The Business Value for Creating a Connector to Apache Kafka®

To connect your software to Kafka, you need a connector. This paper covers the compelling business reasons behind building one and resources to do it.

August 2019
Overview

Today, every business, small or large, across all vertical markets and industries, must continually look for new ways to innovate. Competitive pressures from startups, cloud-native businesses, and digital-first enterprises are greater than ever. New technologies for data centers, managing, analyzing and visualizing the data are being unearthed at a rapid pace due to a large part to the open source software movement. These new innovations are always being evaluated not only by consumers that need the functionality but by existing software and hardware solution providers to enhance their own products. In fact, these new technologies can completely reshape product strategies. One of these new technologies is event streaming, specifically open source Apache Kafka®, which makes up the heart of an event streaming platform.

The focus of this brief is about the value of connecting to Kafka and is aimed at business level readers in companies that develop and sell enterprise software. It outlines why connecting to an event streaming platform or Kafka will be of interest to you.

While this won’t go into depth on what Kafka is - there are good resources available for that - it will be helpful to provide an overview of event streaming and Kafka and why it is so transformative for customers. Skip that section if you already know it.
Event streaming and Apache Kafka®

Databases have been around forever and will continue to be. Databases store data - the most valuable asset for most companies today. New buzz phrases like “Data driven” and “Customer 360” originate from the requirement to innovate and collect all the data you can, analyze and create the company strategy or the customer experience on that data. But consider the idea that looking at data from a database or data warehouse perspective is only a part of the picture.

Events

Equally important to the static data in databases and mainframes are the events that created the data, or the application that uses the data. In this context, the definition of an event is anything that happened! Examples include an ATM transaction, a credit card purchase, an order for a product, the shipping of a product, a car engine reaching an unacceptable temperature.

Figure 1: One advantage of Kafka is connecting data sources and destinations
What event streaming can do

Events are being created all the time and this continually updating stream of events along with all the events that occurred in the past, form the essence of businesses today. Events can be leveraged to deliver new customer experiences, create new business models, deliver huge internal efficiencies - at scale, with contextual relevance and in real-time. This is what event streaming is all about - where data is seen as a continuous stream of events and not just a repository of data.

Here’s an example in banking and the impact of event streaming:

- Without event streaming, if we deposited money in an ATM, or transferred money via a mobile app, we would have to wait hours or days to see that reflected in our account balance. Today, we have real-time account updates across web, mobile, and branch.
- Before, fraud checks, for things like account login fraud or credit card fraud, were all done on a batch basis, giving time for money to be stolen. Today, we have real-time fraud alerting.

Let’s look at retail:

- Without event streaming, we as consumers can order a product online, receive a confirmation email, and then 1 day later get an out of stock email because inventory updates are done on a nightly basis. Today, inventory across channels can be managed in real time.
- Before, we interact with websites and applications and receive a generic experience. Today, we can get real-time product recommendations. Netflix estimates this is worth $1 Billion a year to their business.
- And, without event streaming, it’s not uncommon for major, global retailer to have to wait 1-2 days to know what they’ve sold. Today, those reports can be produced in real-time.
Apache Kafka

Event streaming forms a new paradigm that was originally developed at LinkedIn in 2010 by the founders of Confluent as an open source solution called Apache Kafka®. It has seen massive open source adoption across hundreds of thousands of developers.

Separation of data sources and data sinks

Many application developers today and in the past develop applications that are dependent on a direct connection to data in a database. As companies grow, many other applications or processes will need access to the same data or transactions. Incorporating cloud storage, cloud applications, adding more data sources, or growing globally -- makes this architecture become inefficient, unmanageable and costly.

In an event streaming architecture with Kafka, “sources” of data, such as databases or weblogs, are separated from the applications or databases using the data, called “sinks”. By removing the direct connections and dependencies between sources and sinks, it improves efficiencies and eliminates the complexity.

**Figure 2**: Before Kafka, applications are dependent on a direct connection to data in a database
**Kafka Connect API**

Connecting to Kafka is accomplished with a connector, either a source or a sink connector, to import or export your data. Fortunately, Kafka comes with a Connect API that defines a standard and pluggable framework for handling data in and out of Kafka.

**Kafka Growth**

Organizations today that started as digital companies are building their business on event streaming, and now, 60% of Fortune 100 companies are leveraging event streaming, via open source Apache Kafka, as a foundational technology platform.

**Kafka Benefits**

Companies offering software solutions can harness event streaming and Kafka to offer new real-time services, modernize databases, help their customers enable new functionality like microservices, and transform their business strategies -- by connecting to Kafka.
Connecting to Kafka

What type of technologies are connecting to Kafka

Many types of technologies are connecting to Kafka, examples include database software, RDBMS, IoT applications, visualization, analytics or monitoring tools. These companies connect to Kafka to take advantage of this new event streaming paradigm.

Who creates connectors

Connector development may come from the Kafka community, from software vendors, customers, and Confluent. Many connectors have already been developed, some examples of connectors are from industry prominent software companies such as MongoDB, DataStax, Snowflake, IBM, Microsoft, HP, and Oracle. Examples of community or Confluent developed connectors include technologies like HTTP, S3, JDBC, syslog, and HDFS.

Companies with multiple product lines will most likely need multiple connectors. If a connector is not available, you can build it. There are resources available and a verification program to validate connectors and integrations from Confluent.

Let’s look at some of the business reasons why companies should invest in building a connector.

**Figure 4**: A snapshot of the many technologies that connect to Kafka. Each tile represents a connector for users to download on Confluent.io/hub.
Business value for creating a Kafka connector

**Maintain your customer base**
If you want to keep your customers happy your products and solutions need to continue to add value through innovative and relevant updates that Kafka enables.

**Enjoy new opportunities and enter new markets**
Developing a Kafka connector for your product will open up new markets and use cases due to the growing installed base of Kafka customers who have already adopted it.

**Offer a more robust and complete solution**
A Kafka connector helps to add real value to your product offering, making it more extensible, viable and complete.

**Enable a microservices approach**
Microservices break down the traditional monolithic architectures and software development style into smaller manageable elements that are decoupled. This is what Kafka facilitates and microservices are fast becoming a very popular methodology for achieving availability, efficiency and scalability so that new services can be quickly implemented without the need for massive reconfiguration.

**Create a path to a modern data architecture**
As new technologies emerge, customers using legacy systems will look for ways to incorporate new paradigms for competitive advantage. If your company develops these legacy systems, a connector to Kafka will help you and your customer with modernization of the IT infrastructure and data center.

**Leverage the ecosystem network effects**
Since event streaming is becoming so well adopted, it opens up an ecosystem of other companies that have created connectors to their sources of data and their applications. This effect allows customers to connect to all those other sources and destinations.
Market your connector

An option for software vendors with a verified connector is to become a Confluent partner. Partners can take advantage of various marketing activities with Confluent to help market your connector. This includes posting the connector to the Confluent Hub where many customers come to find connectors. Joint marketing activities such as briefs, webinars, blogs, email campaigns create new opportunities.

Technical benefits

There are technical benefits of using the Kafka Connect API over other frameworks that are not Kafka specific. The Connect API is an easy to use, fault tolerant framework that provides a better user experience and scalability. It includes:

- Schema registry integration
- Standardized source and sink
- In-cloud / Hybrid cloud / on-prem support
- Infrastructure capabilities that will minimize any code rewrites like offset management, decoupled serialization, schema migration and error handling

Oh - Let's not forget that Kafka Connect API is free.
Building a connector to Kafka

The decision to create a connector may or may not be a simple decision. It will require resources to build it and a plan to support it. This could take a few weeks or months, depending upon your product and internal processes. The value to the business is very likely justifiable but there are a few more considerations.

Connector verification

Confluent provides a free service as part of the “Verification Integration Program” to make sure connectors have met the technical requirements outlined in the Connect API. This provides customers an assurance of functionality and compatibility with Kafka.

Avoiding vendor finger pointing

An important aspect to consider if you are a software vendor are the support issues that can come up for a customer when multiple vendors are involved. If there’s a problem, is it my connector, their connector, Kafka or Confluent? TSAnet solves that by providing a multi-vendor case management process that helps to alleviate these issues and is an important benefit provided as part of Confluent’s Verification Integration Program.

The Verified Integrations Program Process

1. Initiated
   - The first step is to discuss the overall project so both companies understand the scope and effort involved.

2. Guidance
   - As you work through building the connector, Confluent will assist and provide the necessary resources to help.

3. Submitted
   - Submit the connector as discussed and outlined in the Verification Guide.

4. Verified
   - Once Confluent verifies the integration we will publish it and work with you on joint assets and demand generation activities.

Figure 6: The Verified Integrations Program Process
Get started

If you are interested in how to start building a connector, start with these resources.

Confluent Hub

Check repositories to see if there is a connector available that fits your needs. Confluent has a list of connectors already created on the Confluent Hub.

Watch the webinar

The webinar is in two parts, one focuses on the business reasons and one for more technical leads: Building Kafka connectors - the Why and How.

Verified Integrations Program

Confluent supports companies in building connectors with the Verified Integrations Program that provides guidance on building a connector and verifies the functionality of the connector prior to posting it to the Hub. Questions are welcomed as well, just fill in the contact form at the bottom of the webpage.

Developer references

Developers can obtain a verification guide and checklist from the Verified Integrations Program page. Confluent documentation also helps developers understand Kafka Connect API further.
About Confluent

Confluent, founded by the original creators of Apache Kafka, pioneered the enterprise-ready event streaming platform. With Confluent, organizations benefit from the first event streaming platform built for the enterprise with the ease of use, scalability, security and flexibility required by the most discerning global companies to run their business in real time. Companies leading their respective industries have realized success with this new platform paradigm to transform their architectures to streaming from batch processing, spanning on-premises and multi-cloud environments. Backed by Benchmark, Index Ventures and Sequoia Capital, Confluent is headquartered in Palo Alto and London with offices globally. To learn more, please visit www.confluent.io. Download Confluent Platform and Confluent Cloud at www.confluent.io/download.