Secure and Scalable Cloud Storage Platform

Cristian Casas, Dylan Solarte

Distrital University Francisco José De Caldas

ABSTRACT

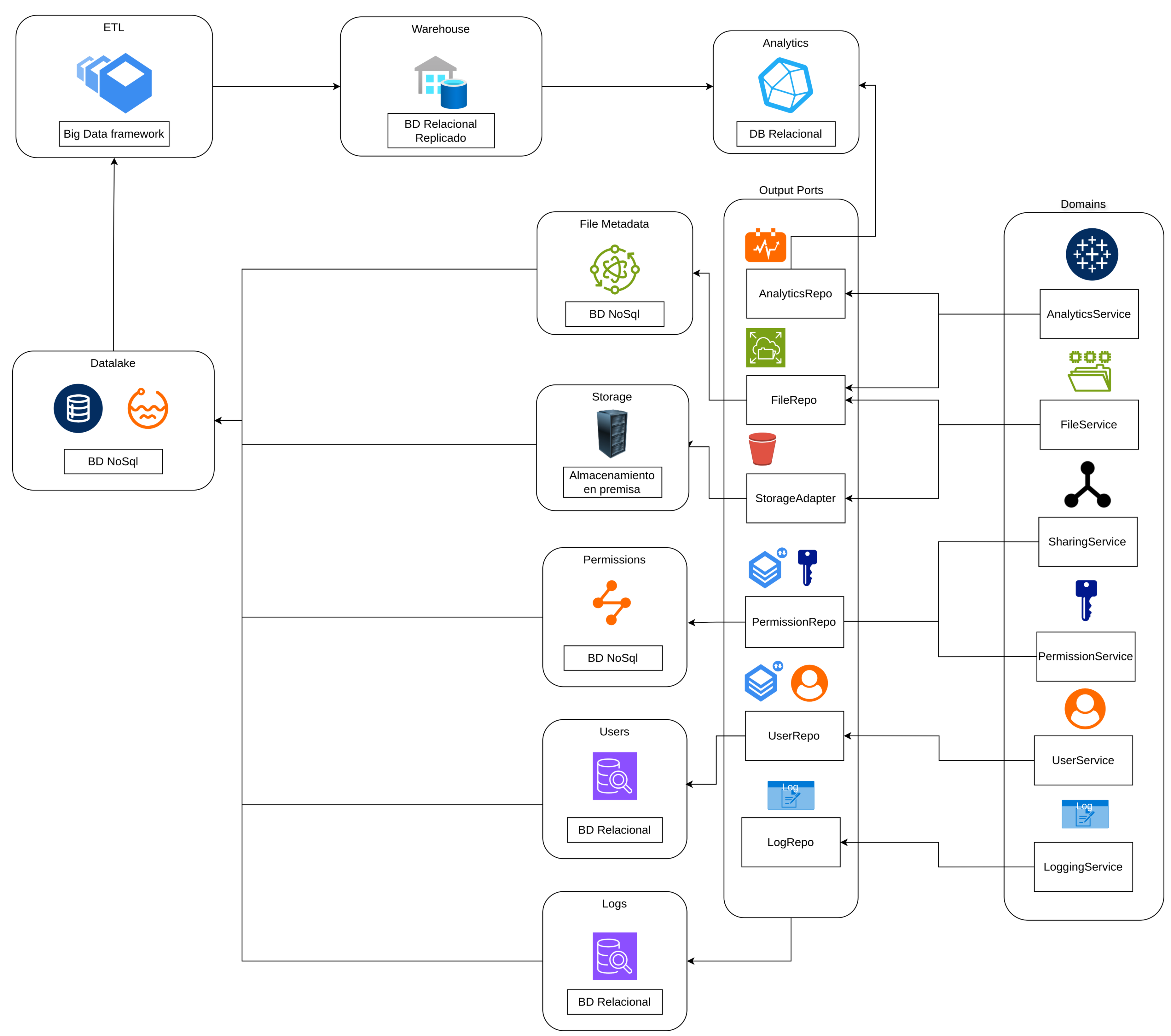
We present the design of a scalable, secure cloud storage platform inspired by Google Drive, focused on modularity and database architecture. The system adopts a hybrid storage model, combining relational and non-relational databases to handle structured user data, flexible file metadata, logs, and analytics. A hexagonal architecture supports maintainability and independent scaling of domain services. Although not fully implemented, the system was validated with synthetic data and performance-tested queries, confirming its technical viability and efficiency under moderate load.

INTRODUCTION

Cloud storage platforms are essential for modern users and organizations, yet most commercial solutions lack transparency, custom analytics, or academic applicability. This project aims to build an open, modular cloud storage system tailored for learning and engineering environments. Key goals include scalable architecture, integrated analytics, and robust metadata handling — all grounded in database design best practices.

METHODOLOGY

The system is built using a hexagonal architecture, separating domain logic from infrastructure. It integrates a hybrid persistence model: PostgreSQL for structured data, MongoDB for file and folder metadata, and Neo4j (planned) for access permissions. Synthetic data was generated for four domains — User, File, Logging, and Analytics — totaling over 150,000 records. The design also considers big data components: a data lake, ETL module, data warehouse, and analytics-focused data mart.



RESULTS

The system was deployed using free-tier cloud infrastructure and populated with realistic data. Query execution times were benchmarked across domains: most relational queries completed under 5 ms, and non-relational queries under 40 ms. While key modules like Neo4j permissions and the analytics engine remain unimplemented, mockups were designed to guide future integration. The architecture and schema demonstrated strong performance, validating the design under simulated workloads.

CONCLUSION

This project establishes a solid foundation for a modular, scalable cloud storage system aligned with academic and engineering needs. The architecture supports domain separation, concurrency control, and big data readiness. While several components remain conceptual, the performance and design quality indicate a high potential for future extension, particularly in analytics, access control, and real-world deployment scenarios.