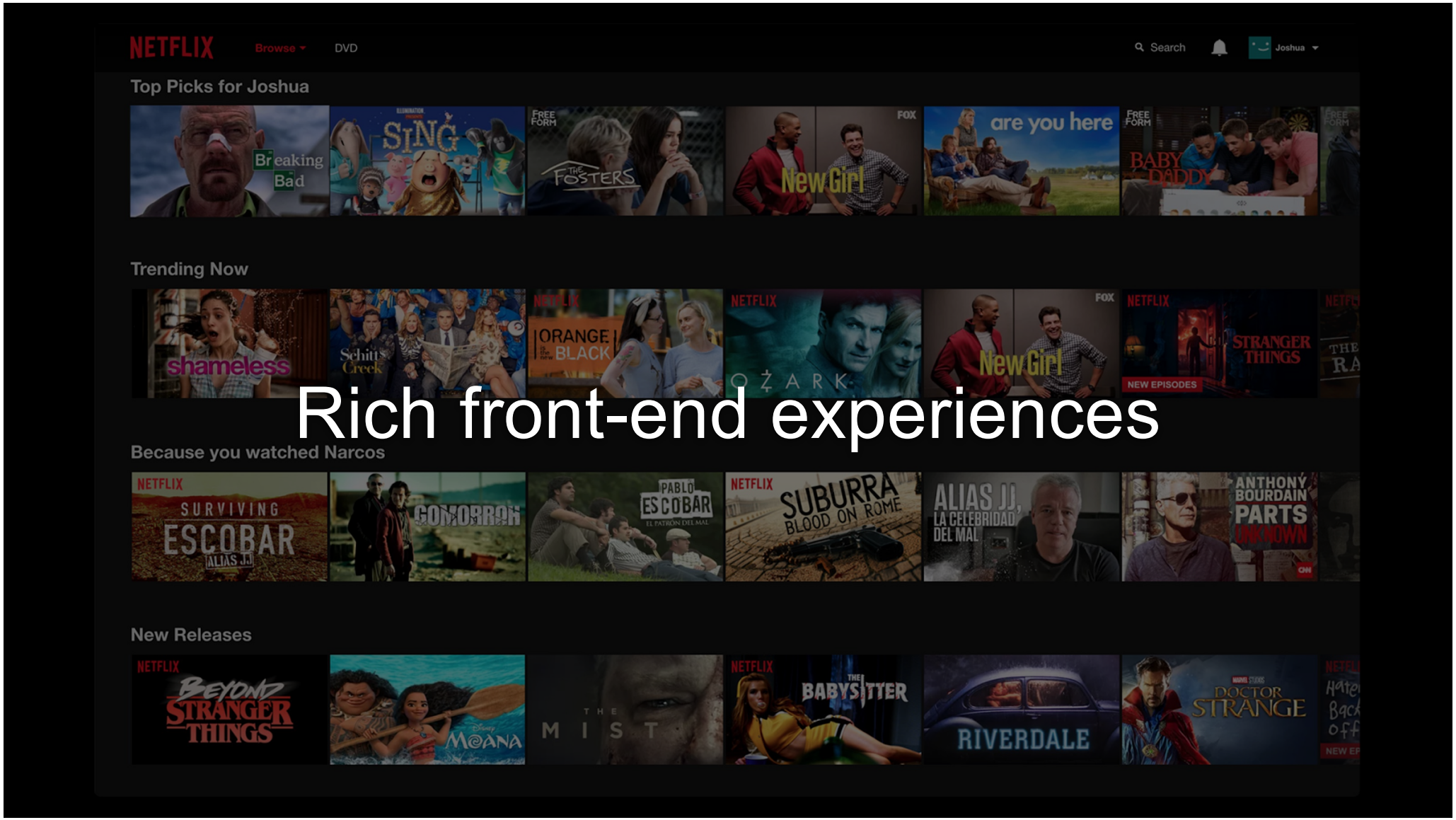
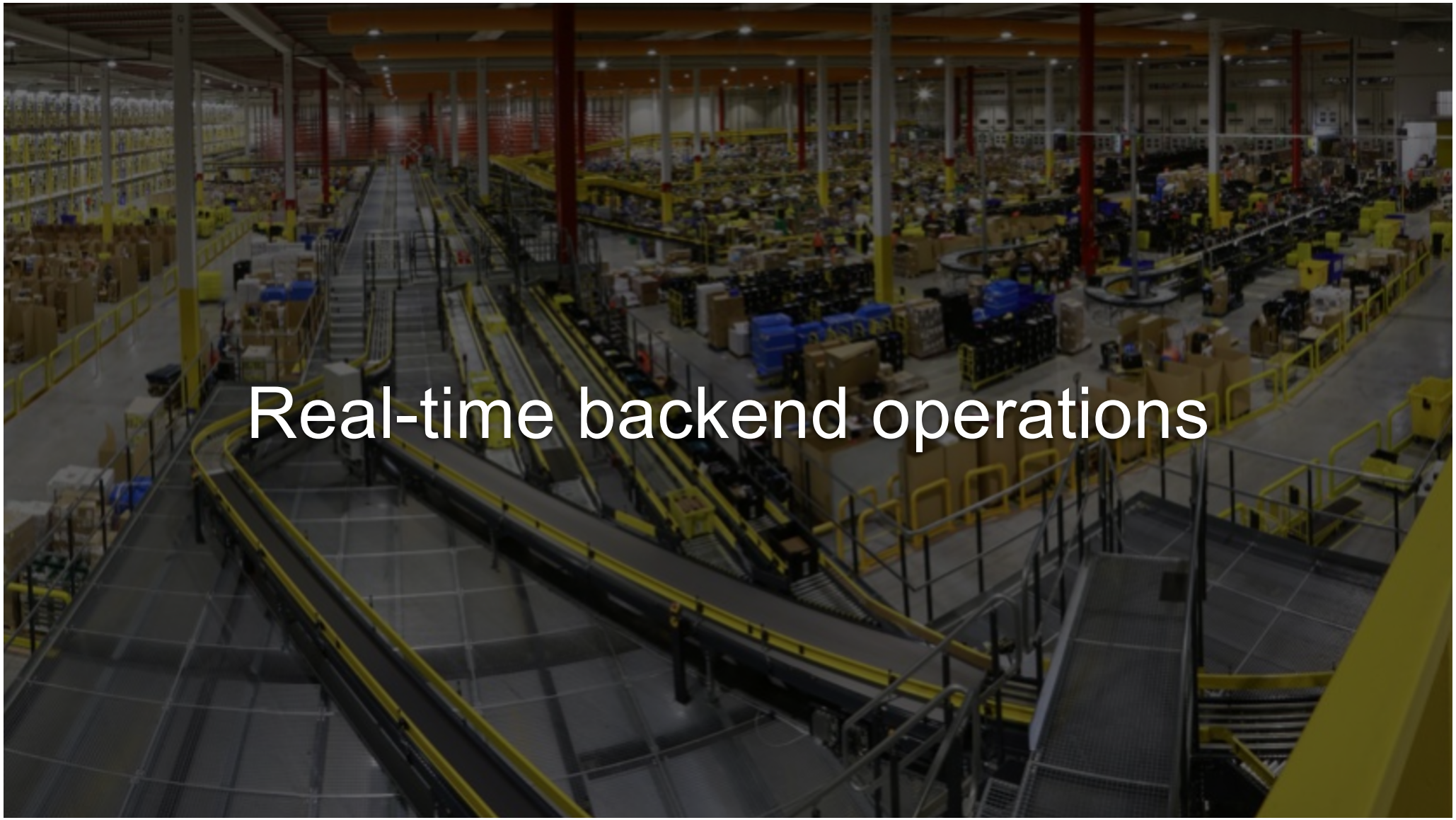
Migrate and Modernize Your Data Platform with Confluent and Databricks

Data is changing the world


NETFLIX
Browse
DVD
Search
Joshua
Top Picks for Joshua
Breaking Bad
SING
THE FOSTERS
New Girl
are you here
BABY DADDY
Trending Now
shameless
Schitt's Creek
ORANGE is the new BLACK
OZARK
New Girl
STRANGER THINGS
NEW EPISODES
Rich front-end experiences
Because you watched Narcos
SURVIVING ESCOBAR
ALIAS JJ
CONTRABANDO
PABLO ESCOBAR
EL PATRÓN DEL MAL
SUBURRA BLOOD ON ROME
ALIAS JJ, LA CELEBRIDAD DEL MAL
ANTHONY BOURDAIN PARTS UNKNOWN
New Releases
BEYOND STRANGER THINGS
MOANA
THE MIST
THE BABYSITTER
RIVERDALE
DOCTOR STRANGE


Real-time backend operations

Shared use cases across industries

Retail

- Customer 360
- Omnichannel optimization
- Supply chain optimization

Improve customer attraction & conversion to increase revenue

Financial Services

- Risk Management
- Trading Operations
- Security Operations & Fraud detection
- Payments Processing
- Customer 360

Improved customer outcomes, increased revenue

Media & Telecom

- Fraud Detection
- Customer churn reduction
- Real-time customer analytics
- Streaming health

Better operational intelligence

Healthcare and Health sciences

- IoT device management and analytics
- Clinical and claims

Improved patient outcomes and cost effectiveness

Manufacturing / Automotive

- Connected and autonomous cars
- Supply chain inventory
- Fleet management

Predictive analytics generate exponential value

CONFLUENT

Companies are moving to the cloud to build modern data platforms

Agility

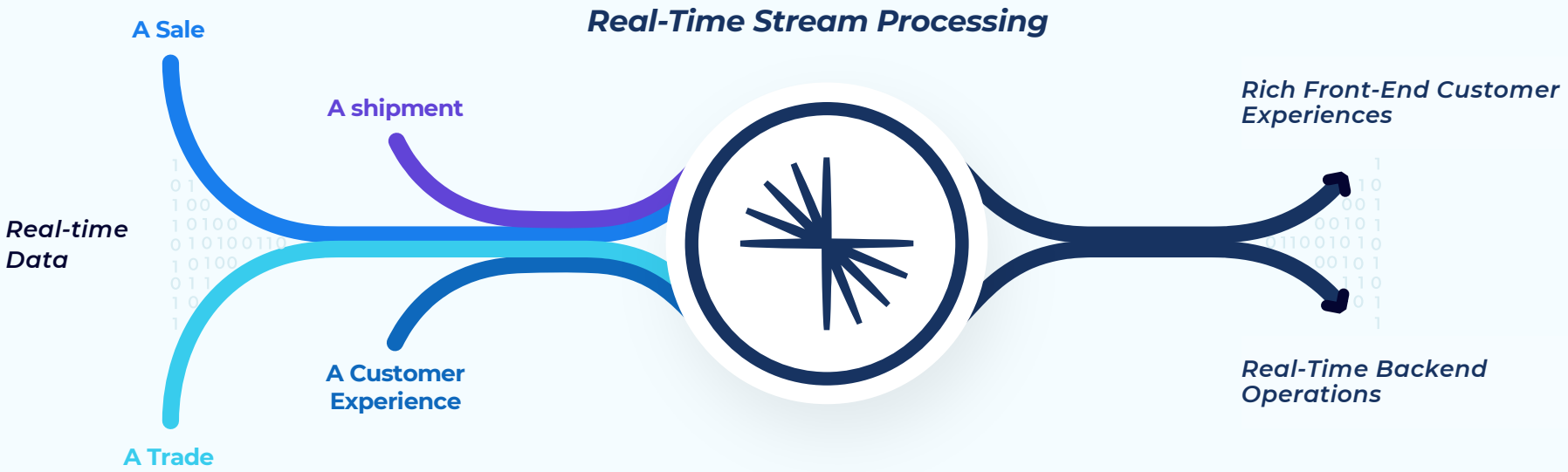
Elastic

Cost-saving

Fast Deployment

"More than 45% of IT spending on infrastructure and application software will shift from on-prem to cloud by 2024." – ***Gartner***

Data streaming ~~and Apache Kafka~~ is required for a modern cloud data platform



> *"We need to shift our thinking from everything at rest, to everything in motion."* — citi

Confluent Cloud

Help enterprises migrate and modernize their data platforms to power real-time analytics and applications

Cloud-Native

Apache Kafka, fully managed and re-architected to harness the power of the cloud

Complete

Go beyond Kafka to build real-time apps on the cloud quickly, reliably, and securely

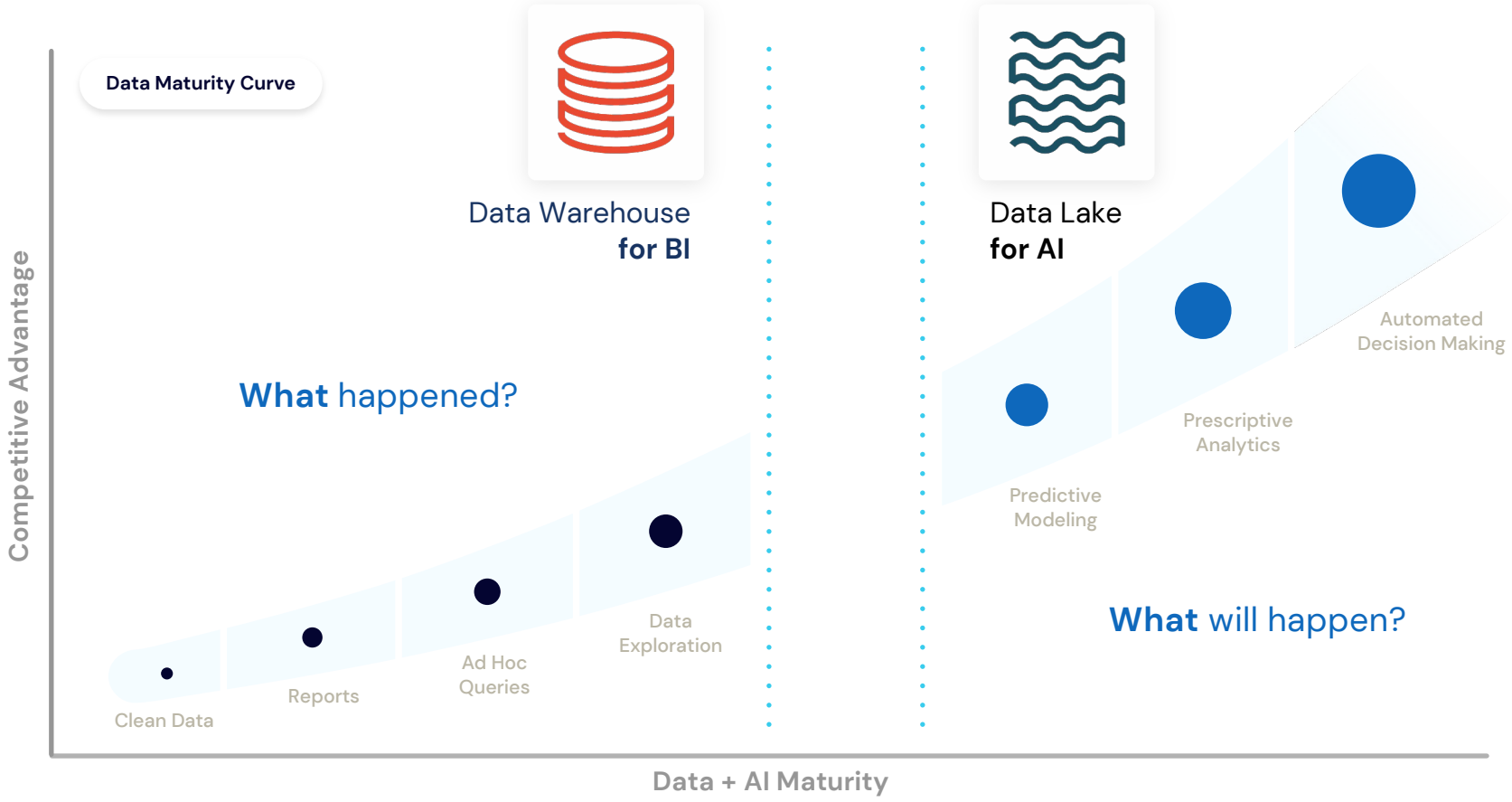
Everywhere

Maintain flexibility with deployments on any cloud and connect across clouds and on-premises that sync in real-time

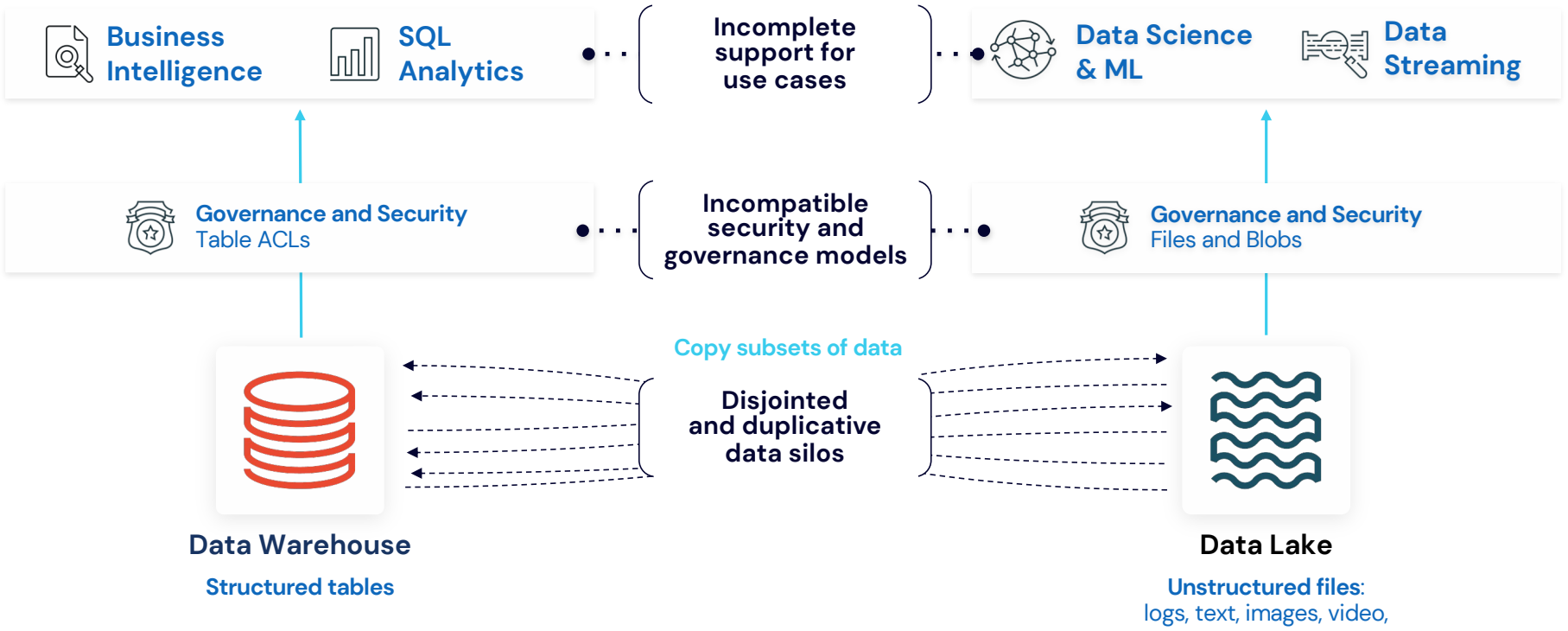
Stream confidently on a highly available platform built by the world's utmost Apache Kafka experts with built-in enterprise-ready security, compliance, and privacy controls

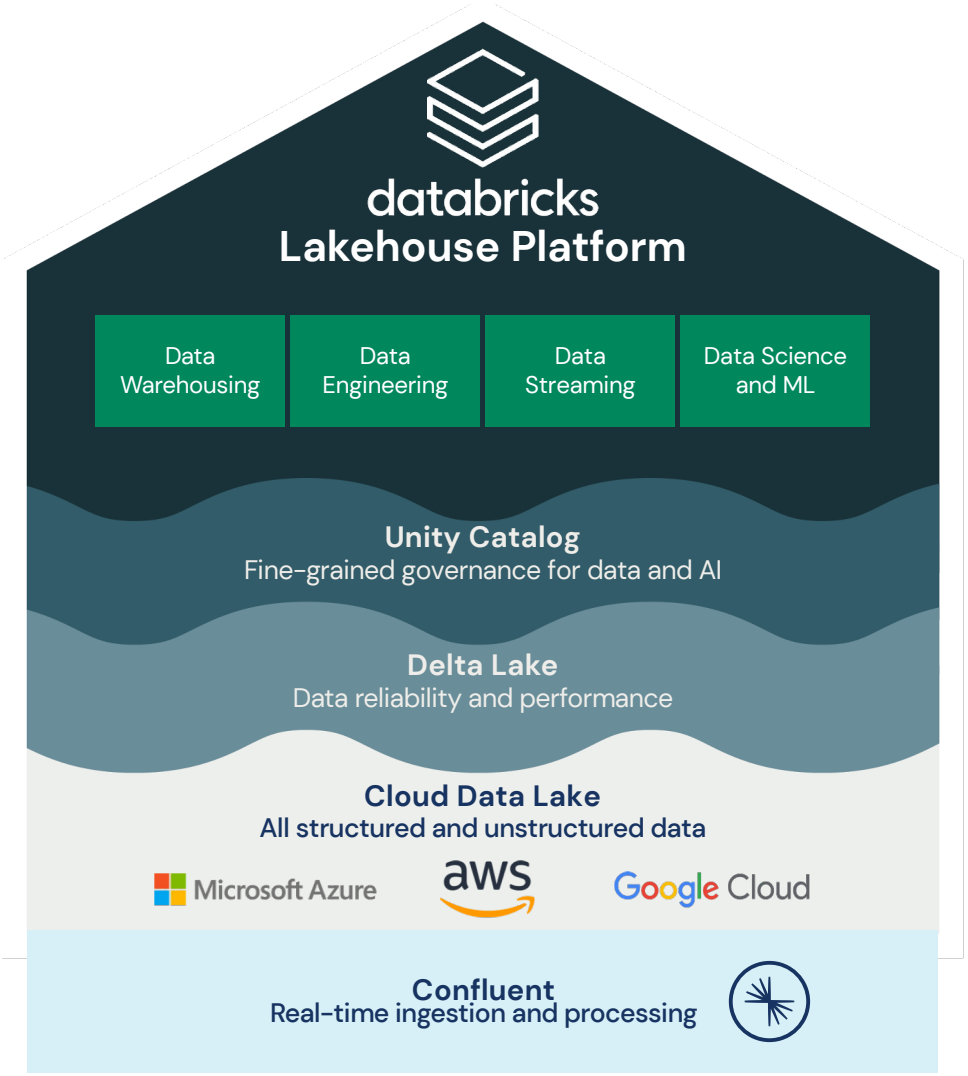
Most enterprises still struggle with data, analytics, and AI

Realizing this requires two disparate, incompatible data platforms



Realizing this requires two disparate, incompatible data platforms





Databricks Lakehouse Platform

Simple

Unify your data warehousing and AI use cases on a single platform

Open

Built on open source and open standards

Multicloud

One consistent data platform across clouds

Joint Value Proposition

Confluent makes analytics on Databricks better by bringing in data in real-time

A complete **real-time** event streaming and **analytics** solution for all data, analytics and AI workloads

CONFLUENT

+

databricks

BI and reporting in the lakehouse have more recent and valuable data for better decision making

ML/AI models can get real-time data from apps, IoT and consumer interactions for better predictions

Delivered at scale with the speed, security, and reliability required by enterprises

From migration to modernization

How Confluent and Databricks can accelerate your journey to real-time analytics in the cloud

Migrating and modernizing your data warehouse is not easy

High Cost

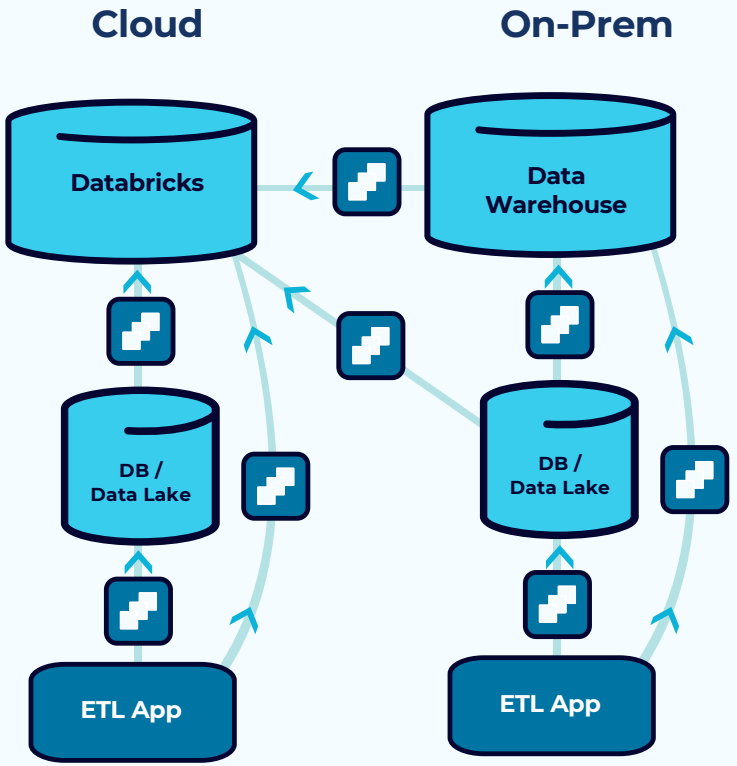
More data systems (e.g. DB) or DW costs to handle volume of data.

Slow time to analytics-ready data

Batch based processes with limited scale for “real-time”. Migrating DW to cloud can take years.

No support for hybrid/multicloud

Having distributed data and systems across environments compounds issues.



Modernize your data warehouse with Confluent and Databricks

✓ **Reduce total cost of ownership**

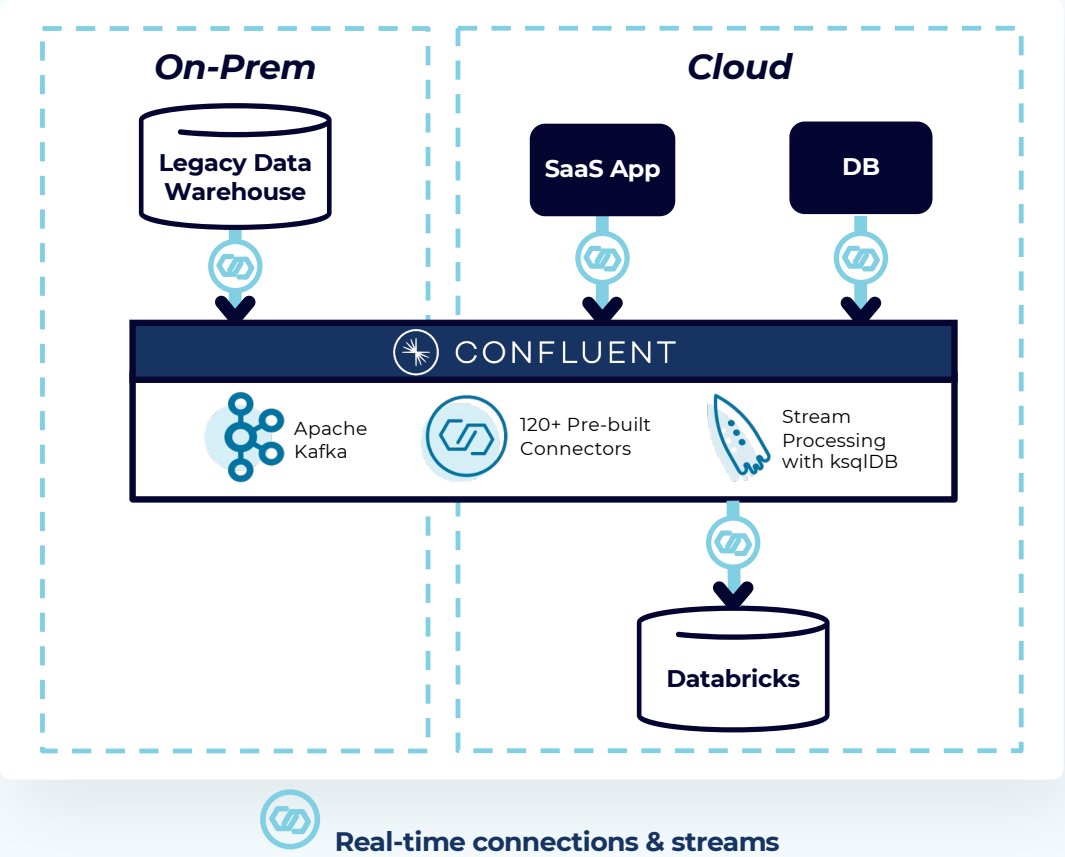
Process data in stream to lower DW costs. Lower data pipeline TCO with fully managed cloud service.

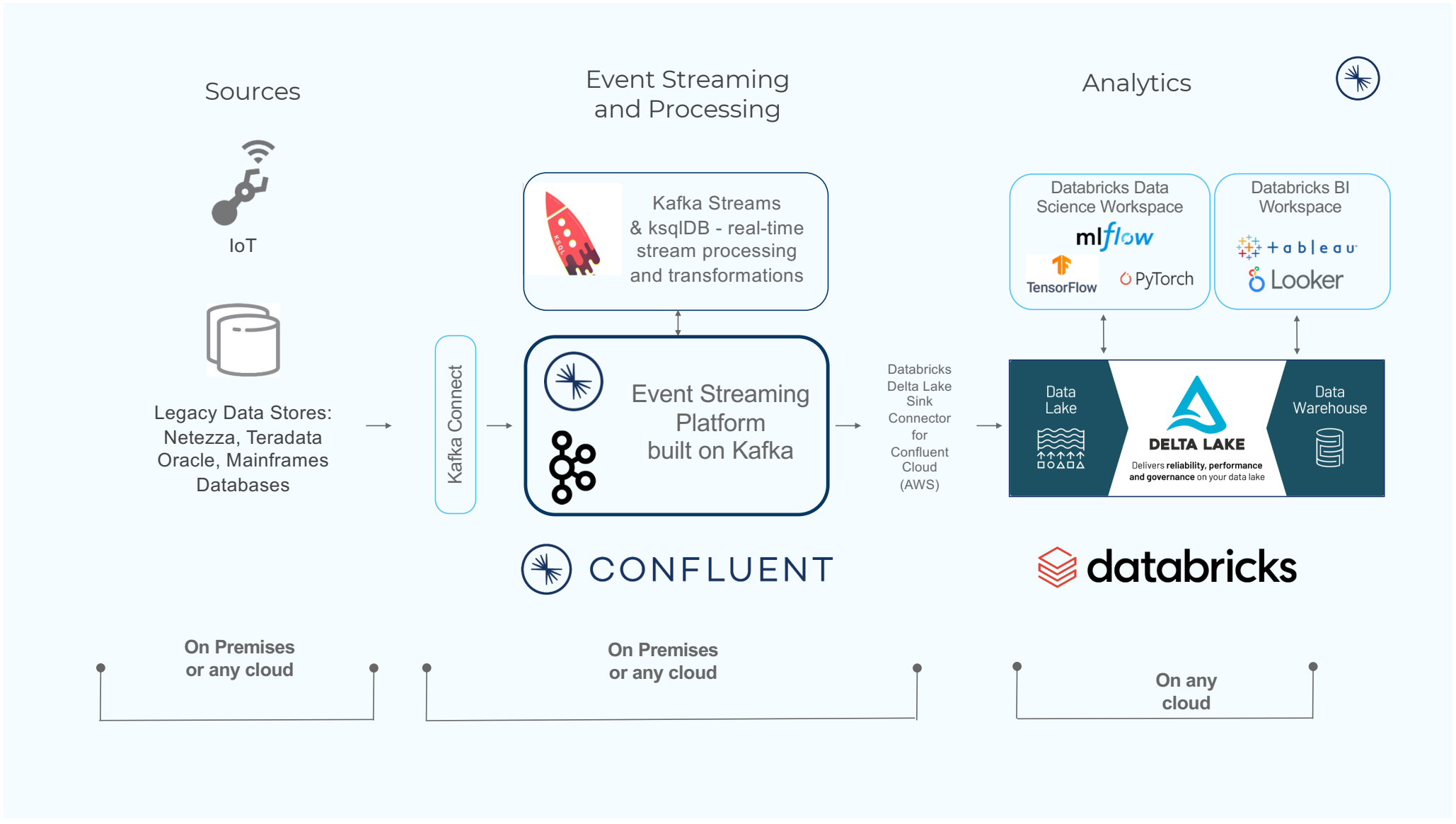
✓ **Power new analytics and apps**

Link on-prem and cloud for easier data movement across environments with **real-time event streaming** to modernize your analytics and applications

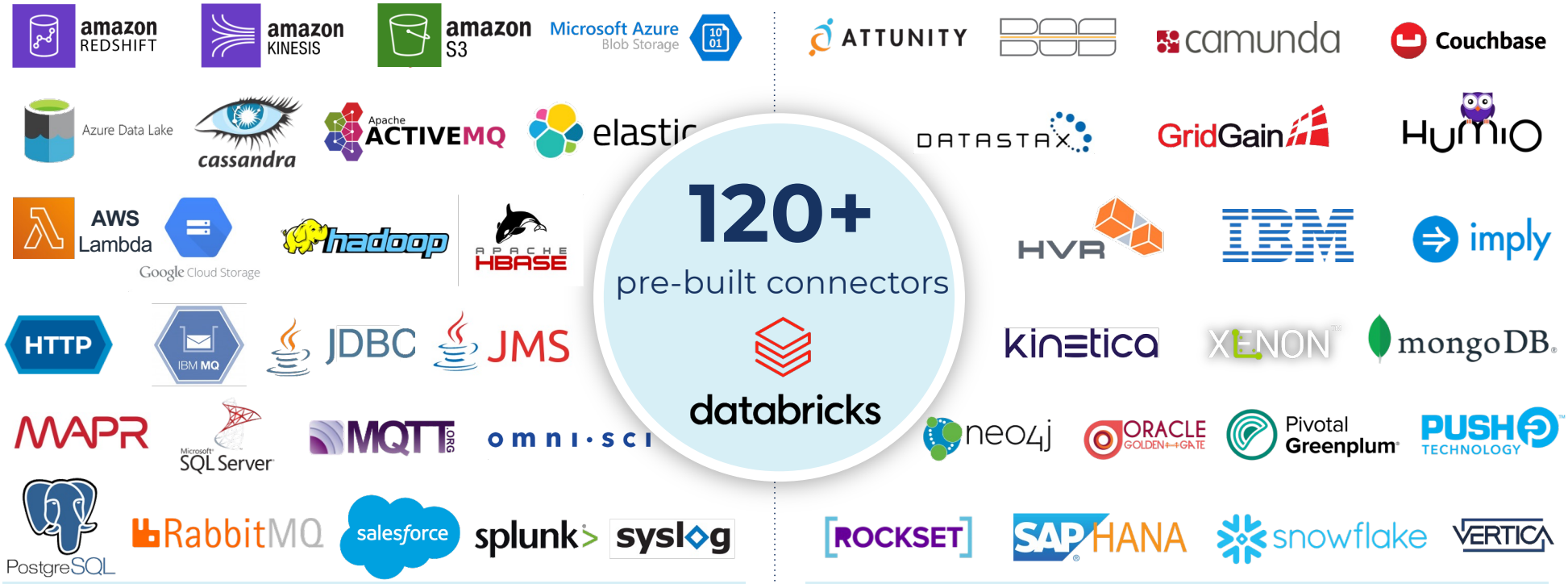
✓ **Get more data to and from your DW**

Break data silos and easily connect your data warehouse to popular sources and sinks using Confluent's 120+ pre-built connectors




Sources
IoT
Legacy Data Stores:
Netezza, Teradata
Oracle, Mainframes
Databases
Event Streaming
and Processing
Kafka Streams
& ksqlDB - real-time
stream processing
and transformations
Kafka Connect
Event Streaming
Platform
built on Kafka
CONFLUENT
Databricks
Delta Lake
Sink
Connector
for
Confluent
Cloud
(AWS)
Analytics
Databricks Data
Science Workspace
mlflow
TensorFlow
PyTorch
Databricks BI
Workspace
+ableau
Looker
Data
Lake
DELTA LAKE
Delivers reliability, performance
and governance on your data lake
Data
Warehouse
databricks
On Premises
or any cloud
On Premises
or any cloud
On any
cloud

Instantly Connect Popular Data Sources & Sinks



100+ Confluent Supported

30+ Partner Supported, Confluent Verified

Confluent: Everywhere

Fully-Managed Service

Confluent Cloud

Apache Kafka Re-engineered for the Cloud

Available on the leading public clouds

aws

Microsoft Azure

Self-Managed Software

Confluent Platform

The Enterprise Distribution of Apache Kafka

Deploy on any platform, on-prem or cloud

3 Modalities of Stream Processing with Confluent

Kafka clients

```
ConsumerRecords<String, String> records = consumer.poll(100);
Map<String, Integer> counts = new DefaultMap<String,
Integer>();
for (ConsumerRecord<String, Integer> record : records) {
  String key = record.key();
  int c = counts.get(key)
  c += record.value()
  counts.put(key, c)
}
for (Map.Entry<String, Integer> entry : counts.entrySet()) {
  int stateCount;
  int attempts;
  while (attempts++ < MAX_RETRIES) {
   try {
     stateCount = stateStore.getValue(entry.getKey())
     stateStore.setValue(entry.getKey(), entry.getValue() +
stateCount)
     break;
   } catch (StateStoreException e) {
     RetryUtils.backoff(attempts);
   }
 }
}
```

Kafka Streams

```
builder
    .stream("input-stream",
            Consumed.with(Serdes.String(), Serdes.String()))
    .groupBy((key, value) -> value)
    .count()
    .toStream()
    .to("counts", Produced.with(Serdes.String(), Serdes.Long()));
```

ksqlDB

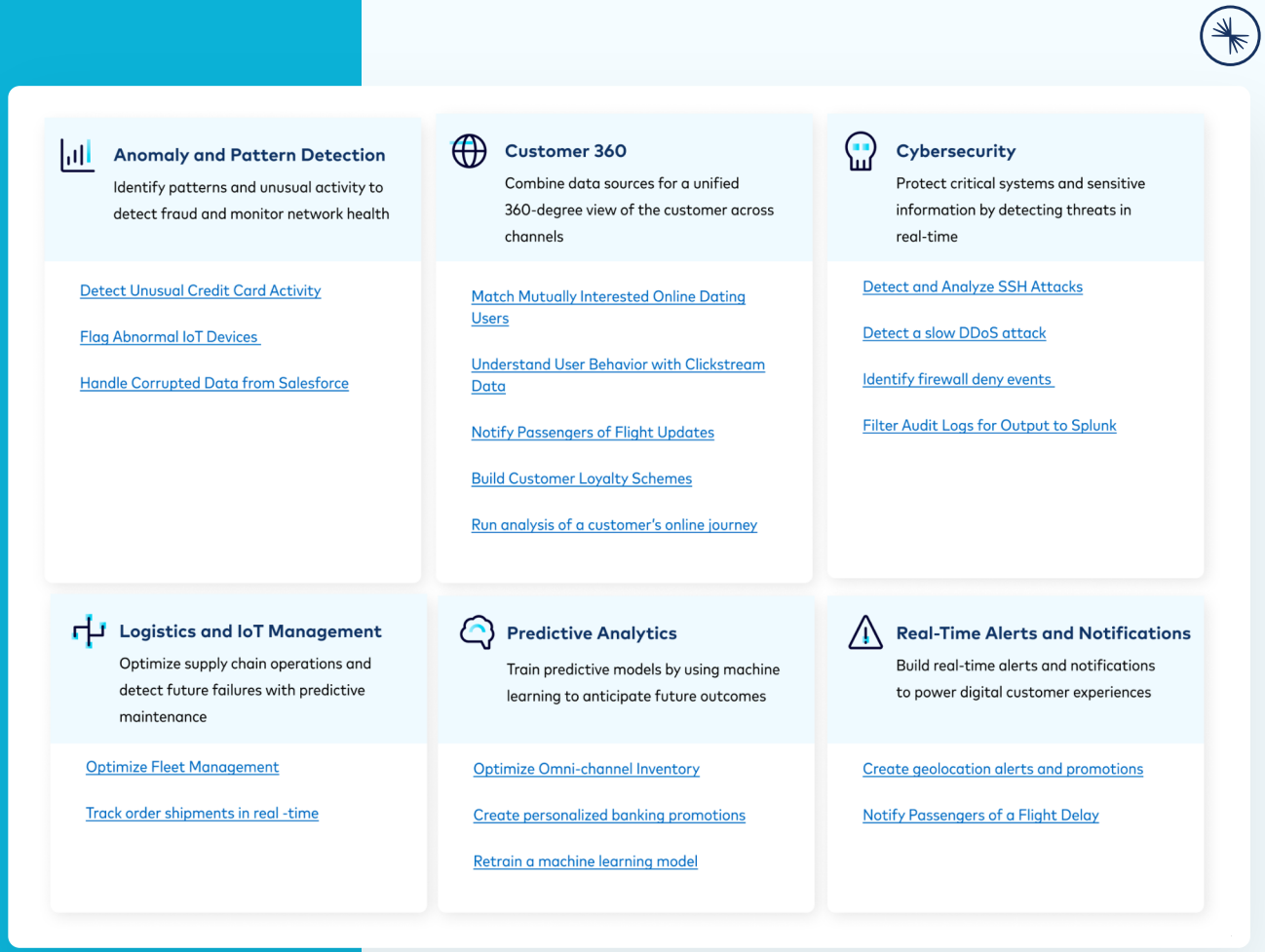
```
SELECT x, count(*) FROM stream GROUP BY x EMIT CHANGES;
```

Deliver value faster with ksqlDB Recipes

Delve into our collection of top stream processing use cases, which provide code samples and step-by-step tutorials to help you from idea to proof-of-concept.

Plus, to make it even easier, there is a one-click button that will launch the SQL script directly within Confluent Cloud.

developer.confluent.io/tutorials



Anomaly and Pattern Detection

Identify patterns and unusual activity to detect fraud and monitor network health

- Detect Unusual Credit Card Activity
- Flag Abnormal IoT Devices
- Handle Corrupted Data from Salesforce

Customer 360

Combine data sources for a unified 360-degree view of the customer across channels

- Match Mutually Interested Online Dating Users
- Understand User Behavior with Clickstream Data
- Notify Passengers of Flight Updates
- Build Customer Loyalty Schemes
- Run analysis of a customer's online journey

Cybersecurity

Protect critical systems and sensitive information by detecting threats in real-time

- Detect and Analyze SSH Attacks
- Detect a slow DDoS attack
- Identify firewall deny events
- Filter Audit Logs for Output to Splunk

Logistics and IoT Management

Optimize supply chain operations and detect future failures with predictive maintenance

- Optimize Fleet Management
- Track order shipments in real -time

Predictive Analytics

Train predictive models by using machine learning to anticipate future outcomes

- Optimize Omni-channel Inventory
- Create personalized banking promotions
- Retrain a machine learning model

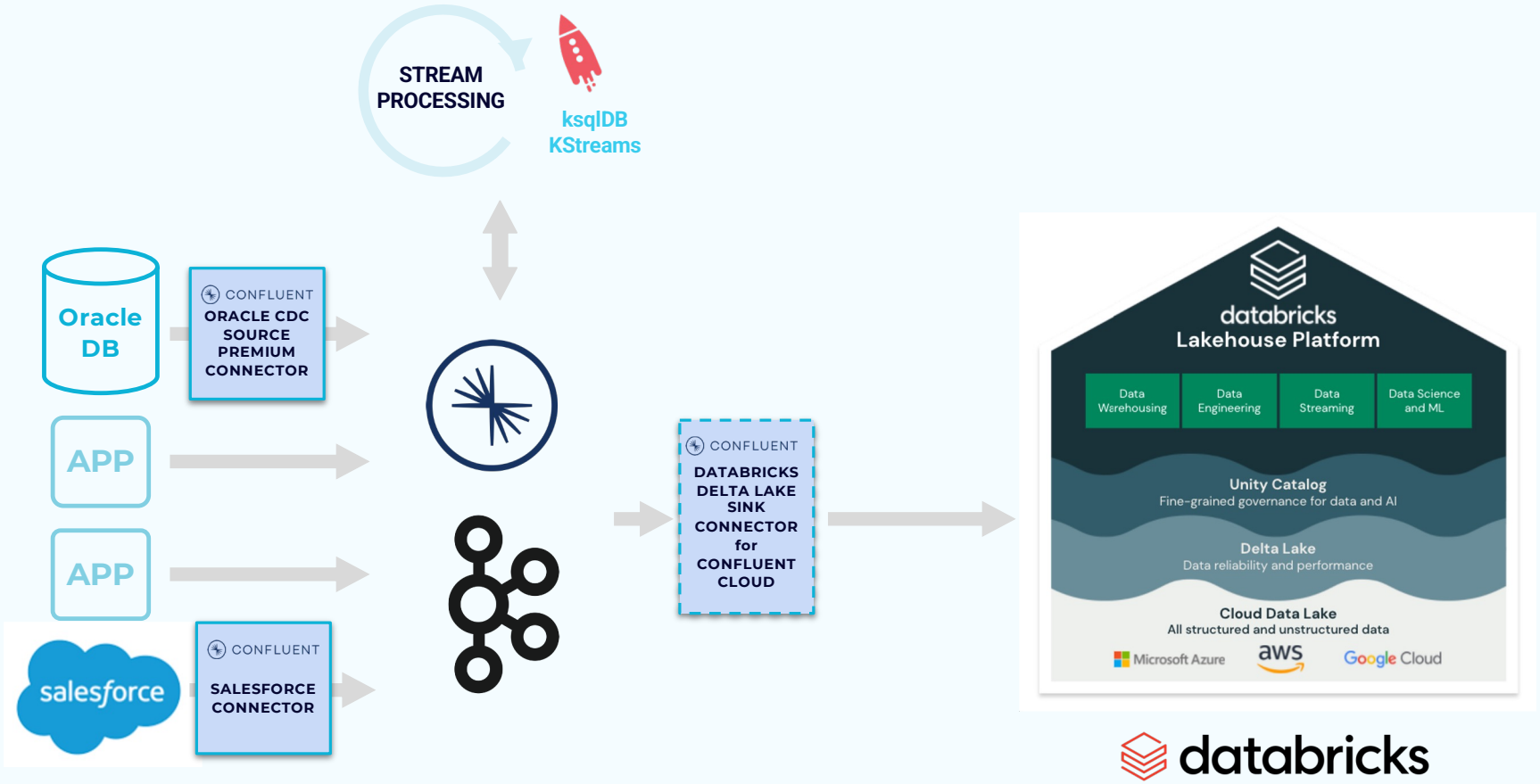
Real-Time Alerts and Notifications

Build real-time alerts and notifications to power digital customer experiences

- Create geolocation alerts and promotions
- Notify Passengers of a Flight Delay

Demo

A financial service company wants to build real-time dashboards and analytics in Databricks but their data is spread across an on-prem Oracle data warehouse and cloud system.


STREAM PROCESSING
ksqlDB
KStreams
Oracle DB
CONFLUENT
ORACLE CDC SOURCE PREMIUM CONNECTOR
APP
APP
salesforce
CONFLUENT
SALESFORCE CONNECTOR
CONFLUENT
DATABRICKS DELTA LAKE SINK CONNECTOR for CONFLUENT CLOUD
databricks
Lakehouse Platform
Data Warehousing
Data Engineering
Data Streaming
Data Science and ML
Unity Catalog
Fine-grained governance for data and AI
Delta Lake
Data reliability and performance
Cloud Data Lake
All structured and unstructured data
Microsoft Azure
aws
Google Cloud
databricks

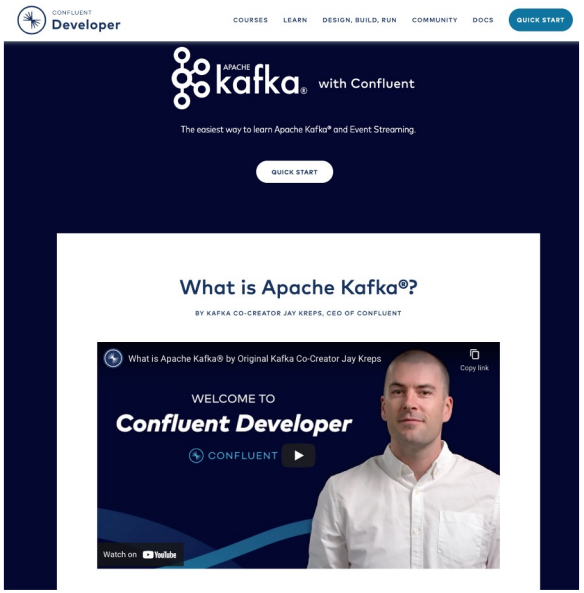
Q&A

Please drop your questions in the chat.

Get started for FREE

$400 free over
2 months from sign-up

confluent.io/confluent-cloud



developer.confluent.io

The new one-stop shop for learning Apache Kafka, event streaming, and associated cloud technologies. The site comes with an extensive library of Kafka-related courses.

Thank you!