

# Streaming Data Integration: Build vs. Buy Analysis

With the wide availability and perceived cost benefits, some organizations consider using open source projects to build custom solutions for data integration and analytics. The cost of building an enterprise-grade streaming integration platform is more than the sum of its components. Significant development investment is required to integrate and maintain a framework combining multiple products in a way that the end-to-end solution is secure, scalable, and reliable.

## Critical Cost Drivers of Building A Streaming Integration Platform

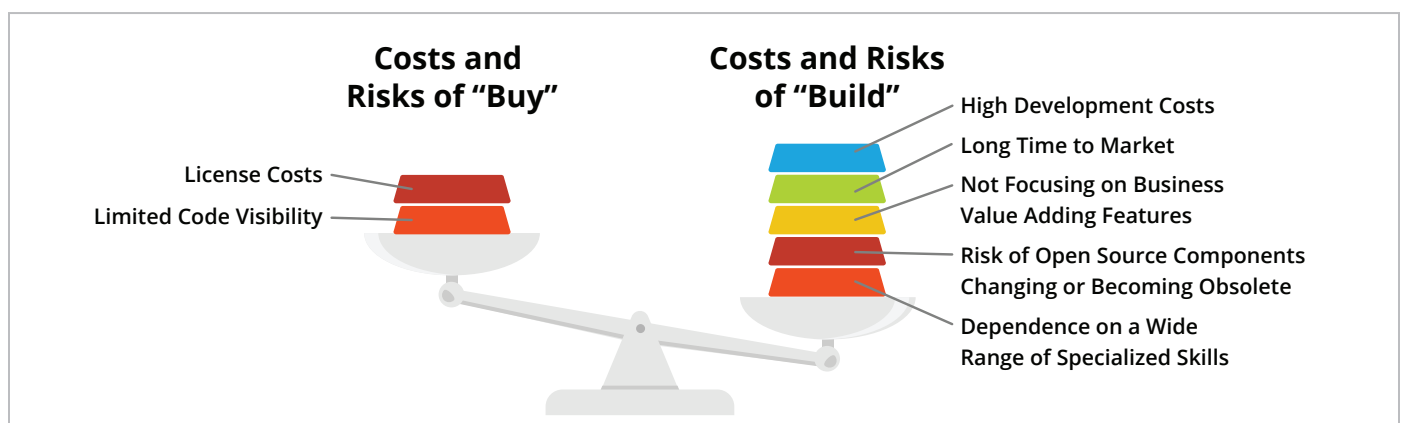
A streaming data integration platform includes multiple functions and requires different components to work together seamlessly. Real-time data ingestion from databases and other sources, high-speed messaging infrastructure, in-flight data processing, distributed in-memory grid, and data delivery to diverse targets are the main elements needed to move high-velocity data in real time. Adding streaming analytics and visualization capabilities, as Striim does, requires yet more components to be chosen and integrated.

The development lifecycle for building a streaming integration solution with open source products consists of the following steps and related costs:

**Evaluation and Design:** For each component mentioned above, there are at least 3-4 open source products that need to be evaluated. Figuring out how they all work together to meet specific requirements can be a very time and resource consuming process.

## Real-World Example

North American retail bank built a cybersecurity solution using a group of open source products (Nifi, Storm, Metron, Kibana, AlertUI, ElastAlert, etc.). They set up the prototype in 6 months, but the platform needs 45 people team to manage it now, "which is not worth it." The management cost can reach \$20-30m over 5-6 years.



The costs and risks of building a streaming integration platform outweighs the licensing costs of a proven, cost-effective commercial software

**Integration of Different Components:** To build an end-to-end streaming data integration platform, developers need to be familiar with the open source components and integrate them, requiring significant time and specialized skill sets.

**Adding Another Code Layer for Enterprise-Grade Features:** To support production solutions for business-critical systems, developers need to code and test for end-to-end security, high availability, reliability, and scalability.

**Adding an Abstraction Layer for the Pipeline Developers:** When users build pipelines, they should not need to know the underlying open source components. Therefore, an abstraction layer with API connectivity needs to be provided. It is unlikely the project has scope for an intuitive UI or other ease-of-use features.

**Testing & Deploying:** Complexities of each platform component and the brittleness of integration between them can delay pinpointing the underlying problems in the code.

**Maintaining the Code:** Possible drastic feature or API changes in open source components may lead to code revisions or filling a gap when one or more open source components become obsolete.

**Support for Production:** Paying for enterprise-level support for multiple open source products increases costs and raises management overhead due to coordinating with various support teams.

## Unique IP and Value of the Striim Platform

Striim uses multiple open source components and supplements them with several unique IP features:

- Real-time, low impact data ingestion from enterprise databases using a log-based change data capture.
- Patented distributed stream processing to support high-volume data environments efficiently.
- Patented real-time replication monitoring for visibility into the health of pipelines.
- Built-in clustering for high-availability and scale-out architecture.
- Pre-integrated in-memory cache to avoid adding latency during processing.
- Intuitive, wizard-based UI with pre-packaged integration applications to develop in seconds.
- Built-in exactly once processing, including when working with time windows, to avoid data loss or data duplicates.

The Striim platform also offers complex event processing and dashboard builder features to enable streaming analytics applications.

## Cost and Risk Drivers of the Build Approach

- High development cost
- Slow time to market with a lengthy development process
- Developers cannot focus on business value adding features
- Dependence on a wide range of specialized skills
- Risk of open source components changing features or becoming obsolete

## Why Striim?

- End-to-end solution with stream processing, monitoring, validation, intelligence
- Built-in enterprise-grade security, scalability, and reliability without having to code
- Pre-configured applications to build real-time pipelines in seconds to minutes
- Patented distributed stream processing to support high-data volumes
- Patented, built-in replication monitoring with real-time metrics
- Intuitive, wizards-based UI with a SQL-based processing language

To learn more, or to download or provision a fully functional trial of the Striim platform, please visit [www.striim.com](http://www.striim.com).