



Apache Kafka Overview

Agenda

- Apache Kafka use cases
- Apache Kafka 101
- Operational Challenges of Apache Kafka

Apache Kafka Use Cases



Apache Kafka



Client



Mobile client



Twitter



IoT sensor

Produce



Databases



Analytics



Dashboards

Consume



Apache Kafka Use cases



Microservice Messaging



Streaming ETL



Log Analytics



Transaction &
Event Sourcing

Apache Kafka 101



Apache Kafka 101

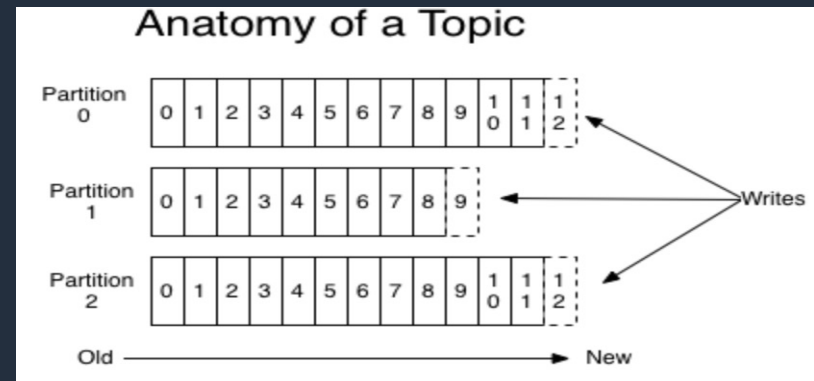
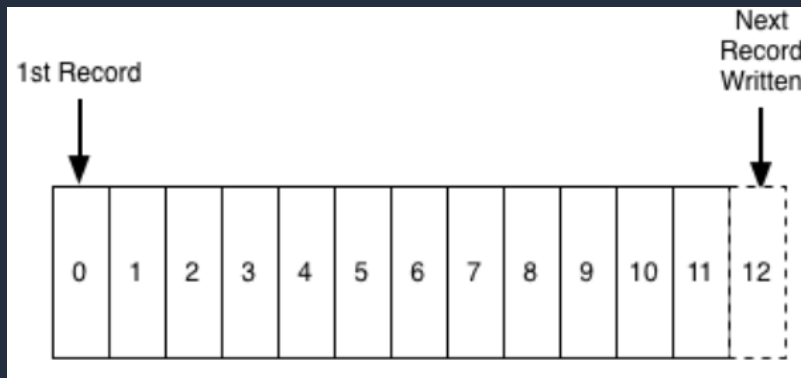
Apache Kafka is a “Distributed Streaming Platform”

Apache Kafka has 3 core capabilities:

- Allows applications to publish and subscribe to streams of records
- Store records in the same way they were received
- Allows applications to process records as they occurred

Apache Kafka 101: The Partition

AN APPEND-ONLY, TOTALLY-ORDERED SEQUENCE OF RECORDS ORDERED BY TIME. IT LOOKS LIKE THIS

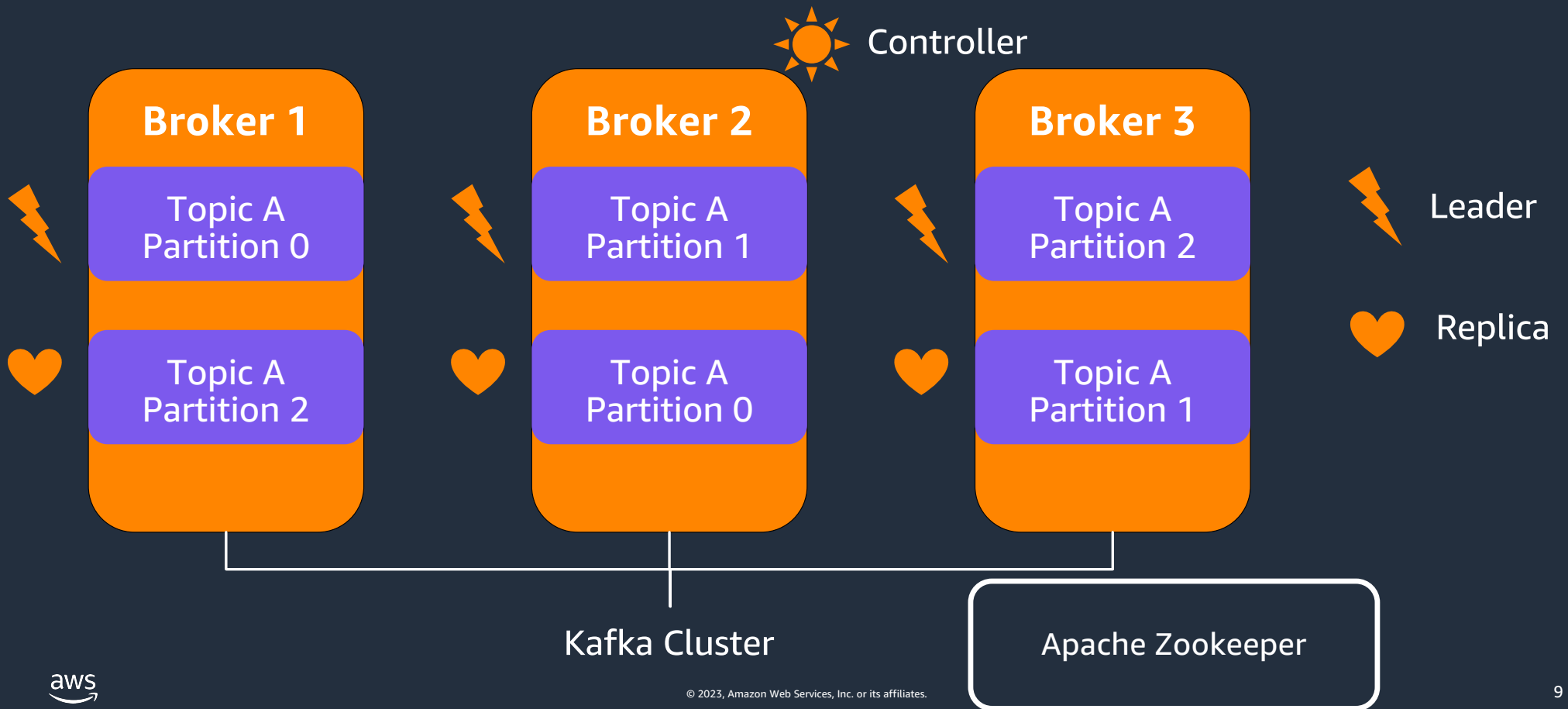


Kafka is run as a cluster, one or more servers, multiple datacenters. (AZ's)

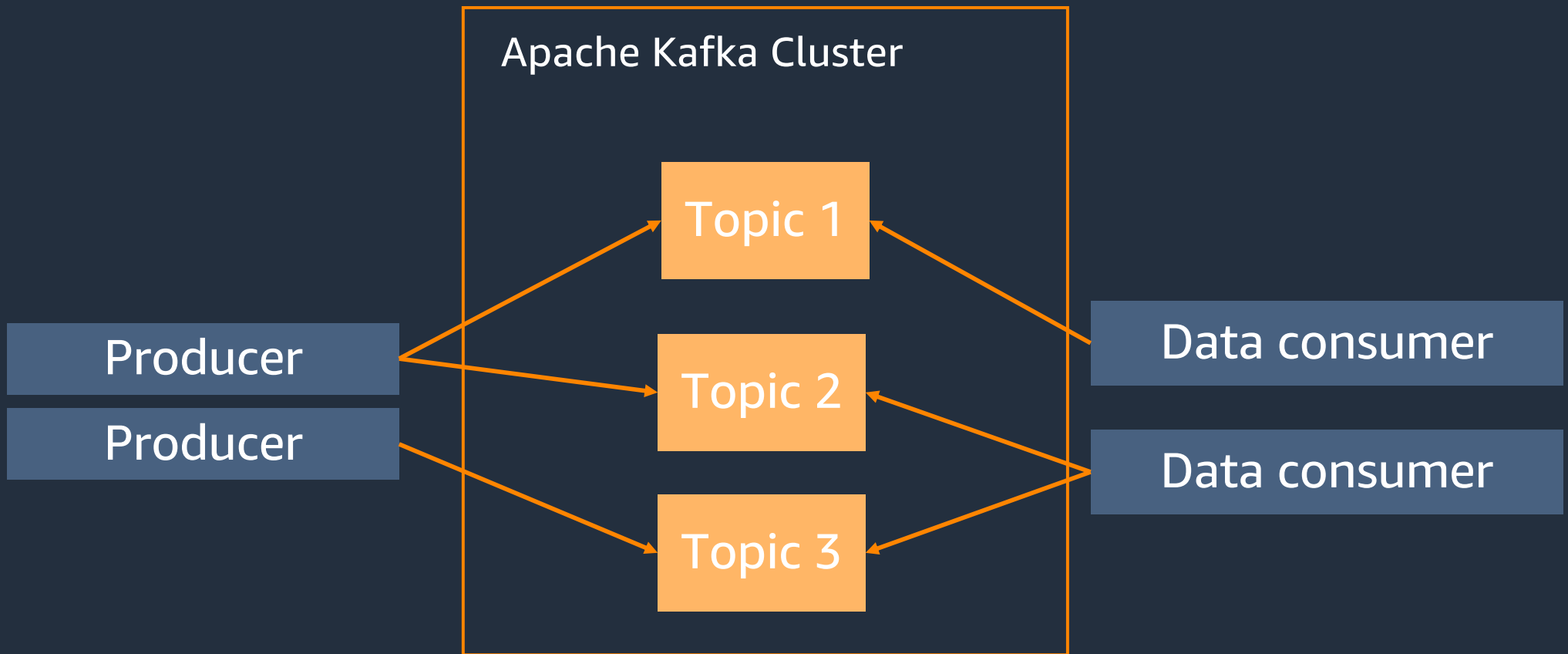
Each record consists of a key, a value, and a timestamp.

The Kafka cluster stores streams of *records* in categories called *topics*.

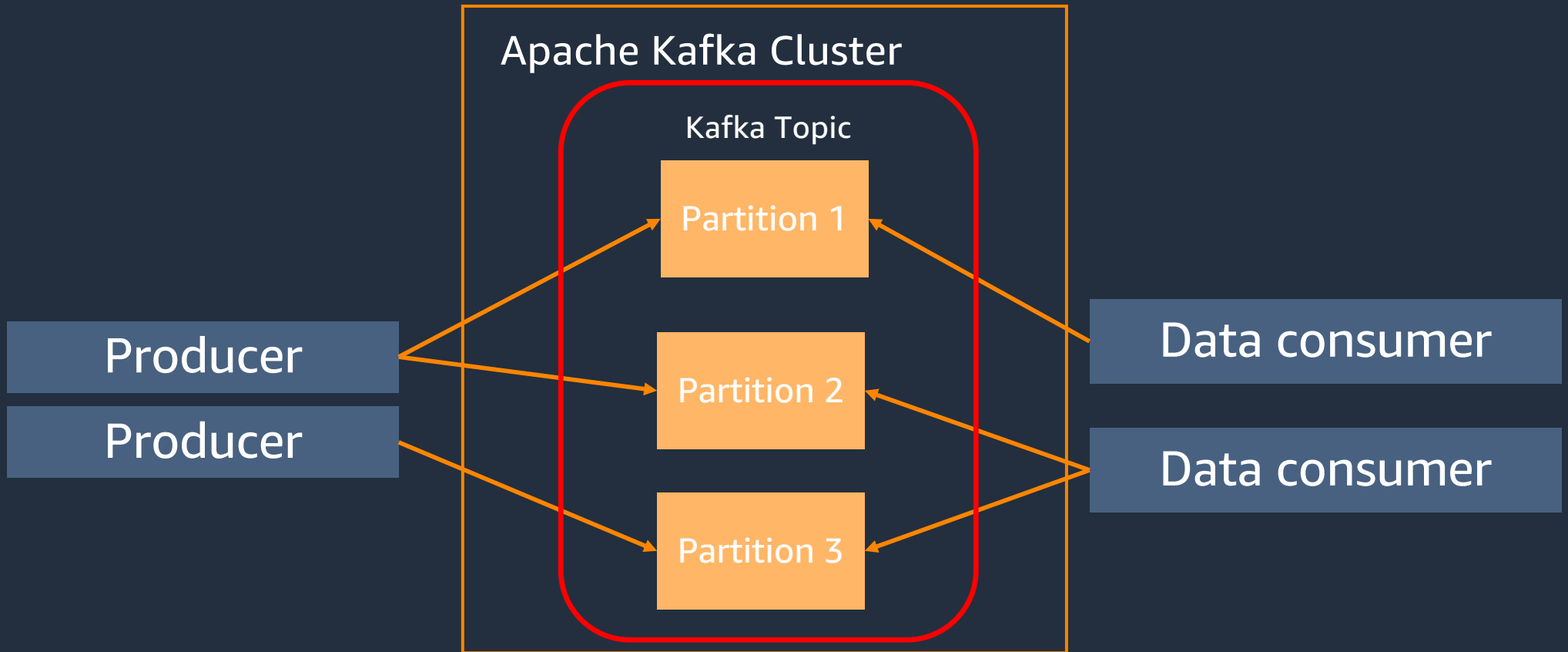
Apache Kafka 101: Cluster



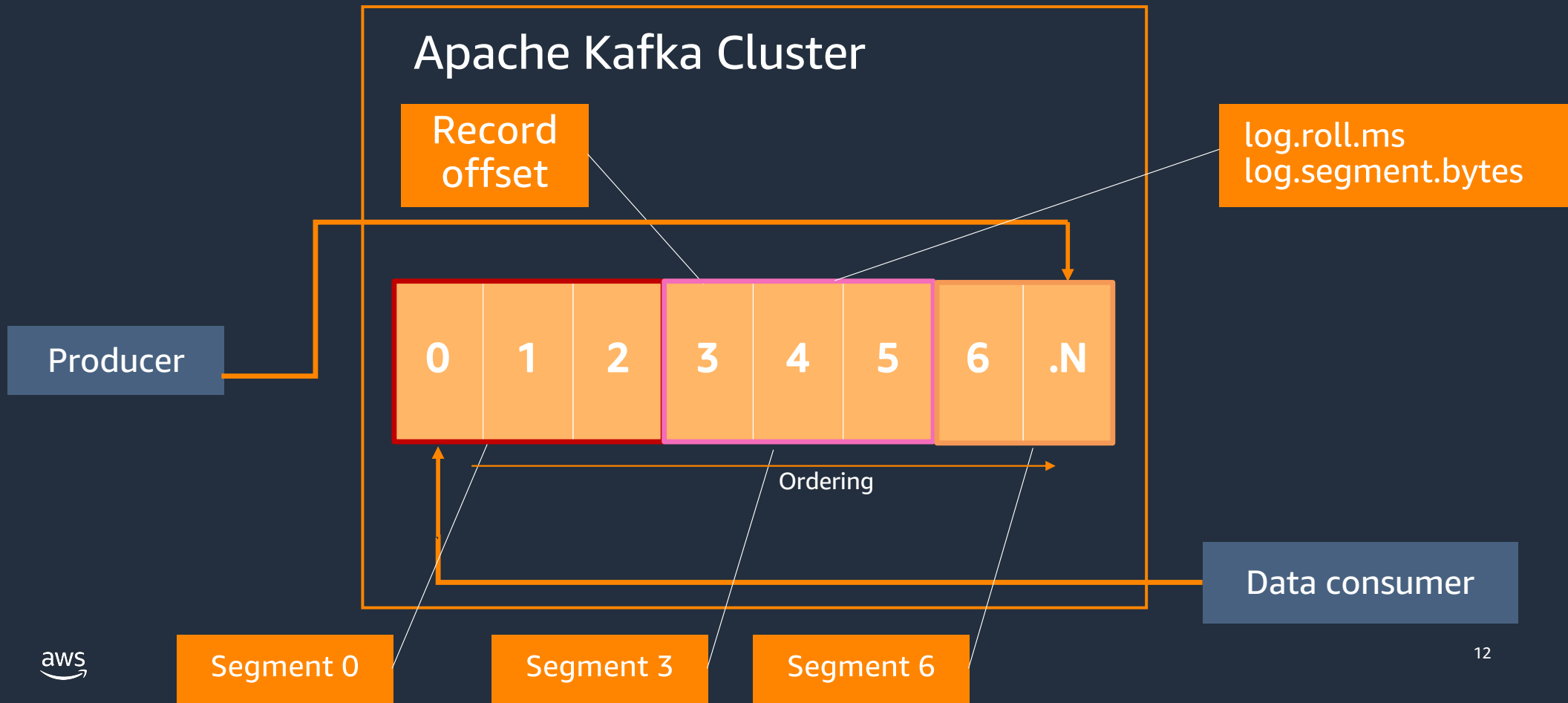
Apache Kafka 101: Topics



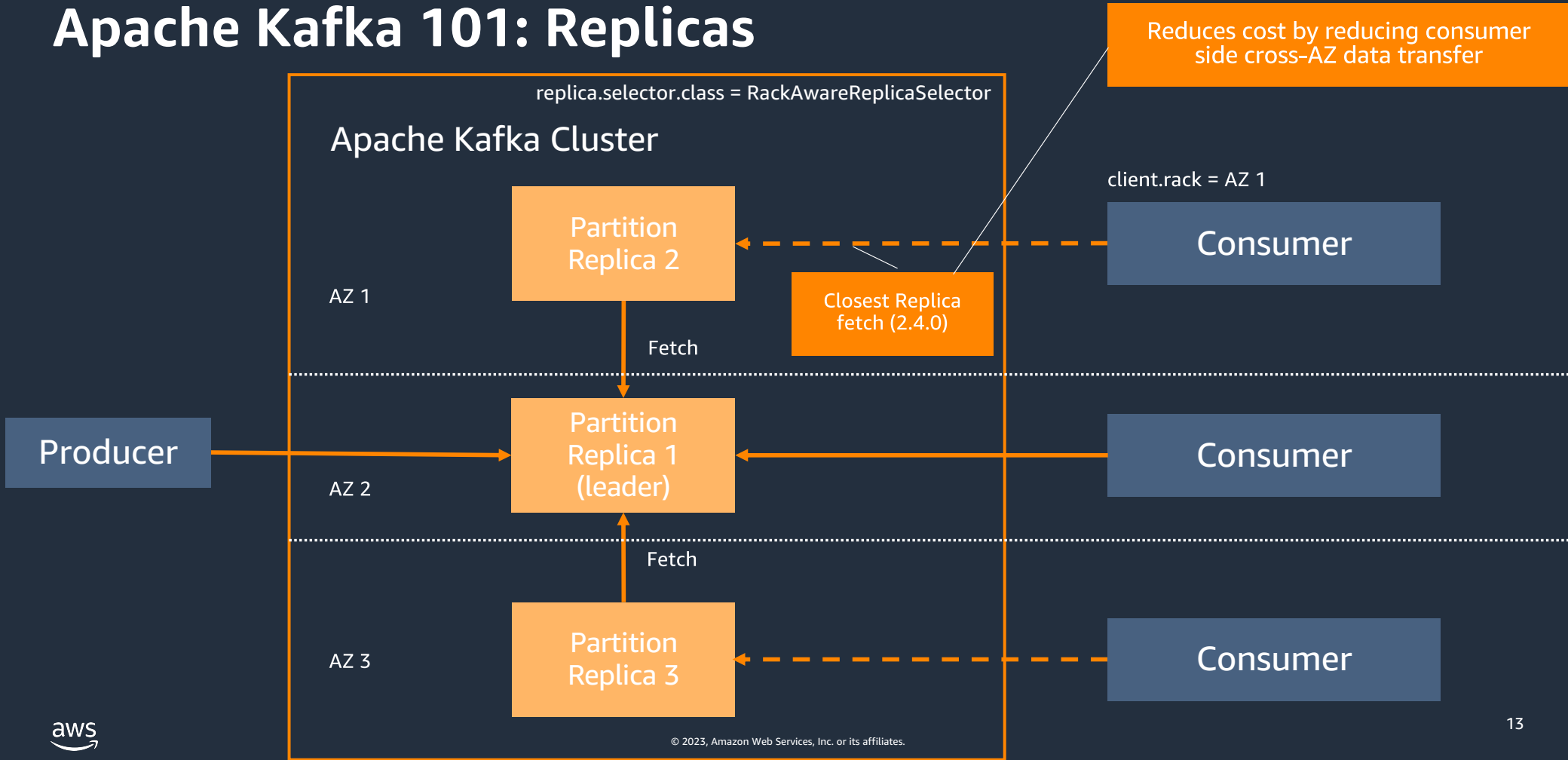
Apache Kafka 101: Partitions



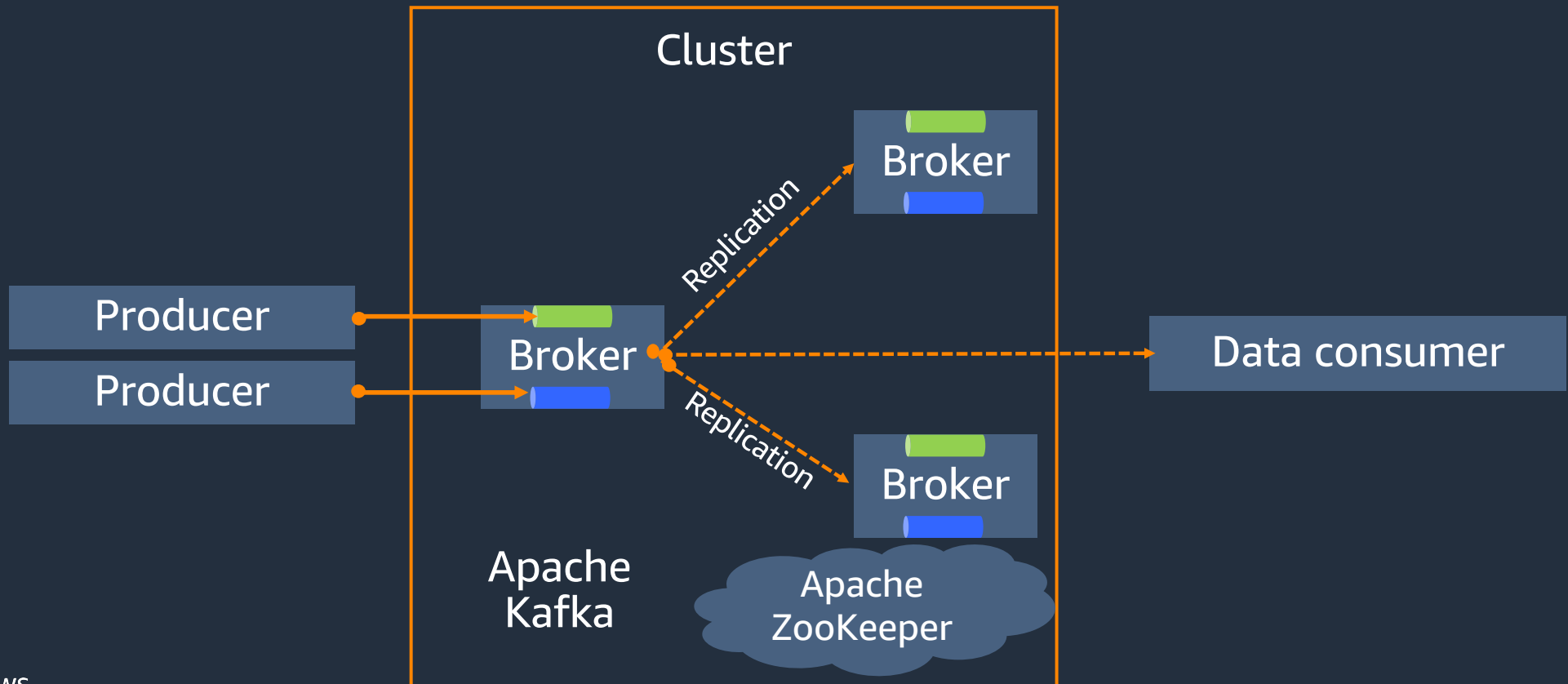
Apache Kafka 101: Offsets



Apache Kafka 101: Replicas

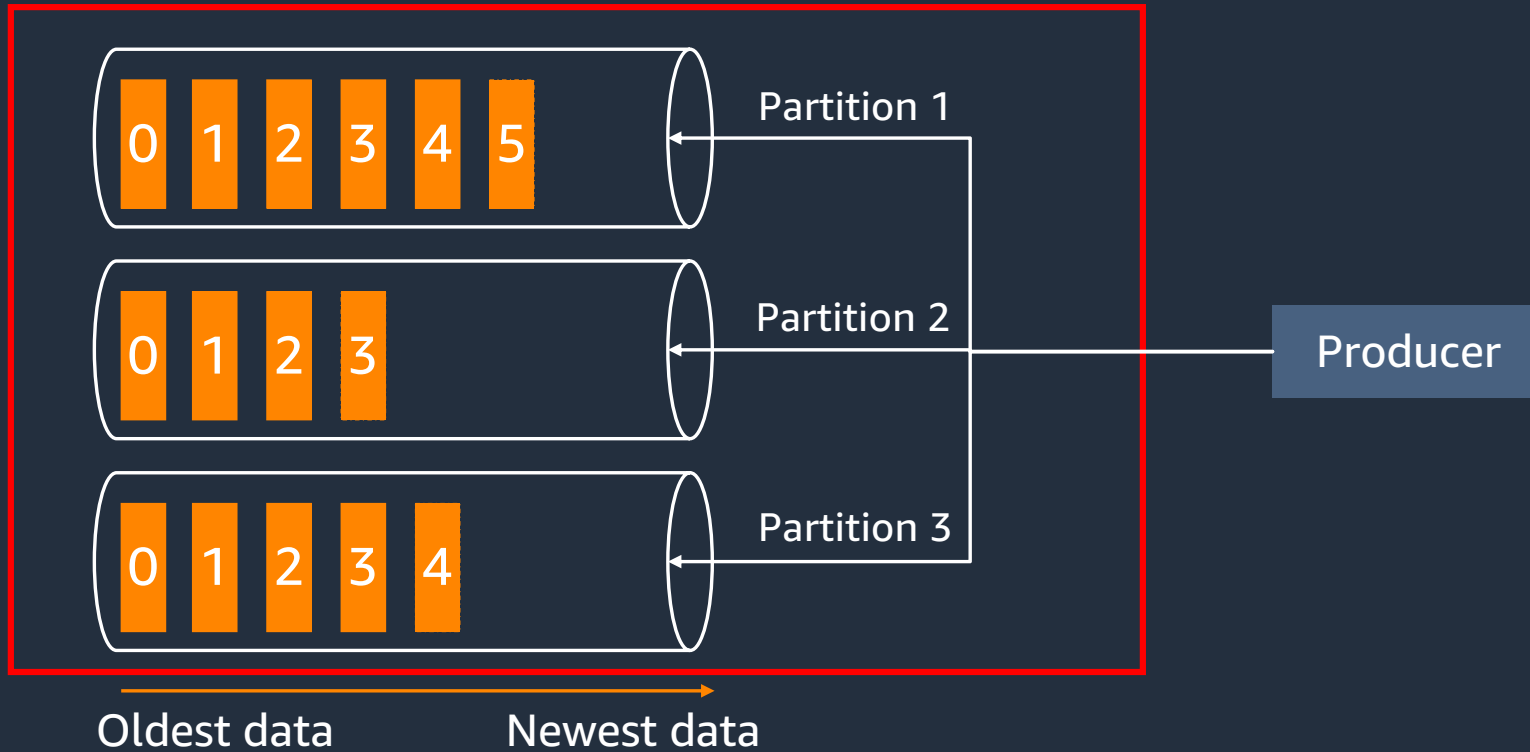


Apache Kafka 101: Architecture

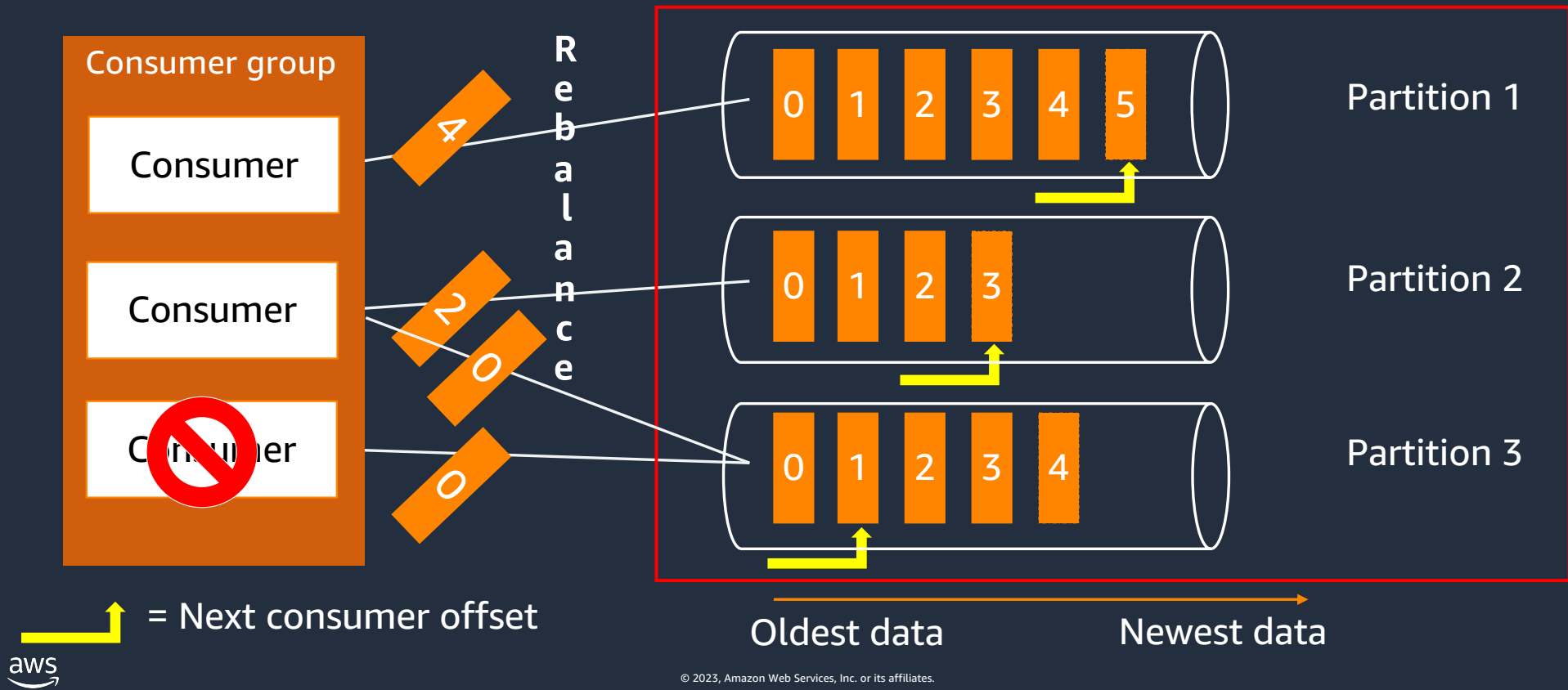


Apache Kafka 101: Writing to partitions

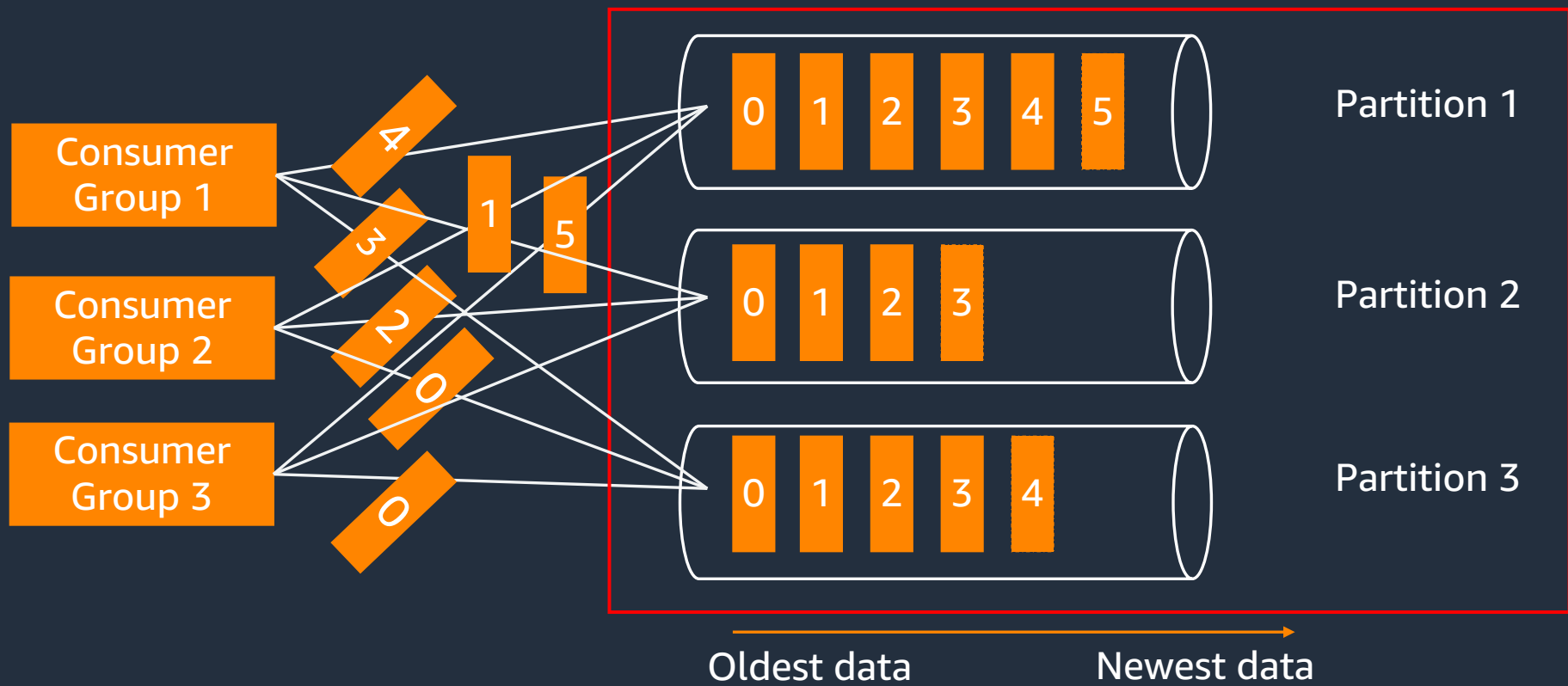
Topic with 3 partitions



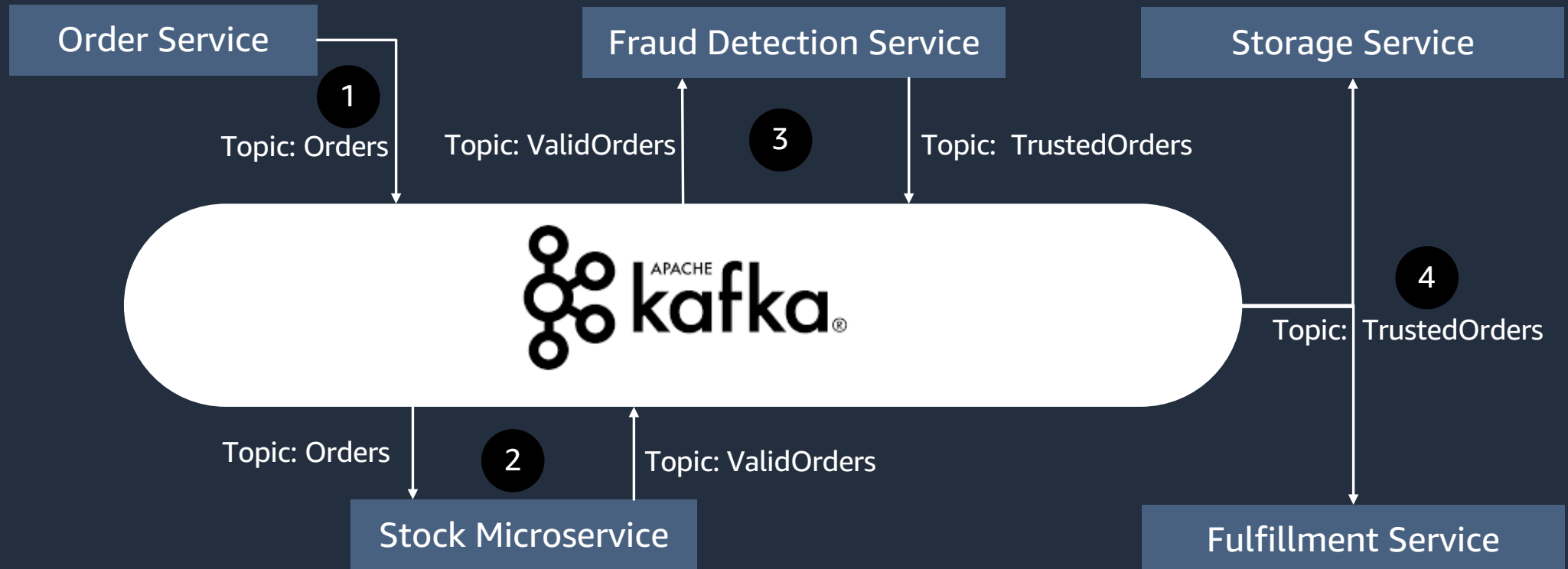
Apache Kafka 101: Reads from partitions



Apache Kafka 101: Pub/Sub



Apache Kafka in event-driven architecture



Challenges operating Apache Kafka



Challenges operation Apache Kafka



Difficult to setup



Tricky to scale



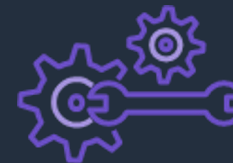
Hard to achieve high availability



Integration required
development



Error prone and complex to manage



Expensive to maintain

Benefits to using Amazon MSK



**“Amazon MSK: A fully managed,
high available and secure service
for Apache Kafka”**

Amazon MSK is committed to improving open-source Apache Kafka

Amazon MSK key benefits



Create a secure, highly available and accessible cluster within minutes



Drives best practices through design, defaults, and automation



Allows developers to focus more on application development and less on infrastructure management



Eliminates right-sizing, scaling, and partition management with MSK Serverless



Thank you!