

PICNIC.APP

# Picnic Modernizes Data Architecture with Confluent for Streaming Analytics



## Headquarters

Amsterdam-Duivendrecht, North Holland, Netherlands

## Industry

Retail/Online Grocery

## Challenge

Address legacy tech-related operational overhead and scalability issues to allow for better customer behavior analytics and improve internal processes.

## Solution

Confluent Cloud to save time and money by reducing operational overhead and allowing for real-time processing and easy scalability of event data.

## Results

- Reduced infrastructure costs by 40%
- Simplified, future-proof data architecture
- Improved infrastructure monitoring for better SLAs and system health
- Elimination of data loss

Picnic is Europe's fastest-growing online-only supermarket, currently operating in France, the Netherlands, and Germany. The company relies heavily on data-driven decisions to provide the lowest-price guarantee to its customers. Picnic processes more than 300 million unique events per week from customer applications (i.e., shopping applications) and internal systems to power predictive analytics in their data warehouse.

Picnic's blazing-fast growth, attributable to consumer demand, industry recognition, and a successful series D funding, underpinned the need to seek out a better performing and more reliable data streaming platform, as well as improved streaming analytics capabilities.

That's when Picnic turned to Confluent.

## The Challenge

### Two quickly growing data pipelines

Picnic has two primary data pipelines. The first one is for their customer-facing application, which collects data on user behavior related to things like product engagement and checkout. Picnic analyzes this data to determine how to improve product recommendations and enhance the application UI.

The second data pipeline processes data for internal backend systems, which help with data warehouse, payment processing, product availability status, and much more. It enables Picnic to plan procurement, optimize warehouse operations, and ensure customers have a positive experience by delivering the right products in a timely manner.

As Picnic grew, they ran into issues using AWS Kinesis for these data pipelines, specifically:

- An inability to store data for more than one week. Replayability and recoverability are key for Picnic to quickly fix malformed events or pipeline failures and also to accomplish complex stream processing.
- Picnic began to require vast tooling around its data pipelines to enable seamless data streaming to other systems without having to create customer services from scratch.
- Last but not least, Picnic needed exactly-once semantics for certain use cases such as dynamic business rules evaluation.

*"Confluent provides exactly what we dreamed of: an ecosystem of tools to source and sink data from data streams. It's provided us not only with great data pipeline agility and flexibility but also a highly simplified infrastructure that's allowed us to reduce costs."*

— DIMA KALASHNIKOV, TECHNICAL LEAD

## Technical Solution

### Easy scaling and data flexibility

Picnic felt that the time for a change had come and sought a solution that could address all of the above issues at once. Their search led them to Confluent Cloud and its Apache Kafka® foundation.

They redesigned the architecture of the data pipelines, which led to a dramatic simplification of their internal-services pipelines, and more efficient processing of customer application data.

Very quickly, they experienced the following benefits by moving to Confluent:

- They can store data in Confluent for as long as they want, since Confluent's Infinite Storage doesn't have limitations on volume or retention periods.
- Confluent offers a rich ecosystem of pre-built, fully managed connectors that are ready to use out-of-the-box and eliminated the operational burden of maintaining them
- Exactly-once semantics is finally possible thanks to Kafka and Confluent's fully managed connectors.

Previously, all data from RabbitMQ was sent into a single Kinesis stream, which put the burden of event separation on the data warehouse and made scalability more challenging.

Now, Picnic uses connectors for RabbitMQ to forward data from RabbitMQ into Confluent Cloud Kafka topics, after which they use fully managed sink connectors to load data seamlessly into Snowflake and Amazon S3 for further analysis by data science teams. They also use Confluent's Data Preview feature on fully managed connectors to help iterate and test a connector's output prior to writing to production.

This new setup has led to greater scalability and established clear streams of homogenous data. It has also benefited from Confluent's managed connectors, which enabled Picnic to more easily monitor APIs for solutions like Prometheus.

## Business Results

### Simplified architecture, saved money

**Reduced infrastructure cost** "Our projected cost-savings improved by 40% and optimizing costs is one of the biggest impacts Confluent has made at Picnic," Dima Kalashnikov, Technical Lead at Picnic said.

**Simplified IT architecture** "The internal-services pipeline shrank from five managed services to only one and the number of maintained services went down by 40% across all managed services," Kalashnikov said.

### Improved infrastructure monitoring for better SLAs

Per Kalashnikov: "Thanks to the way Confluent has simplified our infrastructure, it's allowed us to improve our SLAs on data delivery and also prevent us from having any data loss, which is a major SLA for us."

## What's Next?

Moving forward, Picnic plans to leverage ksqlDB for new streaming analytics use cases like powering real-time reporting dashboards, for streaming ETL, and to enable things like real-time recommendations based on user behavior.

Picnic also wants to improve data delivery to and from its automated fulfillment center and create a self-service platform for data delivery via automated deployment systems for Confluent-based pipelines. Ultimately, the goal of this project is to deliver data from an automated fulfillment center to a data warehouse so analysts are able to analyze it in conjunction with other available data.

## Learn More About Picnic

<https://picnic.app/de/>

<https://blog.picnic.nl/>

*"Our machine learning developers and data scientists can now start to think about and work on in-stream data analysis and real-time models. We now have a vast and thriving ecosystem of Confluent enterprise-grade tooling that allows us to expand our analytics platform to the places where it would have been too much of a burden in the past."*

— DIMA KALASHNIKOV, TECHNICAL LEAD