

The Team

Amanda Baker, CMU Team
Harold Dansu, CMU Team
Sayo Sanu, CMU Team
Sahib Singh, CMU Team
Carlos Velasquez, CMU Team

David Riel, CMU Advisor
George Findling, Walgreens
Faranak Abar, Walgreens
Aaron Bresinger, EY
Sri Prabhakaran, EY

Project Alexandrite

Designing a Flexible Architecture for Real-Time Insights



Research

Project Goals

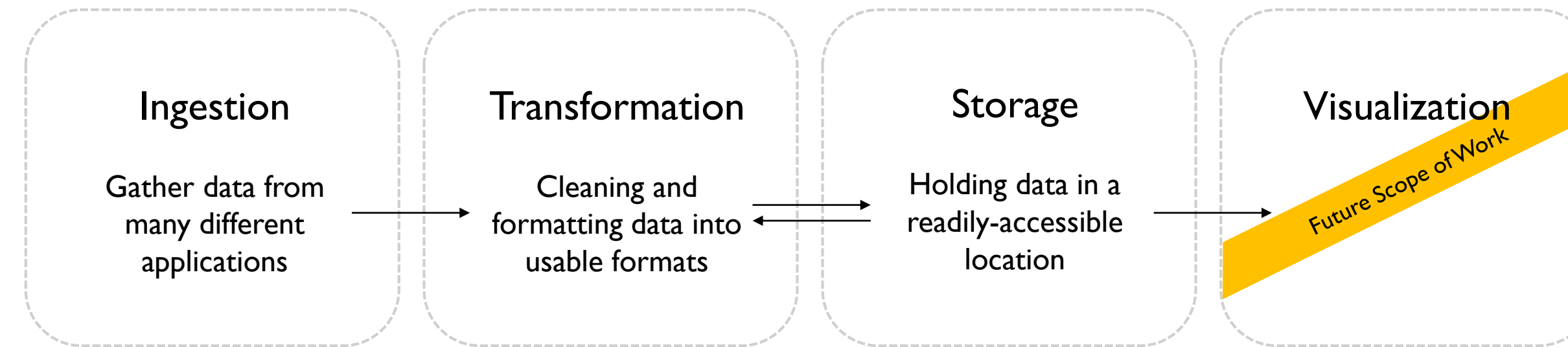
Walgreens and EY have partnered on projects for over 12 years. Together, they asked our team to design an IT architecture that will allow Walgreens executives to monitor and analyze the retail systems health performance Key Performance Indicators (KPIs) of Walgreens stores around the country in near-real time.

Example KPIs

- Average Receipt Print Time:**
Time interval from when payment is processed to when the paper receipt begins printing
- Average Payment Authorization:**
Time interval from when a pin pad auth. request is initiated to when the response arrives back at the POS

Design Methodology

The architecture design process was divided into four phases, each of which required detailed analysis of specific design choices and their implications.



Research Outcomes

Our research resulted in three architectural options that reflect different build choices:

- “Peridot”: Focuses on lighter technology stack to reduce development efforts and costs
- “Amethyst”: Implements architectural choices that balance real-time and historical data
- “Andesine”: Optimized for real-time insights, and carries more development and ownership costs
- Andesine was deemed optimal (see Design Outcomes for details)

* Designs are all named after gemstones, in accordance with the gemstone theme introduced by Walgreens' name for this project.

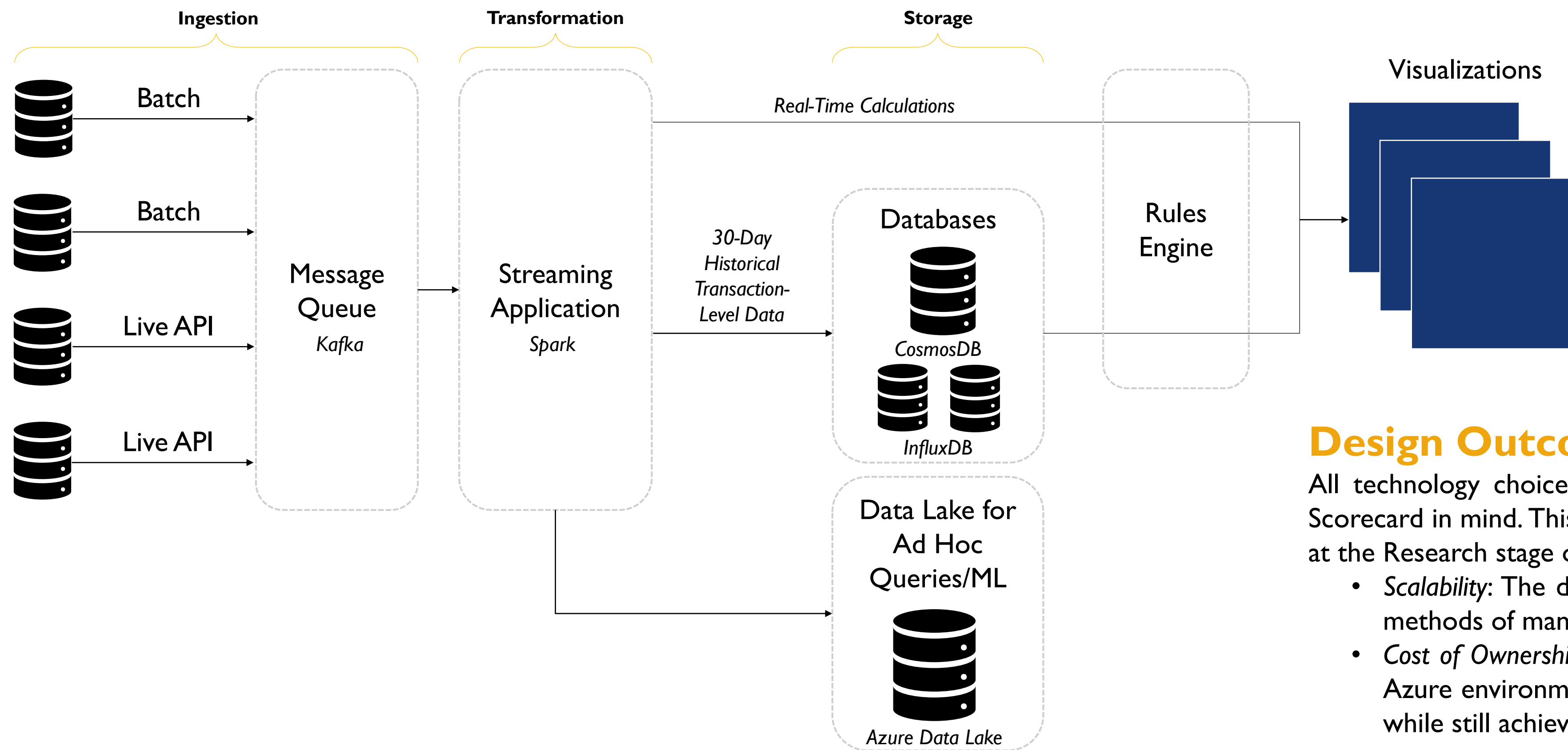
Design

Considerations

At each stage of the architecture, we analyzed key considerations that influenced vendor and tech stack choices. Below are the overarching requirements we considered at each stage.

- Alignment with Walgreens' technology stack**
 - Cloud-enabled solution
 - Microsoft Azure compatibility
- Near-real time analytics**
 - High performance to process high volume data
 - Timeseries data aggregated over time windows
- Scalable and flexible**
 - Ability to integrate with current known and future unknown data formats
 - Support for horizontal scalability in databases
 - Data transformation for structured and unstructured data

The “Andesine” Architectural Design



Evaluation Scorecard

To ensure we developed an optimal architecture for Walgreens, we evaluated the design and selected technologies against a series of requirements based on Walgreens' needs.

Metric	Priority
Scalability/Flexibility	High
Development Effort	Medium
Performance	Medium
Total Cost of Ownership	High
Organizational	High

Design Outcomes

All technology choices were made with the considerations and Evaluation Scorecard in mind. This resulted in a more refined version of designs developed at the Research stage of the project, with a focus on:

- Scalability:** The design allows Walgreens to easily incorporate new input methods of many types
- Cost of Ownership:** Using a combination of open-source tools within the Azure environment allows Walgreens to utilize their current tech stack while still achieving performance, effort, and organizational goals

Prototype

Prototype Goals

Upon completing a high-level architecture, the final phase of the project was to build out a prototype of the Ingestion phase. This prototype functions as a proof of concept for architectural choices, demonstrating the flow of data from various data-generating applications to a Kafka Cluster on Azure. We built the prototype in Azure to ensure compatibility with Walgreens' current technology stack.

Toolbox

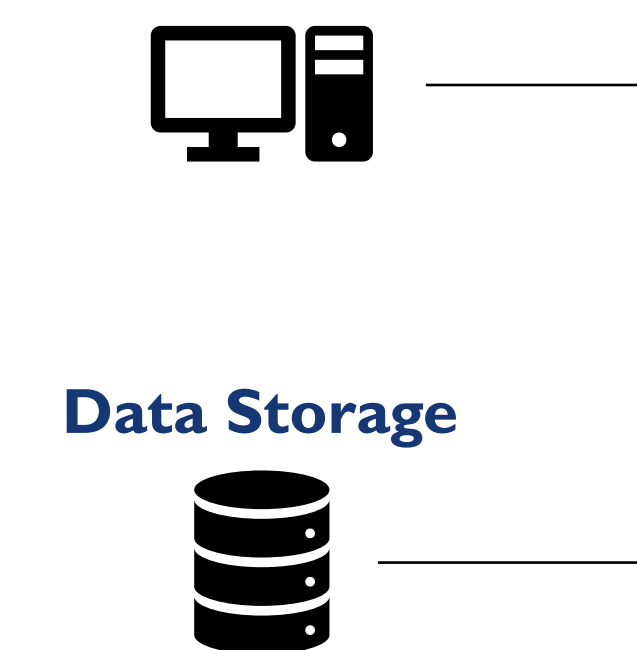
- Sample data from the Walgreens Point of Sales system in JSON
- Kafka
- MongoDB
- Microsoft Azure



Framework

After evaluating many technologies, we chose to use Apache Kafka as a central implementation choice. The Kafka system can deliver in-order, persistent, scalable messaging and enables massively parallel consumption, perfect for Walgreens' needs.

Producer Applications



Consumer Applications

