AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

ARC328-S

SPONSORED BY CONFLUENT

Confluent at the edge on AWS

Joseph Morais
Staff Solutions Architect, AWS Evangelist
Confluent

Justin Lee
Staff Solutions Engineer
Confluent



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Who are we and why are we here?



Joseph Morais
Staff Solutions Architect, AWS Evangelist
Confluent



Justin Lee
Staff Solutions Architect
Confluent

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent at the edge on AWS agenda

01 **Apache Kafka, the event streaming platform**
How does it work at a high level?

02 **Who is Confluent?**
How do we make Apache Kafka easier?

03 **What is the edge?**
A discussion of different edge scenarios

04 **The edge can be difficult**
How can Confluent and AWS make edge more accessible to your organization?

05 **Confluent technology**
A discussion of new technology that makes running data in motion everywhere easier

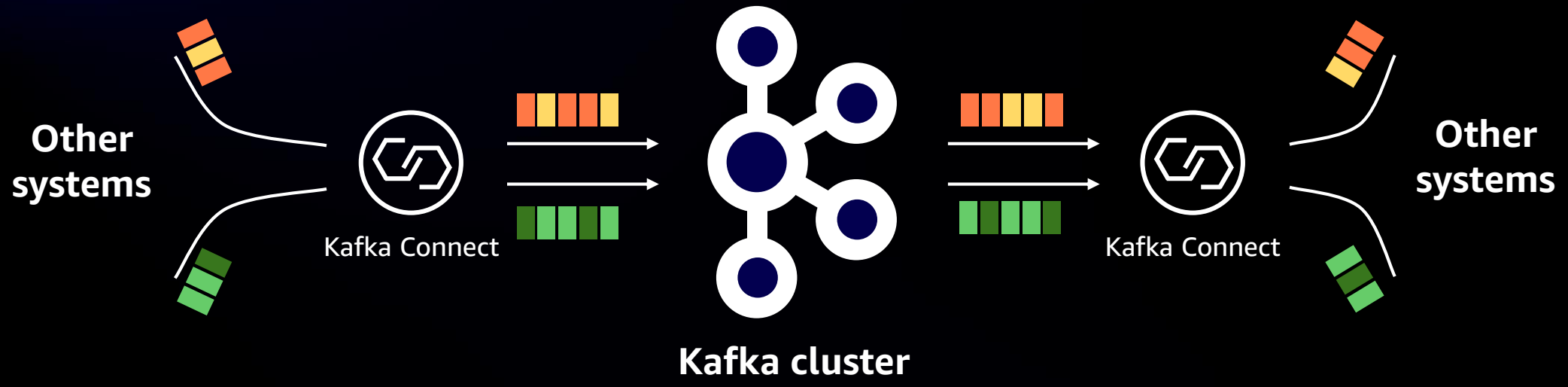
06 **Edge use cases**
Examples of where edge technology is used or desired

07 **Confluent & AWS Wavelength solution demo**

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Apache Kafka is an event streaming platform



Cloud storage services

ORACLE

elastic

Apache Ignite

mongoDB.

Solr

FTP

hazelcast

InfluxDB

hadoop

MQTT.ORG

cassandra

Apache Kudu

ATTUNITY

Apache HBase

Couchbase

Bloomberg

syslog-ng

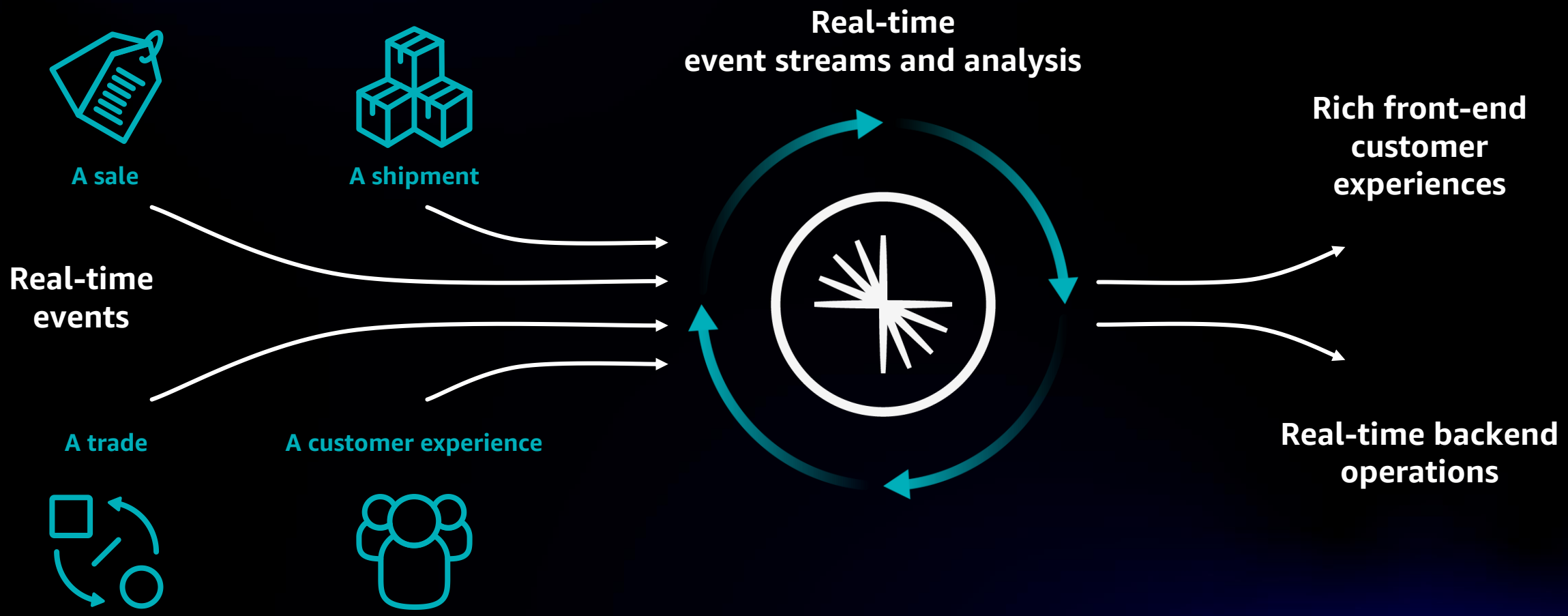
. . . many more

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Event streaming allows us to set data in motion

CONTINUOUSLY PROCESSING EVOLVING DATA STREAMS IN REAL TIME

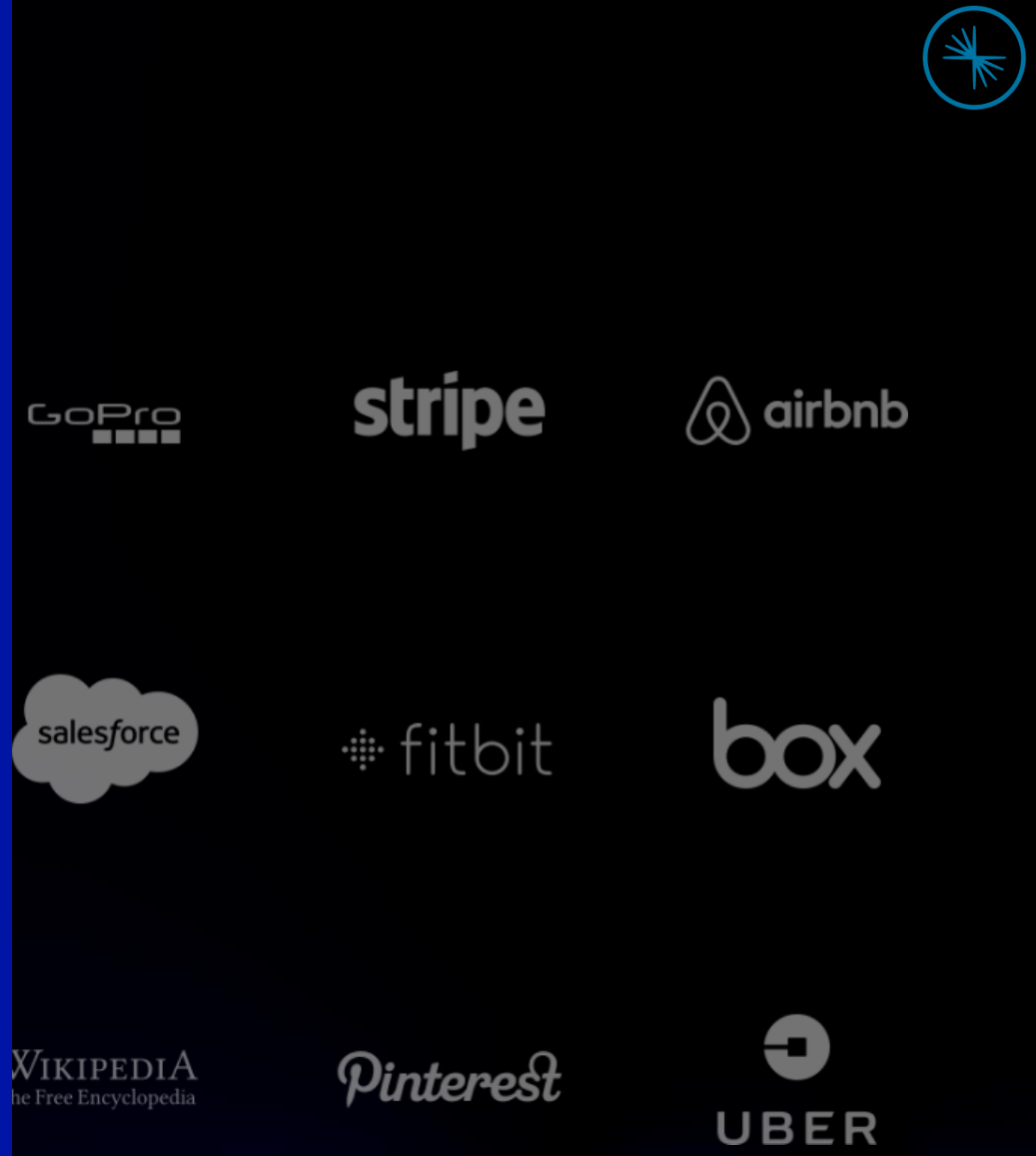


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The rise of data in motion

70%

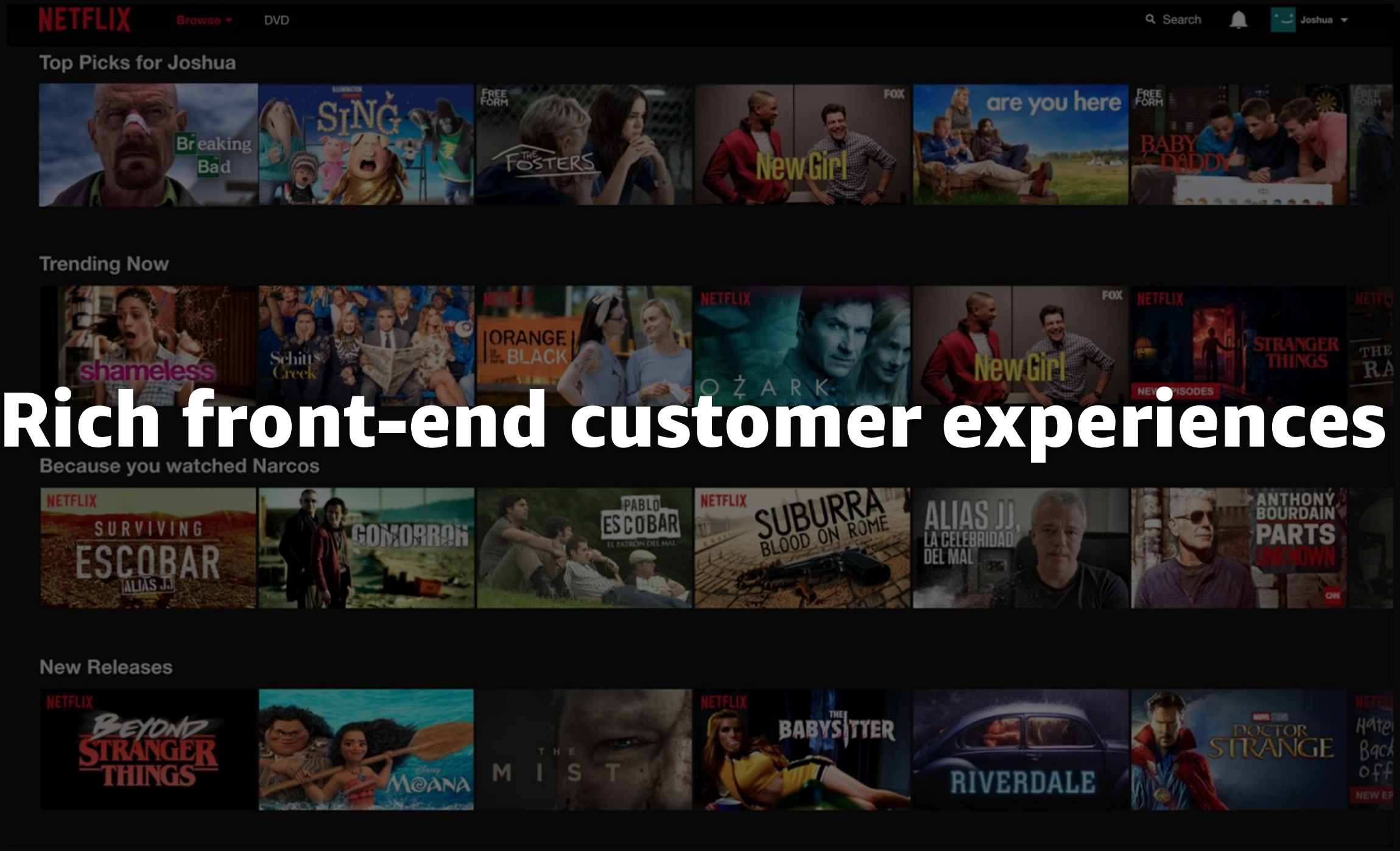
of Fortune 500 companies use Apache Kafka
(majority are Confluent customers)



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Data is changing the world



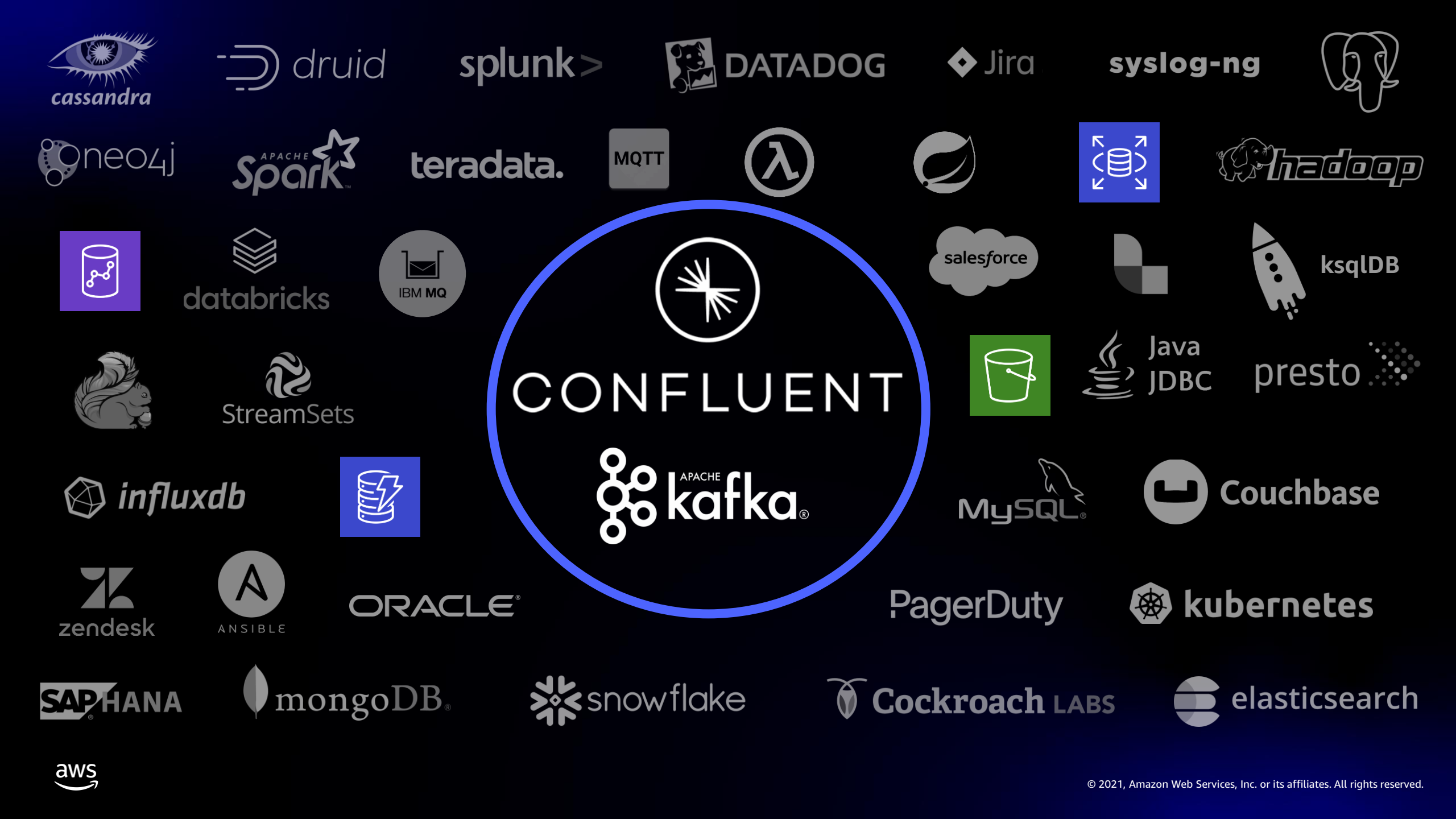
Rich front-end customer experiences


Real-time backend operations

Community

210 meetups	**54,364** meetup attendees	**626** KIPs
10,286 Jira tickets for Apache Kafka	**46,626** Emails sent to Apache Kafka mailing list	**23,363** Stack overflow questions

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent is the only company focused on data in motion

Vision

- Original Kafka creators founded Confluent
- Data in motion pioneers

Category leadership

- 80% of Kafka commits
- 1M+ hours of Kafka technical experience
- Operate 5K+ clusters

Product

- Extends Kafka to be secure and enterprise-ready
- Software or cloud-native service

Value

- Remove risk
- Deploy at scale
- Accelerate time to market

aws

AWARDS

Forbes

THE Cloud100

Morgan Stanley

CTO Innovation Award Winner 2019

JPMorgan Chase & Co.

Hall of Innovation

BANK OF AMERICA

Enterprise Technology Innovation

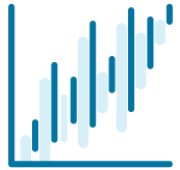
Confluent completes Apache Kafka



Freedom of choice to deploy how and where you want – cloud, on premises, hybrid, or multi-cloud



Connect across key data sources, sinks, and applications



Elastically scale to store and process massive amounts of data



Easiest way to build stream processing applications



Secure sensitive events to keep your business processes and applications compliant



Increase operational efficiency while keeping costs as low as possible

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent: Everywhere

Fully managed service

Confluent Cloud

Apache Kafka reengineered for the cloud

Available on aws

Self-managed software

Confluent Platform

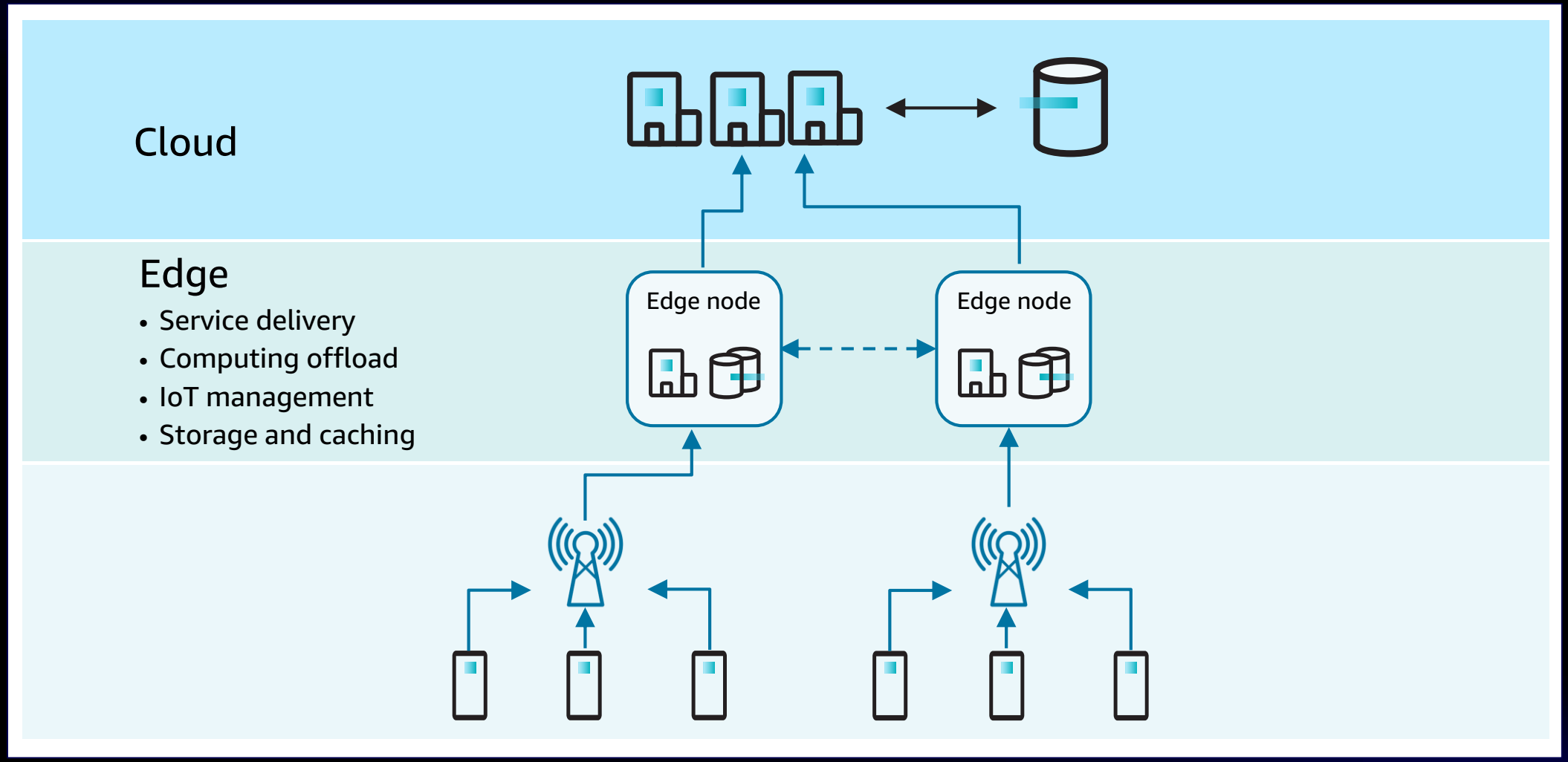
The Enterprise Distribution of Apache Kafka

Deploy on any platform, on premises, or cloud



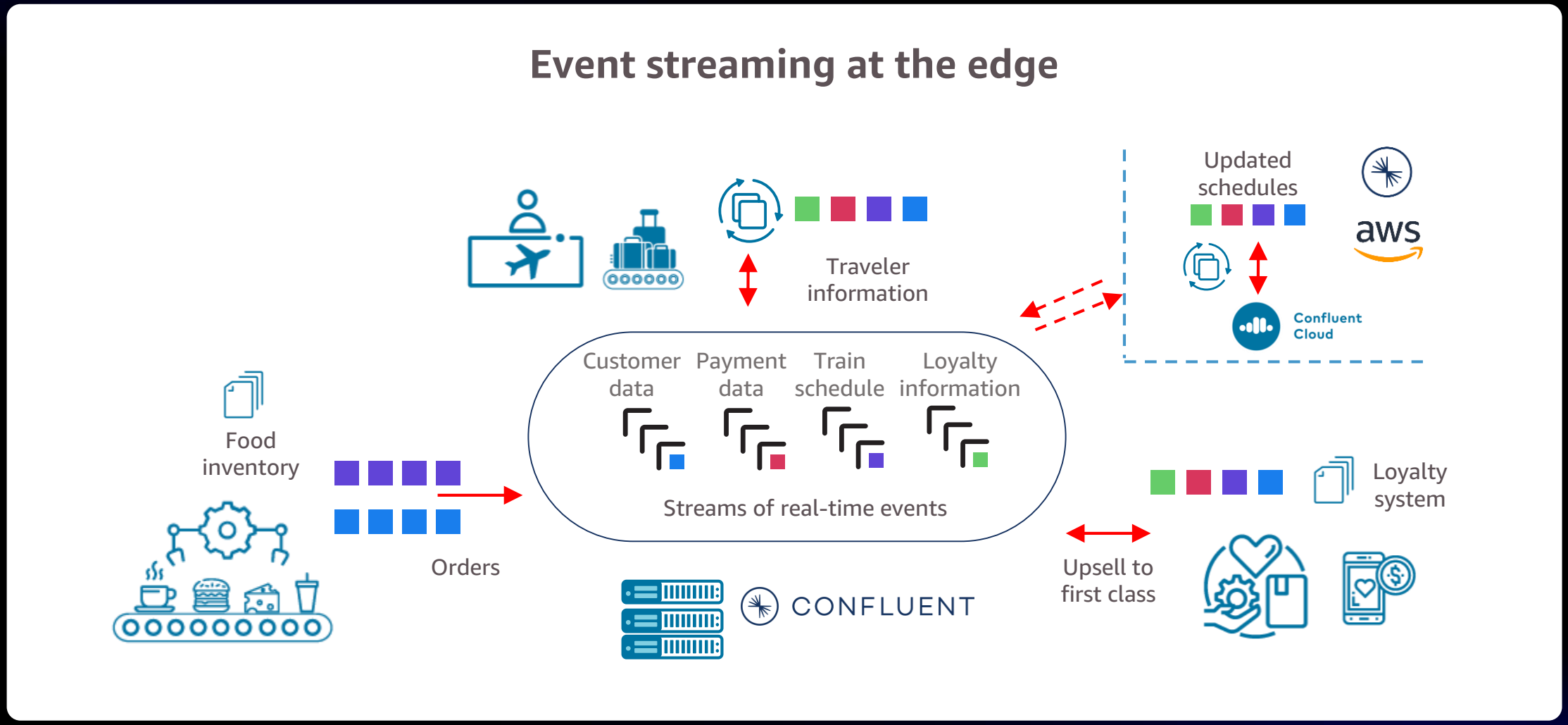
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is the edge?



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

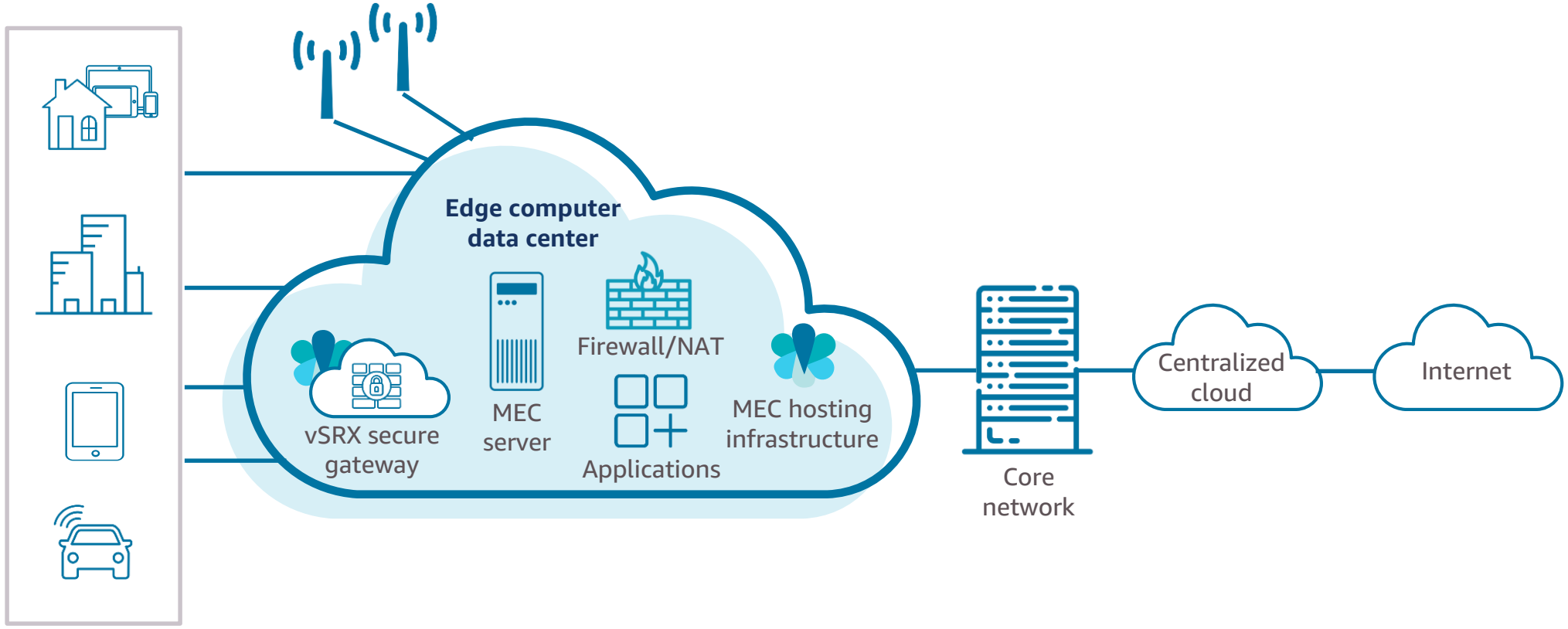
What is the edge?



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Edge types

REMOTE LOCATIONS



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Edge types

PARTIALLY DISCONNECTED

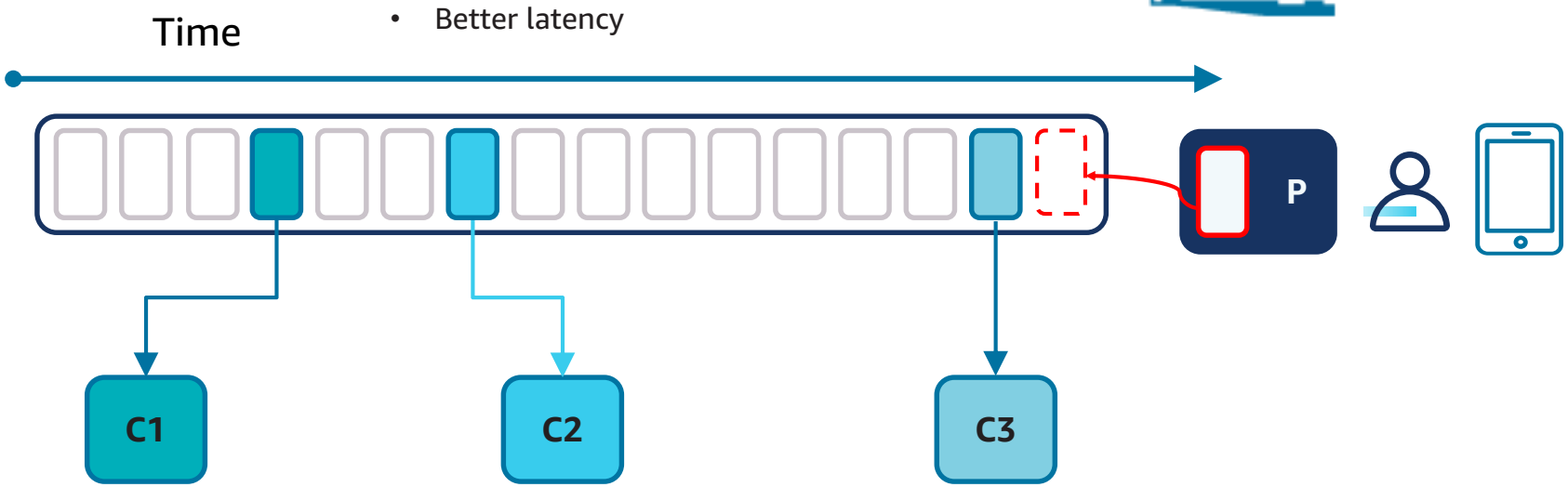
Disconnected edge

- Always on (even when “offline”)
- Replayability
- Reduced traffic cost
- Better latency





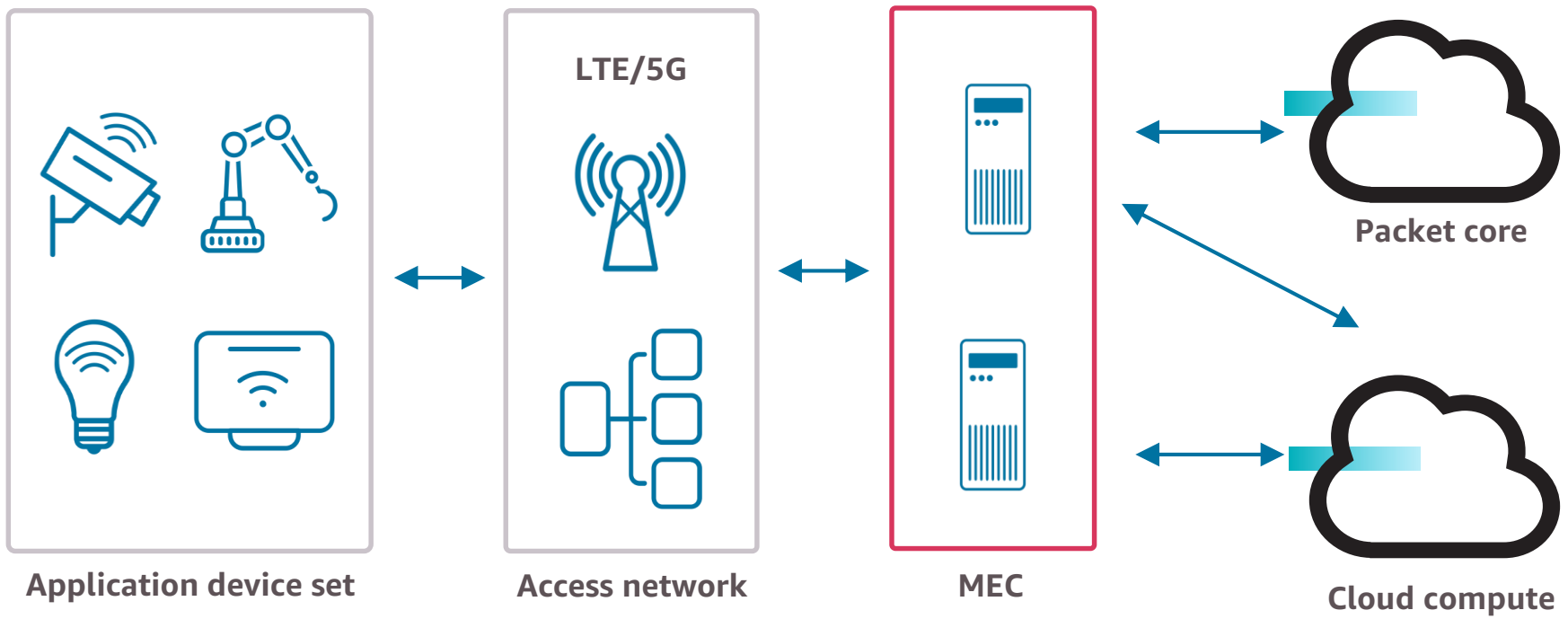




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Edge types

MULTI-ACCESS EDGE COMPUTING (MEC)



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

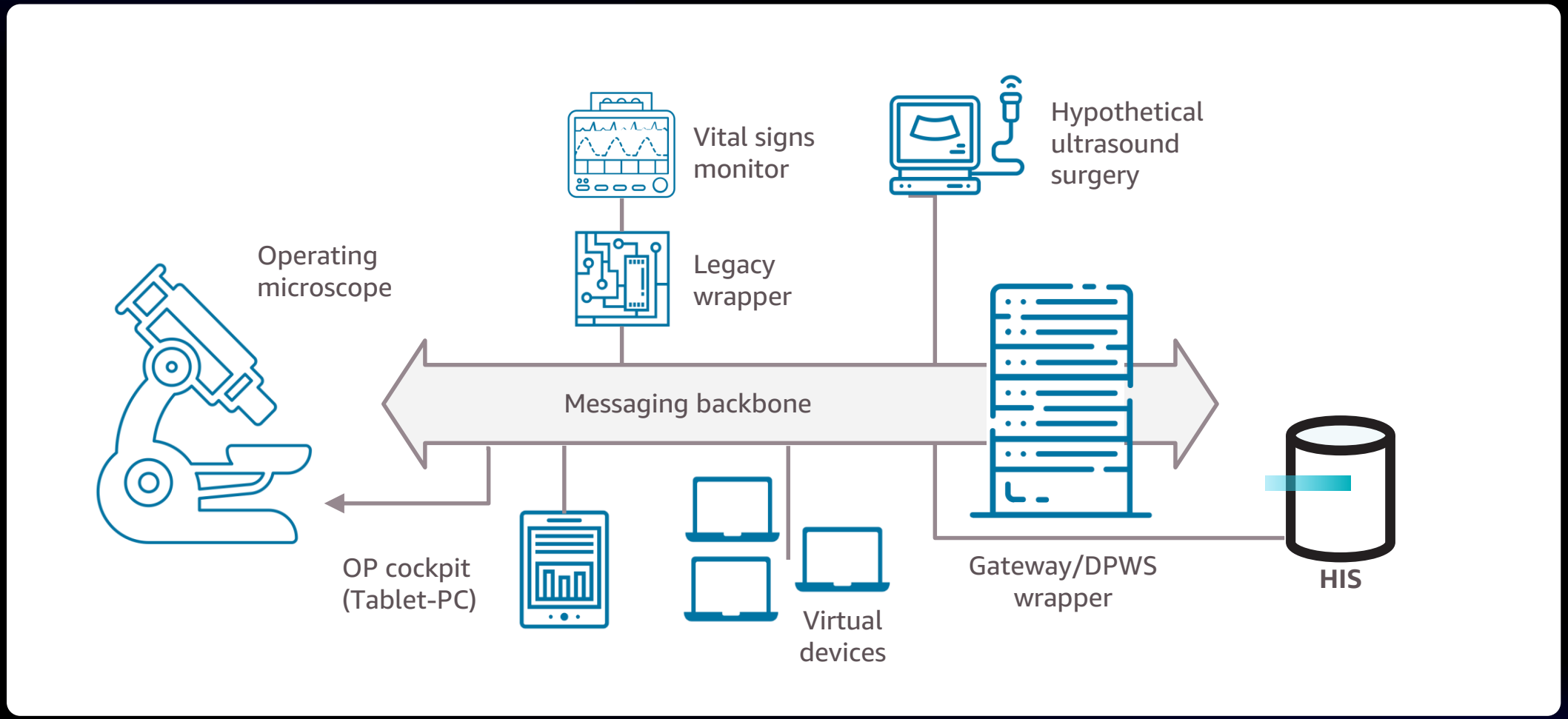

The edge can be difficult
Limited resources – Budgetary


The edge can be difficult
Limited resources – Staffing


The edge can be difficult
Limited resources – Failure recovery

The edge can be difficult

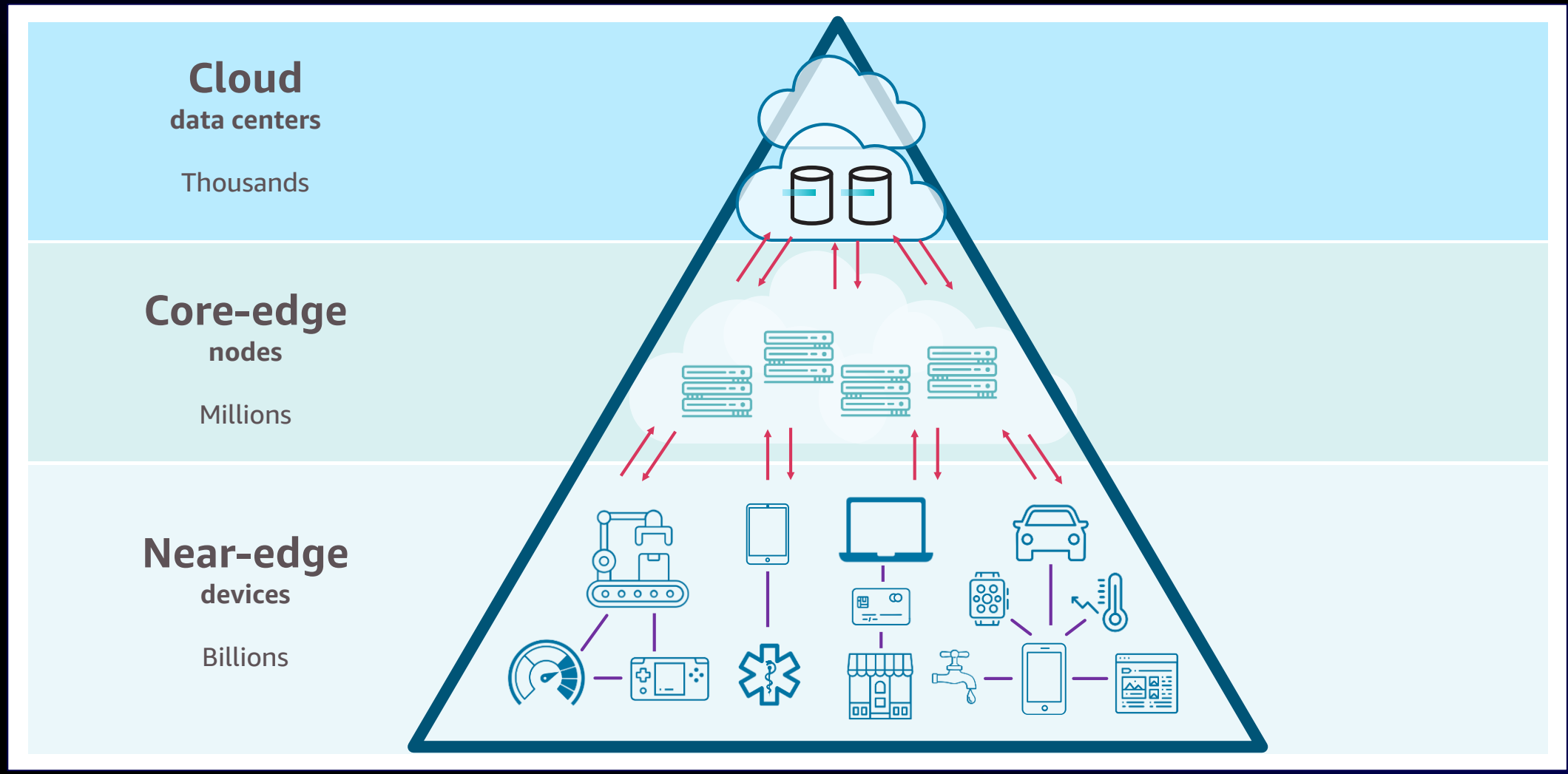
DEVICE INTEGRATION



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The edge can be difficult

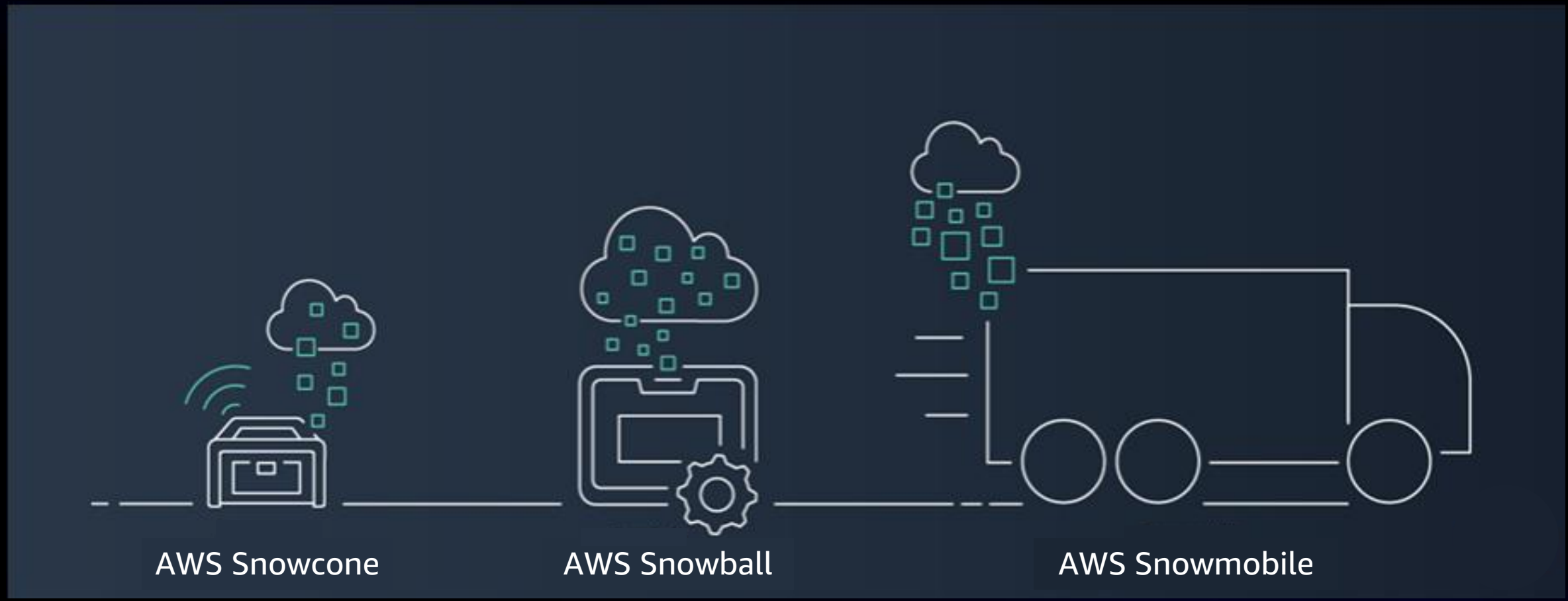
PROCESSING



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS edge computing

MAKE AWS COME TO YOU



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Outposts

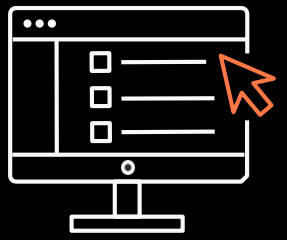
BRINGING AWS ON PREMISES



Same AWS-designed infrastructure as in AWS data centers (built on AWS Nitro System)



Fully managed, monitored, and operated by AWS as if in AWS Regions



Single pane of management in the cloud providing the **same APIs and tools** as in AWS Regions

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

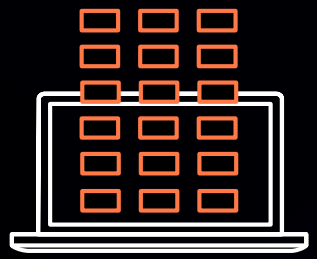
AWS Local Zones

AMAZON EC2 CLOSER TO YOU

- AWS infrastructure deployment that places compute, storage, database, and other services closer to customers
- For applications that require single-digit millisecond latencies



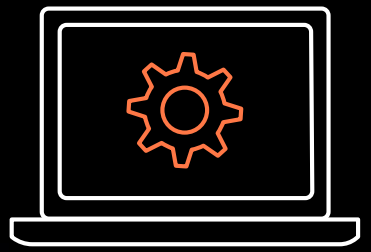
AWS infrastructure at the edge



Local compute, storage, databases, and other services



Connect to services in AWS Regions



Deliver new low-latency applications

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Snow Family

AWS SNOWMOBILE, AWS SNOWBALL, AND AWS SNOWCONE

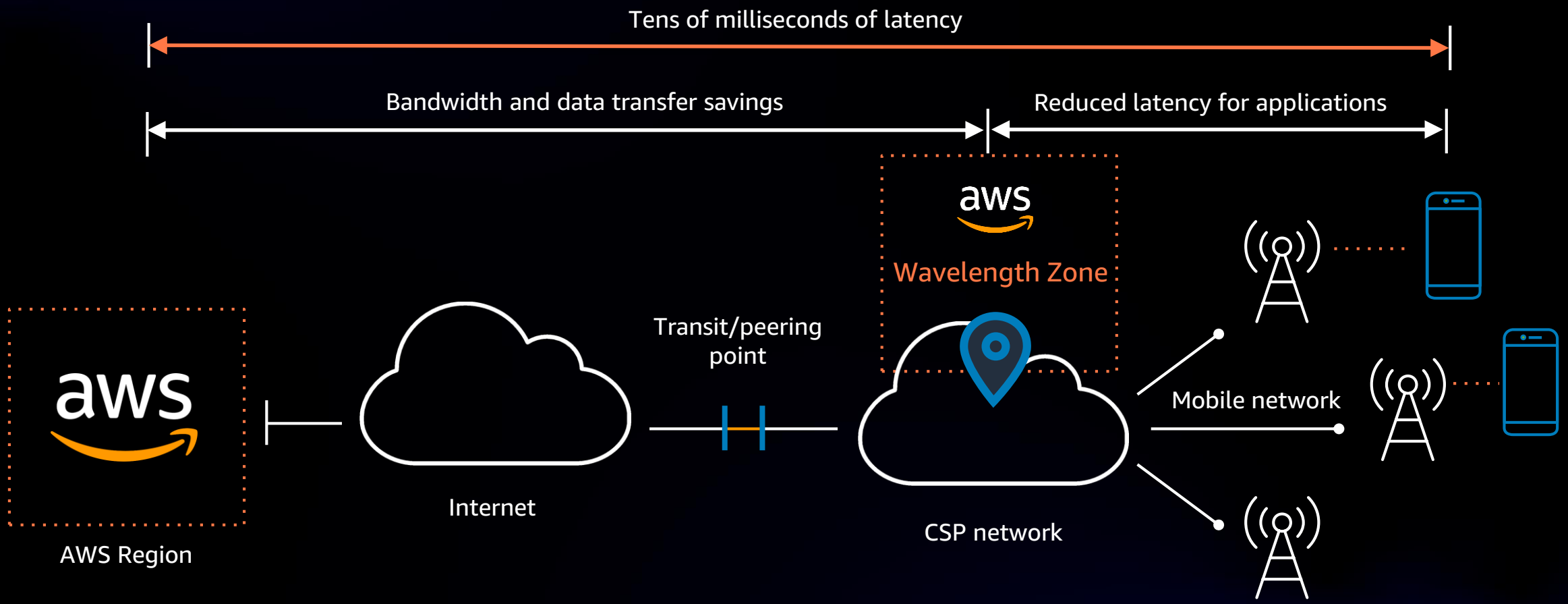


- Set of services and functionality on partially to fully disconnected and secure edge
- Enables data collection, data processing, and data movement from the edge to AWS

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Wavelength

AWS VIA LOW-LATENCY 5G



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Better together

LET AWS HANDLE THE COMPUTE AND CONFLUENT HANDLE THE DATA



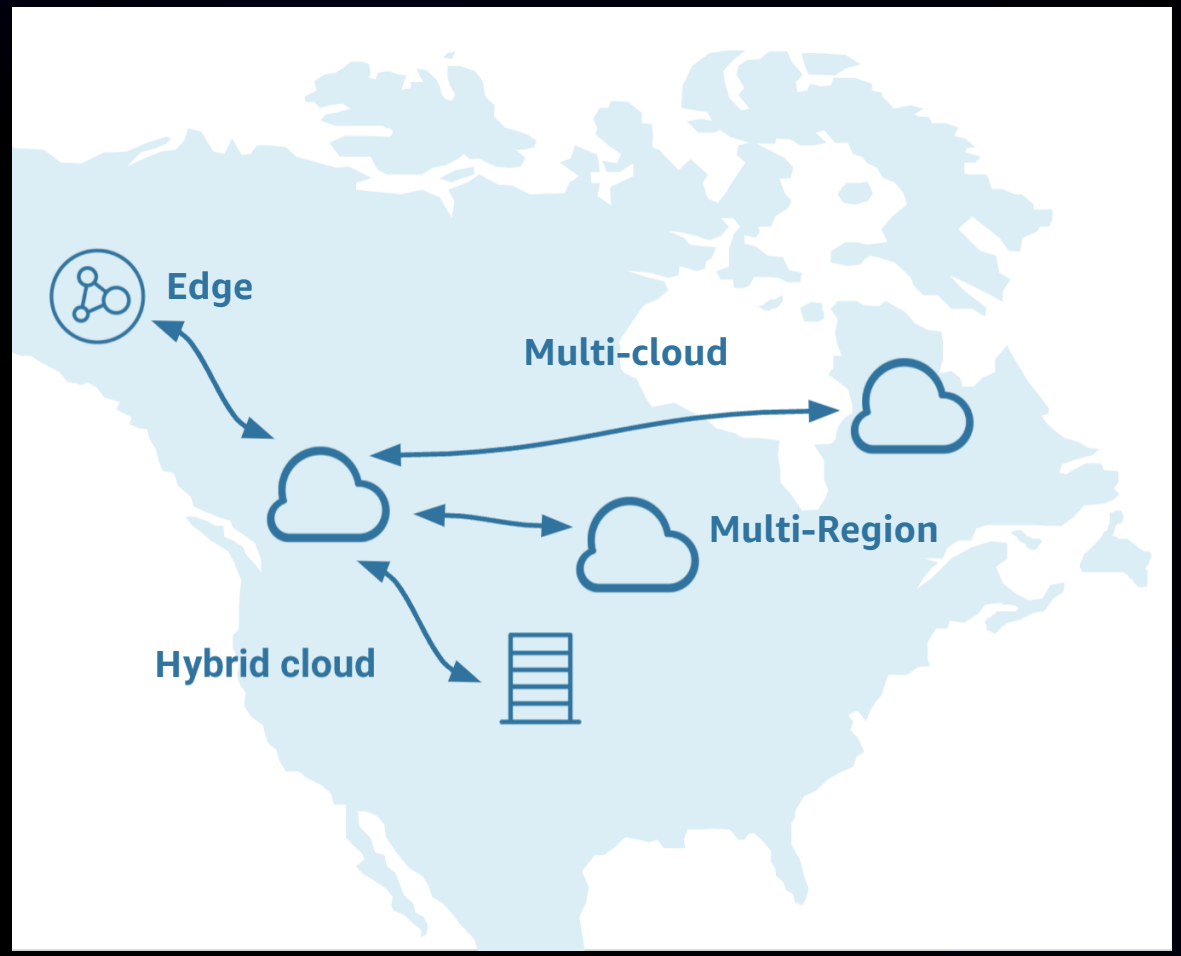


aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Global architectures

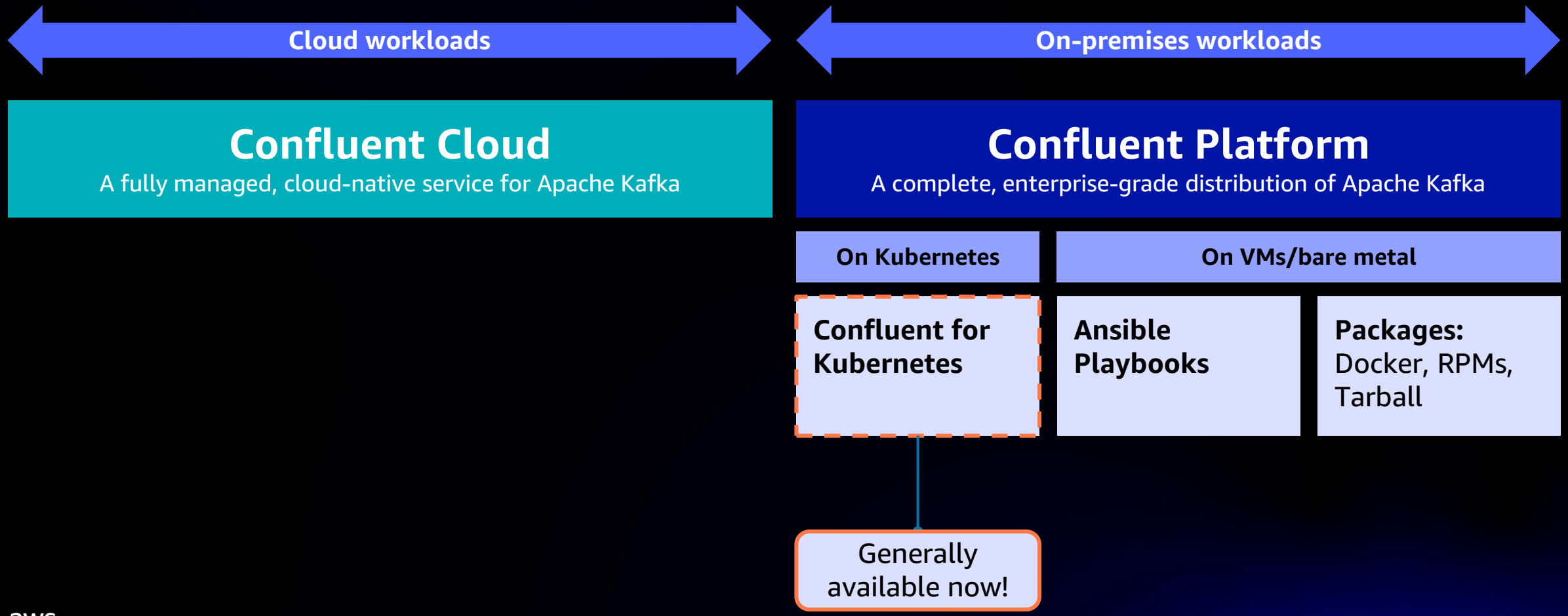
POWERED BY CONFLUENT AND AWS



- Cluster migration
- Global data mesh
- Data sharing between teams, LOBs, organizations
- Higher availability and disaster recovery
- Bridge to cloud
- Edge aggregation

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent for Kubernetes brings cloud-native experience to self-managed software product



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent for Kubernetes

Introduces a declarative API-driven control plane to deploy and manage Confluent in private infrastructures

Declarative API for operating Confluent in production

Manage topics and RBAC policies through infrastructure as code

Integrates with cloud-native system for security, reliability, and DevOps automation



Runs on Kubernetes: the infrastructure runtime for cloud-native architectures

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cloud-native: Accelerate time to value & simplify day-to-day operations with complete, declarative API

Infrastructure as code

Kubernetes-native CLI to manage entire infrastructure

```
$ kubectl apply -f \
mykafka.yaml

$ kubectl scale kafka \ --
replicas=10


$ kubectl apply -f
mytopics.yaml

$ kubectl get topic
```

Declarative API

Use a single deployment spec to manage security

```
mykafka.yaml
---
kafka:
 authentication:
  type: plain
   jaasConfig:
    secretRef: credential
---
```

Opinionated deployment

Enterprise-ready configurations with flexible customization

```
---
kafka:
 configOverrides:
  server:
    enable.feature=true
  jvm: []
  log4j: []
---

---
podTemplate:
  affinity: []
  annotations: []
  labels: []
  envVar: []
  tolerations: []
```

aws

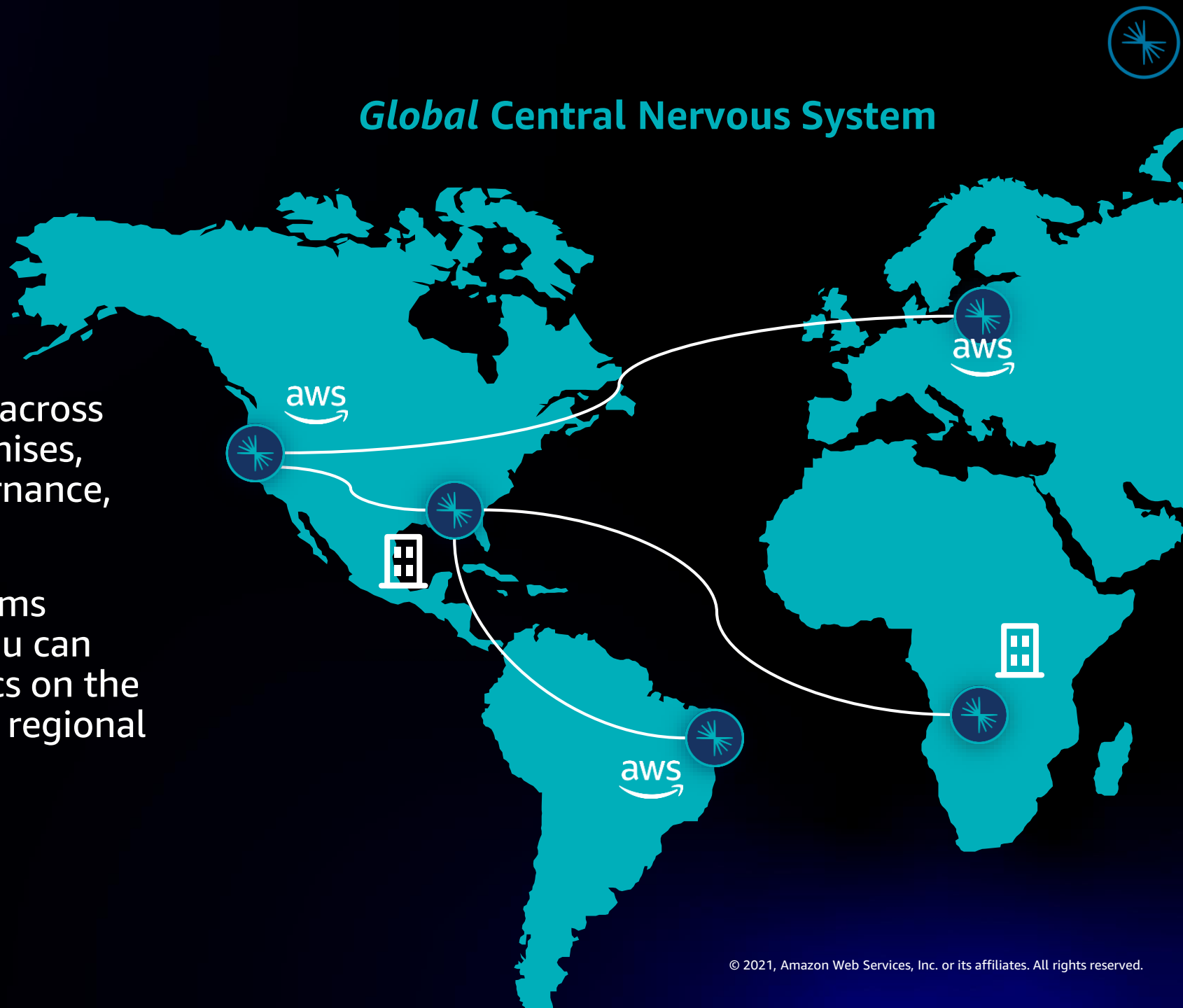
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Everywhere: Cluster Linking

Federated streaming, hybrid and multi-cloud

Data syndication and replication across and between clouds and on-premises, with self-service APIs, data governance, and visual tooling

Reliable and real-time data streams between all customer sites, so you can run always-on streaming analytics on the data of entire enterprise, despite regional or cloud provider outages



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Make Kafka widely accessible to developers

Let developers build with their most productive language and enable use cases and architectures that require multiple programming languages

Confluent clients

Reliable and high performing



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Kafka Connect: Reliable and scalable Kafka integration with other systems

- Centralized management and configuration
- Support for hundreds of technologies including RDBMS, Elasticsearch, HDFS, Amazon S3
- Supports CDC ingest of events from RDBMS
- Preserves data schema
- Fault tolerant and automatically load balanced
- Extensible API
- Single message transforms
- Part of Apache Kafka

```
{
"connector.class": "io.confluent.connect.jdbc.JdbcSourceConnector",
"connection.url": "jdbc:mysql://localhost:3306/demo?user=rmoff&password=foo",
"table.whitelist": "sales,orders,customers"
}
```

docs.confluent.io/platform/current/connect/index.html

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Instantly connect popular data sources & sinks



130+
prebuilt connectors



100+ Confluent supported

30+ partner supported, Confluent verified

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Kafka Streams: Write standard Java apps and microservices to process your data in real time

- No separate processing cluster required
- Develop on Mac, Linux, and Windows
- Deploy to containers, VMs, bare metal, cloud
- Powered by Kafka: elastic, scalable, distributed, and reliable
- Perfect for small, medium, and large use cases
- Fully integrated with Kafka security
- Exactly-once processing semantics
- Part of Apache Kafka

```
KStream<User, PageViewEvent> pageViews = builder.stream("pageviews-topic");
KTable<Windowed<User>, Long> viewsPerUserSession = pageViews
    .groupByKey()
    .count(SessionWindows.with(TimeUnit.MINUTES.toMillis(5)), "session-views");
```

docs.confluent.io/current/streams/

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Stream processing with Kafka

```
val fraudulentPayments: KStream[String, Payment] = builder
  .stream[String, Payment]("payments-kafka-topic")
  .filter((_ ,payment) => payment.fraudProbability > 0.8)
fraudulentPayments.to("fraudulent-payments-topic")
```

```
  config.put(StreamsConfig.BOOTSTRAP_SERVERS_CONFIG, "kafka-broker1:9092,kafka-broker2:9092")

  val streams: KafkaStreams = new KafkaStreams(builder.build(), config)
  streams.start()
}
```

Example: Using **Kafka Streams API** for writing elastic, scalable, fault-tolerant Java and Scala applications

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

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
```

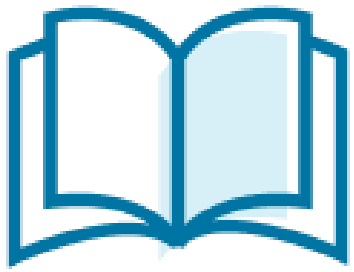
Flexibility — confluent.awsworkshop.io — **Simplicity**

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Discover, understand, and trust your event streams

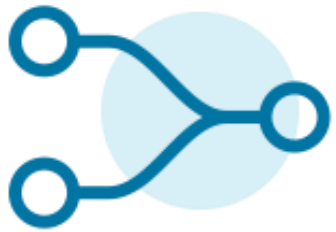
WITH THE INDUSTRY'S ONLY GOVERNANCE SUITE FOR DATA IN MOTION

Stream Catalog



Increase collaboration and productivity with self-service data discovery

Stream Lineage



Understand complex data relationships and uncover more insights

Stream Quality

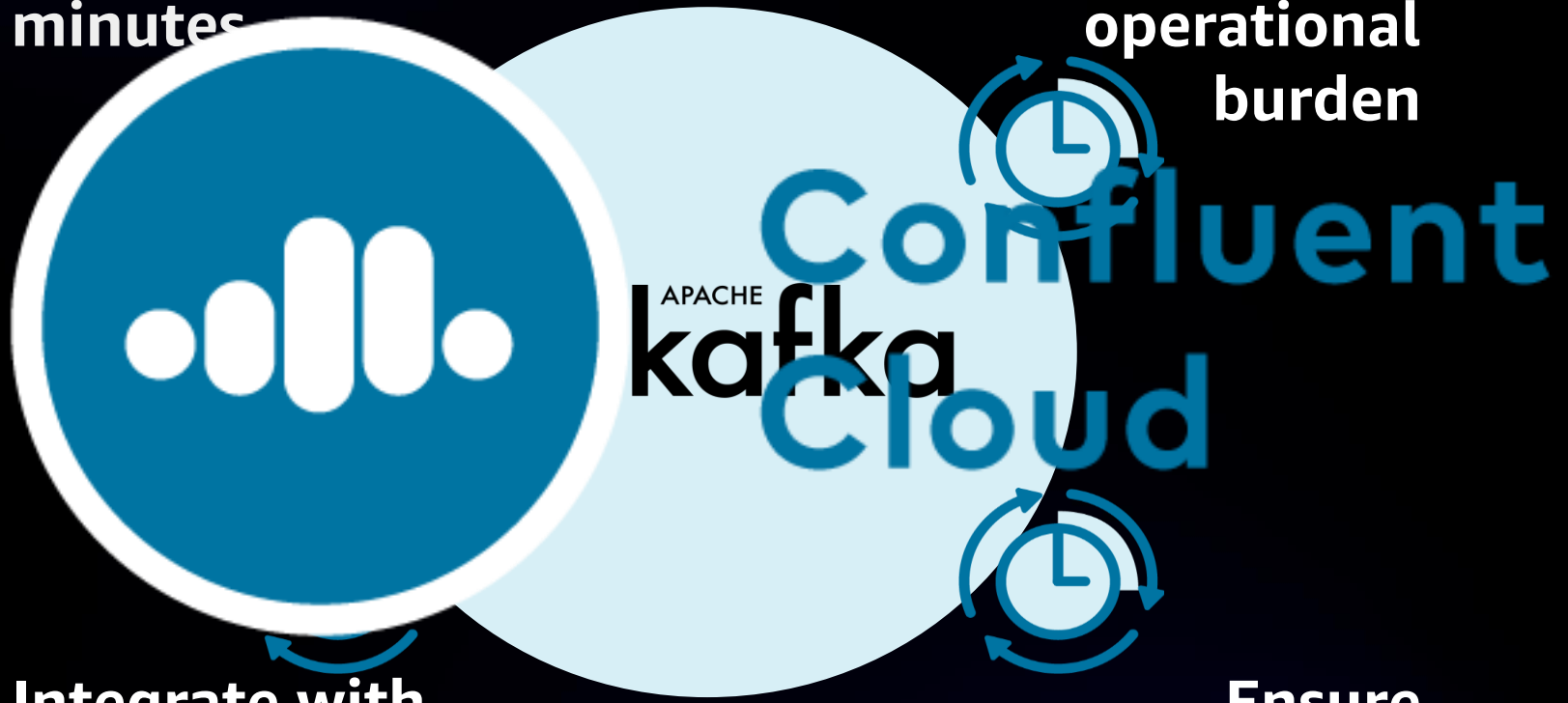


Deliver trusted, high-quality event streams to the business

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data in motion as a cloud-native service



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scale faster with cloud-native experience



Stream with Kafka in minutes

Provision and scale clusters on demand



Pay only for what you use

Start from $0 and pay based on actual usage



Access latest Kafka version

Always work in most recent stable version of Kafka

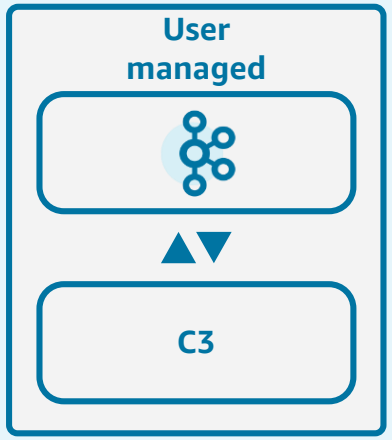
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent support

CONFLUENT Health+

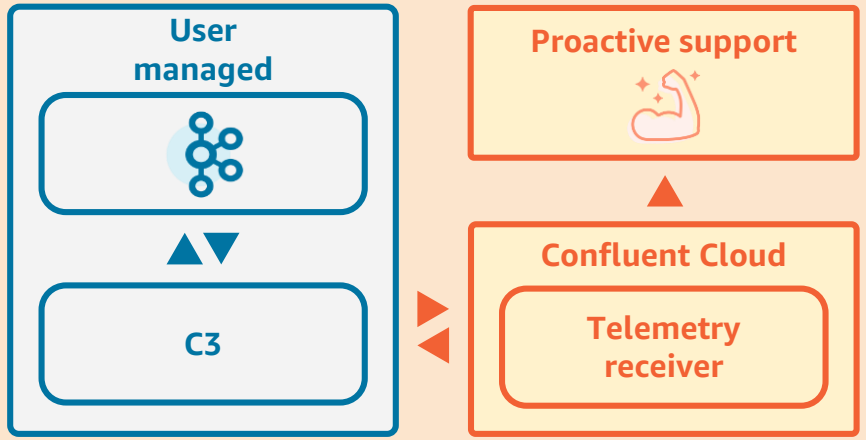
Intelligent alerting and monitoring as a service

Proactive support not available



Users rely on off-the-shelf C3 alerts to troubleshoot – with limited understanding of which alerts are most important

Proactive support enabled



Users can **avoid downtime by customizing and prioritizing alerts based on Confluent's expert-tested rules & algorithms**, developed from running thousands of clusters in Confluent Cloud

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.


Confluent at the edge – Use cases
Retail

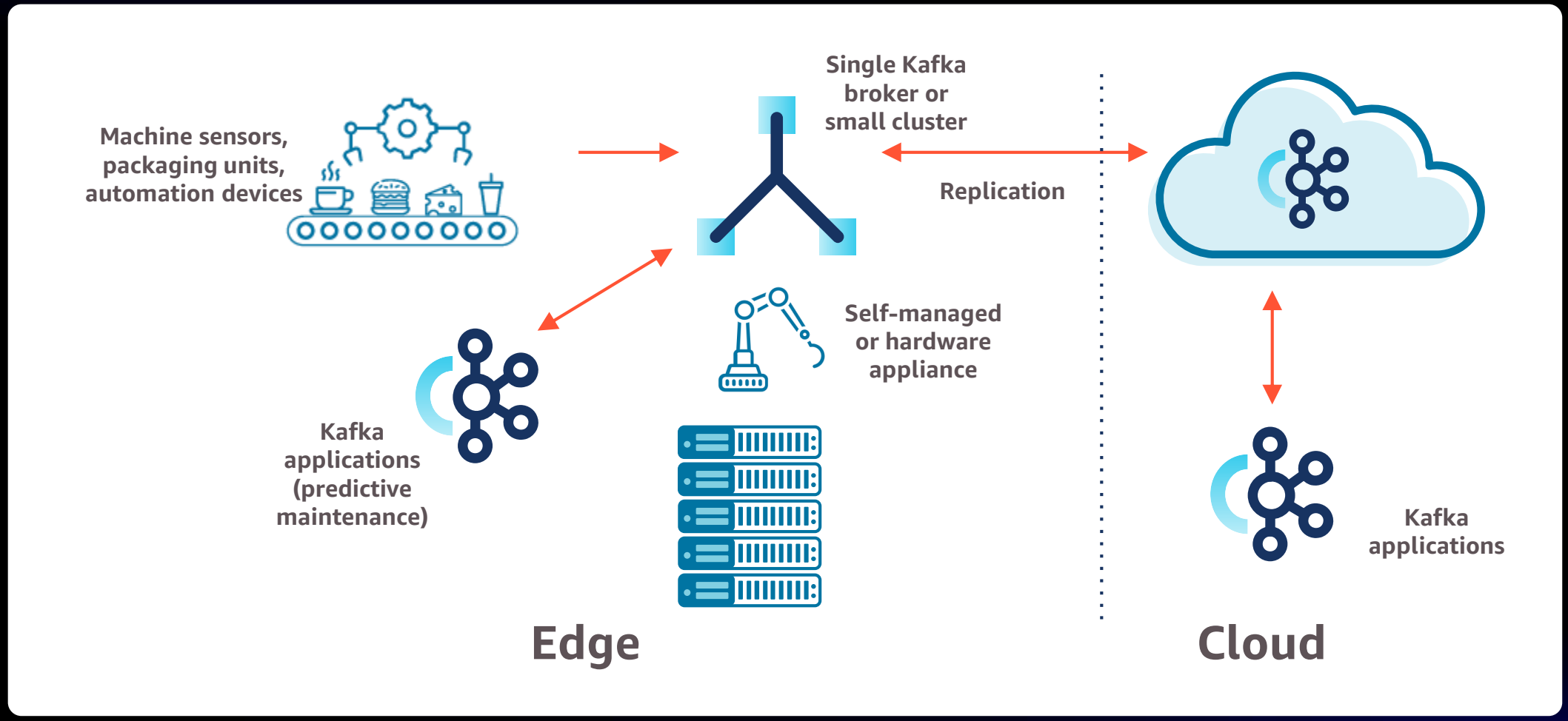


Confluent at the edge – Use cases

Logistics

Confluent at the edge – Use cases

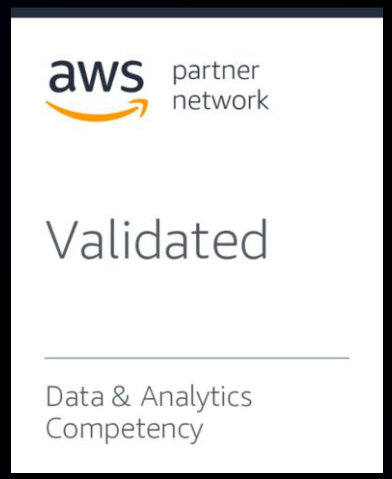
INDUSTRIAL IOT



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent and AWS – Better together









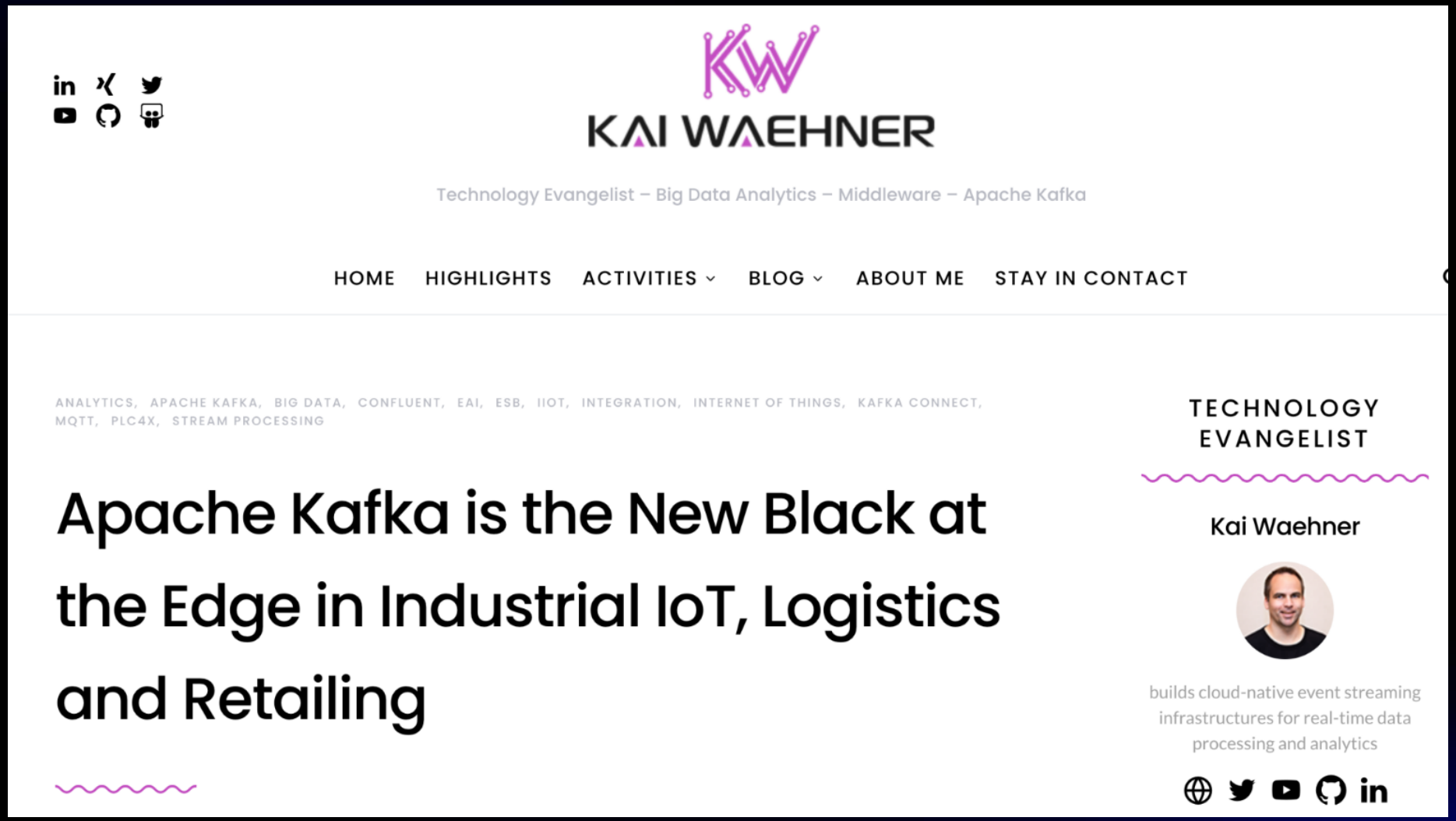




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent on AWS Wavelength solution

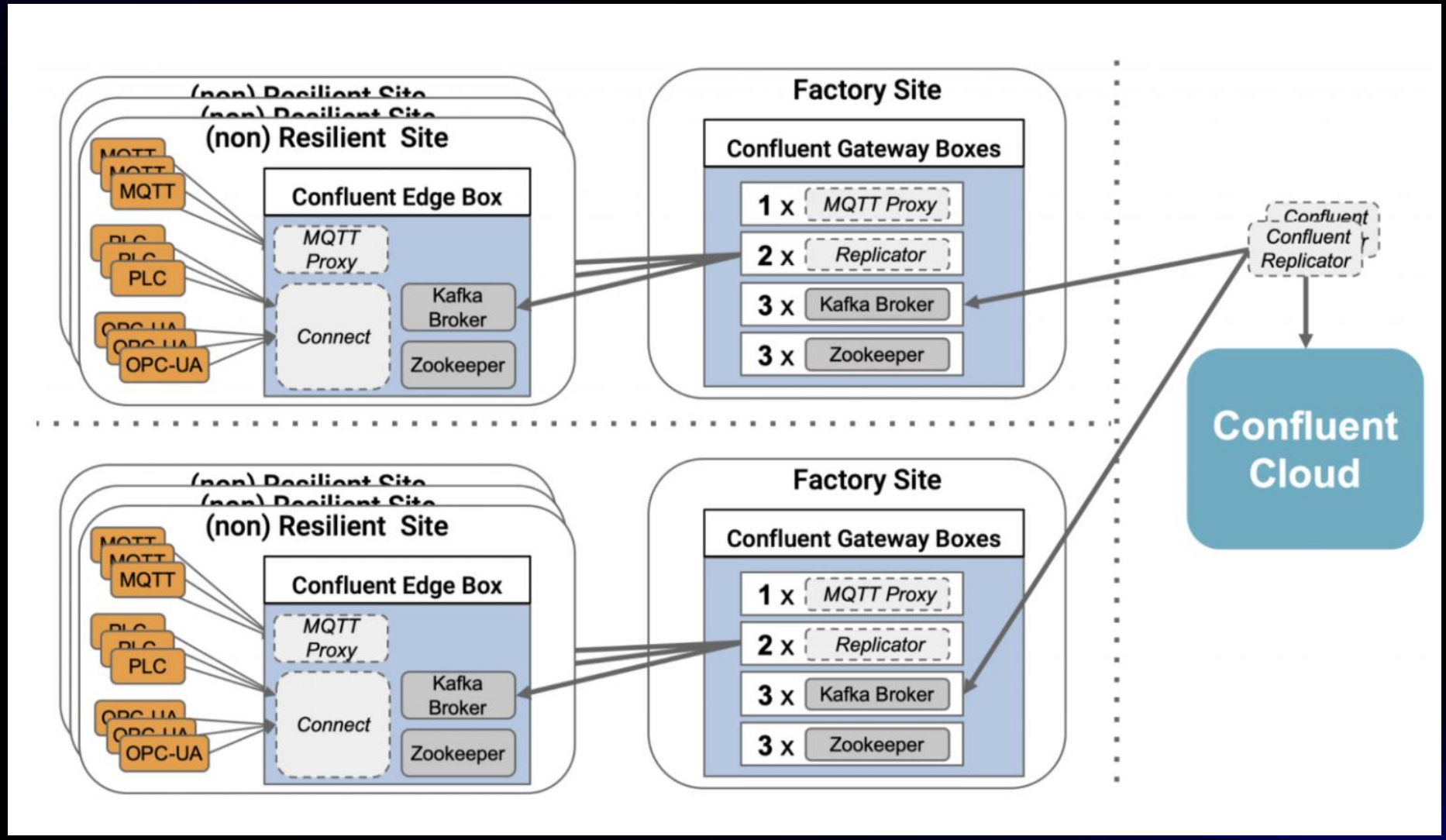
THE ORIGIN STORY



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent on AWS Wavelength solution

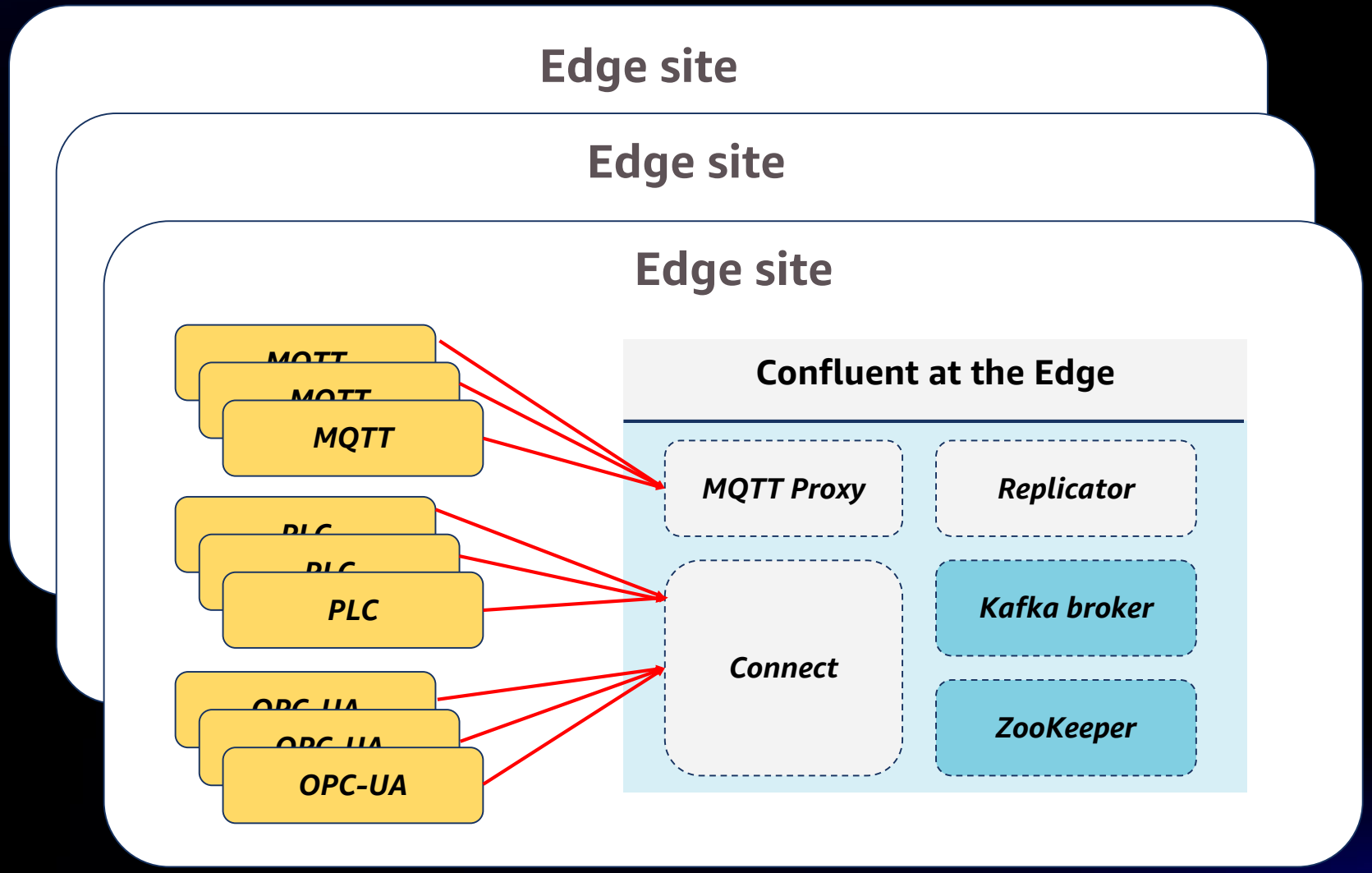
THE ORIGIN STORY



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent on AWS Wavelength solution

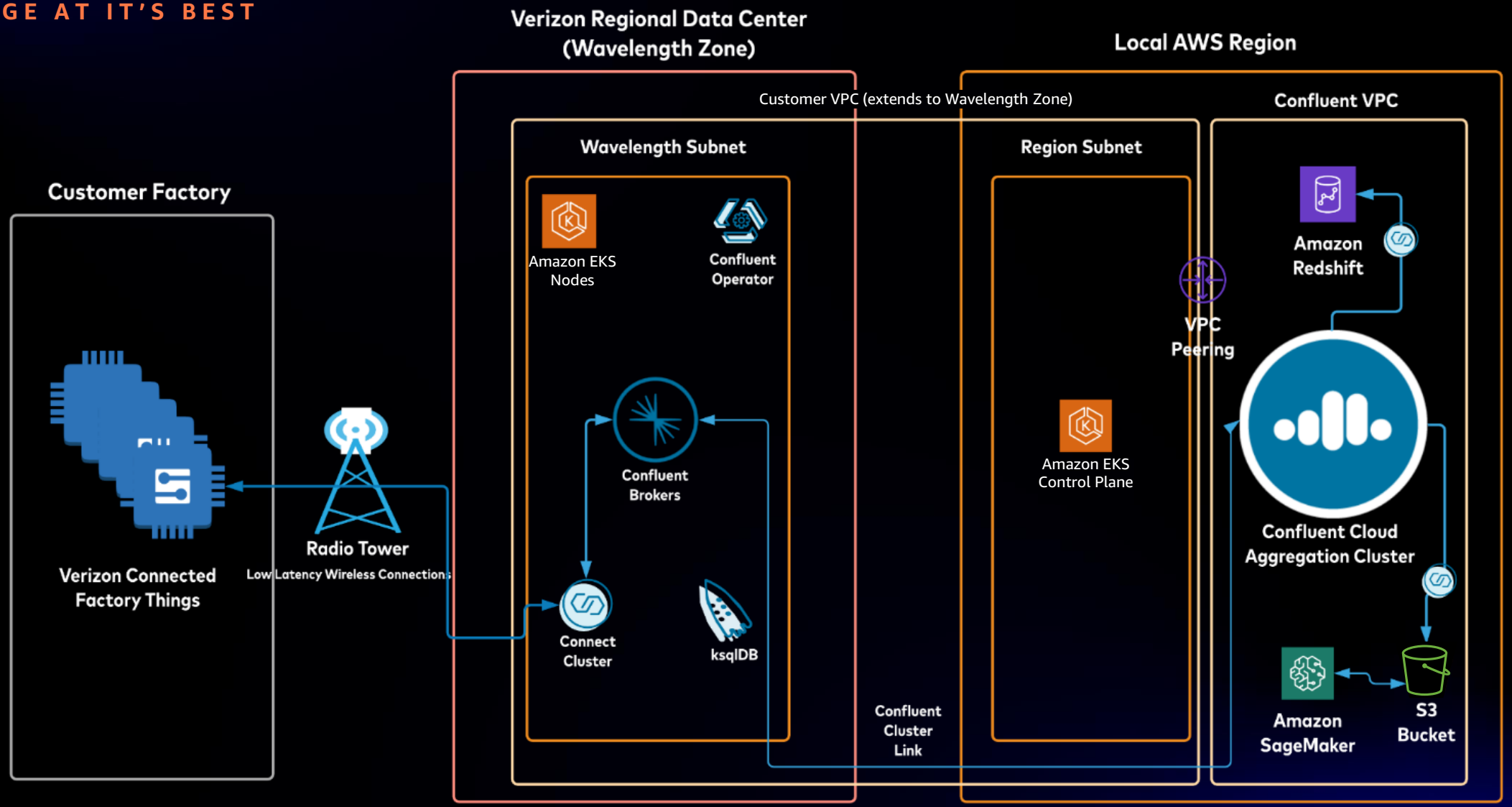
THE ORIGIN STORY



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent & AWS Wavelength solution

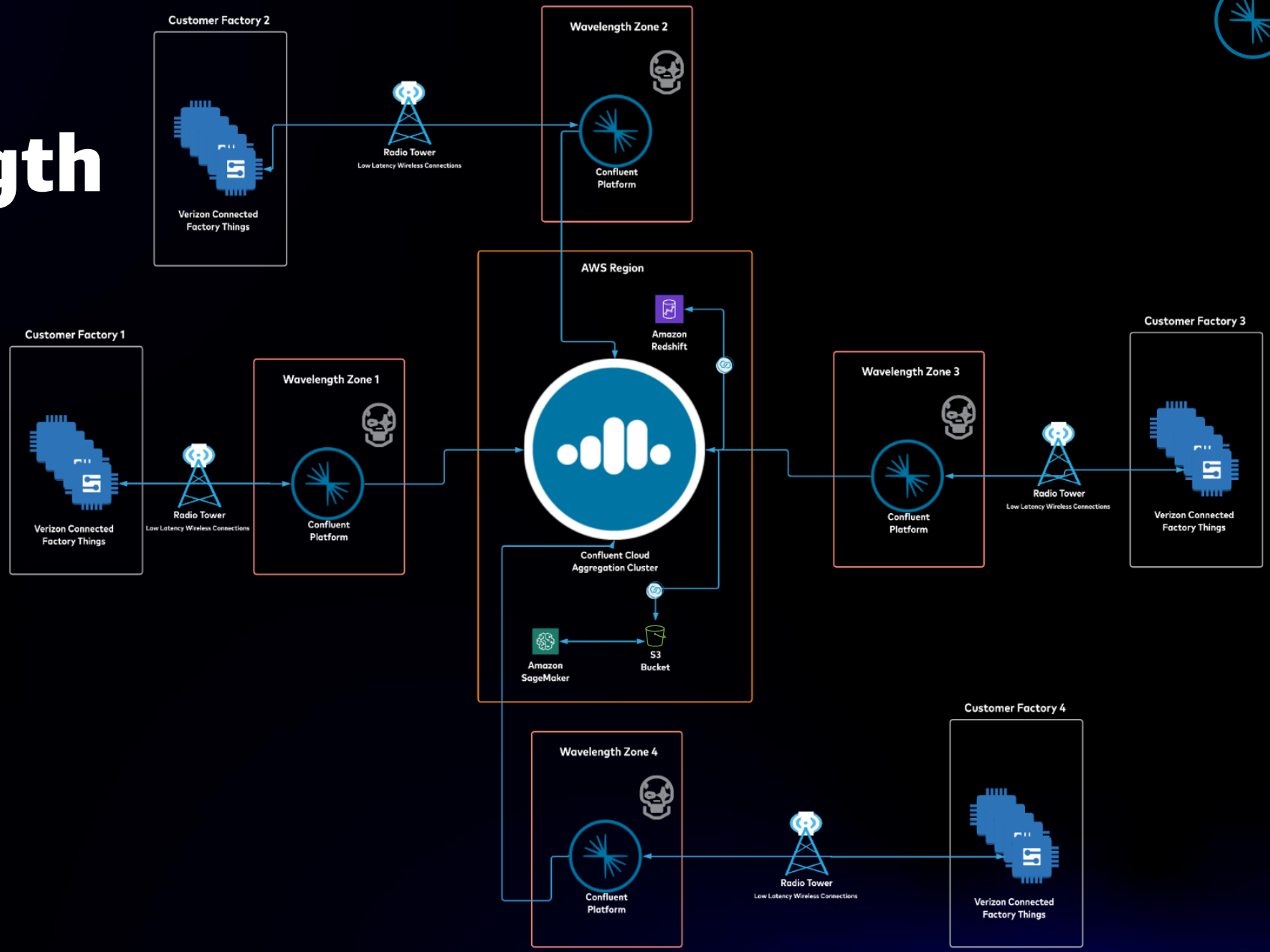
THE EDGE AT IT'S BEST



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent & AWS Wavelength solution

THE EDGE AT IT'S BEST



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent & AWS Wavelength solution demo

HOUSTON, WE HAVE NO PROBLEMS





verizon

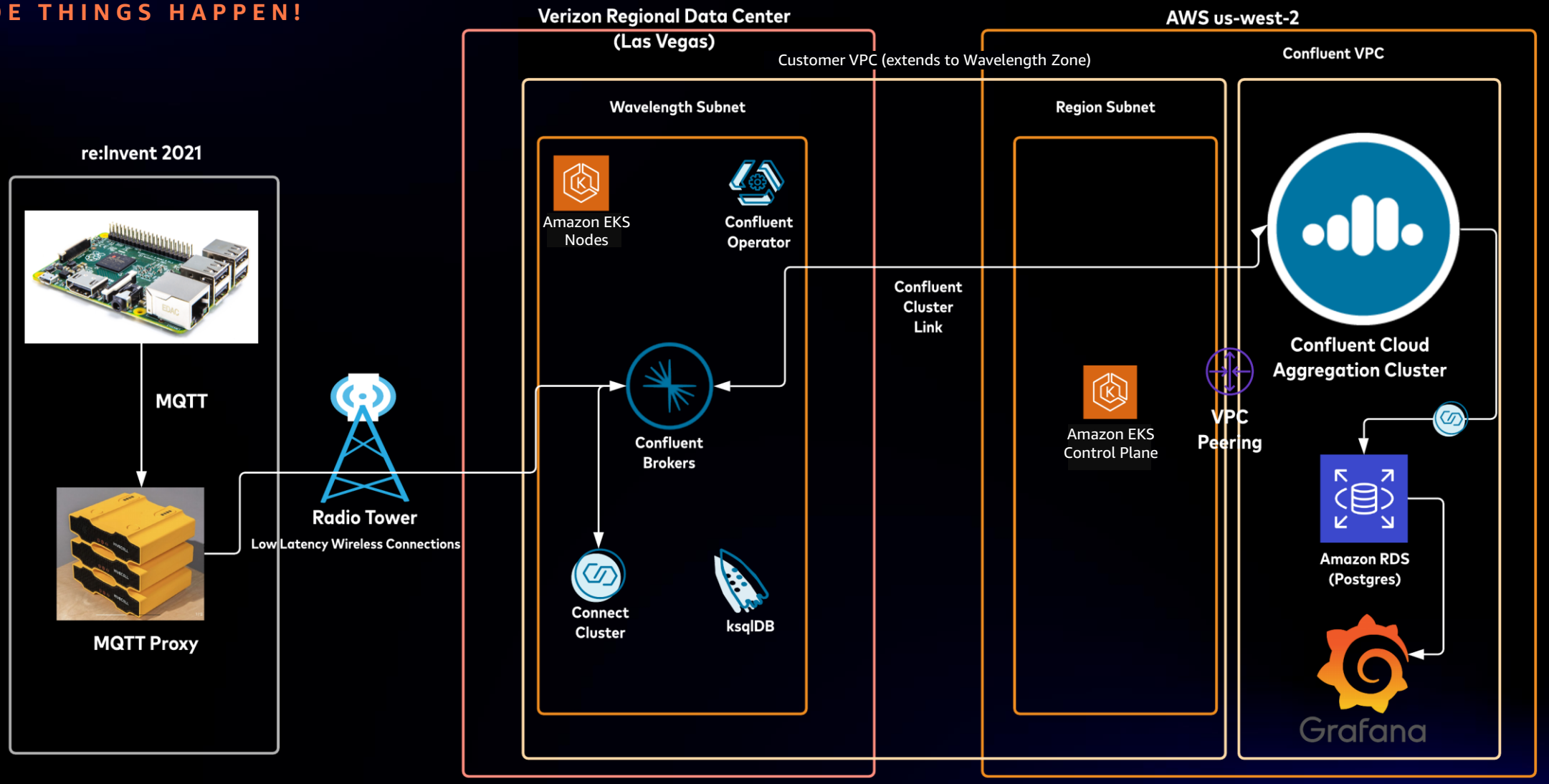
aws

CONFLUENT

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Confluent & AWS Wavelength solution demo

WE MADE THINGS HAPPEN!



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



CONFLUENT

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Joseph Morais

jmorais@confluent.io

Justin Lee

jlee@confluent.io

Other contacts

Sree Karuthody, Senior Manager, Cloud Partners – sree@confluent.io

To get connected to anybody at Confluent – awsteam@confluent.io

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.