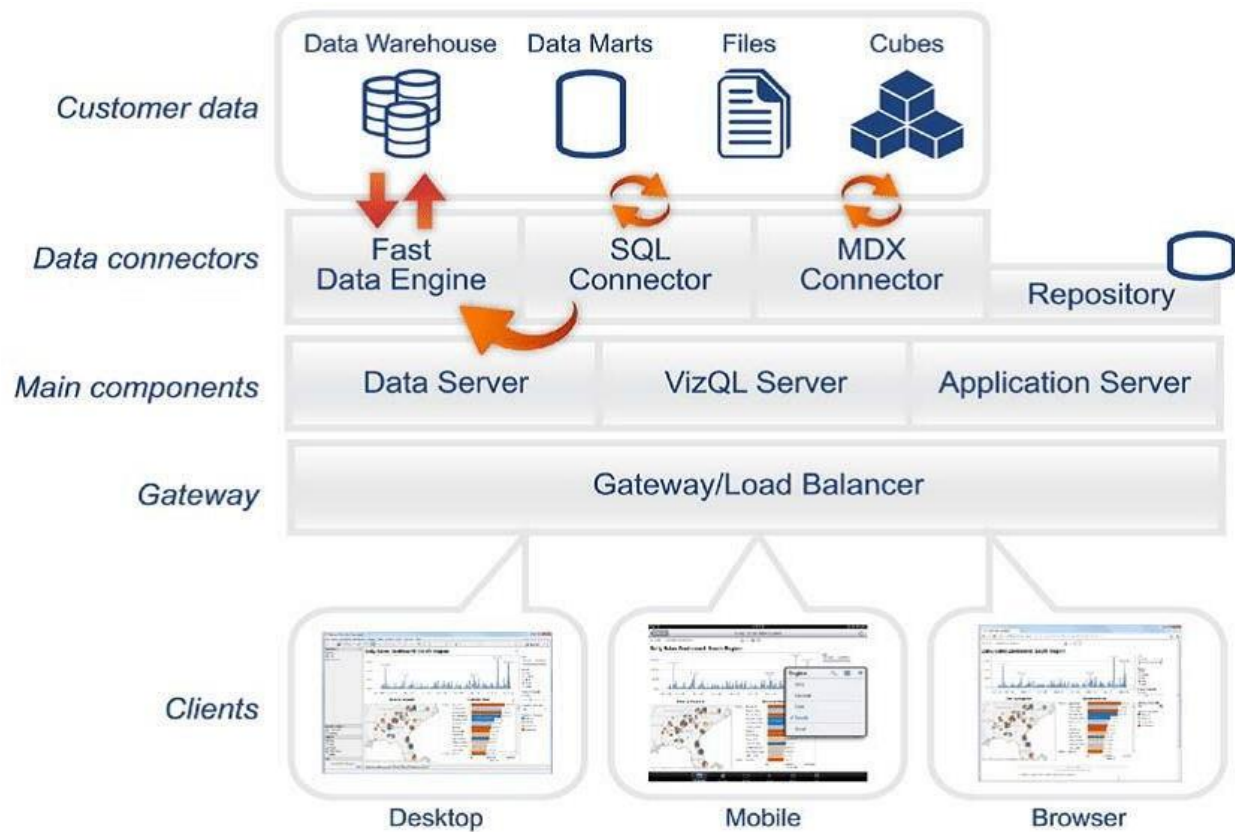
1.2 Scope

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.

The given data is visualized on Tableau Desktop and published on Tableau Public. The dashboard given below can be accessed using the link:

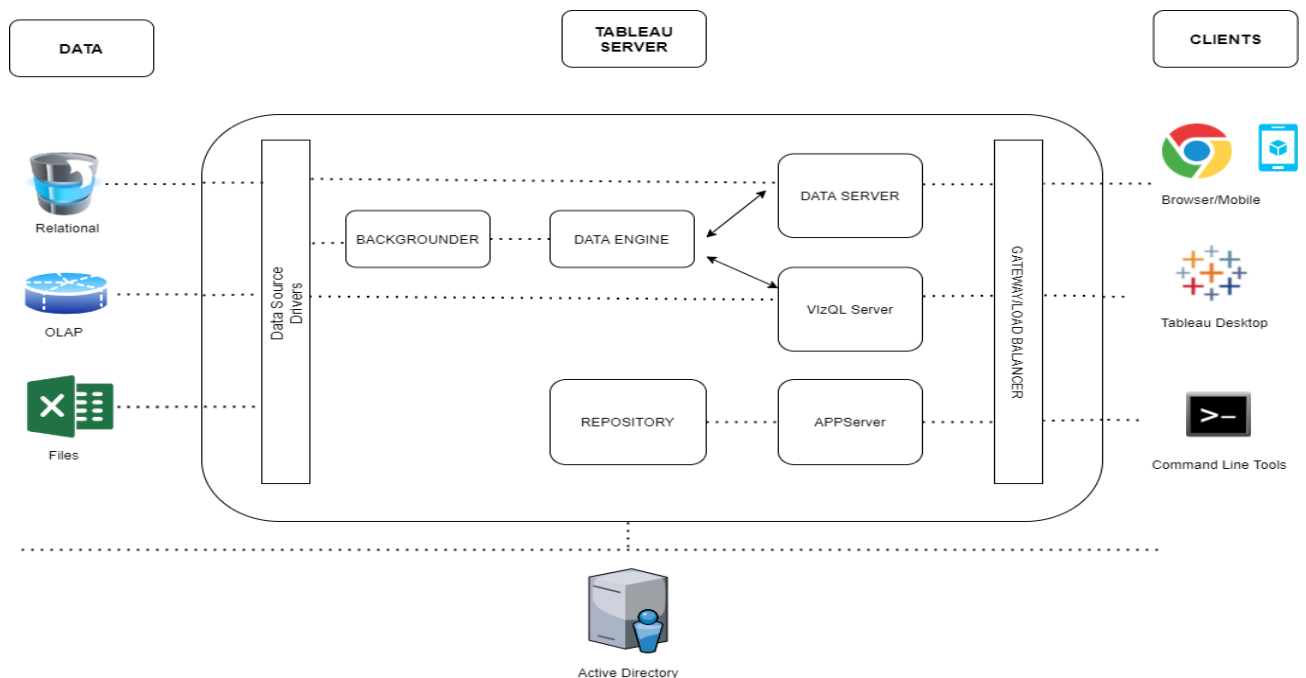
<https://public.tableau.com/app/profile/aishwarya6335/viz/BigGameCensusDashboard/Dashboard1>

1. Architecture



2.1 Tableau Server Architecture

Tableau has a highly scalable, n-tier client-server architecture that serves mobile clients, web clients and desktop-installed software. Tableau Server architecture supports fast and flexible deployments.



The various layers used in the Tableau server are given in the following architecture diagram.

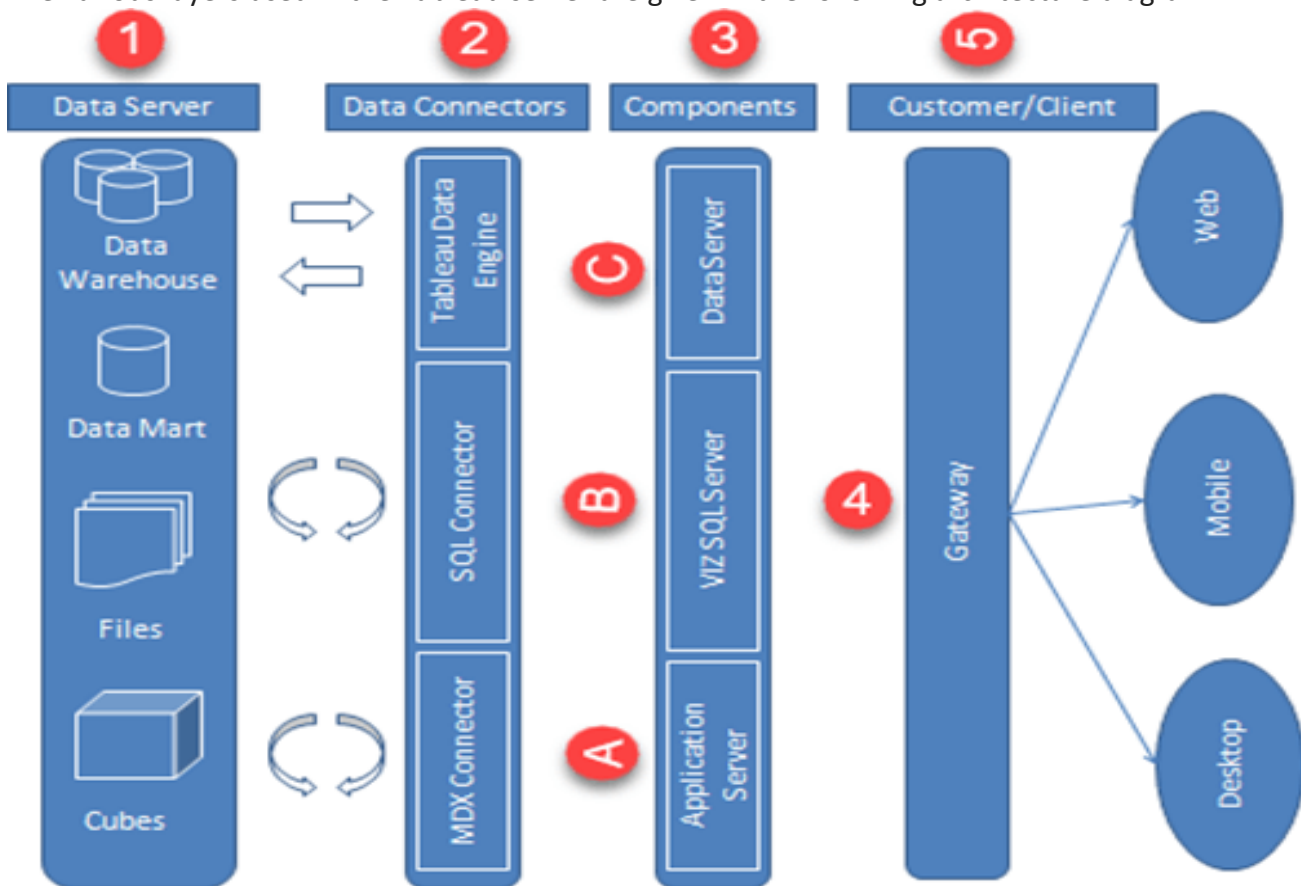


Tableau Architecture Diagram

Let's study the different components of Tableau Architecture

➤ Data Server

The primary component of Tableau Architecture is the Data sources it can connect to it.

Tableau can connect to multiple data sources. These data sources can be on-premise or remotely located. It can connect to a database, excel file, and a web application all at the same time. Tableau can connect data from heterogeneous environments. It can blend the data from multiple data sources. It can also make the relationship between various types of data sources.

➤ Data Connectors

The Data Connectors provide an interface to connect external data sources to Tableau Data Server.

Tableau has in-built ODBC/SQL connector. This ODBC Connector can connect to any databases without using their native connector. Tableau has an option to select both live and extract data. Based on the usage, one can be easily switched between extracted and live data.

- **Live Connection or Real time data:** Tableau can connect to real time data by linking to the external database directly. It uses the infrastructure of existing database system by sending dynamic MDX (Multidimensional Expressions) and SQL statements. This feature can link to the live data with Tableau rather than importing the data. It makes good the investment done by an organization on a fast and optimized database system. In many enterprises, the size of the database is huge and is updated periodically. In those cases, Tableau works as a front-end visualization tool by connecting to the live data.
- **Extracted or In-memory data:** Tableau has an option to extract the data from external data sources. We can make a local copy in the form of tableau extract file. It can extract millions of records in Tableau data engine with a single click. Tableau's data engine uses storage such as RAM, ROM and cache memory to store and process data. Using filters, Tableau can extract few records from a huge dataset. This improves the performance, especially while working on massive datasets. Extracted or in-memory data allows the users to visualize the data offline, without connecting to the data source.

2.2 COMPONENTS OF TABLEAU SERVER

The different components present in a Tableau server are:

- Application Server
- VizQL Server
- Data Server

A) Application Server:

The application server is used to provide the authentications and authorizations. It handles the administration and permission for web and mobile interfaces. It assures security by recording each session id on Tableau Server. The administrator can configure the default timeout of the session in the server.

B) VizQL Server:

VizQL server is used to convert the queries from the data source into visualizations. Once the client request is forwarded to VizQL process, it sends the query directly to data source and retrieves information in the form of images. This image or visualization is presented to the user. Tableau server creates a cache of visualization to reduce the load time. The cache can be shared across many users who have the permission to view the visualization.

C) Data Server:

Data server is used to manage and store the data from external data sources. It is a central data management system. It provides metadata management, data security, data storage, data connection and driver requirements. It stores the relevant details of data set such as metadata, calculated fields, sets, groups, and parameters. The data source could extract data as well make live connections to external data sources.

➤ Gateway

The gateway channelizes the requests from users to Tableau components. When the client makes a request, it is forwarded to external load balancer for processing. The gateway works as a distributor of processes to various components. In case of absence of external load balancer, gateway also works as a load balancer. For single server configuration, one primary server or gateway manages all the processes. For multiple server configurations, one physical system works as primary server while others are used as worker servers. Only one machine can be used as a primary server in Tableau Server environment.

➤ Clients

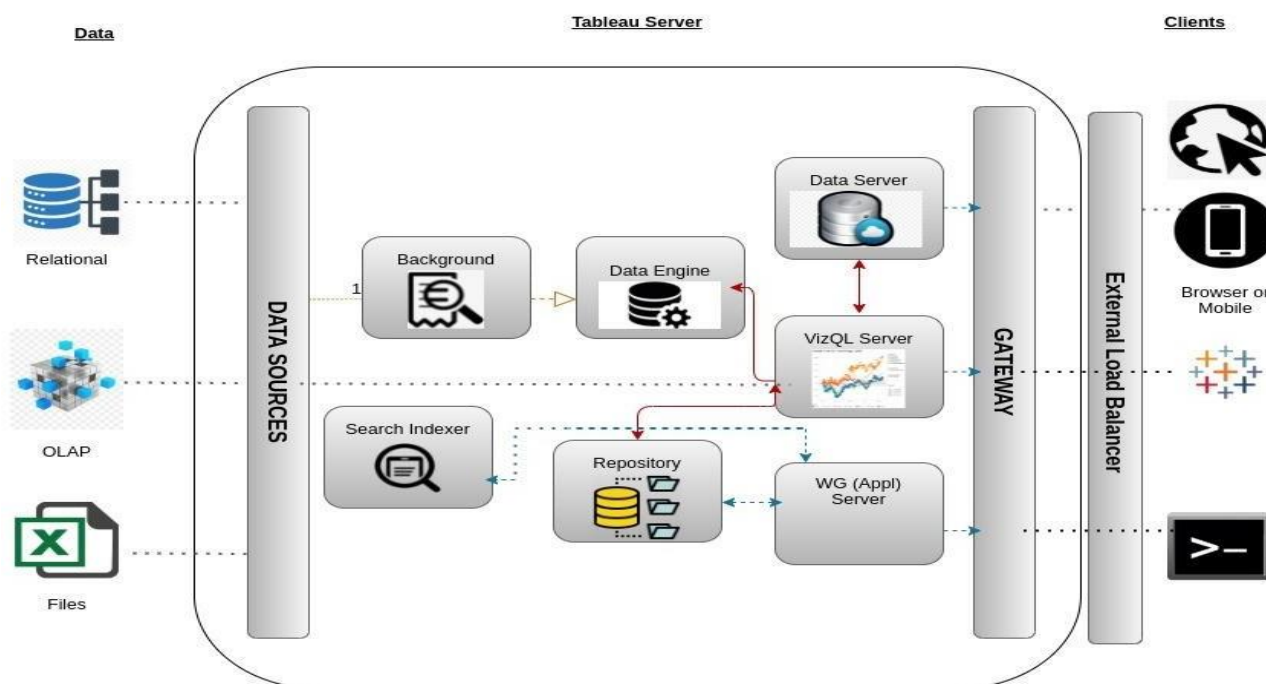
The dashboards and visualizations in Tableau server can be viewed and edited using different clients. The Clients are Tableau Desktop, web browser and mobile applications.

| Clients | Environment |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Tableau Desktop | Tableau Desktop is a business analytics tool. It helps to create, view and publish dashboards in Tableau Server. Users can access various data sources and build visualizations in Tableau Desktop. |
| Mobile | The dashboards from the server can be interactively visualized using mobile browsers and applications. The browser and application can be used to view and edit the contents in the workbook. |
| Web | Web browsers such as Google Chrome, Safari, Firefox and internet explorer support the Tableau server. The contents and visualizations in the dashboard can be edited through these web browsers. |

2.3 Summary

- The architecture of Tableau Server is designed to connect different data sources securely.
- Data Server is the first layer in the architecture. It helps Tableau to connect data in various heterogeneous environments.
- Data connector is the second layer. It helps to connect to various databases using its ODBC connector.
- Tableau can connect the real time live data by connecting the database directly. It can also extract a local copy of data through its in-memory data store for faster processing.
- The components such as Application Server, VizQL Server, and data server act as the third layer.
- The application server is used for authentications and authorizations.
- VizQL is used to convert the SQL query into visualizations.
- Data Server is a centralized data management system used in the architecture.
- A gateway is used to distribute the processes into different components.
- The fourth layer of the architecture are the clients such as Tableau Desktop, web and Mobile.
- Tableau Server is internally managed by the multiple server processes.
- The following diagram shows Tableau Server's architecture:

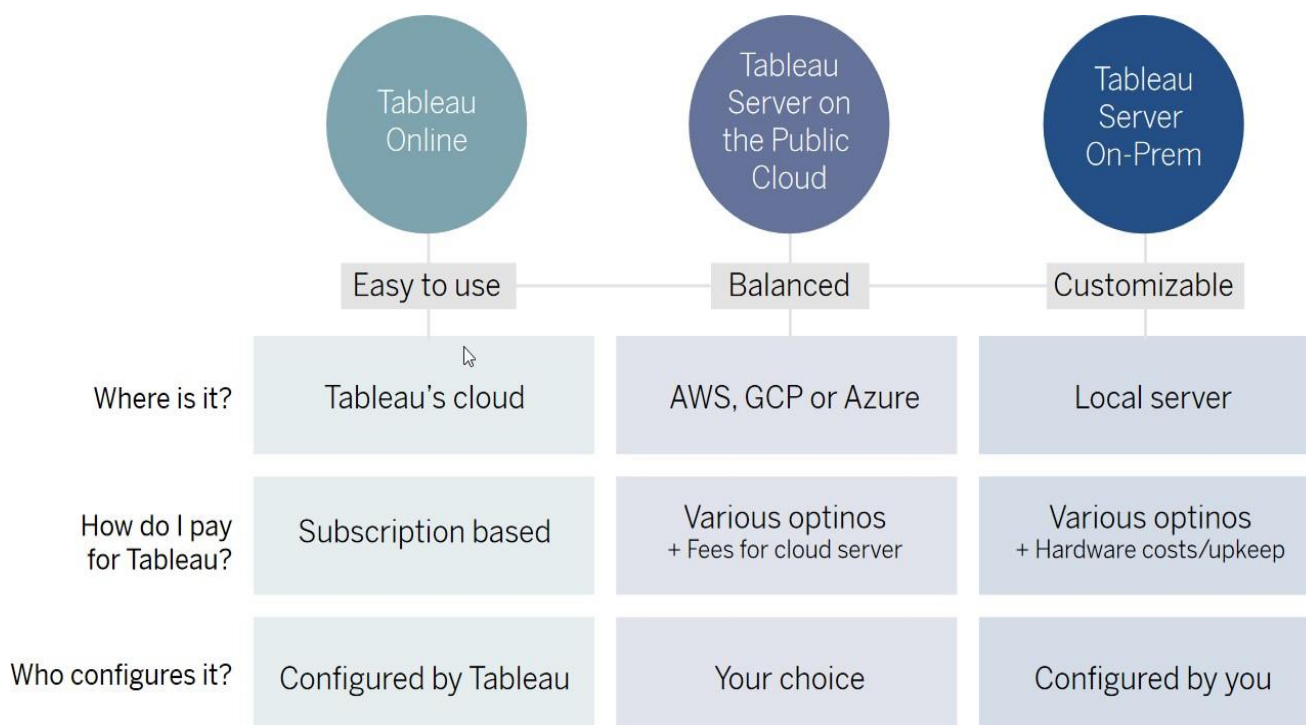
Tableau Communication Flow



3. Deployment Description

3.1 Deployment options in Tableau

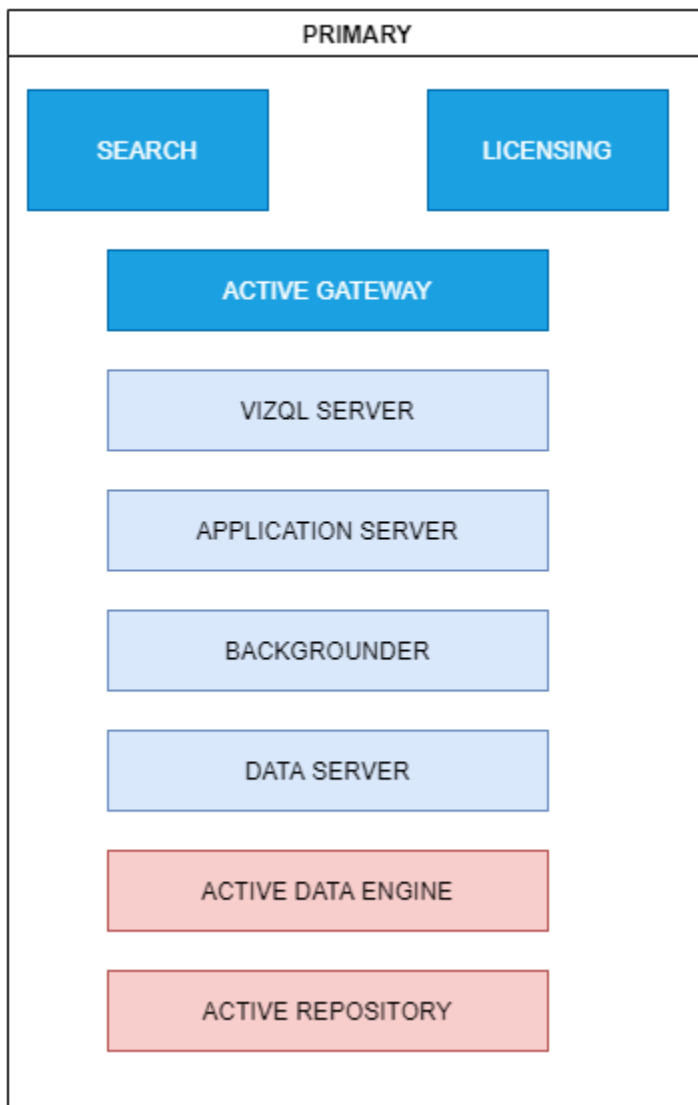
Tableau's analytics platform offers three different deployment options depending on your environment and needs. The below graphic shows each option at a glance:



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** choice of Windows or Linux). Bring your own license or purchase on your preferred marketplace.
4. **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.

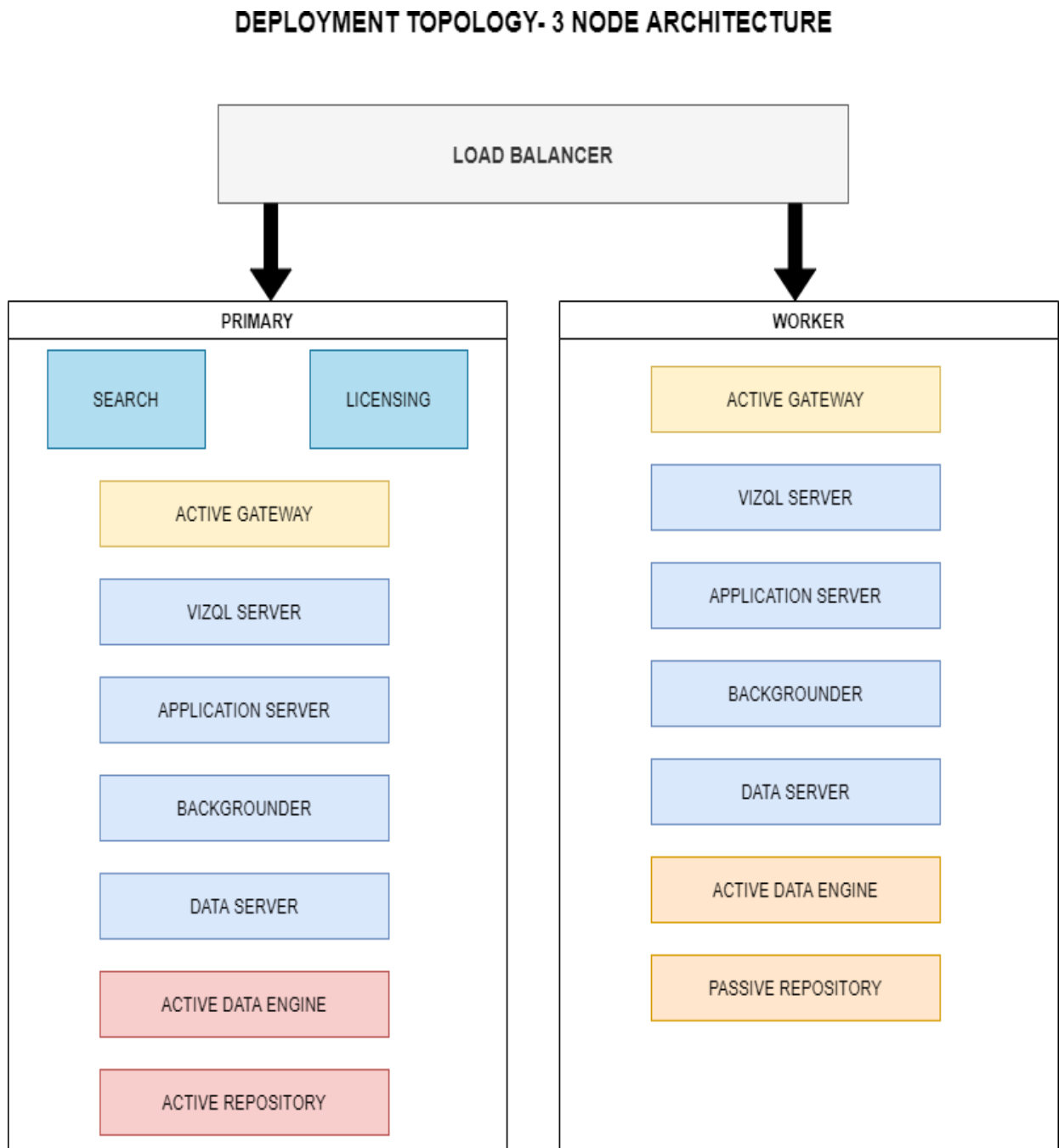
3.2 Single Node Architecture

DEPLOYMENT TOPOLOGY - SINGLE NODE ARCHITECTURE



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

3.3) 3 Node Architecture

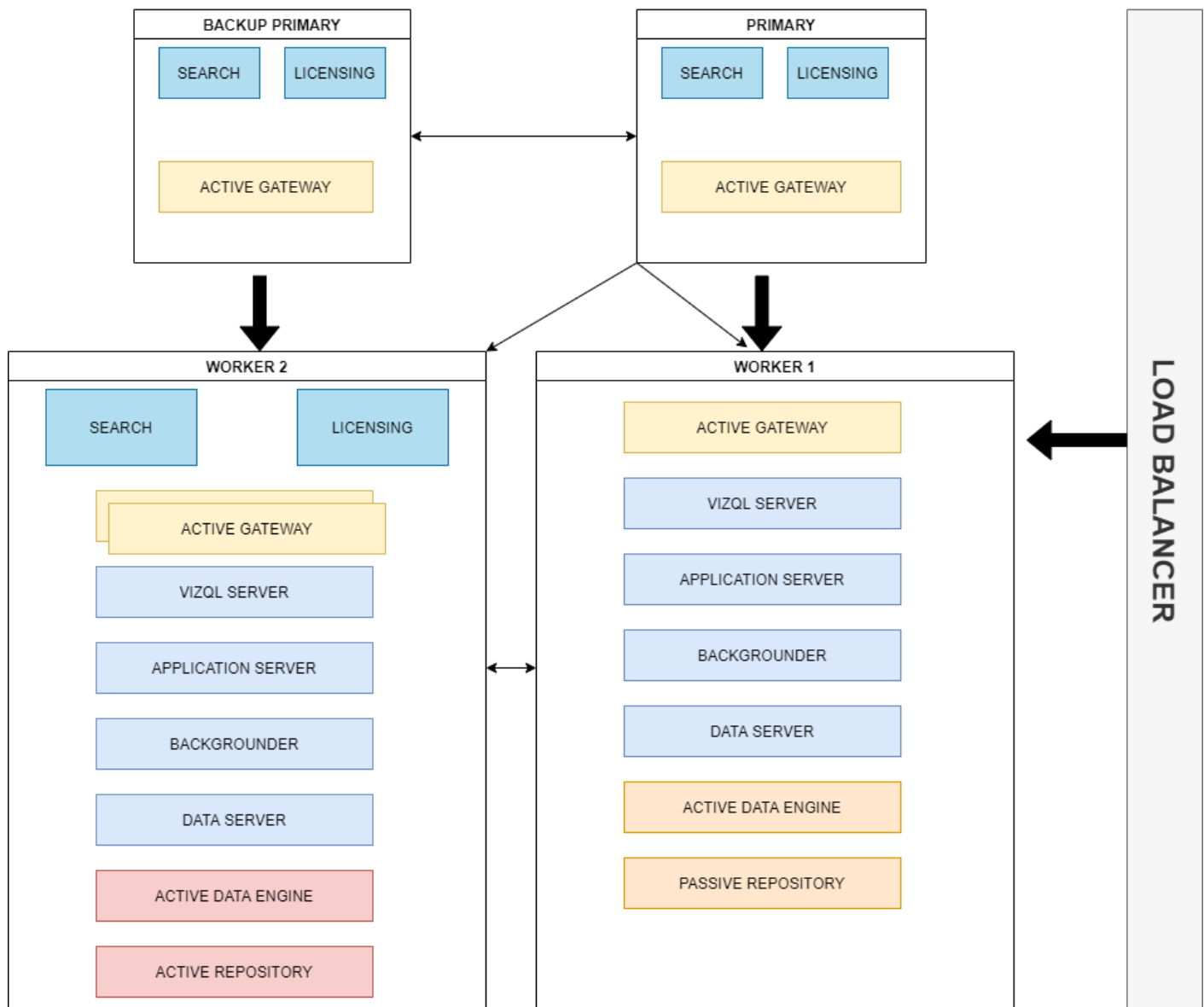


This architecture is a 3 Node Architecture which is more capable to handle concurrent requests.

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.

3.4) 5 Node Architecture

DEPLOYMENT TOPOLOGY- 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.