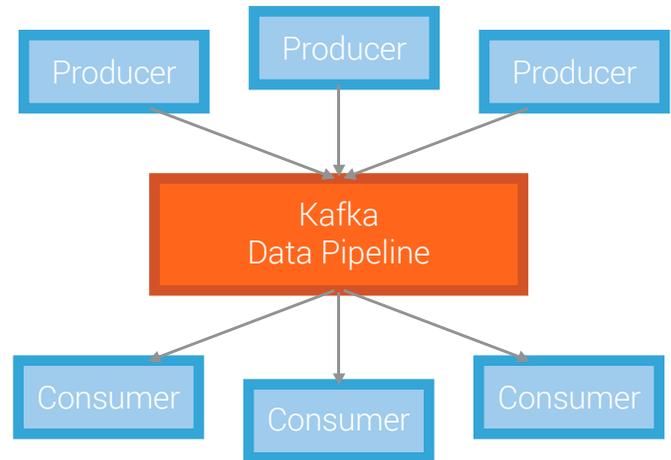# Confluent Platform Datasheet

Apache™ Kafka: the Data Pipeline for Stream Data

At the center of the Confluent Platform is Apache Kafka. Apache Kafka is a modern distributed system for moving data in real time. It allows a single cluster of computers to perform all the data transport between an organization's data systems regardless of their number or size.

At first glance, Kafka looks like a messaging system—clients publish messages to Kafka, which passes them to other systems. In fact, Kafka works more like a distributed database: when you write a message to Kafka, it is replicated to multiple servers and committed to disk. Kafka's consumers perform the database read function while its producers perform the database write function. In Kafka, a topic is a partitioned set of logs. You can distribute these partitions over a cluster of servers, with each partition persisted.

This unique architecture enables Kafka's performance. It also explains Kafka's throughput capability—and its approach to data persistence—which address the needs of real-time operations better than a typical messaging system.

The Benefits of Kafka

All Your Data
Kafka's architecture makes it an ideal centralized service for all the data in your organization. Get all your data from disparate systems flowing through Kafka — unlocking your data to fuel your growth.

Highly Scalable
Kafka scales to trillions of messages a day and thousands of producers and consumers without loss of performance, allowing you to build a real-time streaming platform connecting your entire infrastructure.

Real-Time
Kafka satisfies the millisecond latency requirements of streaming applications, and is scalable enough to handle very high volume log and event data, and fault-tolerant enough for critical data delivery. Kafka allows you to run very large data streams in real time across your organization.
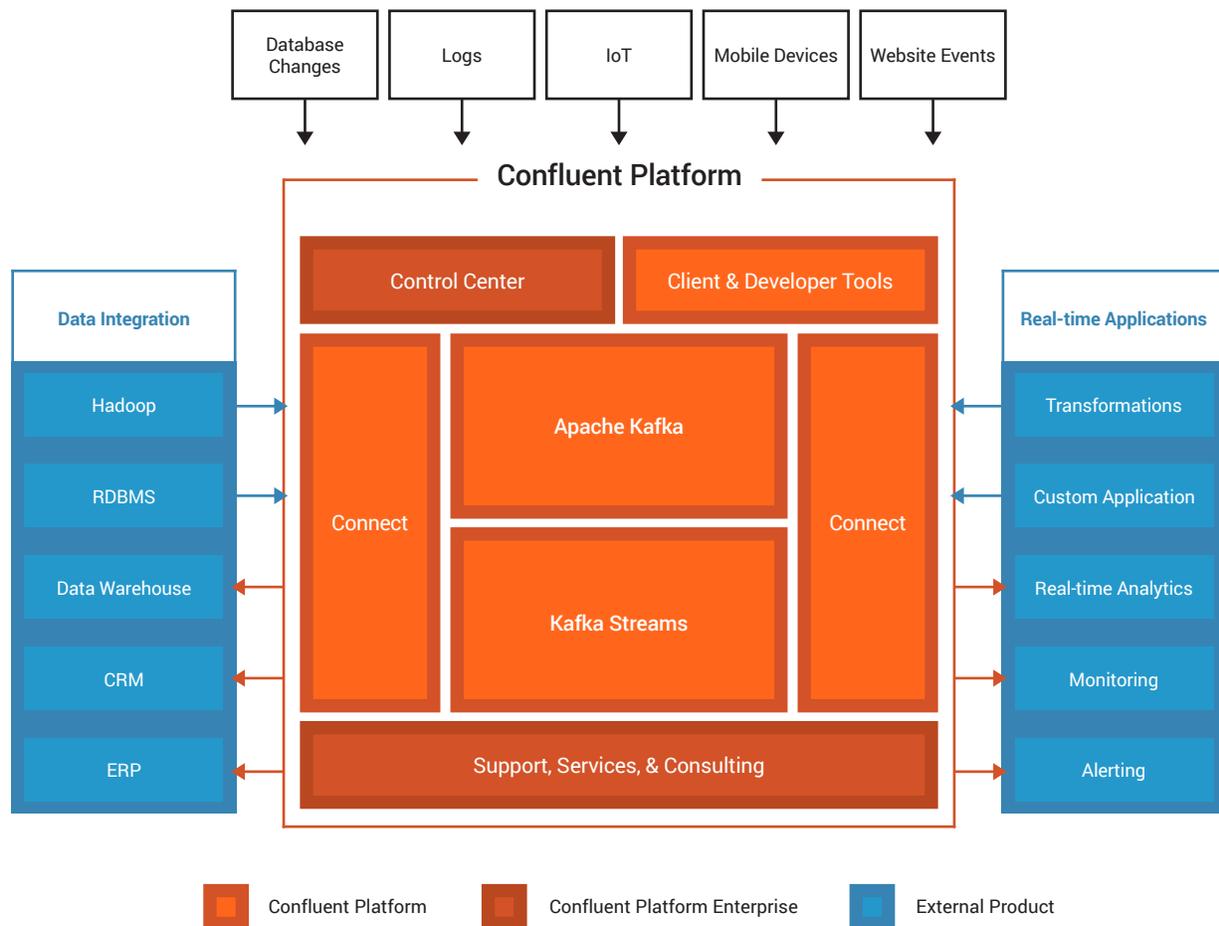
Secure
Kafka includes the encryption, authentication, and authorization capabilities needed to ensure sensitive data is protected and accessible only by authorized parties.

The Confluent Platform

A Complete Solution for Real-time Streams

The Confluent Platform is an open source framework that includes all the components you need to create a scalable stream data platform with Apache Kafka. These components draw on Confluent's experience building some of the largest streaming data pipelines in the world. Confluent makes implementing and managing a Kafka environment easy, reliable, secure, and auditable.

The Confluent Platform includes all components from the Apache Kafka Core along with some Apache-licensed open source additions (client tools, pre-built connectors, and cluster-side services). The platform is also available in an Enterprise Edition, which adds 24/7 support and unlimited use of the Control Center.

✓ Distributed   ✓ Fault Tolerant

✓ Processes Streams   ✓ Stores Messages

Features of Confluent Platform

Confluent Platform is the first Stream Data Platform. Built on the core Apache Kafka components (persistent publish/subscribe messaging, Kafka Streams, and the Kafka Connect framework for client development), the platform enables enterprises of any size to integrate real-time streams into their business processes.

Kafka Streams *

Kafka Streams is an open-source component of Apache Kafka that integrates into an application to build and execute stream processing functions. If you use Kafka for stream data transport, Kafka Streams adds stream processing capabilities to your application without requiring you to invest in a separate processing cluster or another stream processing framework.

Connectors *

Apache Kafka includes Kafka Connect, a framework for easily reading data from external sources into Kafka and easily writing data from Kafka into external destinations. The Confluent Platform includes all of the connectors included with Apache Kafka, plus connectors for:

- Databases that support JDBC (including Oracle, MySQL, and PostgreSQL)
- Hadoop File System (HDFS)
- Elastic Search

Clients *

Apache Kafka includes pure Java clients (both producers and consumers) that you can easily integrate into Java applications or other languages that run on the JVM, including Scala, Python, Clojure, JRuby, and Jython.

The Confluent Platform also includes a native C/C++ Kafka client. This full- featured client includes support for all features of the Java clients including the Schema Registry.

REST Proxy

You can publish data to Kafka or consume data from Kafka using the Confluent Kafka REST Proxy. The REST proxy provides access to Kafka from any network connected devices that can connect with HTTP. Additionally, it provides a way for clients to connect to a Kafka cluster through web proxies or other intermediate services.

Schema Registry

With the Confluent Platform, it is easy to guarantee that data that flows through your data pipeline is well-formed. The Confluent Schema Registry provides a central registry for the format of the data in each Kafka topic and provides a central service that helps to make changes to data format easy and backwards compatible.

The Confluent Schema Registry is based on Apache Avro. It provides a RESTful interface to manage schemas and schema changes. It provides multiple compatibility settings and allows evolution of schemas according to the configured compatibility setting. It provides serializers that plug into Kafka clients that handle schema storage and retrieval for Kafka messages that are sent in the Avro format.

Confluent Control Center

Control Center provides a comprehensive management system for Kafka, enabling the visibility and operational strength needed to manage a Kafka environment. The Control Center delivers real-time analysis of the performance of Kafka, allowing data teams to drill into topics, producers, consumers, and more to understand what's happening with their data pipelines so they can govern a growing ecosystem of stream data applications.

Enterprise Support

Confluent's engineering and support teams include the largest team of contributors to Apache Kafka, with many years of development and operational experience with distributed systems. Confluent provides 24/7 enterprise support, critical bug fixes, indemnification, training, and operational advice to customers.

* Part of Apache Kafka

Commitment to Kafka

Committed to Apache Kafka

You will not find better resources to support your Kafka implementation than those that Confluent can provide. The Confluent team includes the foremost experts in stream data management of all types, particularly Kafka. Confluent was founded by the engineering team at LinkedIn that originally created Apache Kafka, employs many of the active committers to the open source project and continues to invest in a thriving Kafka community.

Get the Confluent Platform

The Confluent Platform is open source, free to use, and includes Apache Kafka. You can download and try the Confluent Platform at www.confluent.io/download, or contact us about our enterprise subscription and services.

Apache Kafka
Open source

- Kafka Connect
- Java Client
- Kafka Streams

Confluent Platform

- *Includes Apache Kafka*
- Additional clients
- REST Proxy
- Schema registry
- Pre-built and certified connectors

Open source

Confluent Enterprise
Proprietary

- *Apache Kafka & Confluent Platform*
- Confluent Control Center
- Enterprise 24x7 support

Feature	Benefit	Apache Kafka	Confluent Platform	Confluent Platform Enterprise
Apache Kafka	High throughput, low latency, high availability, secure distributed message system	●	●	●
Kafka Connect	Advanced framework for connecting external sources/destinations into Kafka	●	●	●
Java Client	Provides easy integration into Java applications	●	●	●
Kafka Streams	Java library that enables scalable, fault-tolerant secure stream application development within the Kafka framework	●	●	●
Additional Clients	Supports non-Java clients; C, C++, Python, etc.		●	●
REST Proxy	Provides universal access to Kafka from any network connected device via HTTP		●	●
Schema Registry	Central registry for the format of Kafka data – guarantees all data is always consumable		●	●
Pre-Built Connectors	HDFS, JDBC and other connectors fully certified and supported by Confluent		●	●
Confluent Control Center	Includes Connector Management and Stream Monitoring			●
Support	Enterprise class support to keep your Kafka environment running at top performance	Community	Community	24x7x365
		Free	Free	**Subscription**