CONFLUENT

Building streaming data pipelines with Google Cloud Dataflow and Confluent Cloud

Elena Cuevas,

Senior Partner Solutions Engineer @ Confluent

August 2021

Agenda

01. Data in Motion
The new Data in Motion Paradigm

02. Product Overview
Introduction to Confluent

03. Demo Use Case and Architecture
How Confluent and Google partner to fulfill the Data in Motion paradigm

04. Demo
Building streaming data pipelines with Google Cloud Dataflow and Confluent Cloud

05. Q&A

Data in Motion

Our Mission
Set Data in Motion

Today, the digital realm is as important as the physical world in how business is transacted.

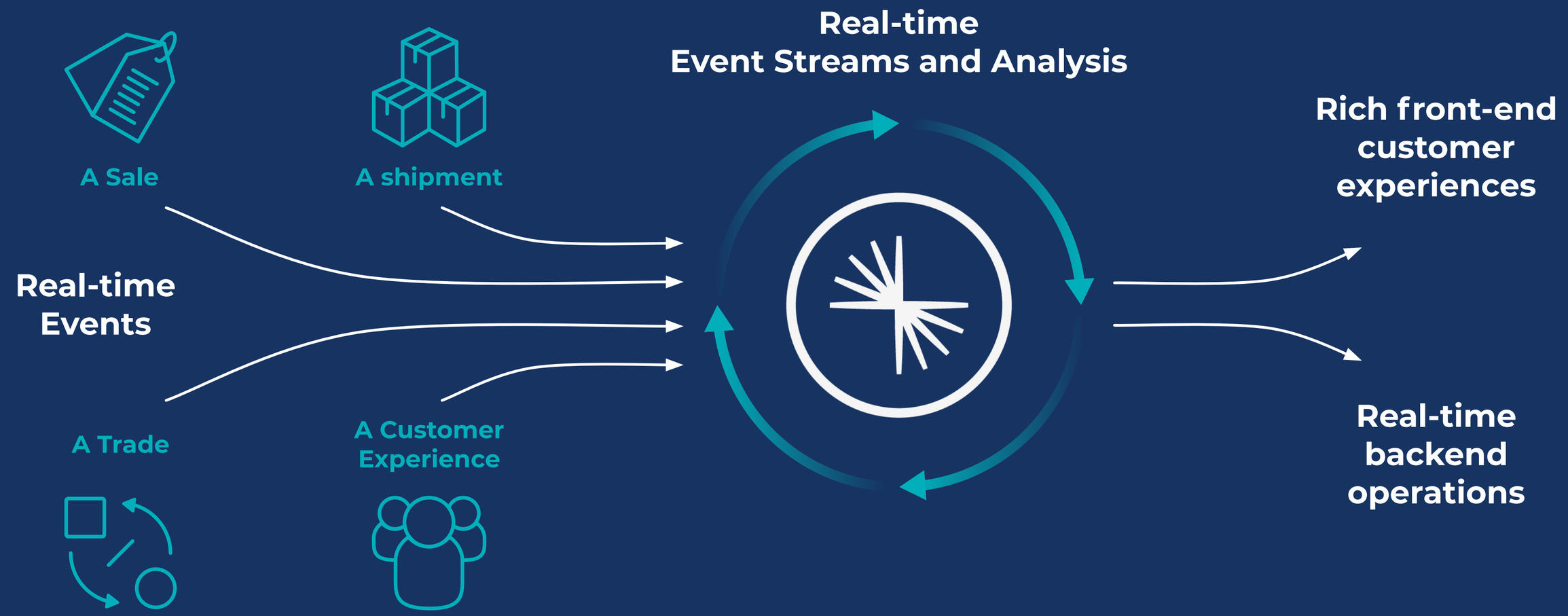
Enterprises require total connectivity and instant reaction, 24x7, anywhere, in real-time.

But they can’t get there with traditional, historical databases filled with data at rest.

They need a complete streaming platform, dually capable of **setting data in motion** and **analyzing** that data in real-time, and which is globally interconnecting the clouds and on-premises data centers.

A New Paradigm is Required for Data in Motion:

Continuously processing evolving streams of data in real-time

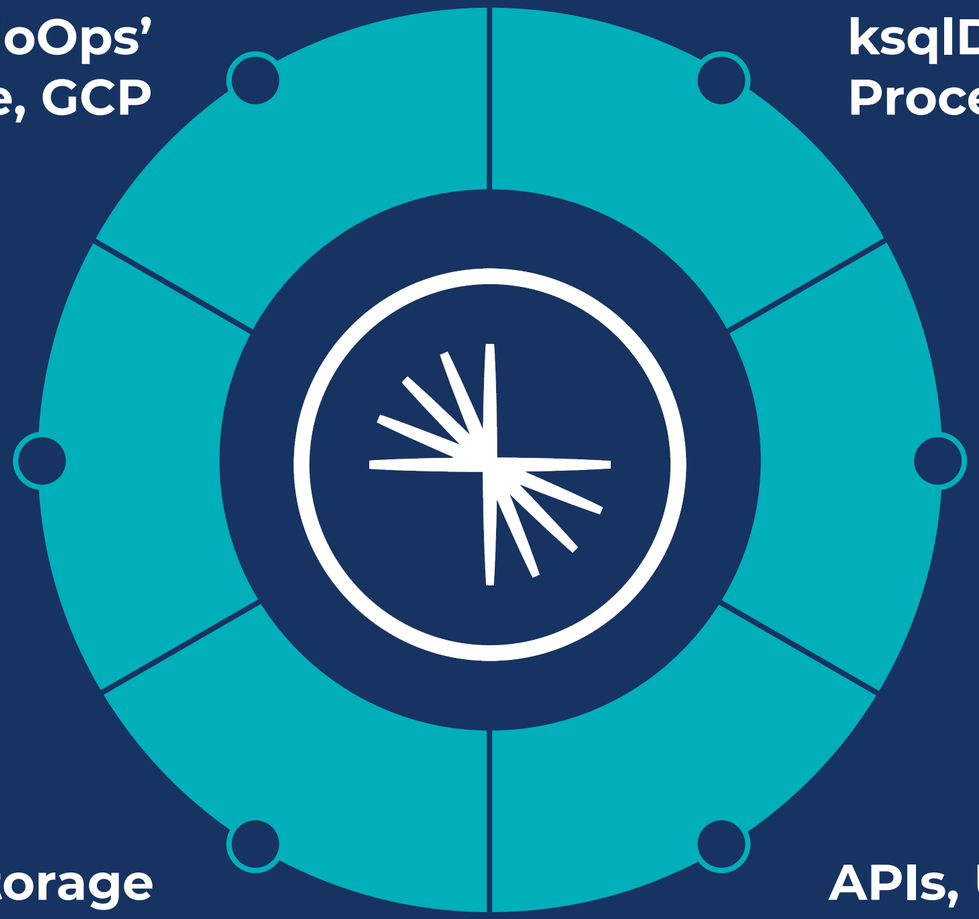


At Confluent, streaming is in our DNA.

We help the world's largest organizations make it part of theirs.

https://www.confluent.io/customers/

Cloud-native, Complete, Everywhere
with Kafka at its core



Product Overview

What makes Confluent unique?

Cloud-Native

Available as a fully managed service that is a serverless, infinitely scalable, elastic, secure, and globally interconnected. Our self-managed service inherits all the work born in the cloud.

Complete

ksqlDB, Connect, Schema Registry, and more

Capable of end-to-end applications

Kafka from the people who made it

Everywhere

Global availability on AWS, Azure, and GCP

Bridge on-prem to cloud with cluster linking

Extend streaming apps across clouds

Confluent: Everywhere

Fully-Managed Service

Confluent Cloud

Apache Kafka Re-engineered for the Cloud

Available on the leading public clouds

aws

Google Cloud

Microsoft Azure

Self-Managed Software

Confluent Platform

The Enterprise Distribution of Apache Kafka

Deploy on any platform, on-prem or cloud

Both: Subscription products where price scales with usage

Large Ecosystem for Event Streaming

Easily connect to 100+ data systems

https://www.confluent.io/hub/

Large Partner Network

Consulting Partners, Cloud Partners, OEM Partners, Tech Partners

accenture
Altair
aws
Couchbase
DATASTAX
DIAMANTI
elastic
Google Cloud Platform
HVR
HIVEMQ
IBM
imply
Infosys
kinetica
LUMINEX
MODERN MAINFRAME DATA SOLUTIONS
mongoDB
Microsoft Azure
neo4j
PRIVITAR
PURESTORAGE
PUSH TECHNOLOGY
Rethink Real-Time™
Qlik
ROCKSET
SCYLLA.
SecuPi
vmware
and more

https://www.confluent.io/partners/

Confluent: Complete

DEVELOPER	OPERATOR	ARCHITECT
Unrestricted Developer Productivity	**Efficient Operations at Scale**	**Production-stage Prerequisites**
Multi-language Development *Non-Java Clients* | *REST Proxy* *Admin REST APIs*	**GUI-driven Mgmt & Monitoring** Control Center | Proactive Support	**Enterprise-grade Security** RBAC | Secrets | Audit Logs
Rich Pre-built Ecosystem Connectors | *Hub* | *Schema Registry*	**Flexible DevOps Automation** Operator | *Ansible*	**Data Compatibility** *Schema Registry* | Schema Validation
Event Streaming Database *ksqlDB*	**Dynamic Performance & Elasticity** Self-Balancing Clusters | Tiered Storage	**Global Resilience** Multi-Region Clusters | Replicator Cluster Linking

Apache Kafka

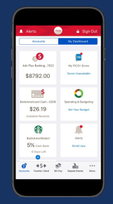
Open Source | Community licensed

Self-managed Software — **Freedom of Choice** — **Fully Managed Cloud Service**

Enterprise Support | **Professional Services** — **Committer-driven Expertise** — **Training** | **Partners**

Demo Architecture

Demo



Real-time user application sends entries

Microsoft Azure

aws

salesforce

Cloud-native Services

...

Confluent Cloud

Schema Registry

Streaming Analytics & Processing (ksqlDB)

Cloud Connectors

Infinite Storage

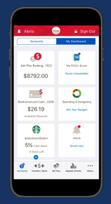
Promotions BigQuery

beam

Process Winners Dataflow

Winners BigQuery

Demo



Real-time user application sends entries

Confluent Cloud

topic: **entries**

topic: **application**

topic: **...**

Cloud Connectors

Microsoft Azure

aws

salesforce

Cloud-native Services

...

beam

Promotions BigQuery

PCollection: **Winners**

PCollection: **Entries**

WithKeys

Window (60s)

Window (60s)

CoGroupByKey

ParDo: **toKV**

ParDo: **toTableRow**

Winners BigQuery

Demo

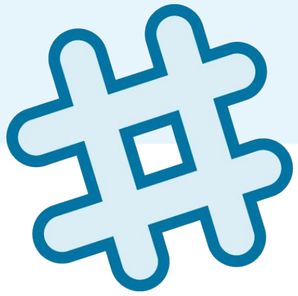
git.io/JROpx

git.io/JROpx

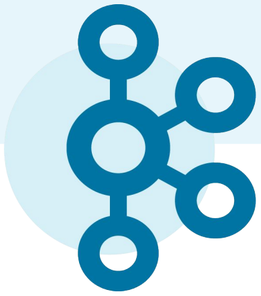
Q&A

Stay in touch!

Confluent Blog
cnfl.io/developer



Community Slack
cnfl.io/slack



Global Training
cnfl.io/kafka-training

Demo Architecture

Confluent Cloud

Customer services & apps on-premises (US)

Connectors

Customer services & apps self-managed in AWS

aws

Connectors

Microsoft Azure aws

salesforce Cloud-native Services ...

Cluster Linking

Schema Registry

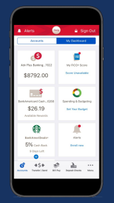
Streaming Analytics & Processing (ksqlDB)

Cloud Connectors

Infinite Storage

Customer services & apps on-premises (EU)

Connectors



Real-time analytics triggered a push alert “Fraudulent transaction!” to the B2C customer

Microsoft Azure aws

... Cloud-native Services mongoDB. Atlas

Thank you!

@elena_cxc

linkedin.com/in/ecuevasc

 **cnfl.io/meetups**

 **cnfl.io/blog**

 **cnfl.io/slack**