



Case Study

Streamlining Data Delivery and Quality Optimization in Enterprise Systems

Client: Premium Apparel Company

Key highlights



Configurable Guardrail
Implemented data filtering and notification system for issue detection.



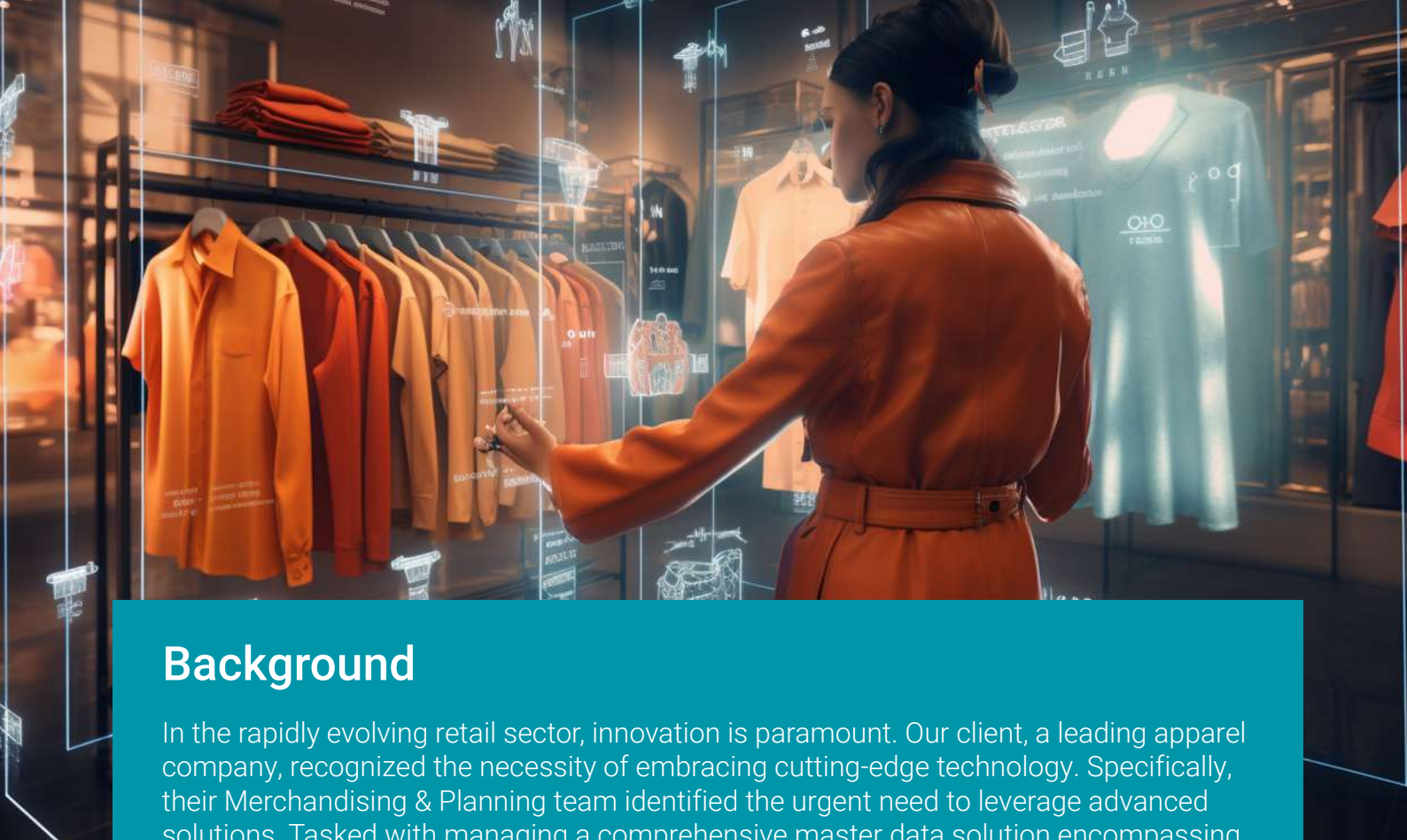
Faster Data Delivery
Reduced overall data processing time from ~ 25 hours to ~6 hours, marking a significant **76%** improvement.



Near Real-Time Data
Achieved near real-time data availability for end customers, reducing the interval from every 3 hours to immediate.



Enhanced Data Tracking
Introduced an end-to-end telemetry dashboard for monitoring, debugging, and altering data movement across the entire pipeline.



Background

In the rapidly evolving retail sector, innovation is paramount. Our client, a leading apparel company, recognized the necessity of embracing cutting-edge technology. Specifically, their Merchandising & Planning team identified the urgent need to leverage advanced solutions. Tasked with managing a comprehensive master data solution encompassing product, location, and vendor data across their enterprise, our client aimed to establish a scalable, efficient, and standardized framework. This framework would serve as the foundation for a robust enterprise data-as-a-service model within a federated scalable architecture.

Seeking a strategic partnership with Altimetrik, they sought assistance in enhancing the reliability, scalability, and simplicity of their master data system.

Our focus was twofold: to enhance data quality and streamline the process of ingesting and validating data, thereby reducing the time required to make it available for consumption. The overarching goal was to elevate the quality of downstream systems, ultimately enhancing the customer experience.

Pain Point

Despite efforts to enforce business rules on data within the existing architecture, the team faced challenges. Frequent rule changes and the constant addition of new business attributes resulted in flawed data passing through, causing significant issues for end-users.

Additionally, the process of incorporating new data was time-consuming, taking approximately 25 hours, while updating existing data required 3 hours. This delay was largely due to a monolithic pipeline generating large files encompassing entire datasets, sometimes reaching gigabytes in size.

Debugging and resolving data movement issues were time-consuming and laborious, leading to customer dissatisfaction. The sequential processing of data also strained server resources, with memory and CPU usage peaking at 90%, further delaying data publication.



Key Objectives

Decrease the overall time for data availability



Implement an additional validation layer



Develop a dashboard for visualising data movement and identifying bottlenecks



Facilitate near real-time data availability



Enhance server efficiency

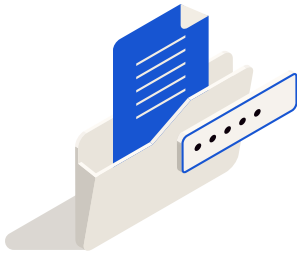


Solution



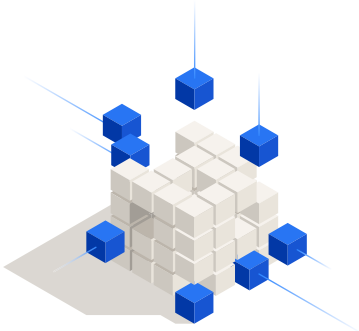
Validation Layer

Integrated an additional data validation layer just before the enterprise data exchange to effectively filter out flawed data and promptly notify the support team for swift resolution.



Parallel Processing

Fragmented the existing monolithic file into multiple smaller files, enabling parallel processing and significantly decreasing outbound file processing time from approximately 10 hours to just around 1 hour.



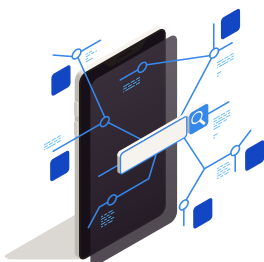
Real-time Data Availability

Transitioned from generating physical JSON files to directly querying data from the MDM system and transmitting it into Kafka every hour, eliminating the time required for file creation, publication, and retrieval.



Enhanced Data Tracking

Deployed a new correlation ID framework to meticulously track the flow of data from inbound to outbound, encompassing third-party systems, thereby furnishing comprehensive visibility and context for each data batch.



Optimized Queries

Enhanced efficiency by refining query logic to execute only a single query encompassing all IDs, resulting in a significant reduction in CPU usage to a mere 20%.

The Outcomes

Accelerated Data Availability

Transitioned from a 24-hour schedule to near real-time availability of master data for end consumers, ensuring they have access to up-to-date information whenever needed. Successfully reduced data errors, guaranteeing the prompt availability of accurate data to end consumers, thereby enhancing decision-making processes.

Enhanced Technical Infrastructure

Implemented architectural and design changes that effectively reduced TechDepth in the MDM system, resulting in improved data quality and optimized infrastructure utilization. These enhancements have led to improved efficiency and reliability in data processing and transmission.

Improved Data Quality

Implemented measures to enhance the quality of data transmitted to downstream systems, thereby facilitating accurate business decisions, and streamlining workflow processes. These measures have significantly contributed to the overall improvement of data integrity and reliability across the organization.

Alignment with Long-term Vision

Continued efforts are aimed at establishing a scalable, efficient, and reliable enterprise data-as-a-service system, with plans for further enhancements and adaptations to meet evolving business needs. These ongoing initiatives demonstrate our commitment to long-term success and innovation in data management and utilization.

About Altimetrik

Altimetrik is a pure play data and digital engineering services company focused on delivering business outcomes with an agile, product-oriented approach. Our digital business methodology provides a blueprint to develop, scale, and launch new products to market faster. Our team of 5,500+ practitioners with software, data, cloud engineering skills help create a culture of innovation and agility that optimizes team performance, modernizes technology, and builds new business models. As a strategic partner and catalyst, Altimetrik quickly delivers results without disruption to the business.