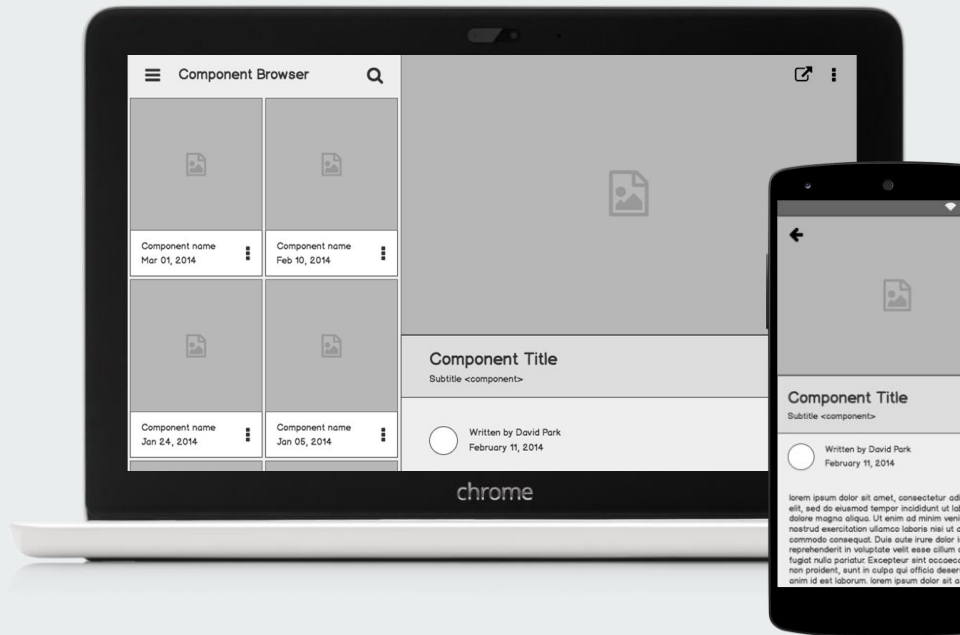




Architecture Design

Amazon sales Data Analysis





Document Overview

Document Version: 1.0

Date: 29-09-2023

Author: Amar Shilvanth

Status: Final

Audience: Business Stakeholders

Table of Contents

- Executive Summary
- Architecture Overview
- Data Flow
- Security
- Scalability
- Monitoring and Maintenance
- Backup and Recovery
- Performance Optimization
- Conclusion



Executive Summary

This document outlines the architecture design for our Tableau implementation, which aims to provide a robust and scalable data analytics platform to support decision-making processes within our organization. The design ensures data security, performance, and usability while accommodating future growth and changes.



Architecture Overview

Our Tableau architecture consists of the following key components:

Data Sources: These include various structured and unstructured data stores, databases, and APIs that feed data into Tableau for analysis.



Architecture Overview

Tableau Server: The central server where Tableau workbooks and dashboards are published and shared with users. It handles authentication, authorization, and user access.

Tableau Desktop: The authoring tool for creating and editing Tableau content, including workbooks and dashboards.



Architecture Overview

Tableau Mobile: The mobile application for accessing Tableau content on smartphones and tablets.

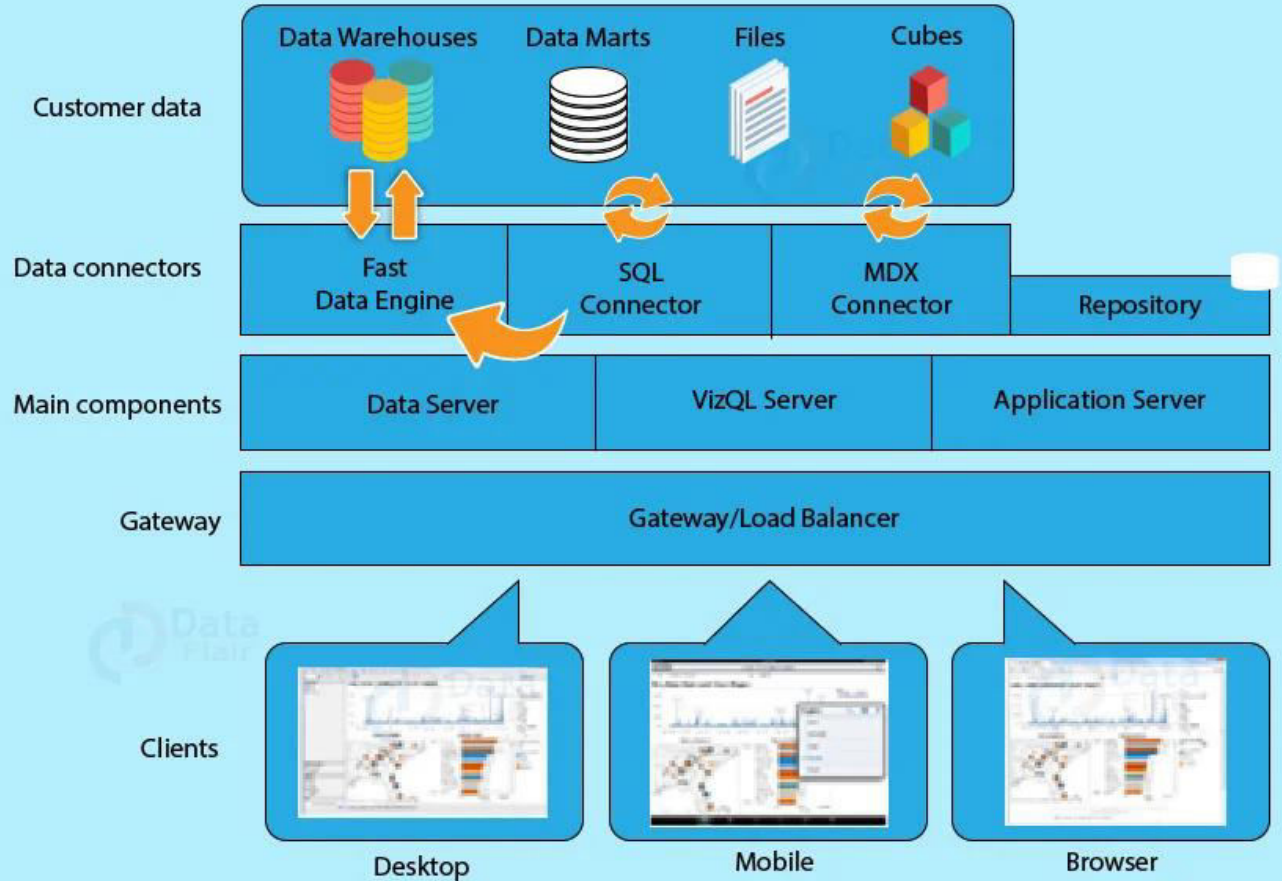


Data Flow



Tableau

Architecture





Security

Authentication: Implement strong authentication mechanisms, such as multi-factor authentication (MFA), to ensure that only authorized users can access Tableau.

Authorization: Set up role-based access control to restrict access to sensitive data and dashboards.

Data Encryption: Encrypt data at rest and in transit to protect against data breaches.

Audit Logs: Enable and regularly review audit logs to monitor user activity and identify potential security threats.



Scalability

Hardware Resources: Continuously monitor hardware resources like CPU, RAM, and storage. Scale up as needed to accommodate growing data and user loads.

Cluster Configuration: Consider deploying Tableau Server in a clustered environment to distribute workloads and improve performance.

Load Balancing: Implement load balancing to evenly distribute requests among server nodes.



Monitoring & Maintenance

Performance Monitoring: Set up monitoring tools to track server performance, identify bottlenecks, and proactively resolve issues.

Regular Updates: Keep Tableau Server and associated components up to date with the latest patches and updates.

Data Source Maintenance: Regularly update and optimize data sources to ensure data accuracy and performance.



Backup & Recovery

Regular Backups: Schedule regular backups of Tableau Server configurations, workbooks, and data sources.

Disaster Recovery Plan: Develop a comprehensive disaster recovery plan to quickly restore Tableau Server in case of data loss or system failure.

Testing: Periodically test the backup and recovery processes to ensure they work as expected.



Performance Optimization

Query Optimization: Optimize data queries and calculations within Tableau to improve dashboard load times.

Data Extracts: Use data extracts (TDEs) to speed up performance for large datasets.

Caching: Utilize caching mechanisms to reduce the load on the database and improve response times.

Indexing: Ensure that your data sources are properly indexed for faster query execution.

Conclusion

User Training: Provide training and resources to Tableau users to ensure they can leverage the platform effectively.

Feedback and Improvement: Continuously gather feedback from users to identify areas for improvement and new feature requests.

Documentation: Maintain up-to-date documentation for Tableau configurations, data sources, and best practices.

Scalability Planning: Develop a long-term scalability plan to accommodate the organization's growing data and user needs.