
AWS
re:Invent

OPN202

App modernization on AWS with Apache Kafka and Confluent Cloud

Kai Waehner
Technology Evangelist

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



KAI WAEHNER

Email: contact@kai-waehner.de
LinkedIn: @KaiWaehner
www.confluent.io
www.kai-waehner.de

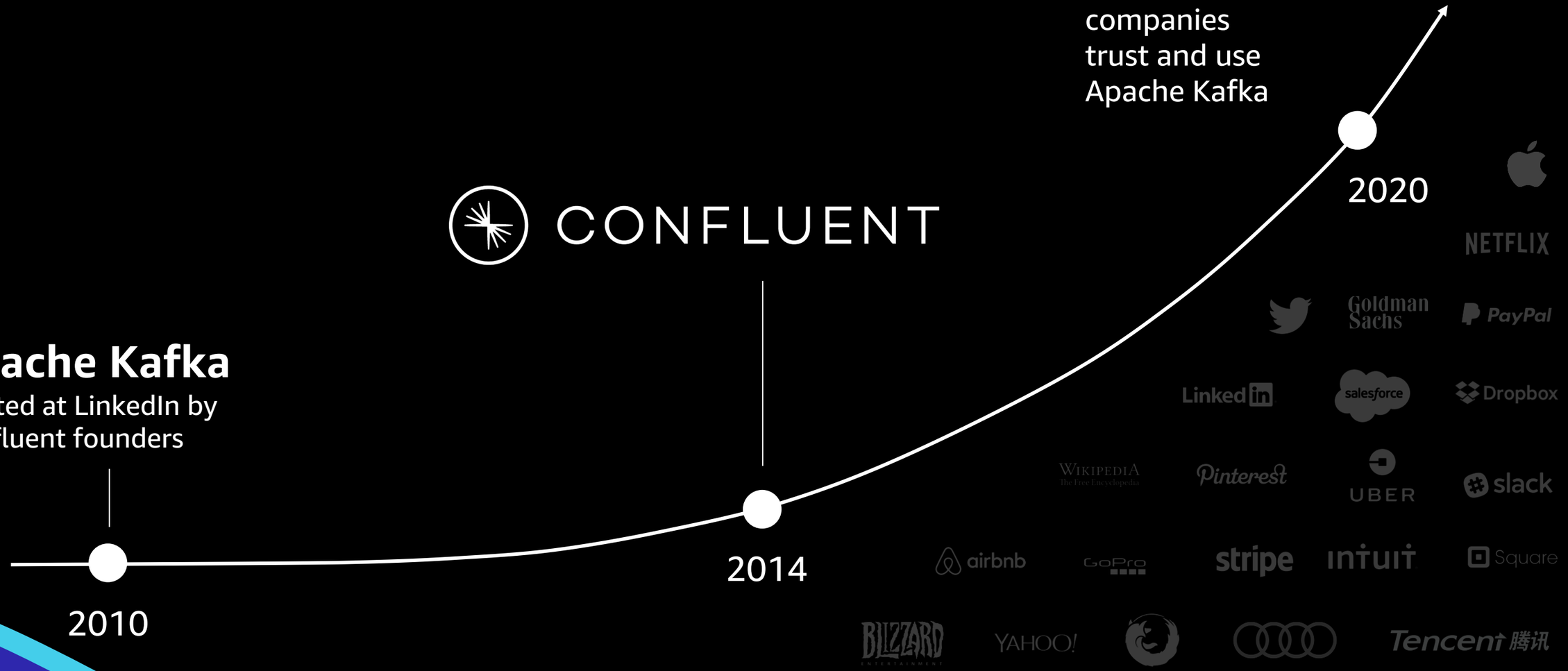
Event streaming with Apache Kafka and Confluent



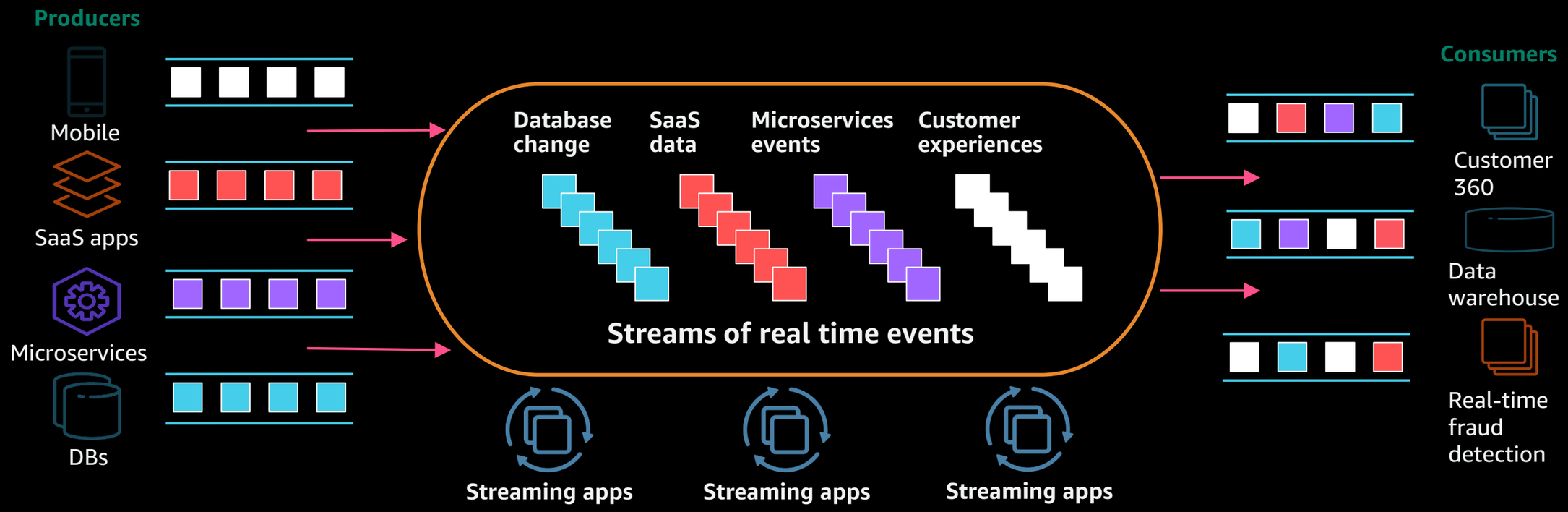
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

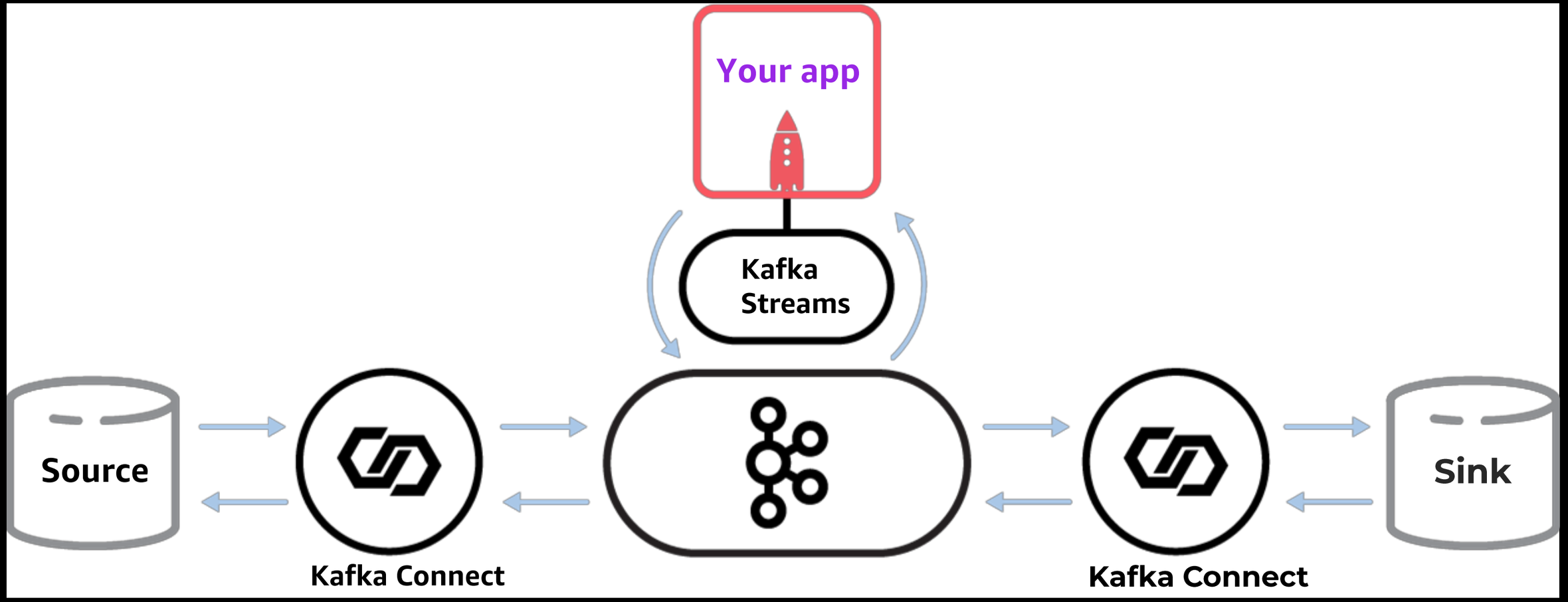
The rise of event streaming



An event streaming platform is the underpinning of an event-driven architecture



Kafka Connect and Kafka Streams



Confluent pioneered event streaming

Confluent founders are **original creators of Kafka**

Confluent team wrote **80% of Kafka software** commits, has over one million hours technical experience with Kafka, and operates 5,000+ clusters

Confluent Platform completes Apache Kafka and turns it into a secure, enterprise-ready platform

Confluent Cloud is the only multi-cloud, fully managed, pay-as-you-go event streaming service in the world

AWARDS

Forbes

THE Cloud100



Morgan Stanley

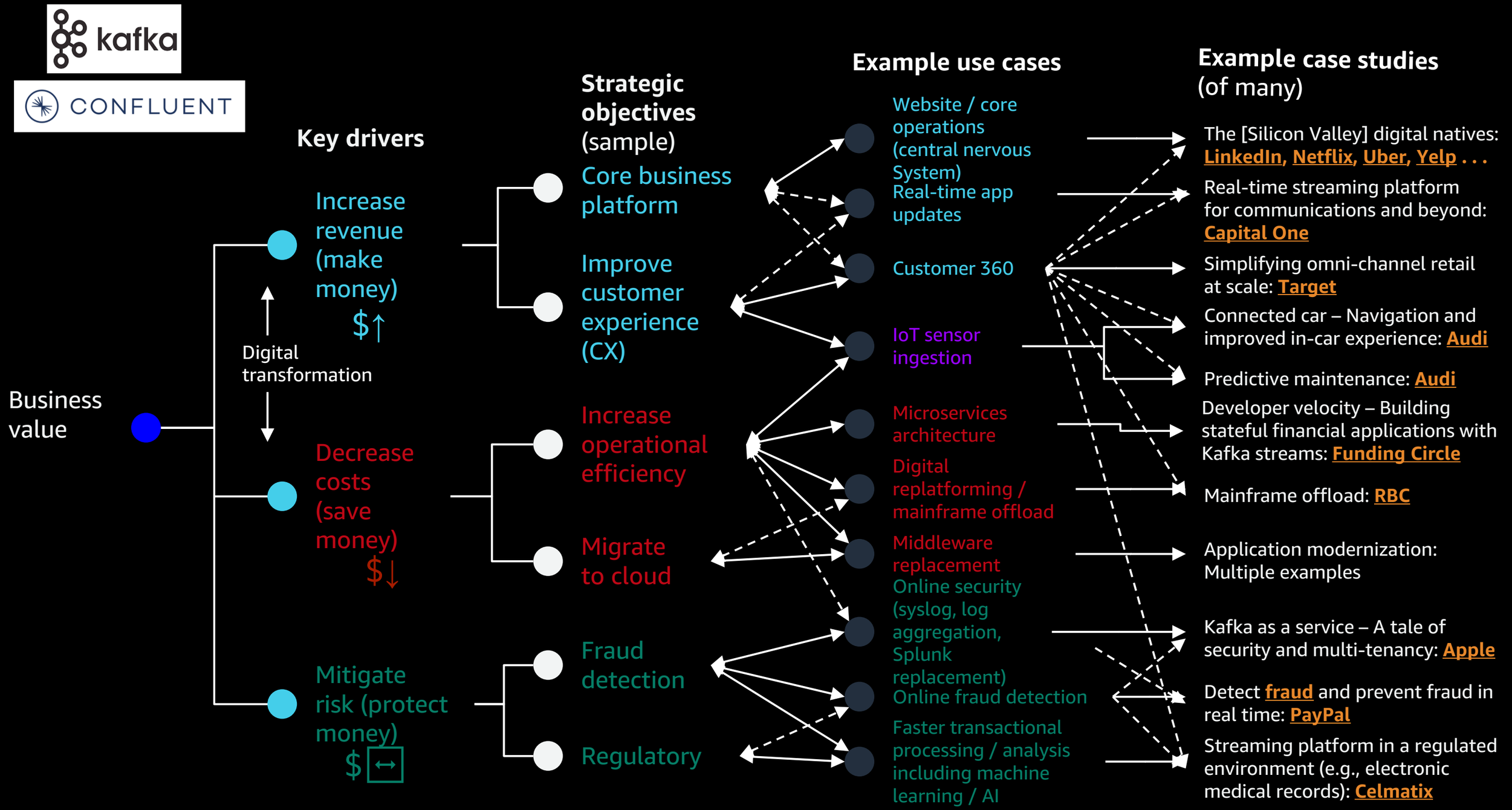
CTO Innovation Award Winner 2019

JPMorgan Chase & Co.

Hall of Innovation

BANK OF AMERICA

Enterprise Technology Innovation


kafka
CONFLUENT
Key drivers
Strategic objectives (sample)
Example use cases
Example case studies (of many)
Business value
Digital transformation
Increase revenue (make money) $↑
Decrease costs (save money) $↓
Mitigate risk (protect money) $↔
Core business platform
Improve customer experience (CX)
Increase operational efficiency
Migrate to cloud
Fraud detection
Regulatory
Website / core operations (central nervous System)
Real-time app updates
Customer 360
IoT sensor ingestion
Microservices architecture
Digital replatforming / mainframe offload
Middleware replacement
Online security (syslog, log aggregation, Splunk replacement)
Online fraud detection
Faster transactional processing / analysis including machine learning / AI
The [Silicon Valley] digital natives: LinkedIn, Netflix, Uber, Yelp . . .
Real-time streaming platform for communications and beyond: Capital One
Simplifying omni-channel retail at scale: Target
Connected car – Navigation and improved in-car experience: Audi
Predictive maintenance: Audi
Developer velocity – Building stateful financial applications with Kafka streams: Funding Circle
Mainframe offload: RBC
Application modernization: Multiple examples
Kafka as a service – A tale of security and multi-tenancy: Apple
Detect fraud and prevent fraud in real time: PayPal
Streaming platform in a regulated environment (e.g., electronic medical records): Celmatix

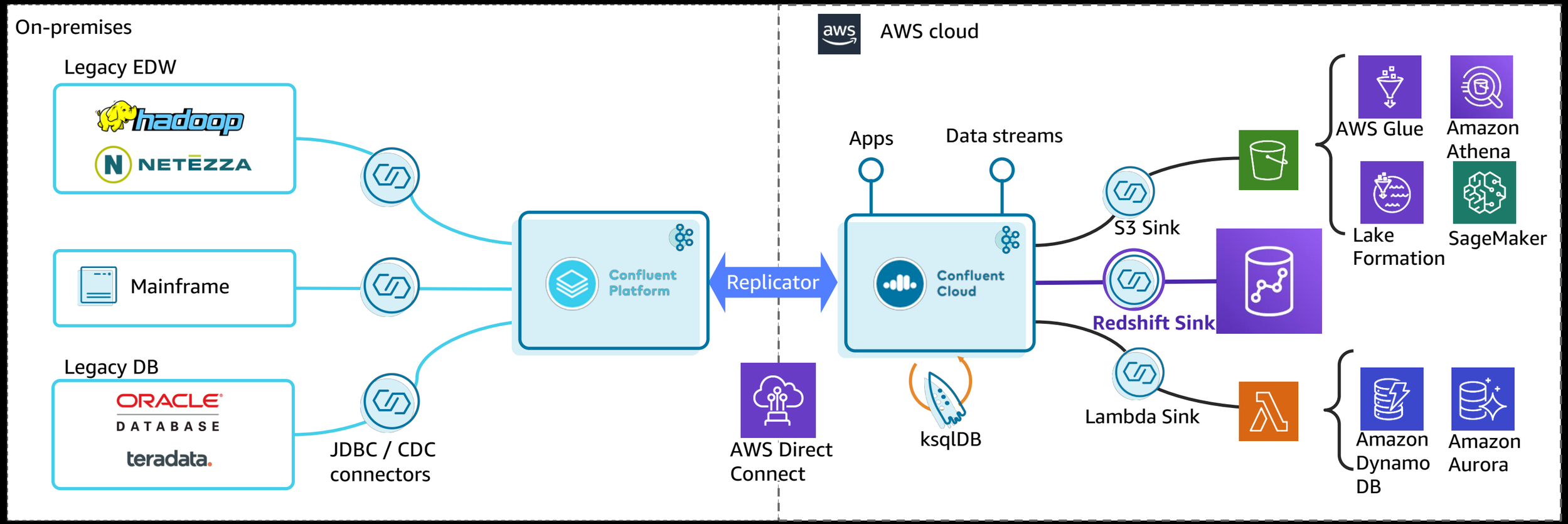
Modernization with hybrid architectures using Kafka, Confluent, and AWS



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Hybrid event streaming



Data integration with Kafka Connect




Amazon Redshift


Amazon Kinesis


Amazon S3

Microsoft Azure Blob Storage


Azure Data Lake








Google BigQuery


Google Cloud Storage

































100+ prebuilt connectors

























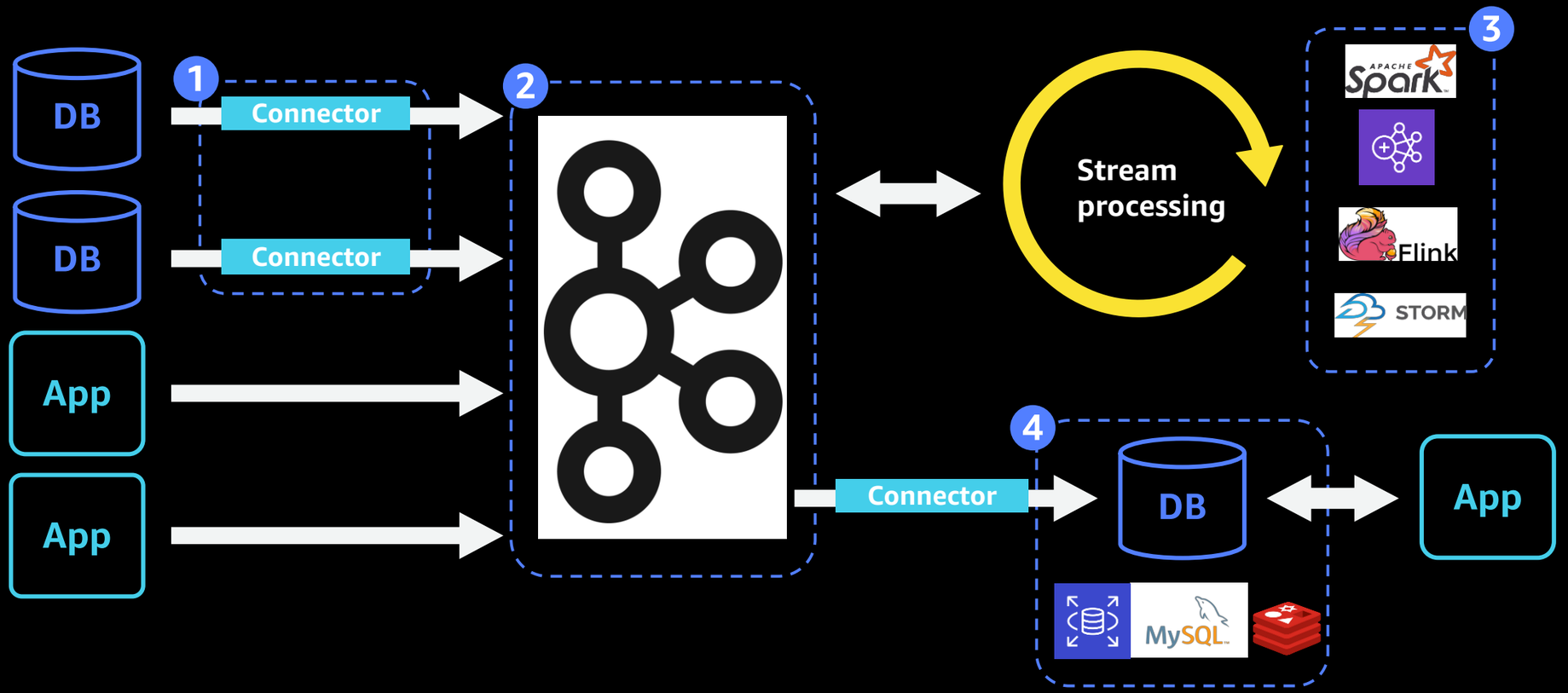




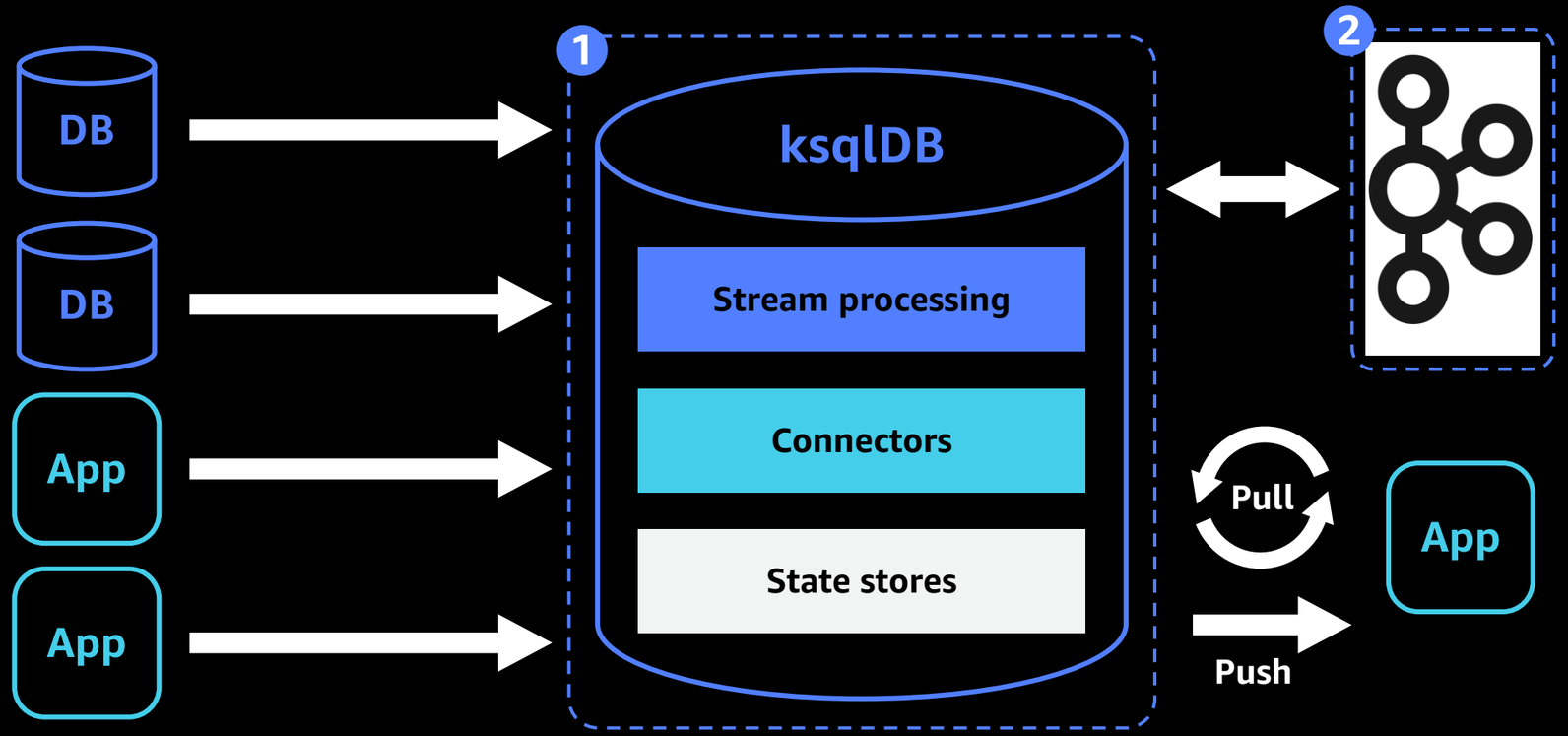
80+ Confluent supported

20+ Partner supported, Confluent verified

A common stream processing architecture



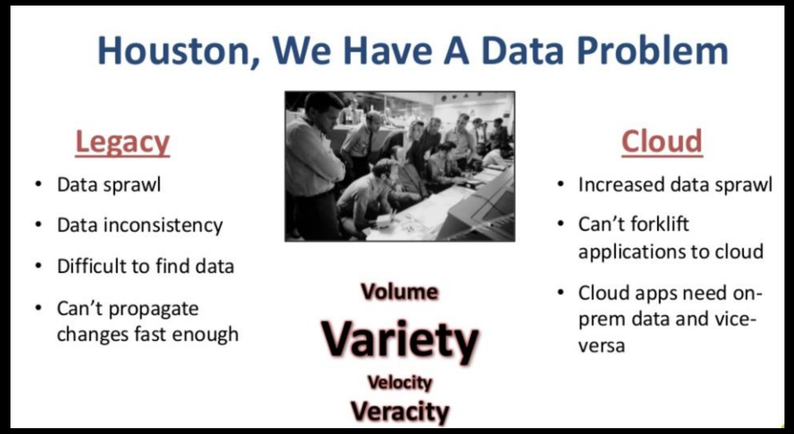
A simplified stream processing architecture with ksqlDB

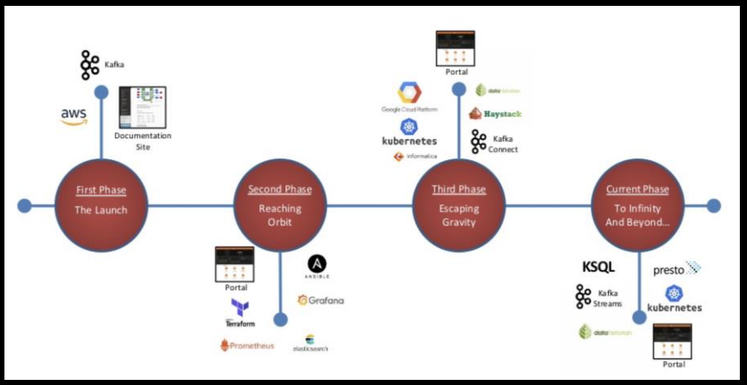


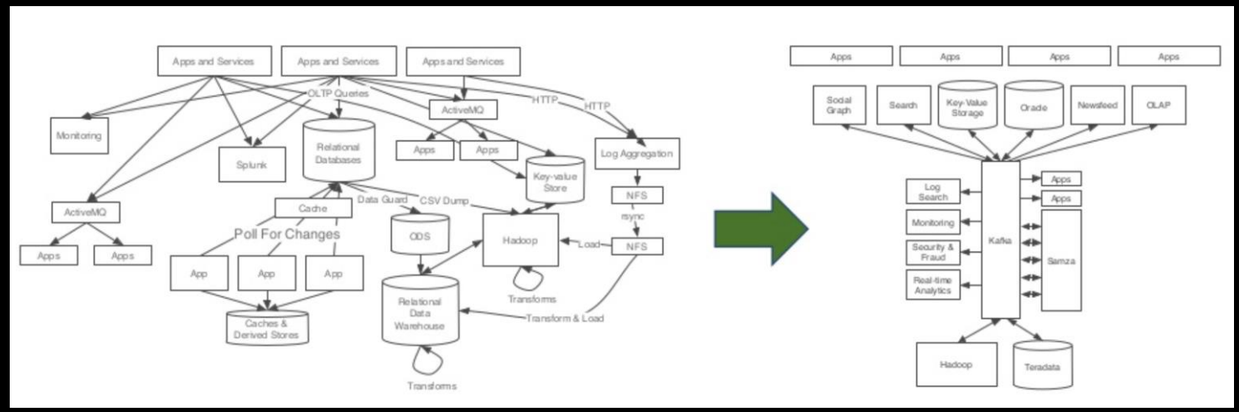
Bayer AG

CLOUD-FIRST STRATEGY AND STARTED A MULTI-YEAR TRANSITION TO THE CLOUD WITH A KAFKA-BASED CROSS-DATA CENTER DATA HUB

Connectivity
+
On premises and cloud
+
Hybrid real time replication at scale







Real-life use cases and architectures for event streaming with Apache Kafka and Confluent

https://www.confluent.io/kafka-summit-sf18/bringing-streaming-data-to-the-masses

Modernization and cloud migration



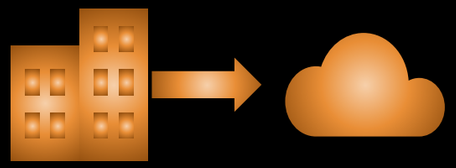
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Journey from mainframe to hybrid and cloud

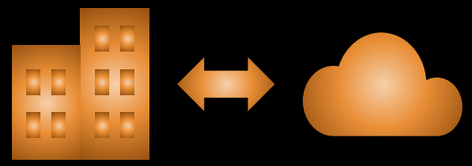


aws



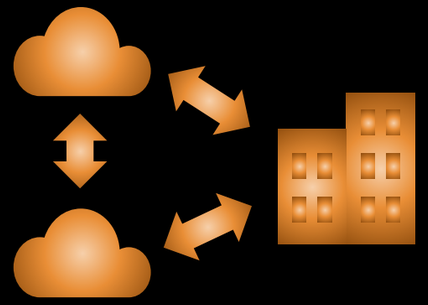
Cloud adoption

Phase 1



Hybrid cloud

Phase 2

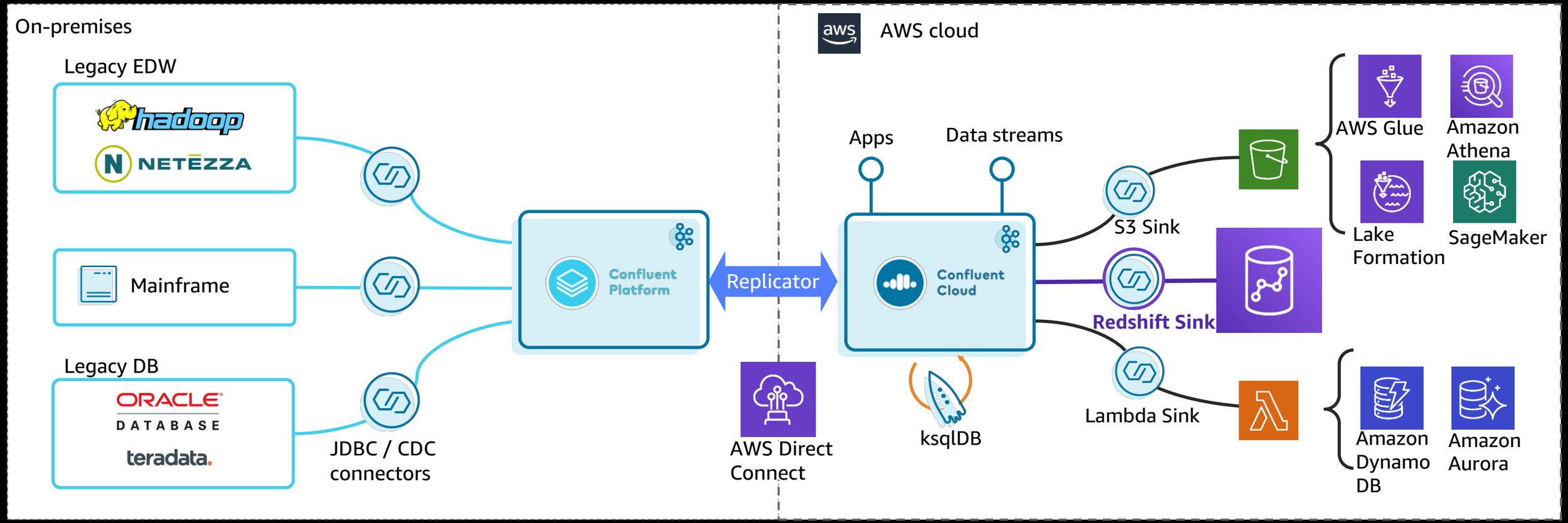


Cloud-first development

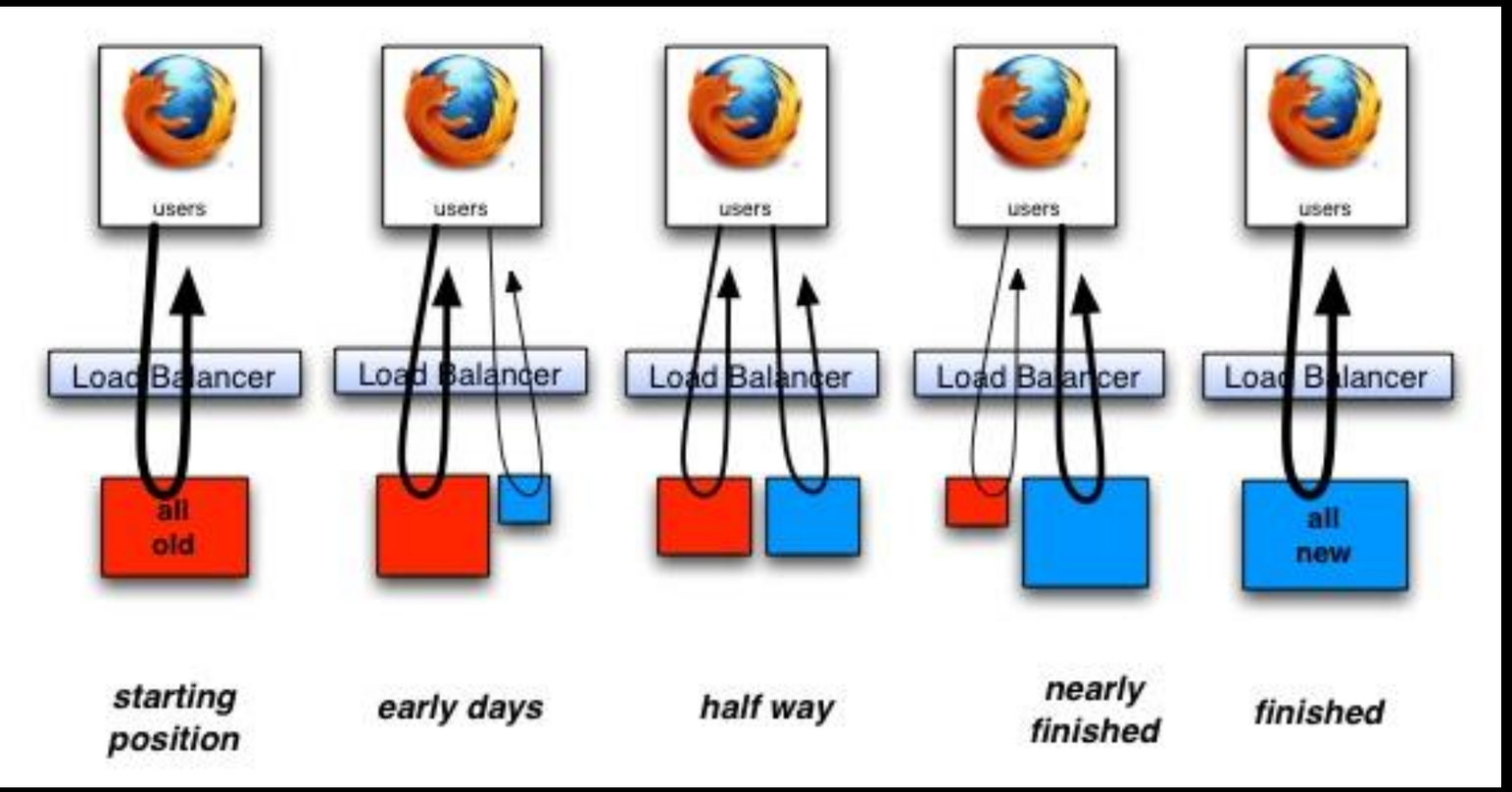
Phase 3

Case study – Bank CEO
“This is the last 5-year $20M IBM contract. Get rid of the mainframe!”

Hybrid event streaming



Strangler design pattern



https://paulhammant.com/2013/07/14/legacy-application-strangulation-case-studies/
https://martinfowler.com/bliki/StranglerFigApplication.html

"The most important reason to consider a strangler fig application over a cut-over rewrite is **reduced risk.**"

Martin Fowler

Year 0: Direct communication between mainframe and app

Core banking ‘1970’
(mainframe)

Date	Amount
1/27/2017	$4.56
1/22/2017	$32.14

On-premises only

Application



1) Direct legacy mainframe communication to app

Year 1: Kafka for decoupling between mainframe and app



Core banking ‘1970’
(mainframe)

Date	Amount
1/27/2017	$4.56
1/22/2017	$32.14

On-premises

Cloud

aws

Mainframe integration
- Change data capture (IIDR)
- Kafka Connect (JMS, MQ, JDBC)
- REST Proxy
- Kafka client
- Third-party CDC tool

Application

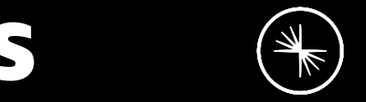


kafka

CONFLUENT

1) Direct legacy mainframe communication to app
2) Kafka for decoupling between mainframe and app

Year 2 to 4: New projects and applications



Core banking ‘1970’
(mainframe)

Date	Amount
1/27/2017	$4.56
1/22/2017	$32.14

On-premises

Cloud

aws

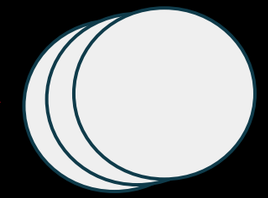
Mainframe integration
- Change data capture (IIDR)
- Kafka Connect (JMS, MQ, JDBC)
- REST Proxy
- Kafka client
- Third-party CDC tool

Application



kafka

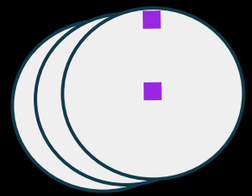
CONFLUENT



External solution

Big data project (Elastic, Spark, AWS Services . . .)

1) Direct legacy mainframe communication to app
2) Kafka for decoupling between mainframe and app
3) New projects and applications



Microservices

Agile, lightweight (but scalable, robust) applications

Year 5: Mainframe replacement



Cloud only

aws

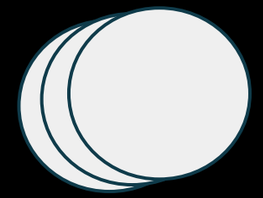
Application



Core banking ‘2020’
(modern technology)

Date	Amount
1/27/2017	$4.56
1/22/2017	$32.14

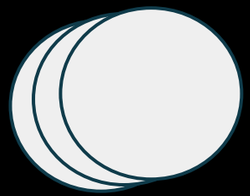
kafka
CONFLUENT



External solution

Big data project (Elastic, Spark, AWS Services . . .)

1) Direct legacy mainframe communication to app
2) Kafka for decoupling between mainframe and app
3) New projects and applications
4) Mainframe replacement



Microservices

Agile, lightweight (but scalable, robust) applications

Why Confluent?



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent Cloud: What does fully managed mean?



Most Kafka as a service offerings are partially managed

Infrastructure management (commodity)	• Upgrades (latest stable version of Kafka) • Patching • Maintenance	**Infra as a service**
Kafka-specific management	• Sizing (retention, latency, throughput, storage, etc.) • Data balancing for optimal performance • Performance tuning for real-time and latency requirements • Fixing Kafka bugs • Uptime monitoring and proactive remediation of issues • Recovery support from data corruption	**Platform as a service** **Harness full power of Kafka** **Mission-critical reliability**
Scaling	• Scaling the cluster as needed • Data balancing the cluster as nodes are added • Support for any Kafka issue with less than 60-minute response time	**Future proof** **Evolve as you need**

An event streaming platform needs more than Kafka Core
→ Fully managed data integration, stream processing, data governance, security . . .

CONFLUENT PLATFORM

Confluent Is Now Certified Ready on AWS Outposts



JOSEPH MORAIS



JOBIN GEORGE

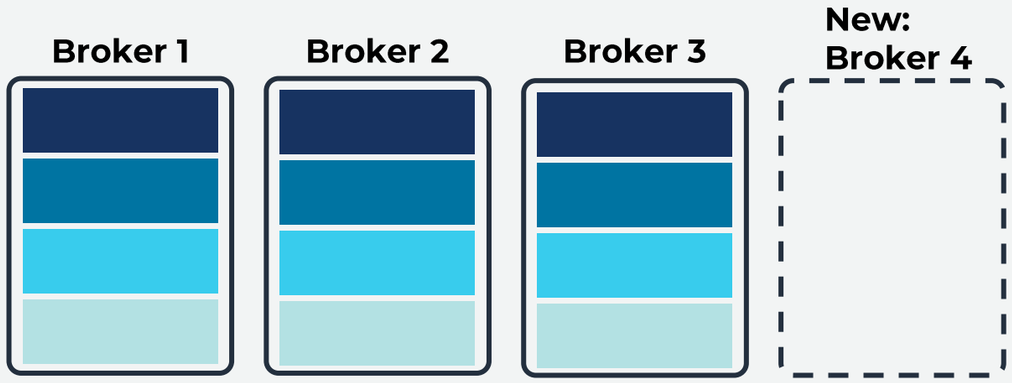
SEPTEMBER 15, 2020

Are you looking for a way to run AWS services on premises in your own datacenter? I am excited to share today that we have completed validation of support for running Confluent Platform on AWS Outposts. AWS Outposts delivers fully managed, configurable compute and storage racks built with Amazon Web Services (AWS) designed hardware that allows you to run instances as you would on Amazon EC2 locally in your datacenter or co-location facility. You get all these resources locally while seamlessly connecting to a broad array of cloud-based AWS services.

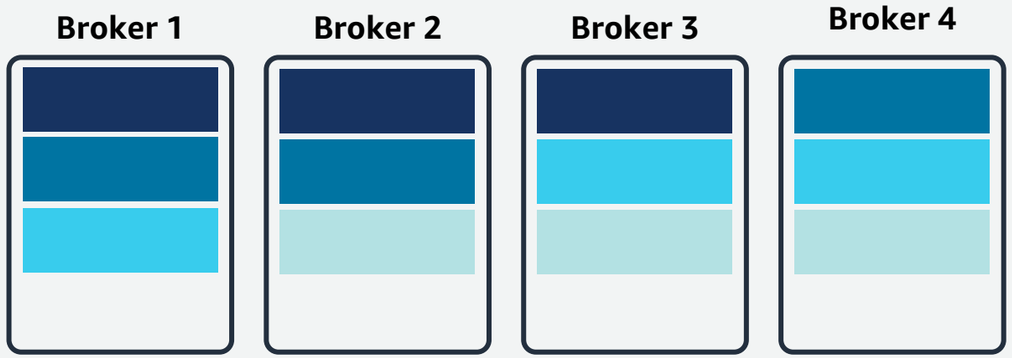
https://www.confluent.io/blog/confluent-platform-is-now-certified-ready-on-aws-outposts/

Tiered storage and self-balancing clusters for elastically scalable and cost-efficient Kafka

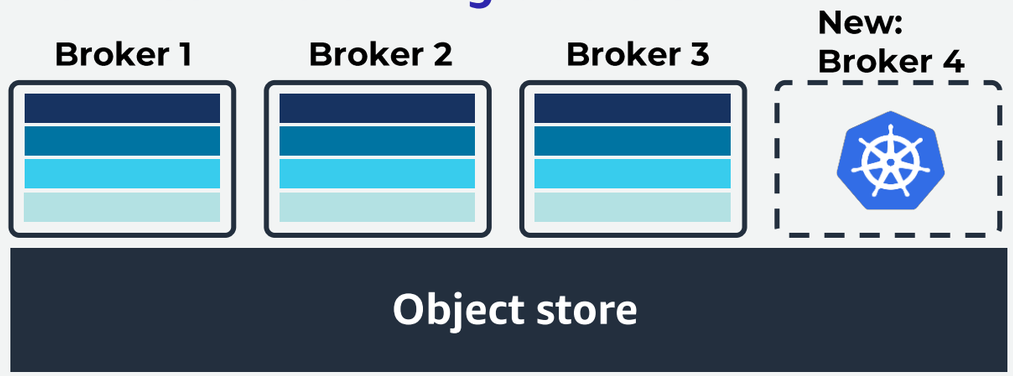
Inelastic scaling process



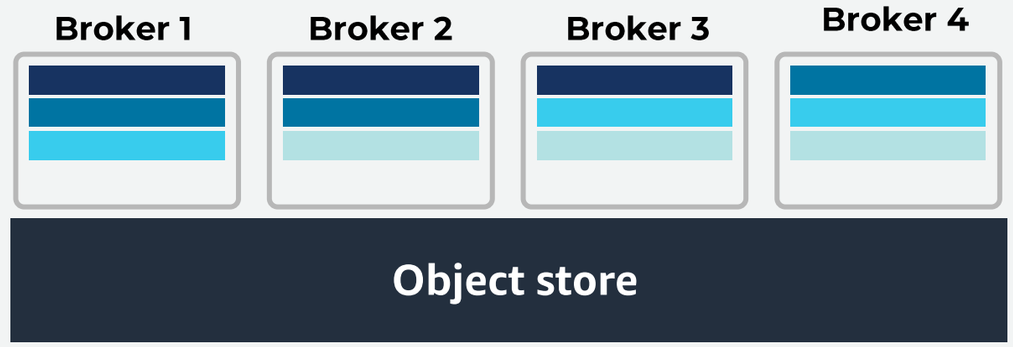
Lengthy, manual process for reassigning large topic partitions



Scaling with tiered storage and self-balancing clusters

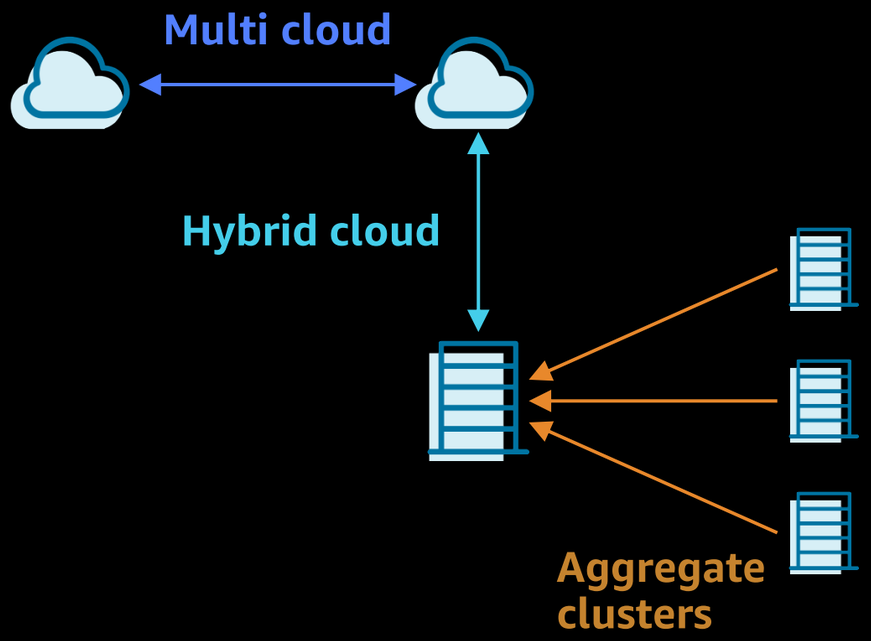


Fast, automated process for reassigning small topic partitions



Migrate data across multi-cloud environments

Confluent replicator



Make clusters globally available

Replicate clusters or a subset of topics across any distance

Aggregate or migrate clusters anywhere

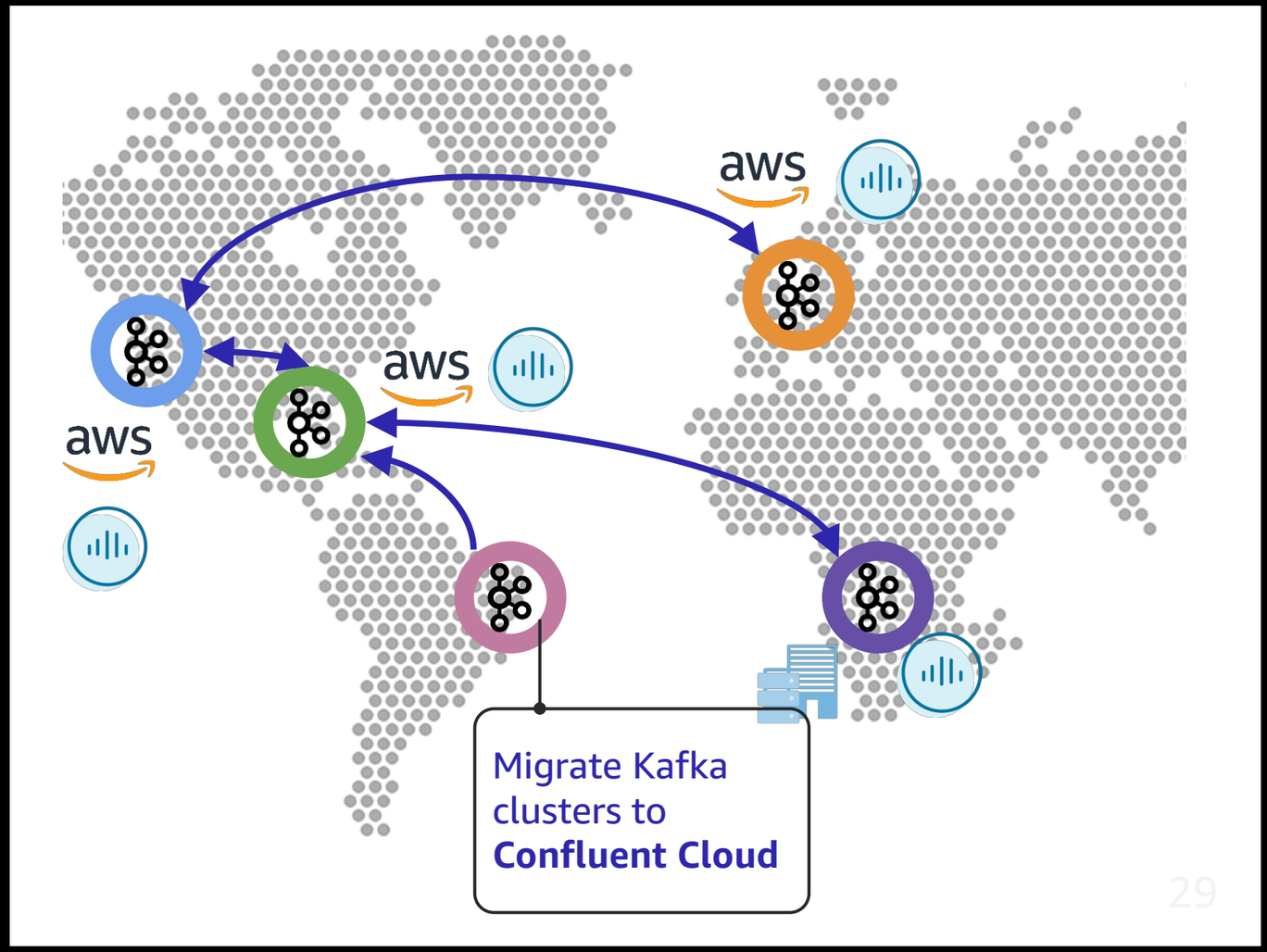
Aggregate many clusters together or migrate entire clusters to a preferred environment

Bridge self-managed clusters to a fully managed Kafka service

Enable hybrid cloud deployments with **Confluent Cloud**

Cluster linking uses the Kafka protocol

- Uses the Kafka protocol
- Requires **no** additional infrastructure (such as MirrorMaker)
- Preserves offsets



Thank you!



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.