WELCOME

MongoDB SI Certification Presentation

Name: Antonio Abate

Date: 06/11/2024

Agenda

Use-case Description

Problem Statement

Current Architecture

Proposed Architecture

3 Whys

Modernization Scorecard

Competition Comparison Analysis

Customer Case Study

USE CASE DESCRIPTION

The main purpose of the project is to have a Business Intelligence application, in order to obtain value and insight from customer's data. The main focus is to allow authorized users to view dashboards via browser and to enable data analysis and queries, both for time series and predictive analysis. Currently, the application uses relational databases as data sources and an open-source tool to perform data exploration and data visualization tasks.

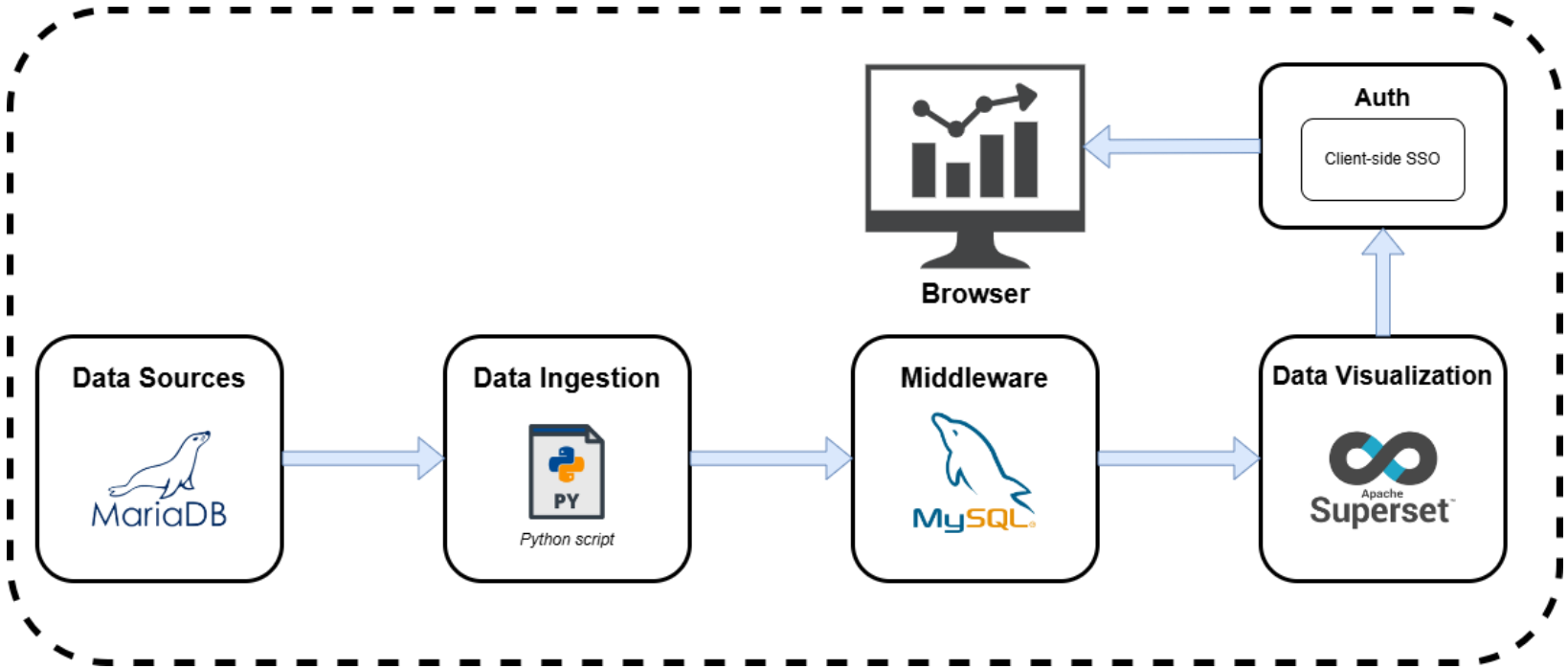
PROBLEM STATEMENT

1. Scalability & Performance Issue
    - Relational databases are not optimal for handling large, fast-growing dataset, leading to slow data retrieval and lagging dashboards.
2. Data Flexibility
    - The evolving nature of application data requires a flexible schema that can evolve easily.
3. Data Redundancy and Fault Tollerance
    - The existing architecture does not support data redundancy and automated backups.

PROBLEM STATEMENT

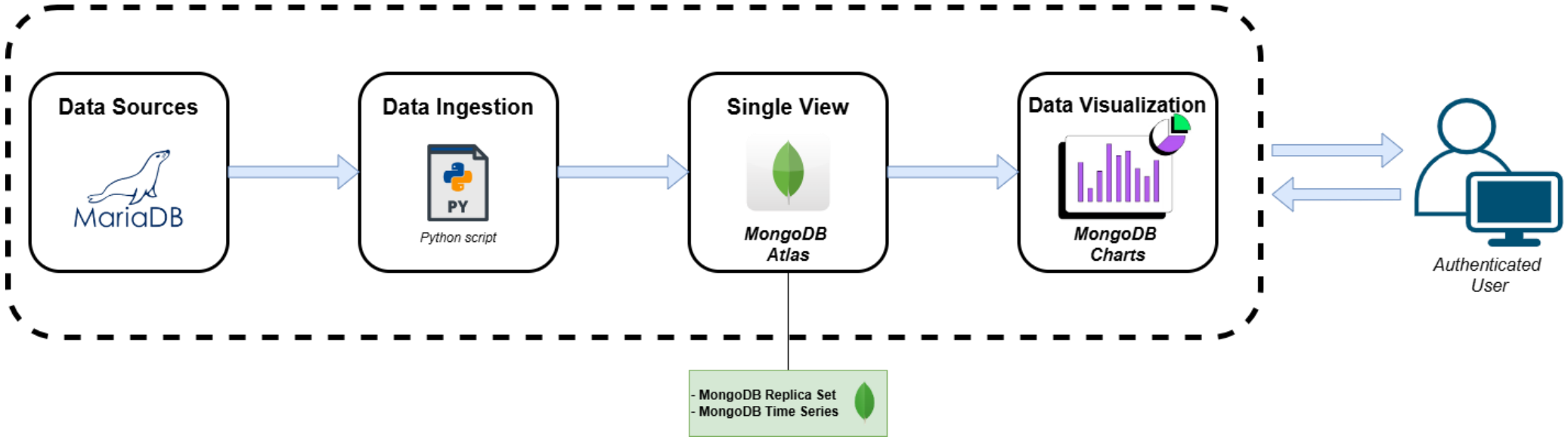
4. Application Maintenance
    - Application components requires manual intervention, for backups or patching, increasing the risk of downtime.
5. Time Series and Predictive Analysis
    - Relational databases do not handle time-series data efficiently. This is critial for the application, as the client needs to track patterns in data and predict future trends.

CURRENT ARCHITECTURE



The application uses **MariaDB** to store data, and through a Python script, data is sent to a **MySQL** database that serves as a *Middleware* to perform BI analysis using **Apache Superset**.

PROPOSED ARCHITECTURE



PROPOSED ARCHITECTURE

To modernize the existing architecture, we recommend transitioning to a Single View database in cloud, using MongoDB Atlas, configured with Replica set and combined with MongoDB Time Series. Additionally we propose using MongoDB Charts as a data visualization tool.

- **MongoDB Atlas** provides a fully managed, cloud-based database service with high availability, automated backups, and easy scalability, that allows our customer to scale its infrastructure as data volume grows, reducing overhead for infrastructure management and enhancing reliability.
    - **Replica Set** ensures redundancy and data reliability by maintaining multiple copies of data across distributed nodes, in order to enhances fault tolerance and obtain an uninterrupted access to BI data.

PROPOSED ARCHITECTURE

- **MongoDB Time Series** collection is optimized for storing and querying time-stamped data, helps tracking changes in metrics over time and implementing predictive analytics.

- **MongoDB Charts** as a powerful tool for creating and sharing data visualizations directly from MongoDB data, without additional ETL steps, provides an intuitive interface for building interactive dashboards and sharing insights across our customer's organization.

3 WHYs

Why do anything?

- **Performance improvement:** MongoDB offers horizontal scalability and high-performance querying capabilities, which are essential for applications with growing datasets and concurrent user demands.

- **Data flexibility and Schema evolution:** MongoDB's schema-less design provides flexibility in handling changes to data structure without the need for rigid schema definitions typical of relational databases.

- **Application Maintenance:** by leveraging MongoDB Atlas, as a managed cloud database solution, the customer reduces the need for in-house database maintenance and updates.

- **Better Time-Series Analysis:** the application's need for time-series analysis is well-supported by MongoDB's dedicated Time Series collections, optimized for handling and querying time-stamped data efficiently. This feature is crucial for tracking metrics, usage patterns, and other temporal data.

- **Improved Data Visualization Integration:** transitioning to MongoDB Charts aligns data visualization closely with the MongoDB data source, reducing dependency on additional middleware and enabling a more direct, efficient pathway for visual data exploration.

3 WHYs

Why Now?

- **Costs and resources allocation:** efforts required to manage and maintain outdated systems may pull resources away from more strategic, value-adding projects, impacting overall efficiency.
- **Analytical Accuracy and responsiveness:** as data volumes grow, the current relational database setup will struggle to keep up with real-time or near real-time demands, slowing down decision-making and impairing the ability to adapt quickly to evolving business needs.
- **Maintenance overhead:** maintaining a complex, multi-database setup not only requires additional effort from IT teams but also introduces more points of failure, which can lead to downtime or data inconsistencies that disrupt business intelligence processes.
- **Inflexibility for new Requirements:** as the business environment evolves, a rigid architecture will be less adaptable to new BI features and functionalities that are crucial for predictive analytics, sustainability tracking, and digital transformation analysis.
- **Missed insights:** without a robust, modern BI solution, the customer risks missing key insights from their data.

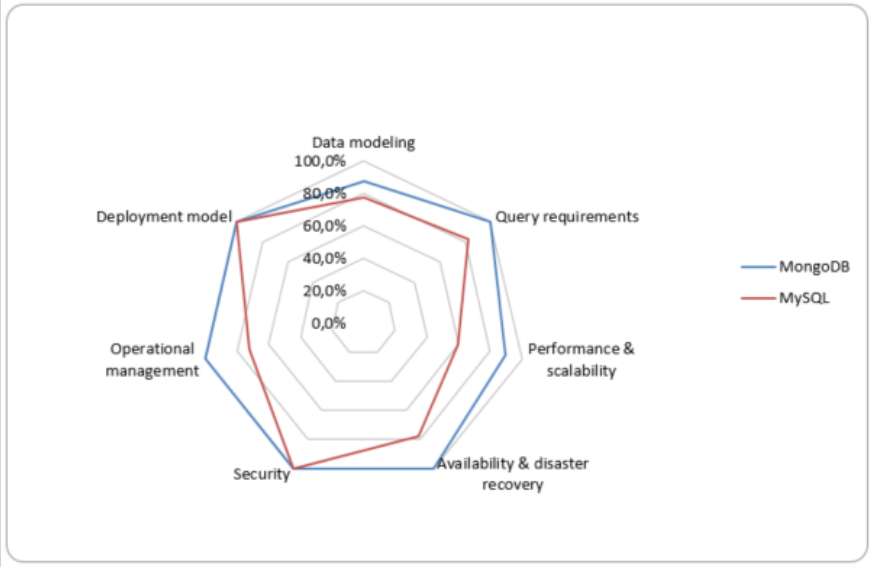
3 WHYs

Why MongoDB?

MongoDB is an ideal choice for modernizing the client's Business Intelligence platform because it offers a flexible, scalable, and modernized architecture that aligns well with the data needs and growth objectives. Its document model supports diverse data types, making it easier to manage complex and evolving information without rigid schemas. MongoDB's built-in time series capabilities enable efficient handling of time-based data, allowing to track trends in data over time while providing near real-time insights. With MongoDB Atlas, the client benefits from automated maintenance, enhanced security, and cloud-native scalability, which reduce operational overhead and allow the focus to remain on actionable insights. Additionally, MongoDB Charts provides integrated data visualization, simplifying the creation of interactive dashboards and enhancing data accessibility for authorized users.

MODERNIZATION SCORECARD

BI Platform - MongoDB Modernization Scorecard Results					
Category	**Maximum**	**MongoDB**	**MySQL**	**MongoDB**	**MySQL**
Data modeling	120	105	93	87,5%	77,5%
Query requirements	51	51	42	100,0%	82,4%
Performance & scalability	30	27	18	90,0%	60,0%
Availability & disaster recovery	27	27	21	100,0%	77,8%
Security	48	48	48	100,0%	100,0%
Operational management	33	33	24	100,0%	72,7%
Deployment model	9	9	9	100,0%	100,0%
Total	**318**	**300**	**255**	**94,3%**	**80,2%**



COMPETITON COMPARISON

Feature	Data Model	Horizontal Scalability	Vertical Scalability	Query Language	ACID Support	Licensing	Cost	Scalability Approach	Indexing	Data Consistency	Security
MongoDB	Schema-free, NoSQL, Document-oriented	Excellent	Good (Adding more resources to a single node)	MQL (MongoDB Query Language)	Configurable	SSPL (Server Side Public License)	Generally more cost-effective for horizontal scaling	Automatic sharding for horizontal scaling, suitable for large datasets	Dynamic indexing with automatic and manual indexing options	Eventual consistency by default, configurable for stronger consistency	Role-based access control (RBAC), SSL/TLS encryption, authentication mechanisms, and auditing
MySQL	Relational	Limited (Master-Slave Replication)	Good (Adding more resources to a single node)	SQL (Structured Query Language)	Strong (ACID guarantees)	Open-source (Community and Enterprise versions)	Free (Community version), Paid (Enterprise version)	Primarily vertical scaling	Static (Indexes need to be defined upfront)	Strong consistency (Data is always consistent across nodes)	Role-Based Access Control (RBAC), Encryption

CUSTOMER USE CASE

Meltwater

Challenge:

The competitive landscape for media listening platforms has evolved significantly, prompting Meltwater to transition to a cloud-first strategy using MongoDB. This shift involved moving from a monolithic infrastructure to a microservices environment, aimed at enhancing their product suite in response to market demands.

Solution:

The Meltwater team utilizes MongoDB Charts to analyze user journeys, enhancing their architecture and feature design while improving the overall user experience. Insights from this data enable developers to make informed, data-driven decisions, from roadmap strategies to granular choices like dropdown result displays. The integration of MongoDB Charts with MongoDB Atlas allows product owners direct access to metrics without needing to request data from other teams, facilitating efficient visualization and eliminating the need for additional software or data duplication.

Result:

By moving to a DevOps approach supported by MongoDB Atlas, Meltwater is releasing new features in a much faster and data-driven manner, thanks to the insights it can quickly obtain through MongoDB Charts. As a result:

- Countless hours saved
- Creating data visualizations and providing new insights in minutes
- Efficiently deliver insights that drive business decision

THANK YOU