tableau

For visualizing and analysing data

Tableau Server

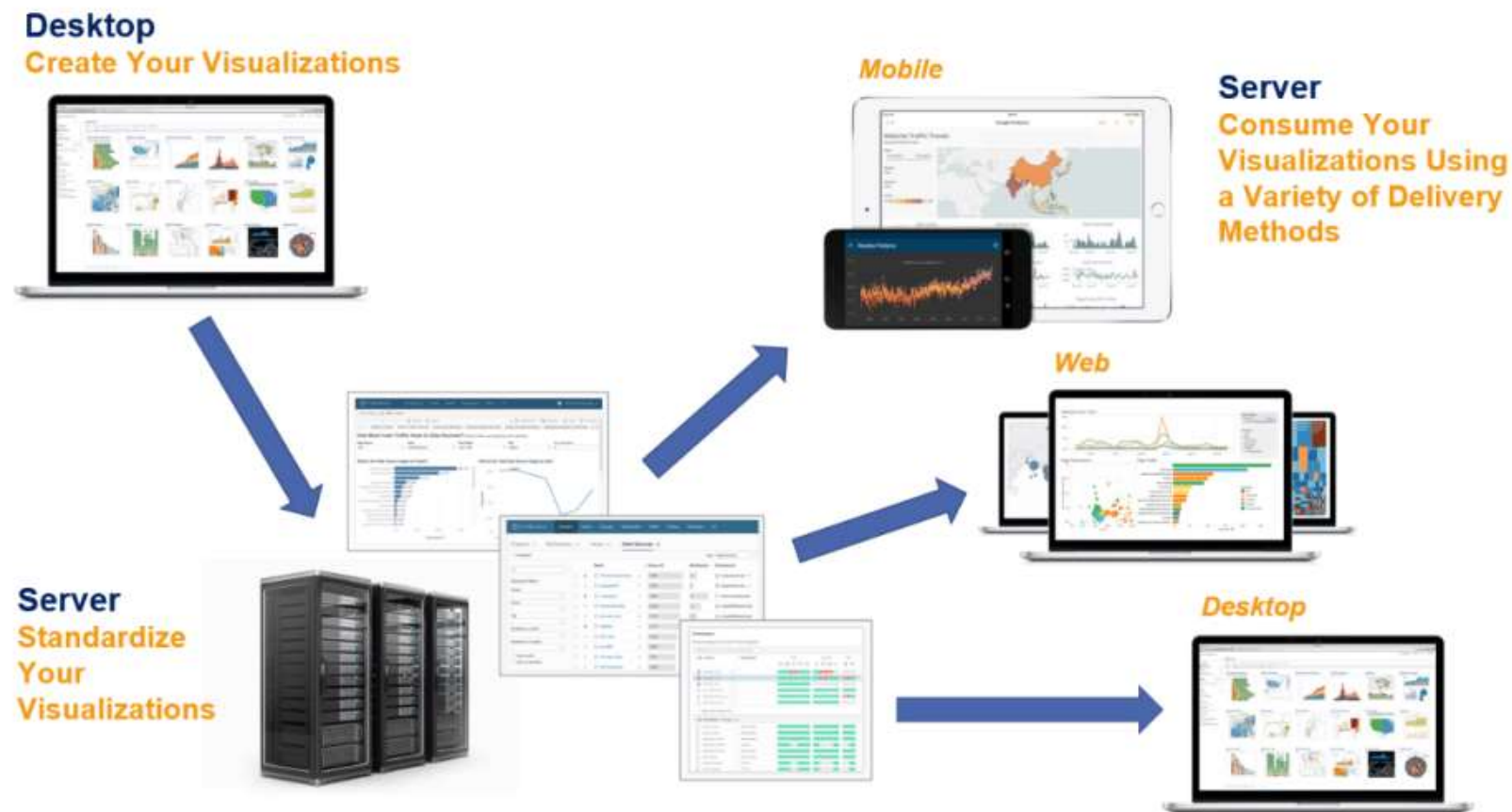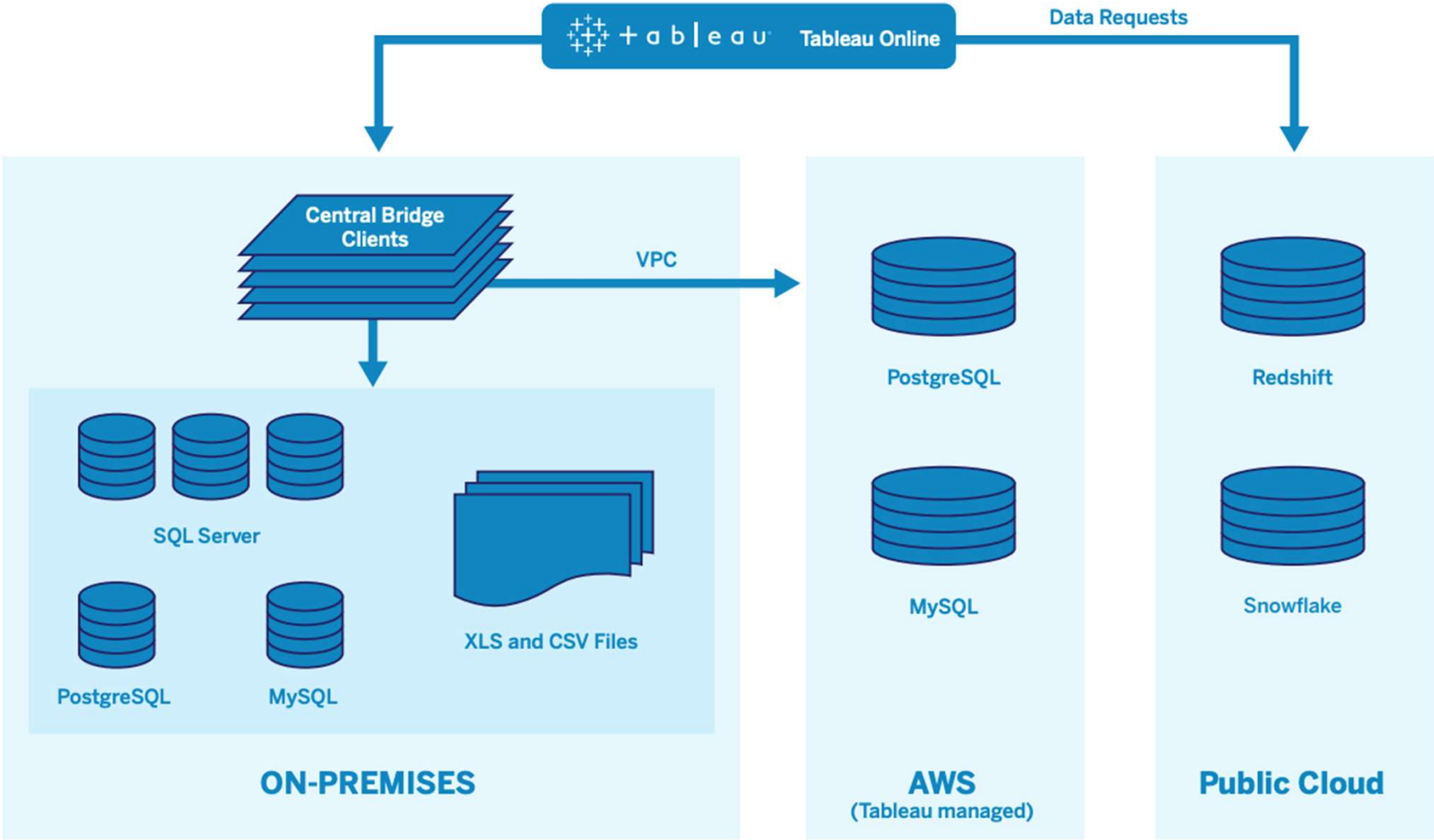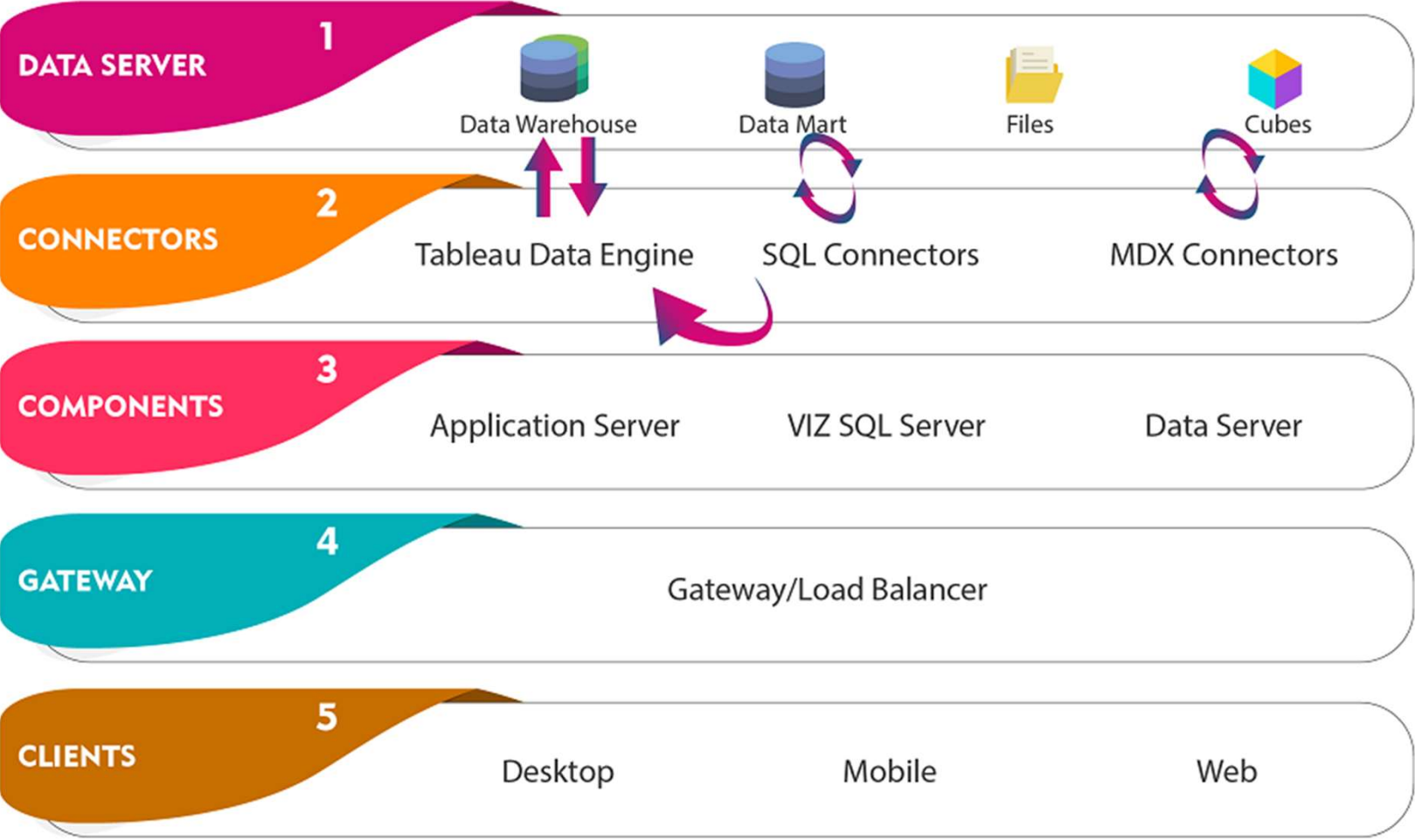# Tableau Server

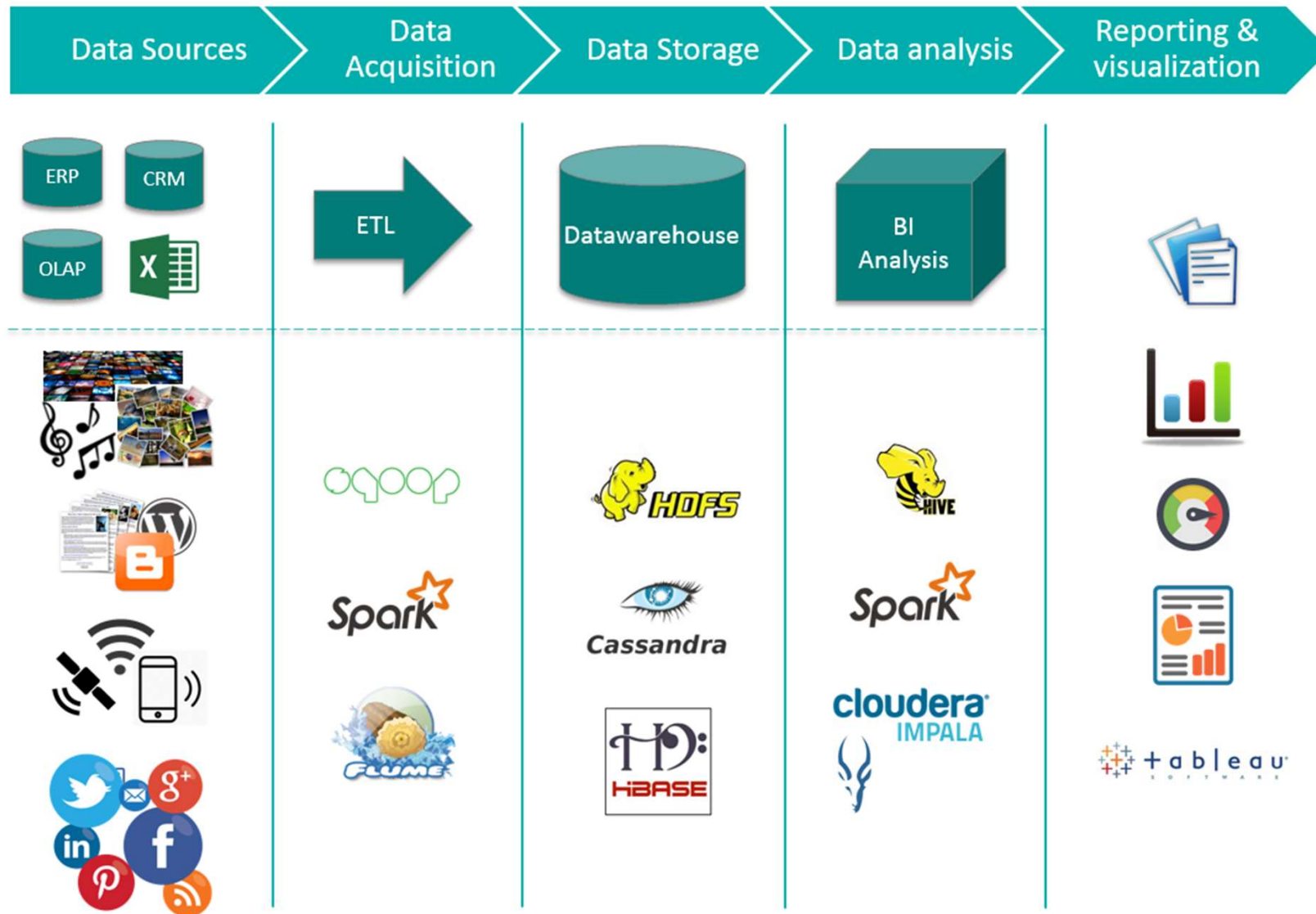# Tableau Server Components

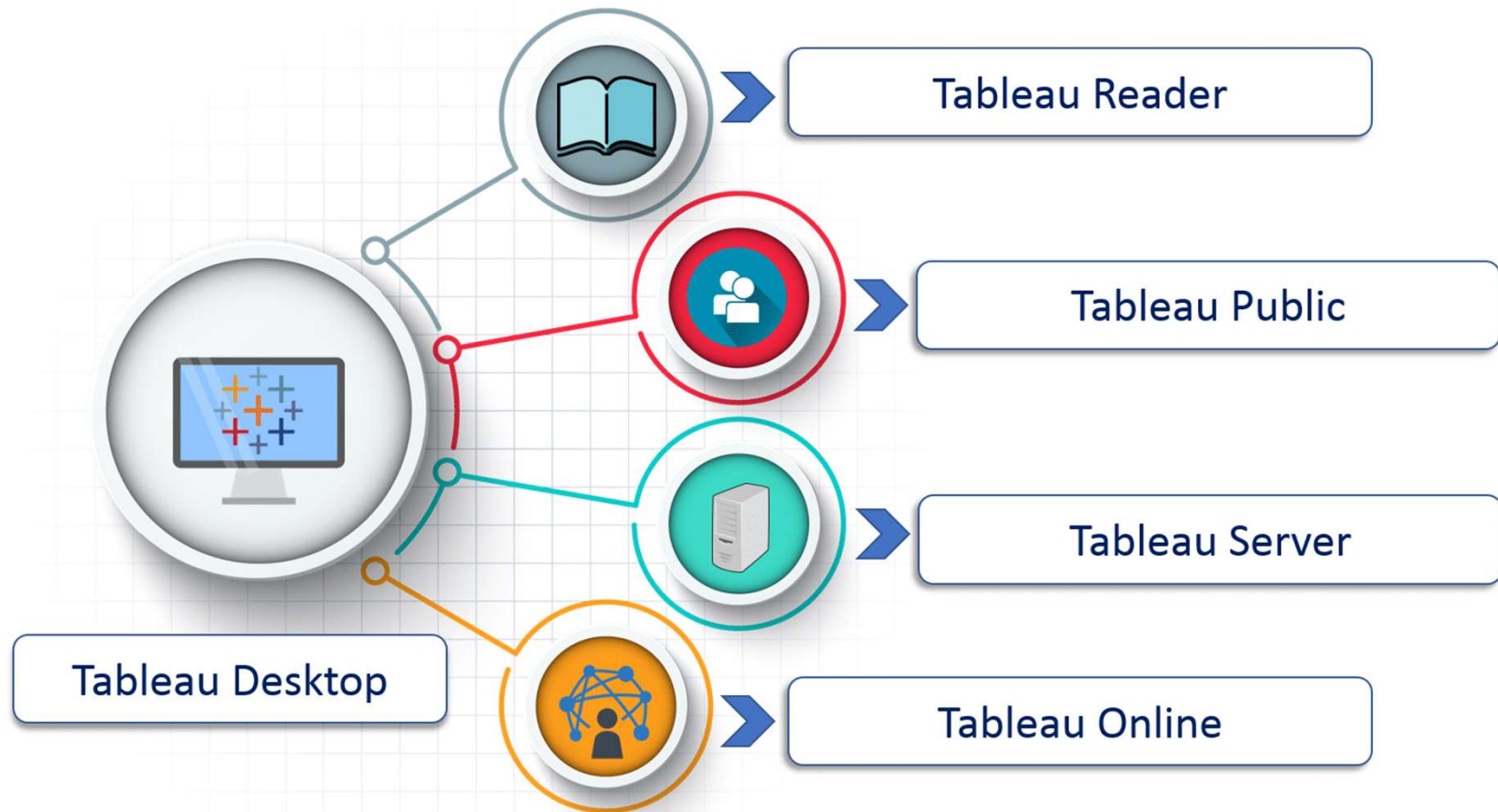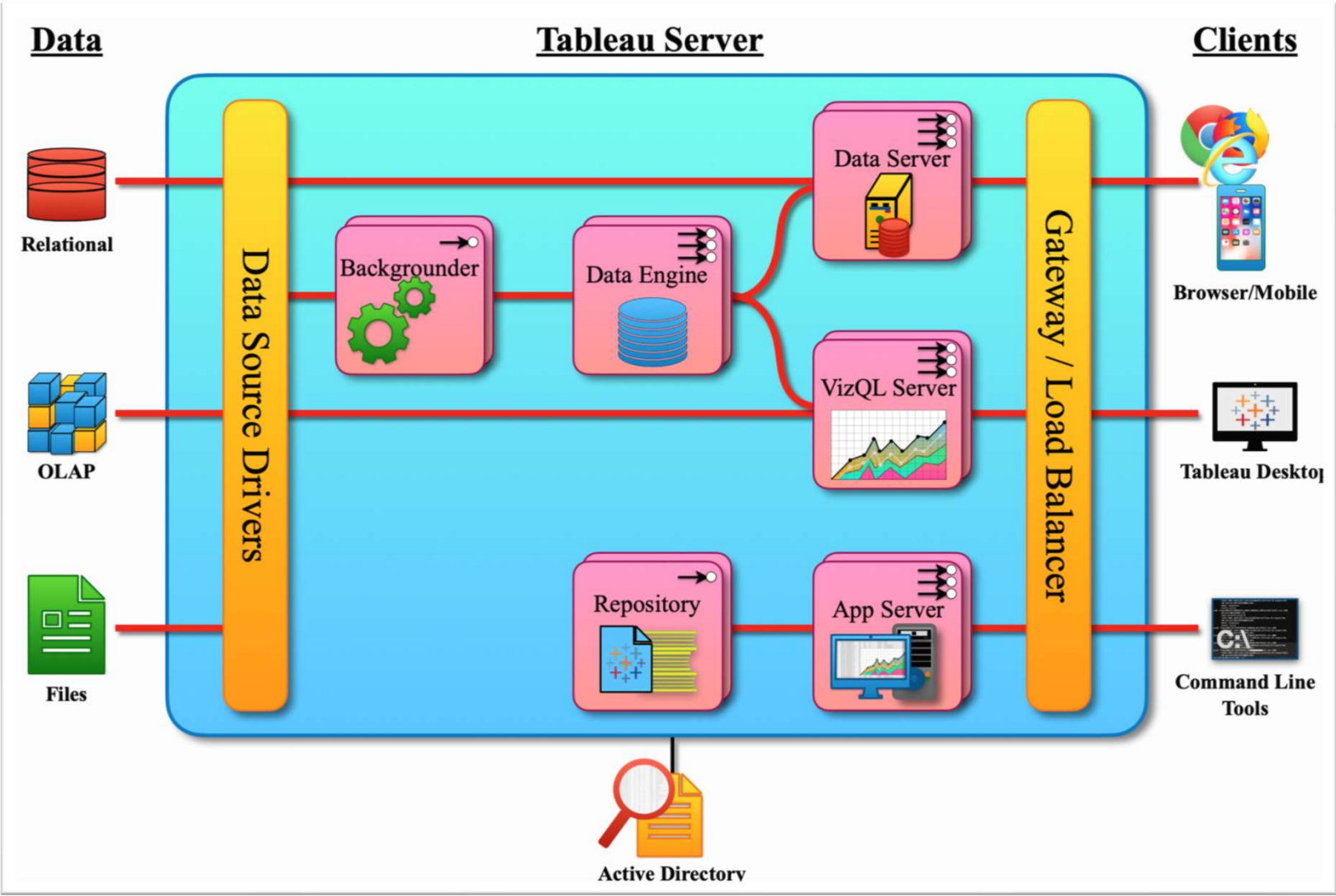# Big Data Ecosystem- Spark and Tableau

Tableau Products

Tableau Reader

Tableau Public

Tableau Server

Tableau Desktop

Tableau Online

Tableau Server Architecture and Components

Install Tableau Server

- Visit: https://www.tableau.com/support/releases/server/2023.1.4

- # Update Ubuntu and install gdebi
  - sudo apt-get update -y
  - sudo apt-get -y install gdebi-core
- # Download and install Tableau Server latest version
  - wget https://downloads.tableau.com/esdalt/2023.1.4/tableau-server-2023-1-4_amd64.deb
  - sudo gdebi -n tableau-server-2023-1-4_amd64.deb
  - sudo /opt/tableau/tableau_server/packages/scripts.20231.23.0711.1149/initialize-tsm --accepteula -f
- # Open below URL on browser and initialize Tableau server
  - https://localhost:8850

Gateway/ Load Balancer

- It is the primary Tableau Server that trails requests to other components

- Requests which come in from the client firstly strike the gateway server and are then routed to the appropriate procedures

- Once the request processing is completed, it collects the output and sends it to the Clients

- If multiple procedures are configured for any component, the Gateway will work as a load balancer and share the requests with the procedures
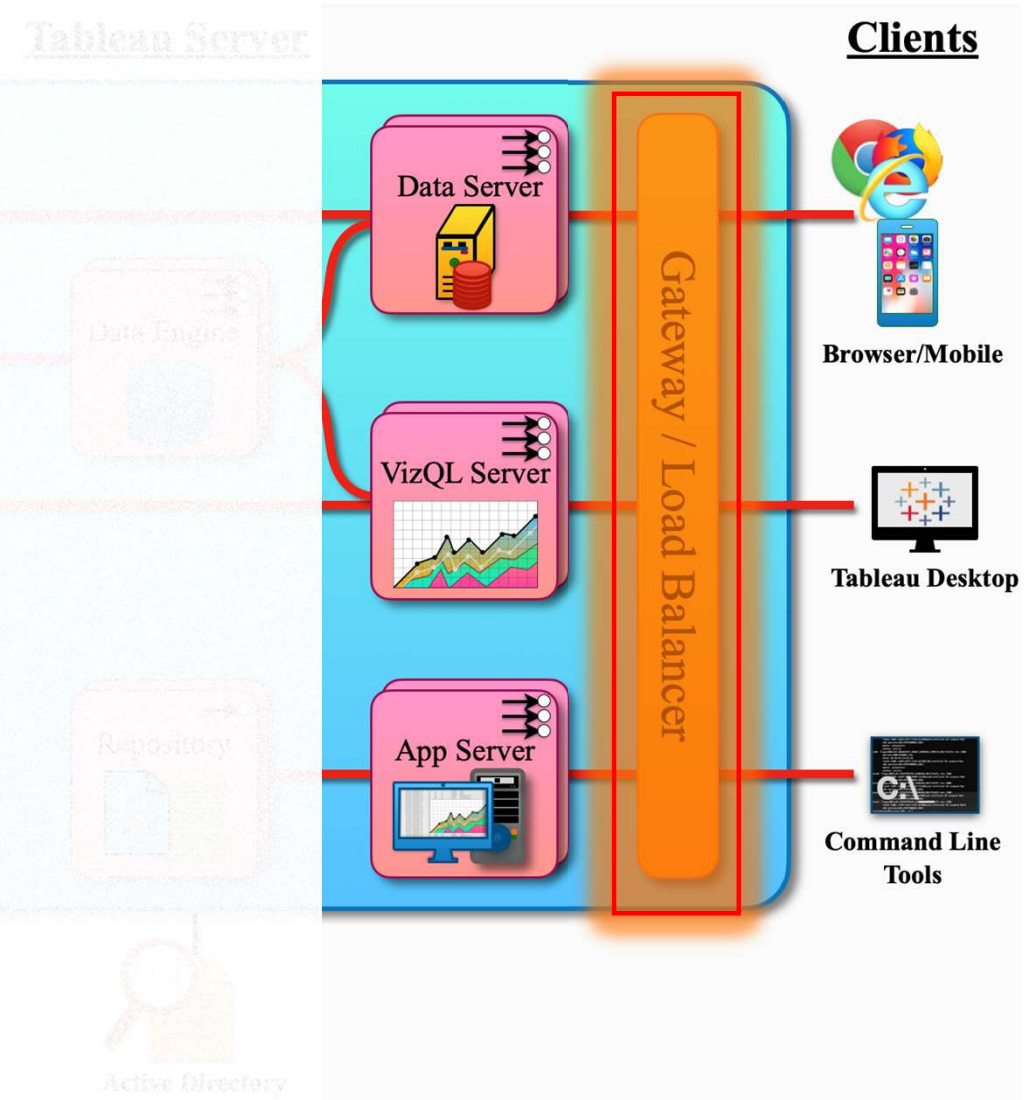
Application Server

- The Application Server (VizPortal) handles the web application and REST API calls.

- When a user opens a view in a client device, that user begins a session on Tableau Server

- Then the Application Server thread begins and verifies the permissions for that user and that view

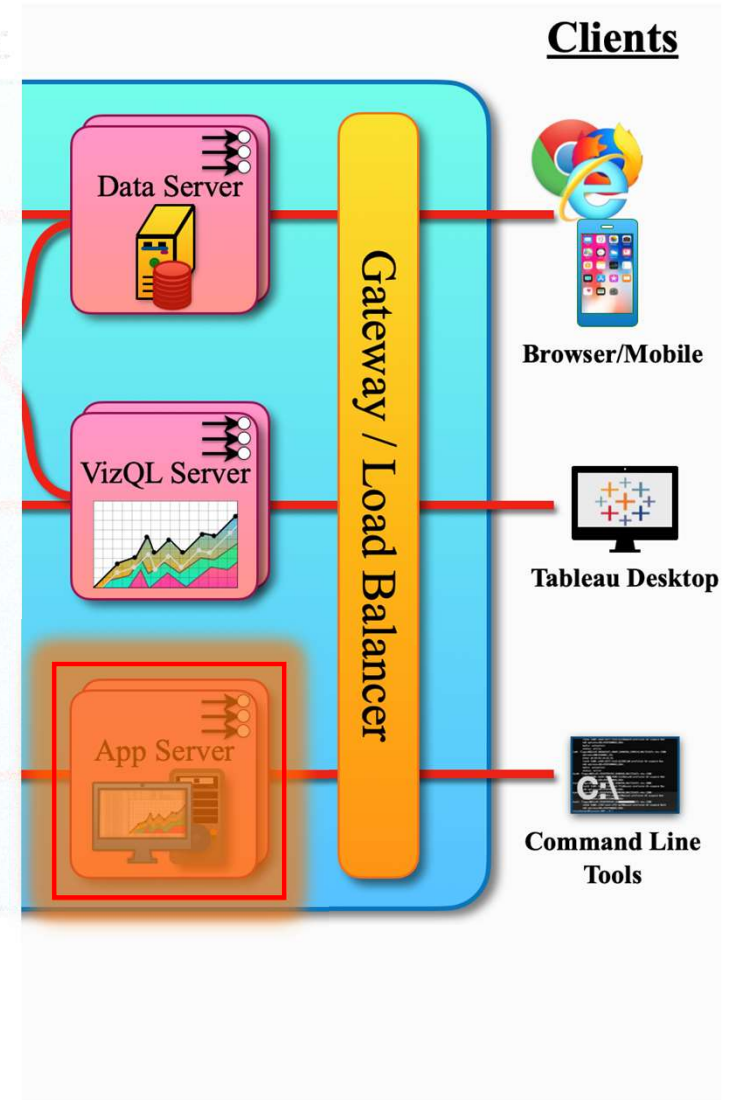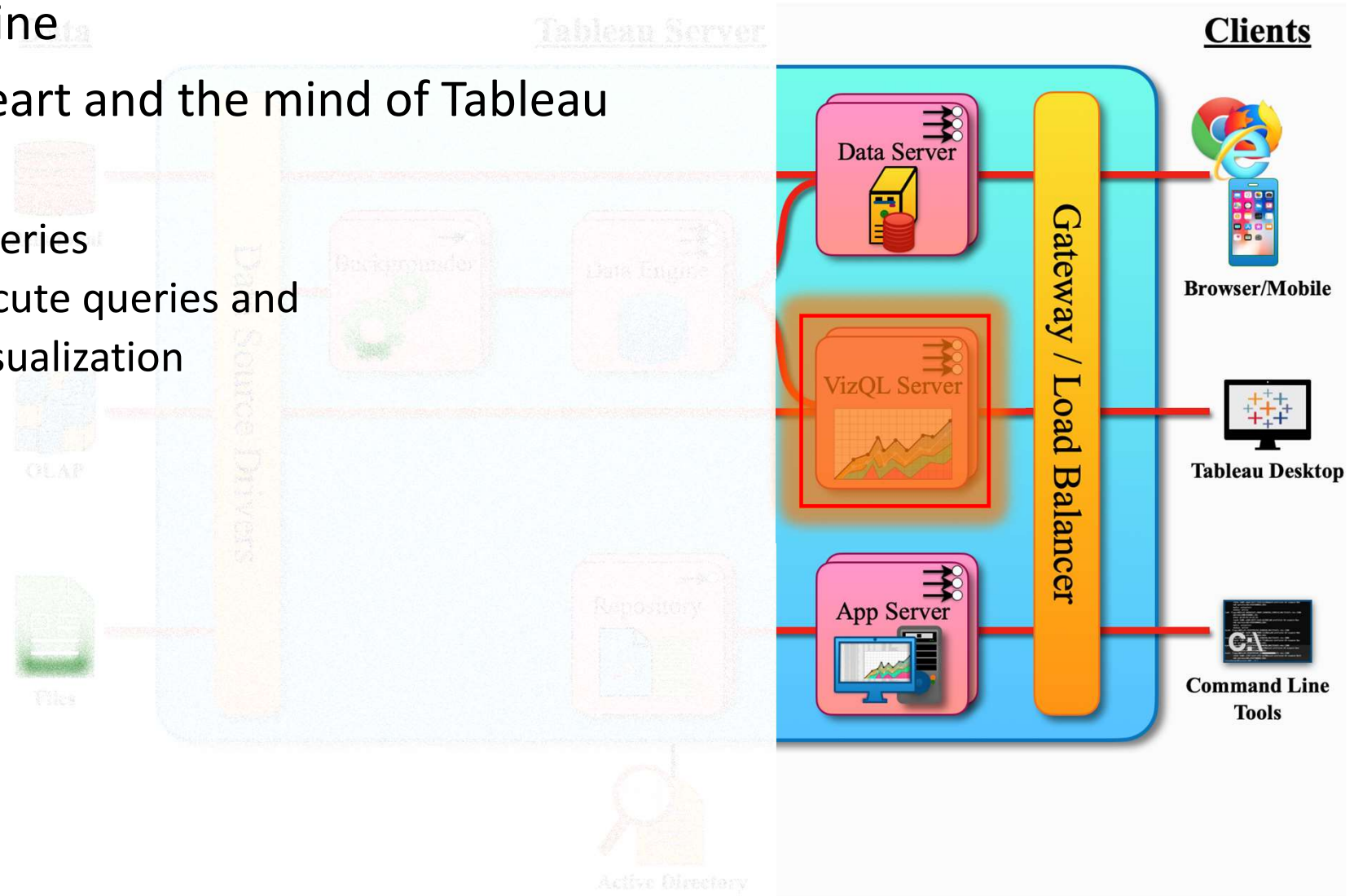- It also performs the content management, searches, and the process of the Virtualization
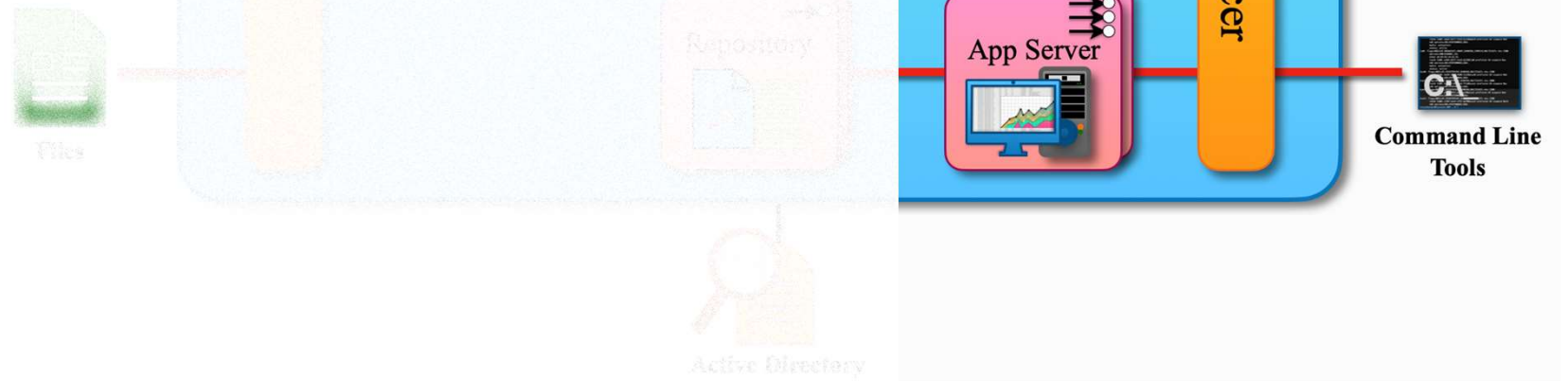
VizQL Server

- Visualization engine
- Also called the heart and the mind of Tableau
- Used to
  - Generate the queries
  - Process and execute queries and
  - Generate the Visualization

Data Server

- Works as a proxy for data sources and provides centralized management of database connection and Tableau Data Extracts

- The management of Tableau Data Extracts and database connection is done within the Tableau server and distributed with the workbooks

- The users who are using the Tableau Desktop can connect with these Data Extracts and run the required queries

Clients

Data Server

VizQL Server

App Server

Gateway / Load Balancer

Browser/Mobile

Tableau Desktop

Command Line Tools

Backgrounder
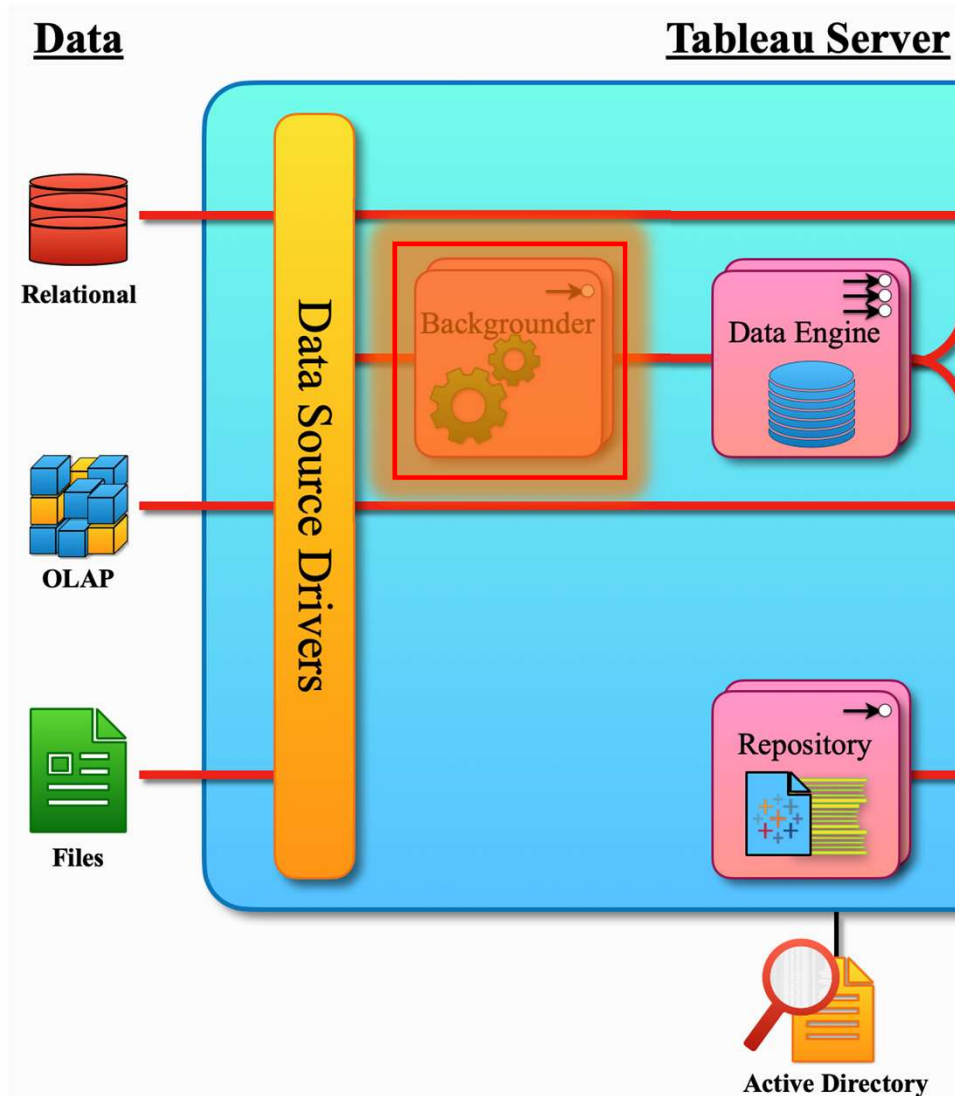
- **Refresh the Tableau extracts**
  - Extract is a compressed snapshot of data stored on disk and loaded into memory as required to render a Tableau viz
- **Compile the .tde and send it to data engine for execution**
  - The TDE data file format is used by Tableau Desktop
  - It may contain data extracted from external sources
  - TDE files are used for storing external data in a format recognized by the Tableau software

Data Engine
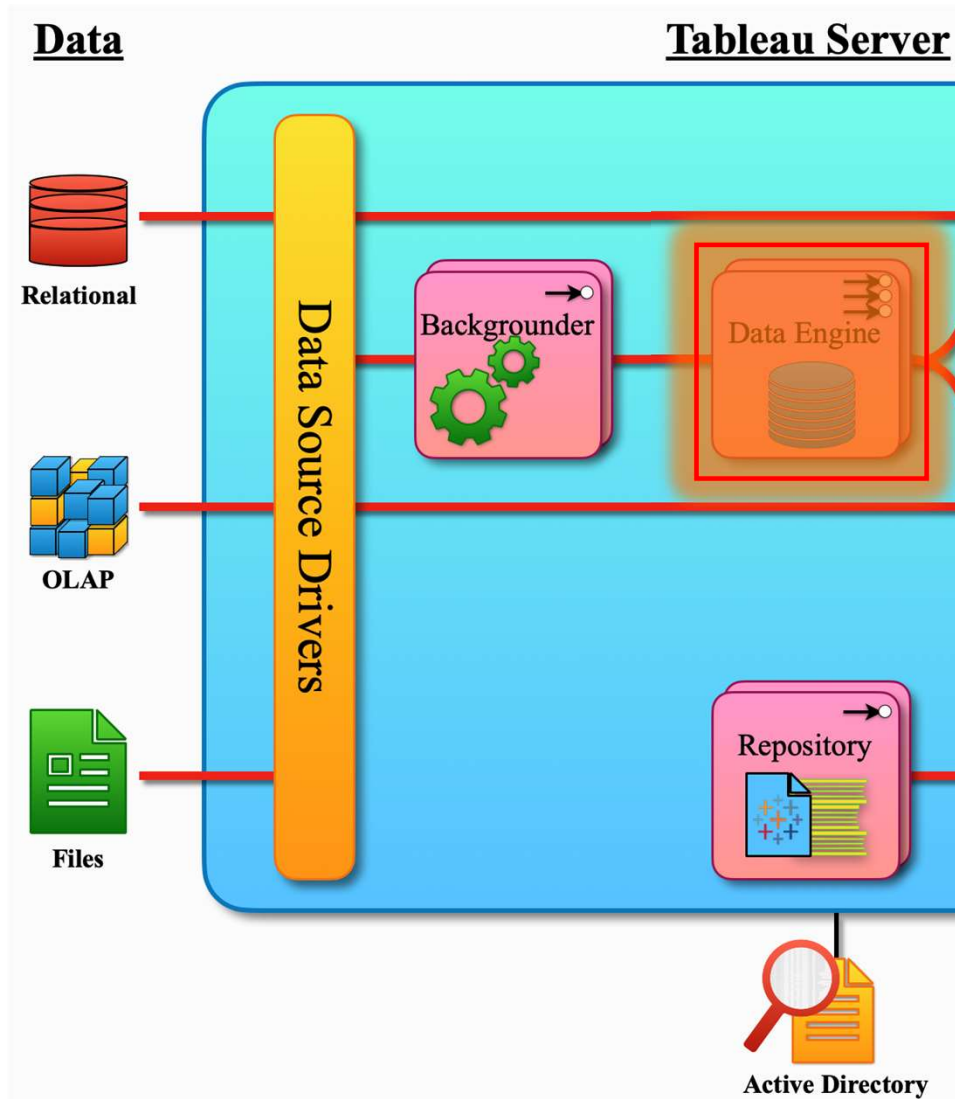
- Used to store the data extracts and work hand to hand with the backgrounder to process the queries
- It is used while creating, refreshing, or querying the extracts

Repository
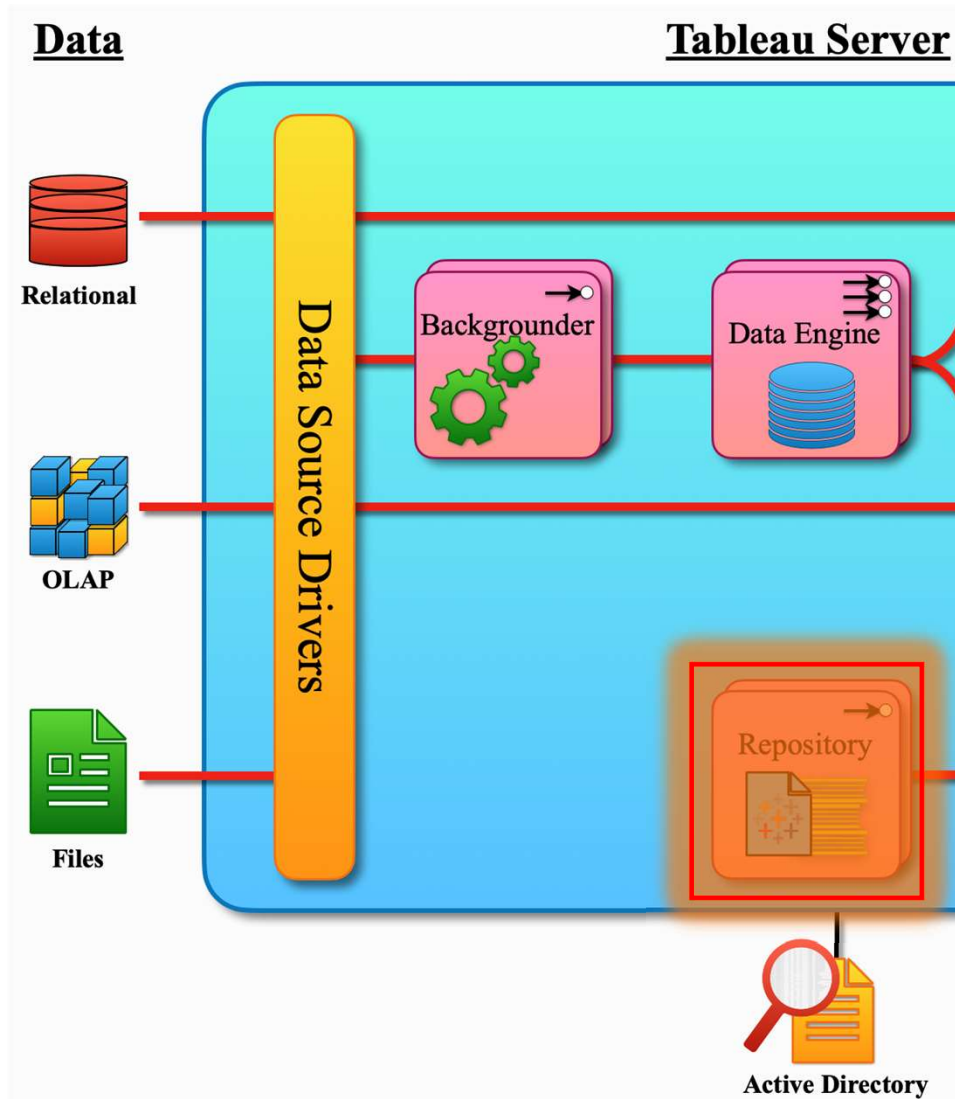
- The PostgreSQL database
- Stores the Tableau metadata and Server detail such as
  - Users
  - Groups assignments
  - Data sources
  - Permissions
  - Refreshes information, and
  - The extract information
- The Application Server and the VizQL Server talk to the Tableau Server Repository to verify the permission before accessing the views.

Search

- Used to perform the fast and filter search
- It is based on the Apache SOLR
- Works with the Application Server component of Tableau Server
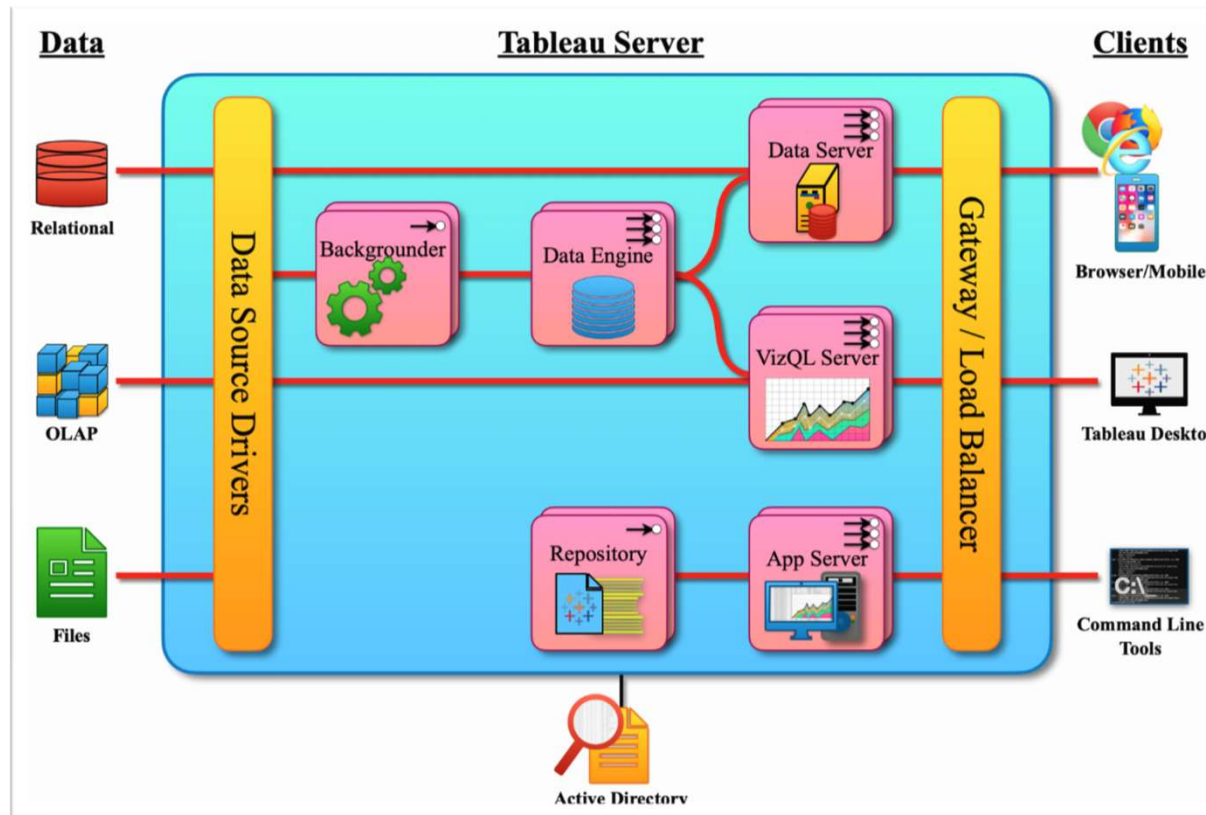
Licencing

- Verifies the registration of users with Tableau under the agreement term
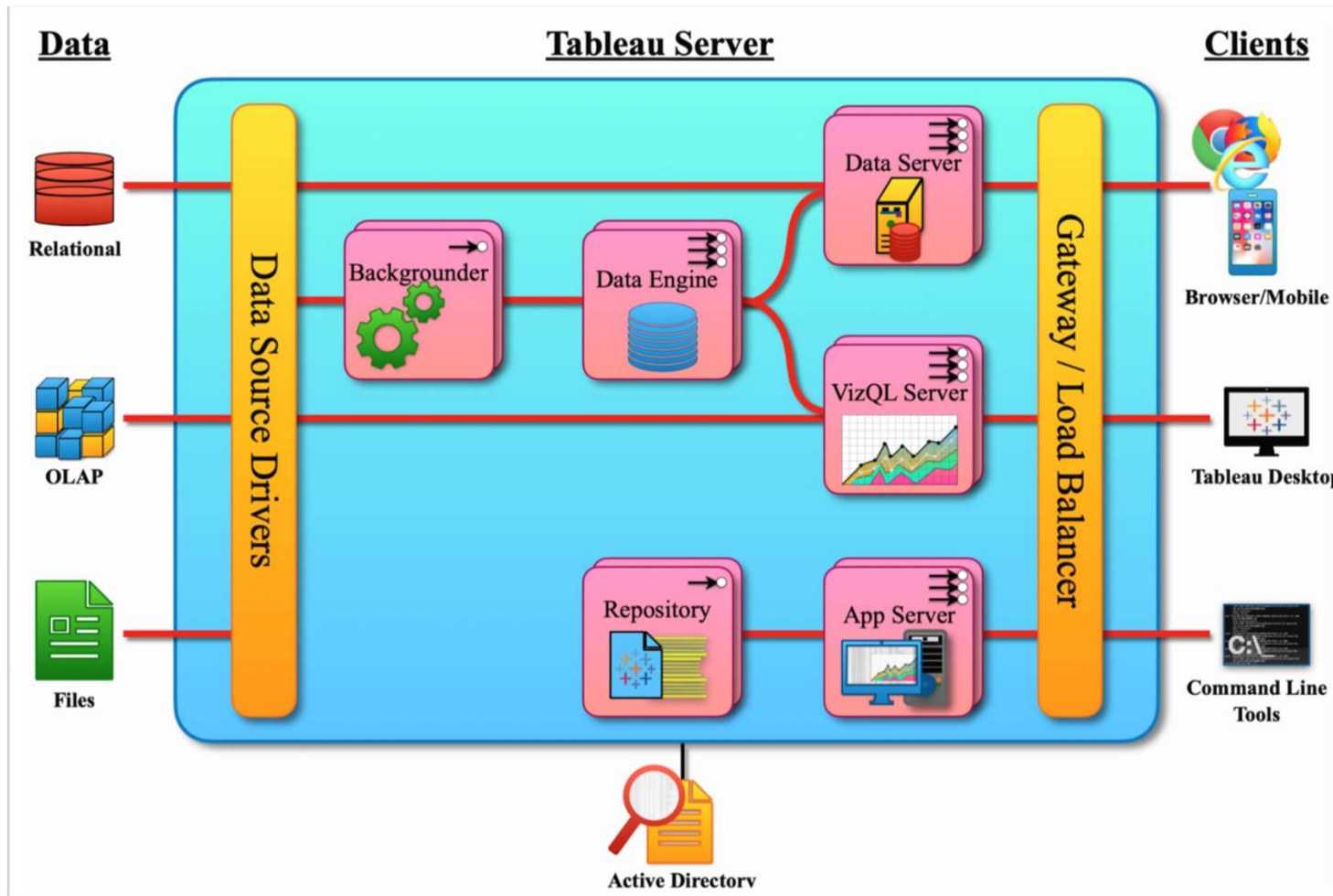
Clients (Web Browsers and Mobile Apps)

- Tableau Server supports:
  - Web browsers
  - Mobile Safari
  - iPad app
  - Android app
  - Android browser
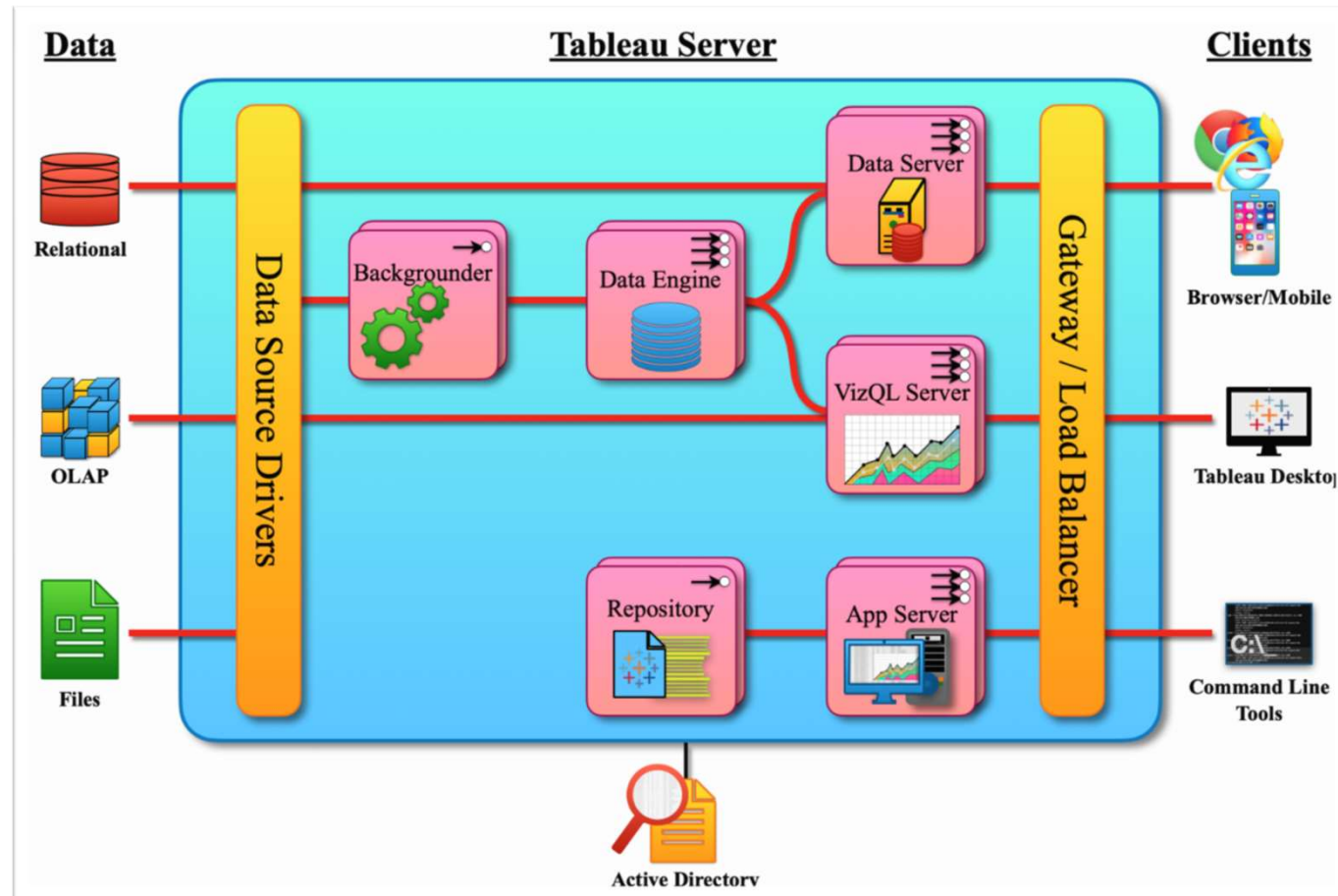
Clients (Tableau Desktop)

- To generate and publish sights, reports, and dashboards to Tableau Server

- Using this, a report author can
  - Tie with multiple data sources
  - Explore relationships
  - Create dashboards
  - Modify metadata and finally
  - Issue a completed workbook or data source to Tableau Server
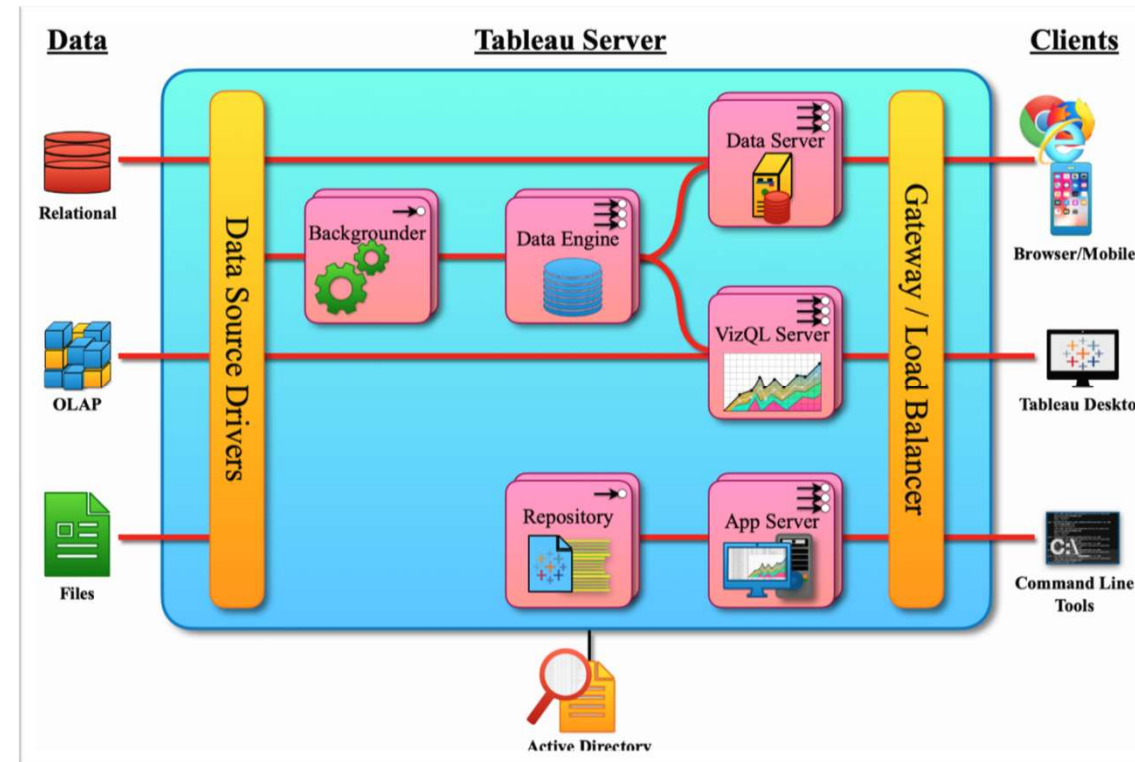
Tableau Server Processes

Licensed processes

Some of the processes that are installed as a part of Tableau Server are "licensed" processes

Licensed processes need a valid Tableau Server license in order to run

This has the following impact:

- Every licensed process needs to regularly contact the Tableau Server License Manager service to verify they are licensed
- If they cannot confirm there is a valid license the process will not run and Tableau Server may not function properly or reliably

Tableau Microservice Container Processes

These processes are automatically added when the first instance is added to a node

If all microservices are running, the container process has a status of running

If all microservices are stopped, the container process status is error

If one or more microservices is running while others are not, the container service has a status of degraded
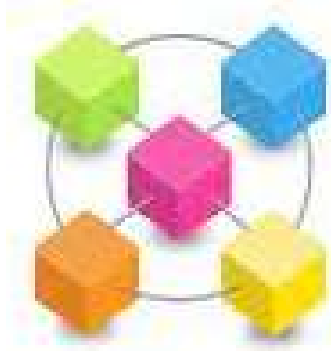
Tableau Services Manager (TSM) Processes

TSM is used to manage installation and configuration of Tableau Server.

These processes have a status of running once TSM has been initialized

Tableau Server Maintenance Processes

- There are three TSM Maintenance Services that are installed on every node of the cluster
  - Database Maintenance
  - Backup/Restore

Process workflow

- Allows individuals to generate actionable insights from data and share them with other business teams
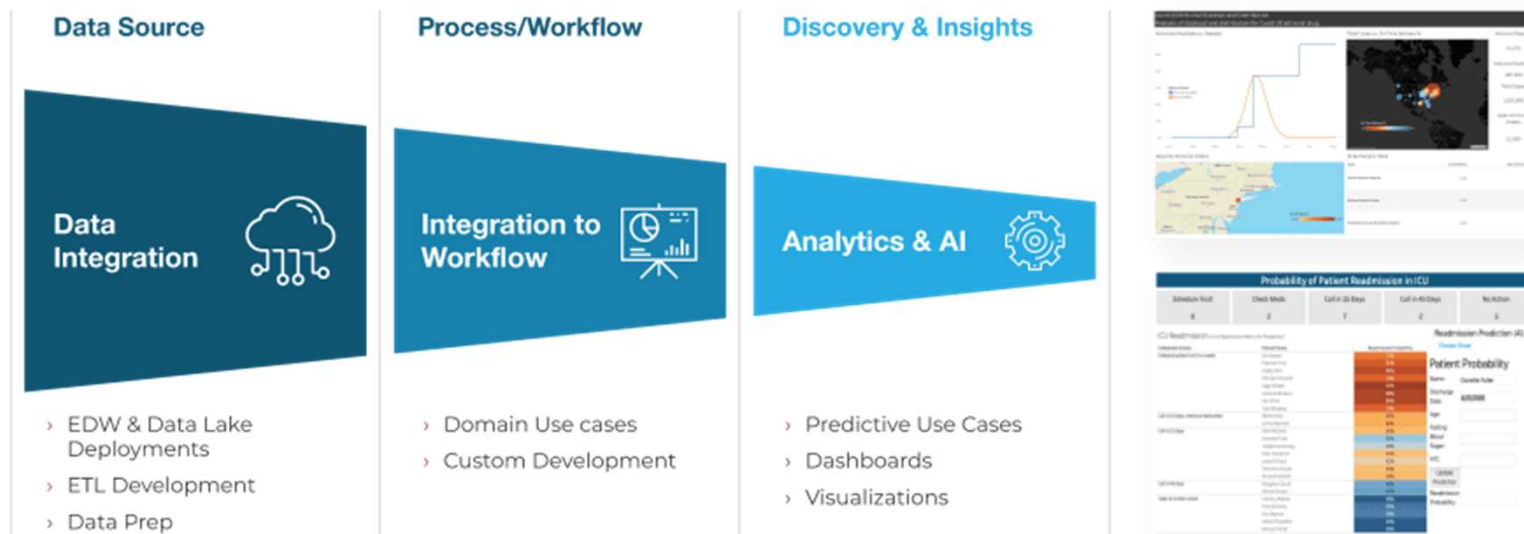
What is Tableau Prep?

- Makes it easier to combine, clean, shape, and share your data
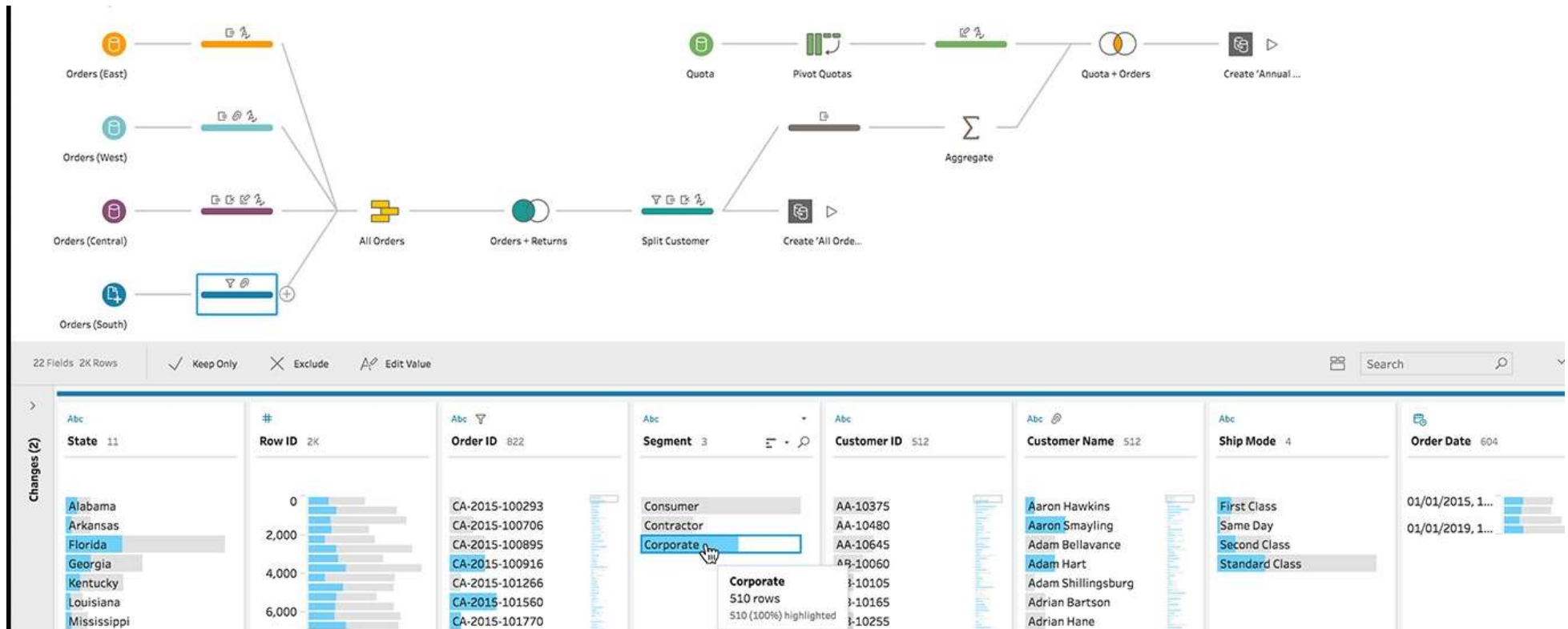
What is Tableau Desktop?

• Leading data visualization tool used for data analysis and business intelligence
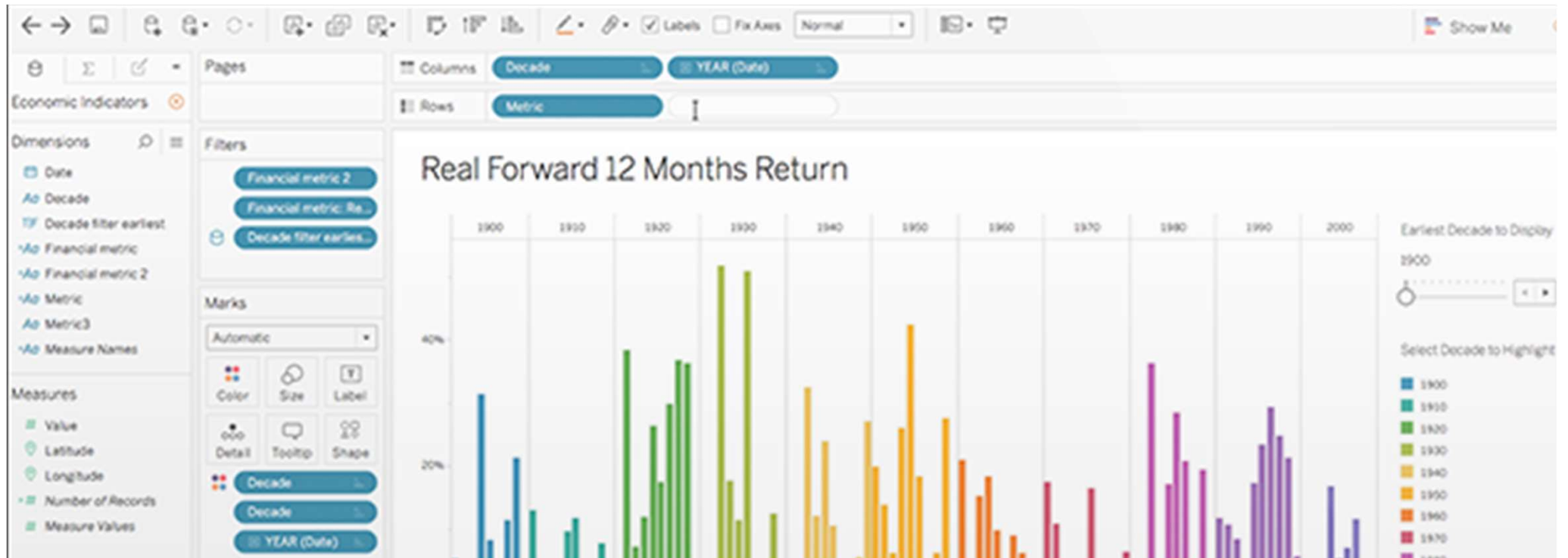
Tableau Desktop vs Tableau Prep

- Tableau Desktop is a data visualization tool for creating interactive dashboards and reports, while
  - https://www.tableau.com/products/desktop/download

- Tableau Prep is a data preparation and ETL tool for cleaning and prepping data for analysis.
  - https://www.tableau.com/products/prep/download

Tableau Server Portal

- Online platform

- Allows you to host and manage
  - Tableau data sources
  - Workbooks
  - Reports and
  - Dashboards created on Tableau Desktop
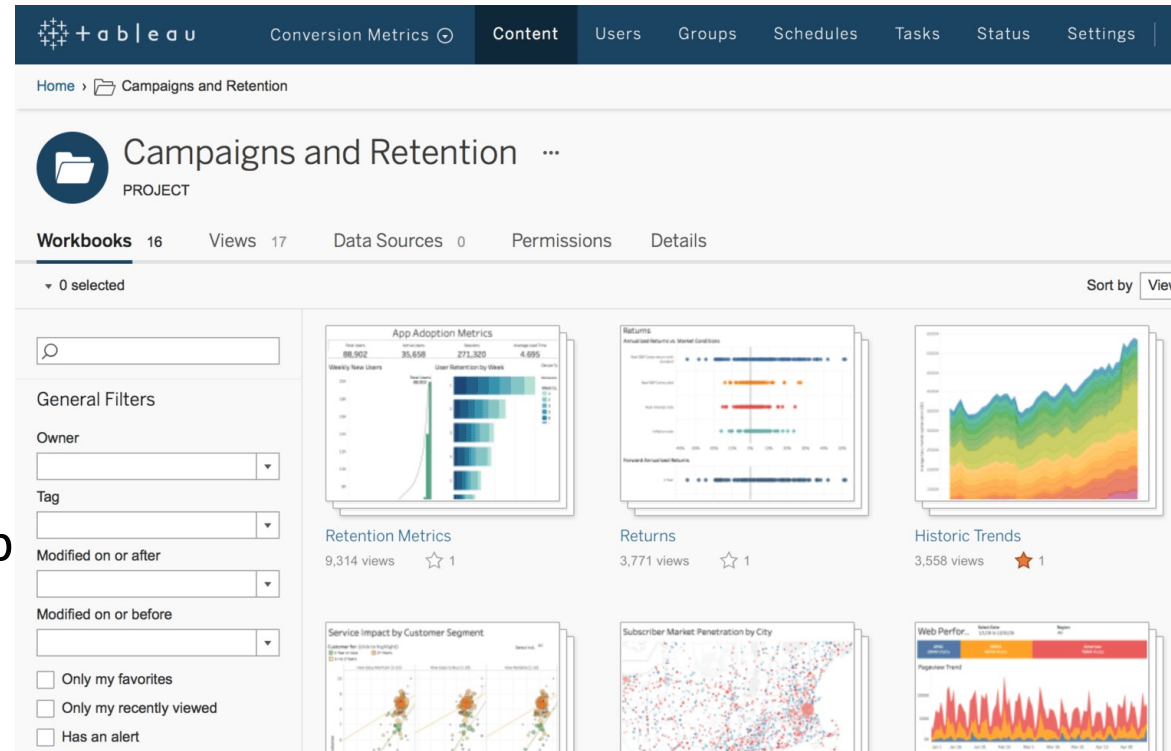
- Can access Tableau Server Portal from a web browser

- To create
  - New workspaces
  - Publish reports and dashboards
  - Share them with other users

TSM (Tableau Server Manger)

- Comprehensive tool for installing and managing Tableau Server
- Contains an enhanced user experience for server administrators with both
  - CLI and
  - Web interface

Extract

Compressed snapshots of data

Optimized for aggregation and loaded into system memory

To be quickly recalled for visualization

There are two options to refresh the data in the extract:

- Full Extracts
  - Rewrites the existing data extract
- Incremental Refresh
  - Help to add new records that have been added since the last extract has been created

LIVE connection

- The data source that contains a direct connection to underlying data, relies on a database for all queries

Data source connections (Live vs Extract)

- Pros of Live
  - Real-time updates

- Cons of Live
  - Databases are not always optimized for fast performance
  - Stress
    - On databases
  - Other factors can affect speed
    - Poor network
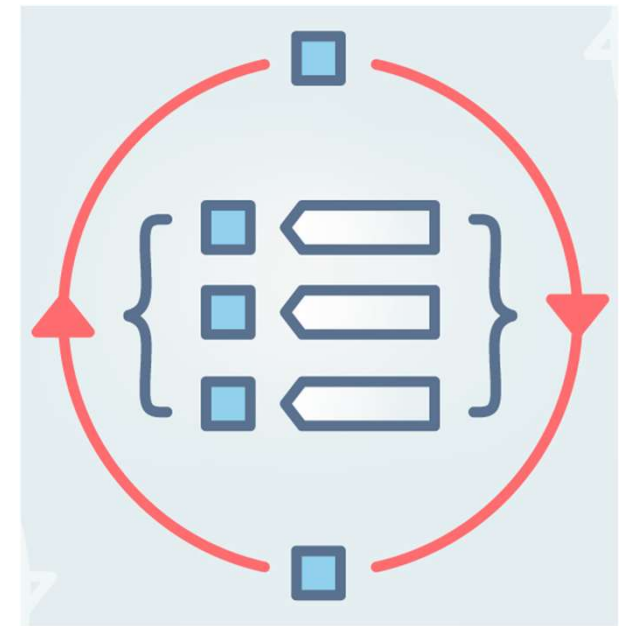
Data source connections (Live vs Extract)

- Pros of Extract
  - Speeds up the workbook through optimization
  - Offline
  - Portability
  - Reduce Load
- Cons of Extract
  - Snapshot
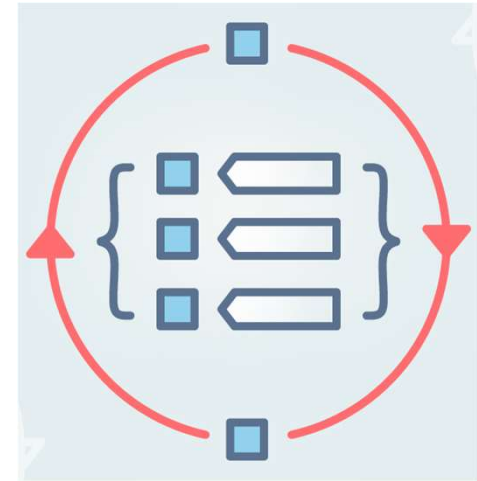   - Because data is extracted from the data source, the data will remain the same unless it is refreshed.
  - Size/Structure
    - Extracts can become slow to refresh and query depending on the data structure. E.g: many columns and structures

Data source connections (Live vs Extract)

Factor	TDE	Live Connection
Performance and Preference	Workbook performance is the preference	Live and fresh data is important i.e If business decisions need to be predicated upon real-time data
Time-based usage	When you want to visualize daily/monthly/weekly trends	When you want every minute update
Publishing to Tableau Public and Tableau Online	Tableau Public only supports TDEs.	Tableau Online supports Live connection
Dependency on Database	Doesn't rely on database queries after extract is created	Rely on database queries
Recommendation acc. to Data Source	Recommended when there is a multi-connection data source	Not recommended where there is a multi-connection data source because it slows down the performance
Refresh	Need to schedule refresh	No need to schedule refresh, it's a real-time update connection
Use Cases	If hospitals visualize the daily/weekly trends of patient data.	Hospitals monitoring incoming patient data need to make real-time decisions

Tableau Server Logs and Log File Locations

- Most of the Tableau Server logs are written to the data directory
  - /var/opt/tableau/tableau_server/data/tabsvc/logs/

- Subdirectories are created for each instance of a service
  - /var/opt/tableau/tableau_server/data/tabsvc/logs/backgrounder

Configuration file locations

- /var/opt/tableau/tableau_server/data/tabsvc/config

Logs that are not written in the primary location

- The TSM log
  - The tsm.log file is located in <home dir>/.tableau/tsm
- The install log
  - The app-install.log file is located in /var/opt/tableau/tableau_server/logs
- The upgrade log
  - The app-upgrade.log file is located in /var/opt/tableau/tableau_server/logs
- Bash script logs
  - Most Tableau Server bash scripts located in the /scripts directory (/opt/tableau/tableau_server/packages/scripts.<version>) generate their own logs
  - These are written to the /var/tmp directory each time a script is run

How content in log files look like?

- A typical Tableau log looks as follows:
  - {"ts":"2021-02-19T16:22:57.834","pid":10221,"tid":"5a1628","sev":"info","req":"-","sess":"-","site":"-","user":"-","k":"open-log","v":{"path":"/Users/tfoldi/Documents/Tableau Repository/Logs/log.txt"}}

- a ts (timestamp)
- pid (process id)
- tid (thread id within the process)
- sev (severity)
- req (request identified)
- sess (Server vizql session/Desktop session)
- site (site in Server)
- user (user in Server)
- log keys (k)
- (v) values.

Tableau Server Scripts

- /var/opt/tableau/tableau_server/packages/scripts

Thanks