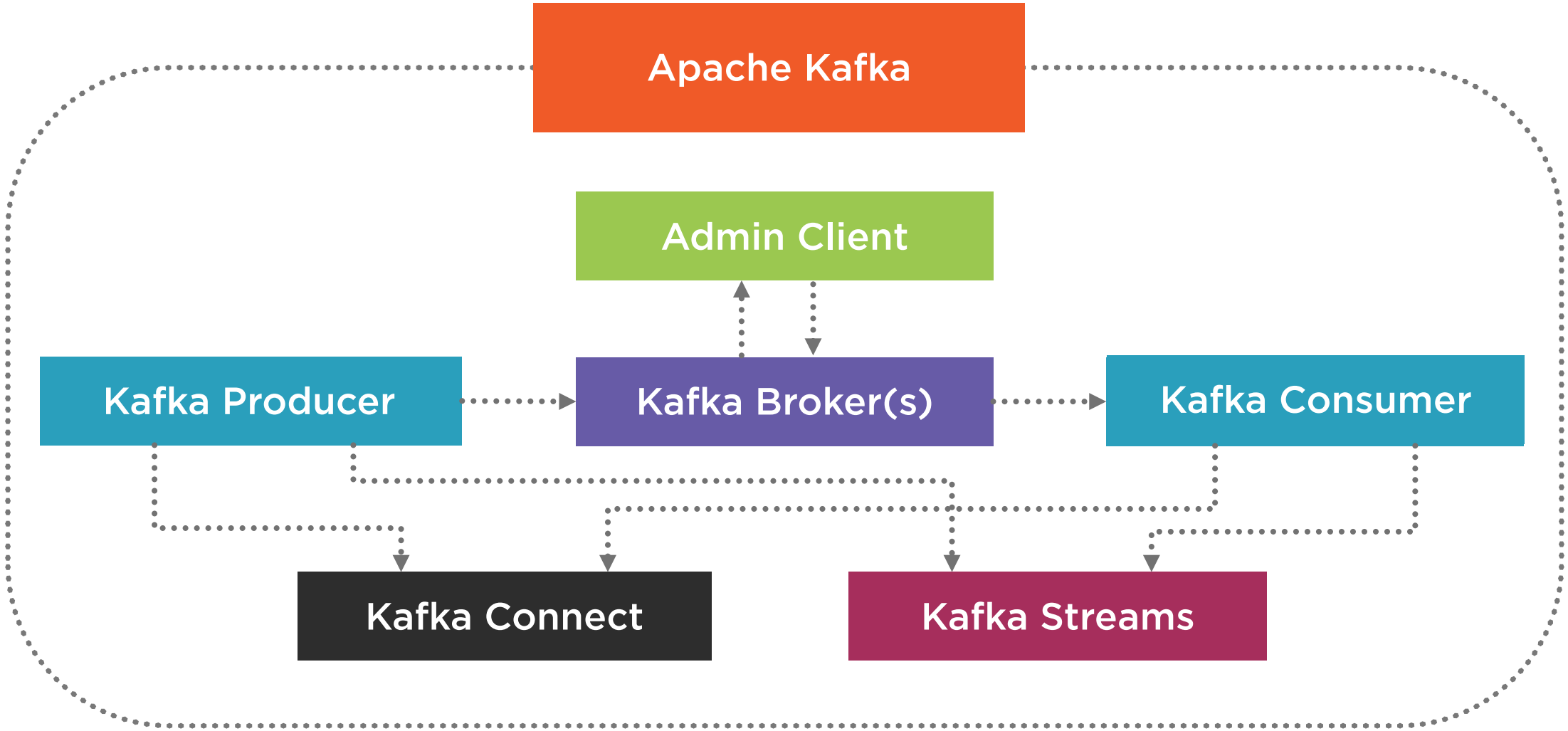
Producing Data to Kafka

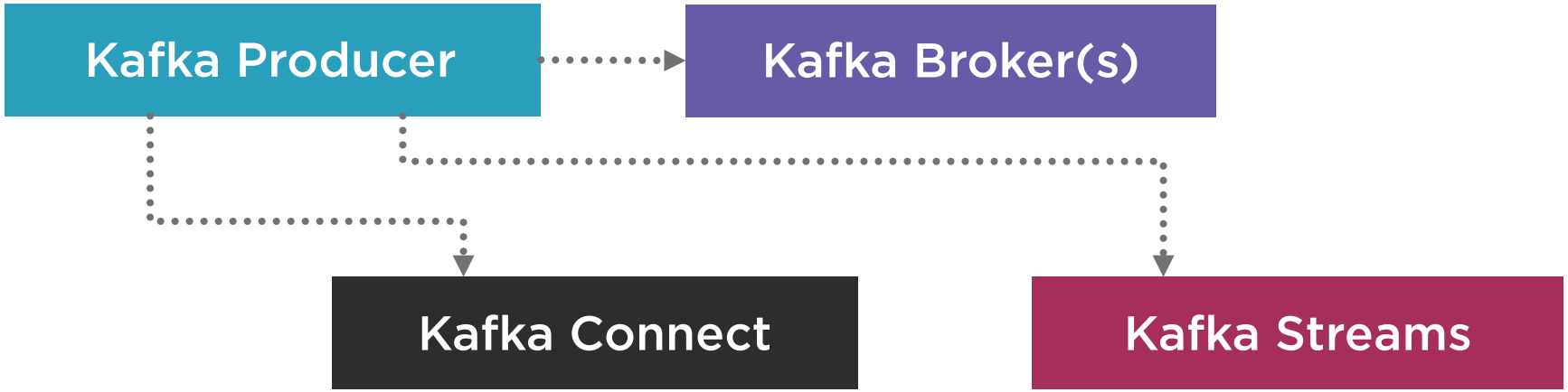
Bogdan Sucaciu
SOFTWARE ENGINEER

@BSucaciu bsucaciu.com

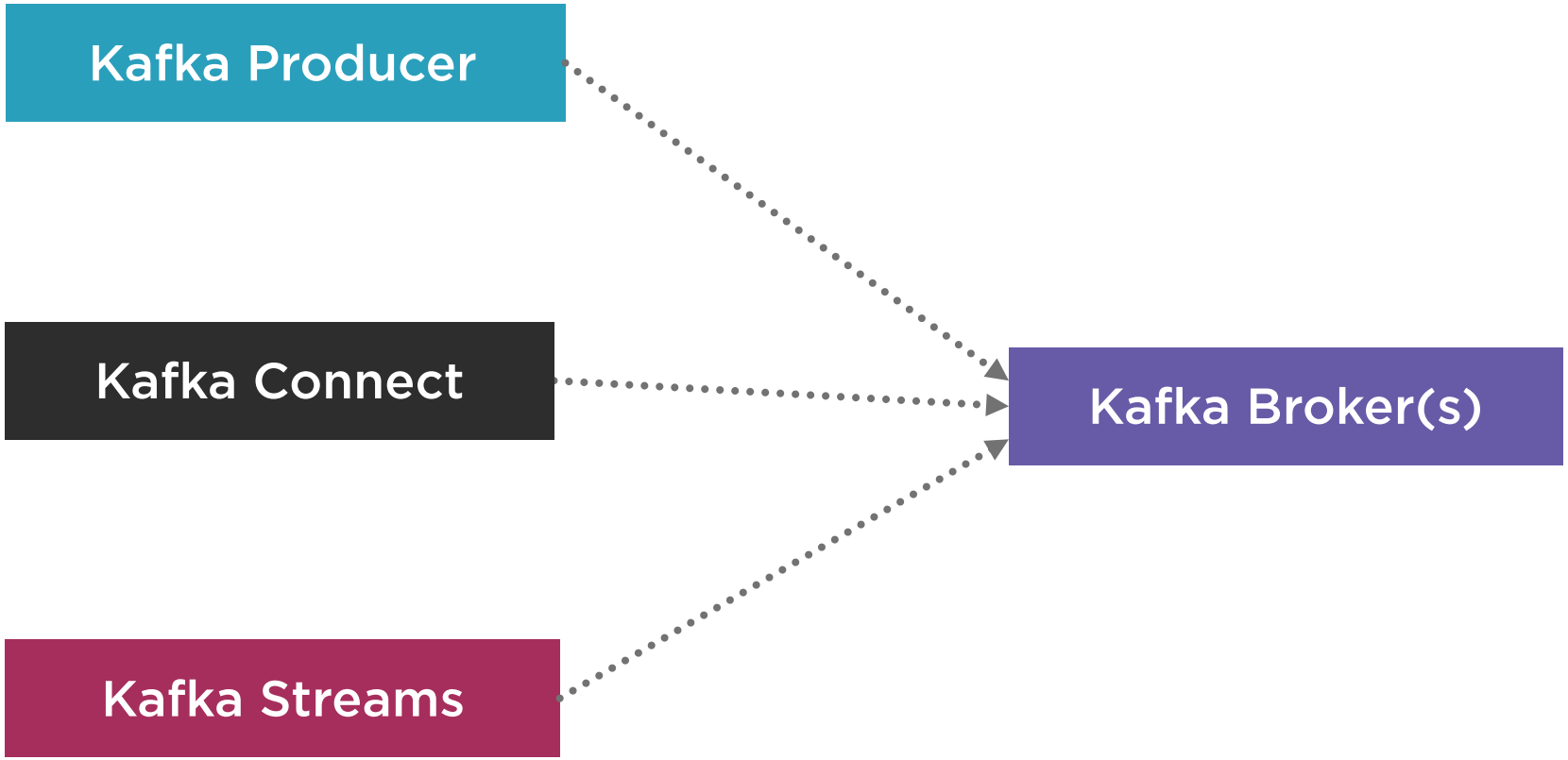
Producing Data to Kafka



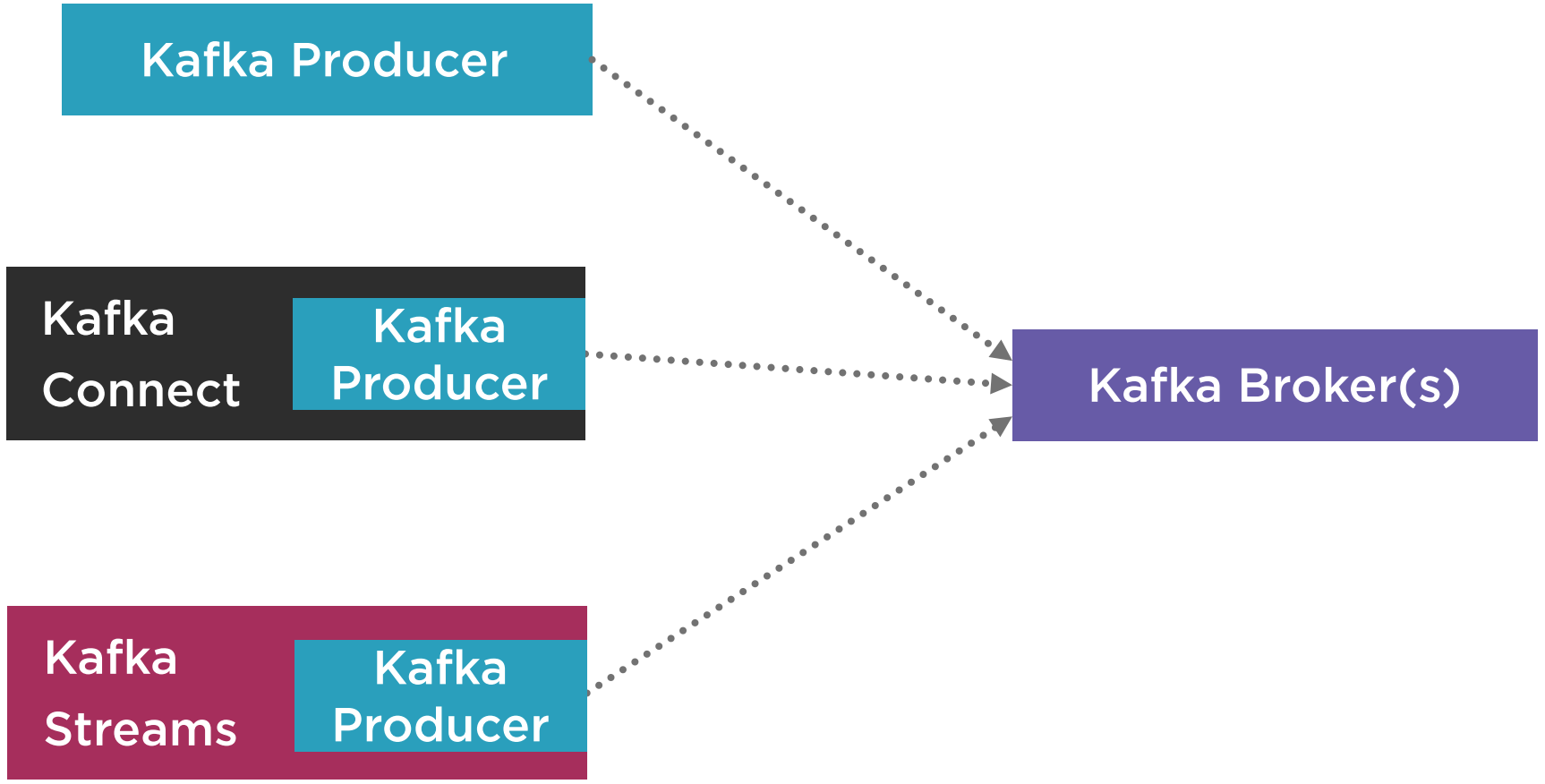
Producing Data to Kafka



Producing Data to Kafka



Producing Data to Kafka

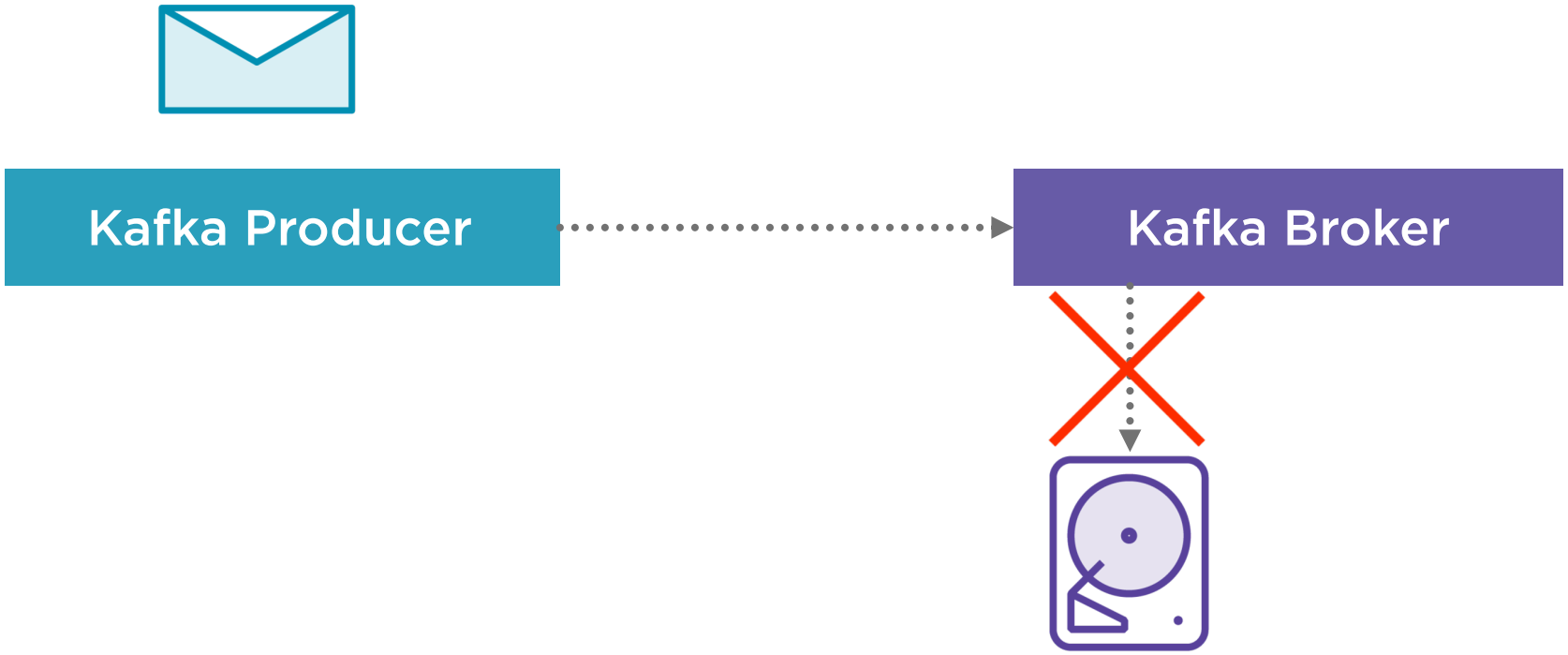


- Java
- C/C++
- Python
- Go (Golang)
- Erlang
- .NET
- Clojure
- Ruby
- Node.js
- Perl
- PHP
- Rust
- Storm
- Scala
- Swift

Message Durability

Message Durability

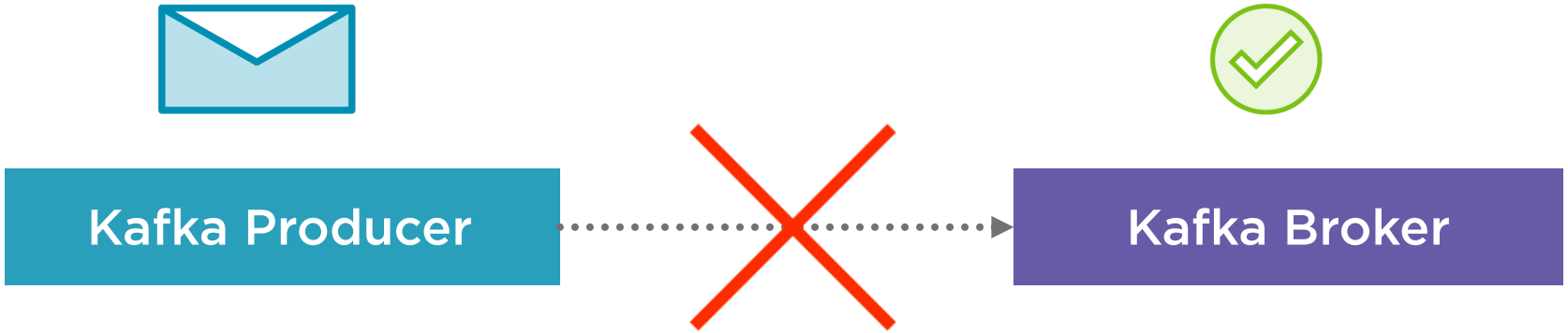
Message Durability



Message Durability

acks

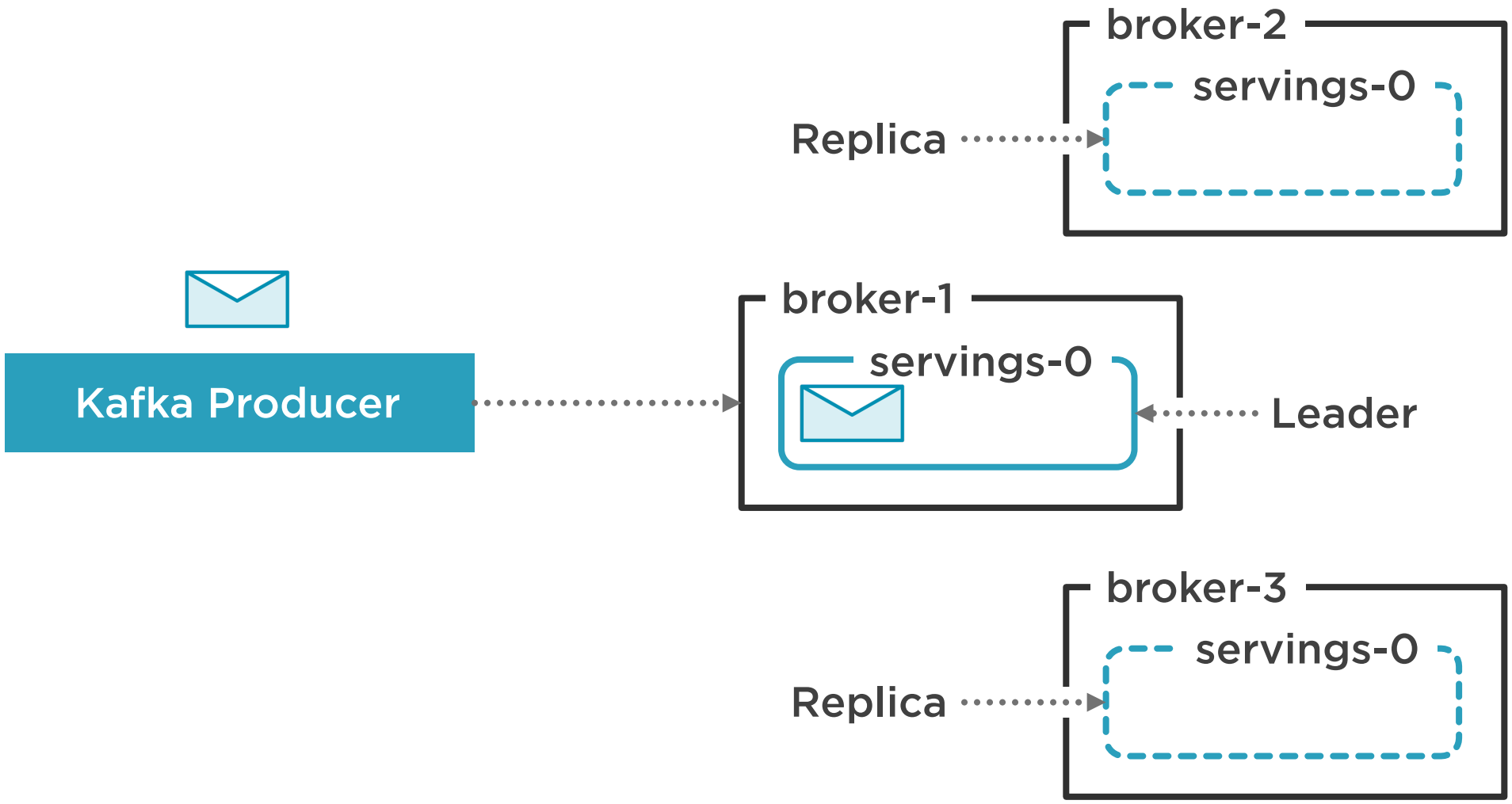
Message Durability



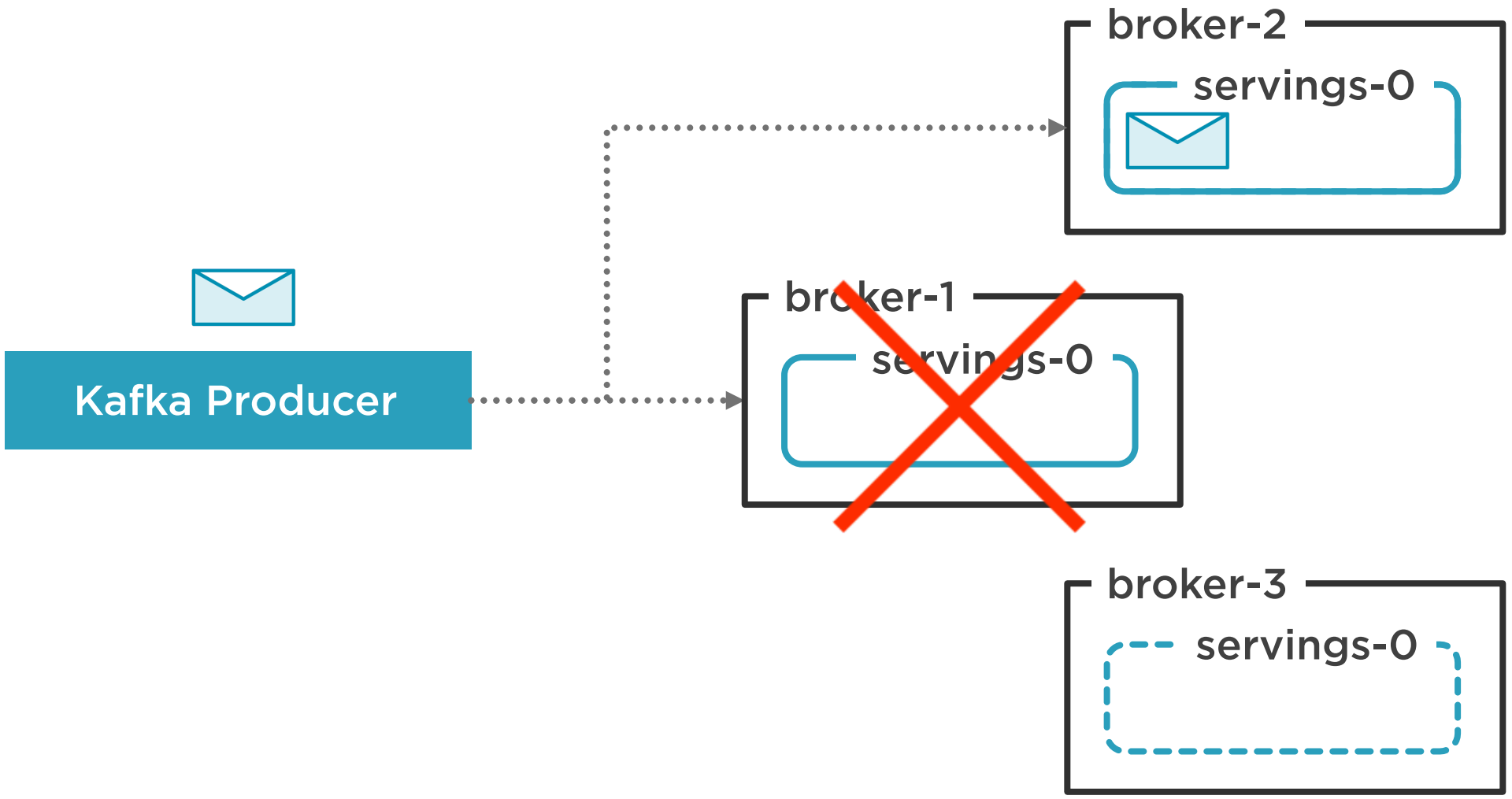
acks

retries = 5

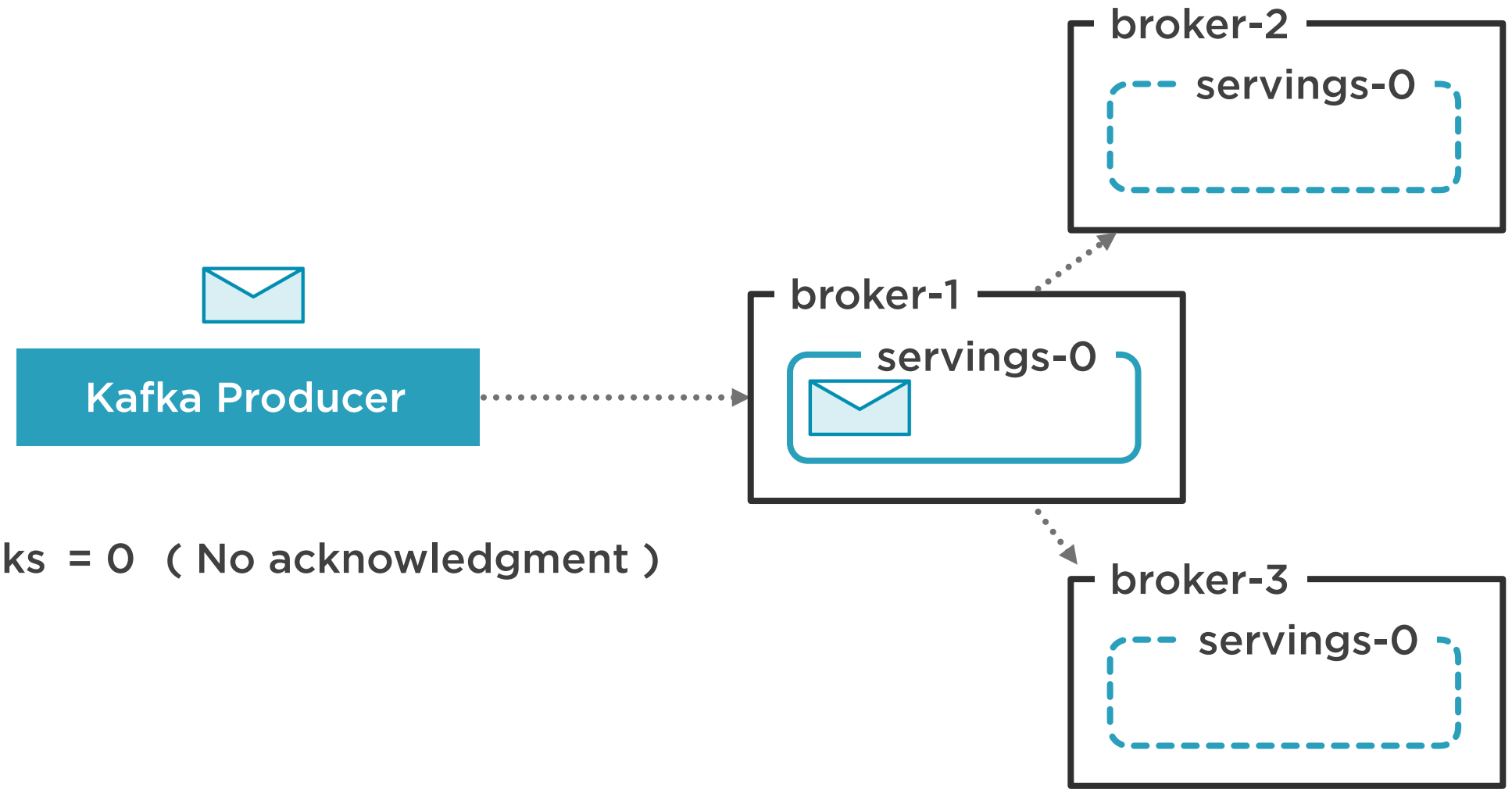
Replicas


broker-2
servings-0
Replica
Kafka Producer
broker-1
servings-0
Leader
broker-3
servings-0
Replica

Replicas

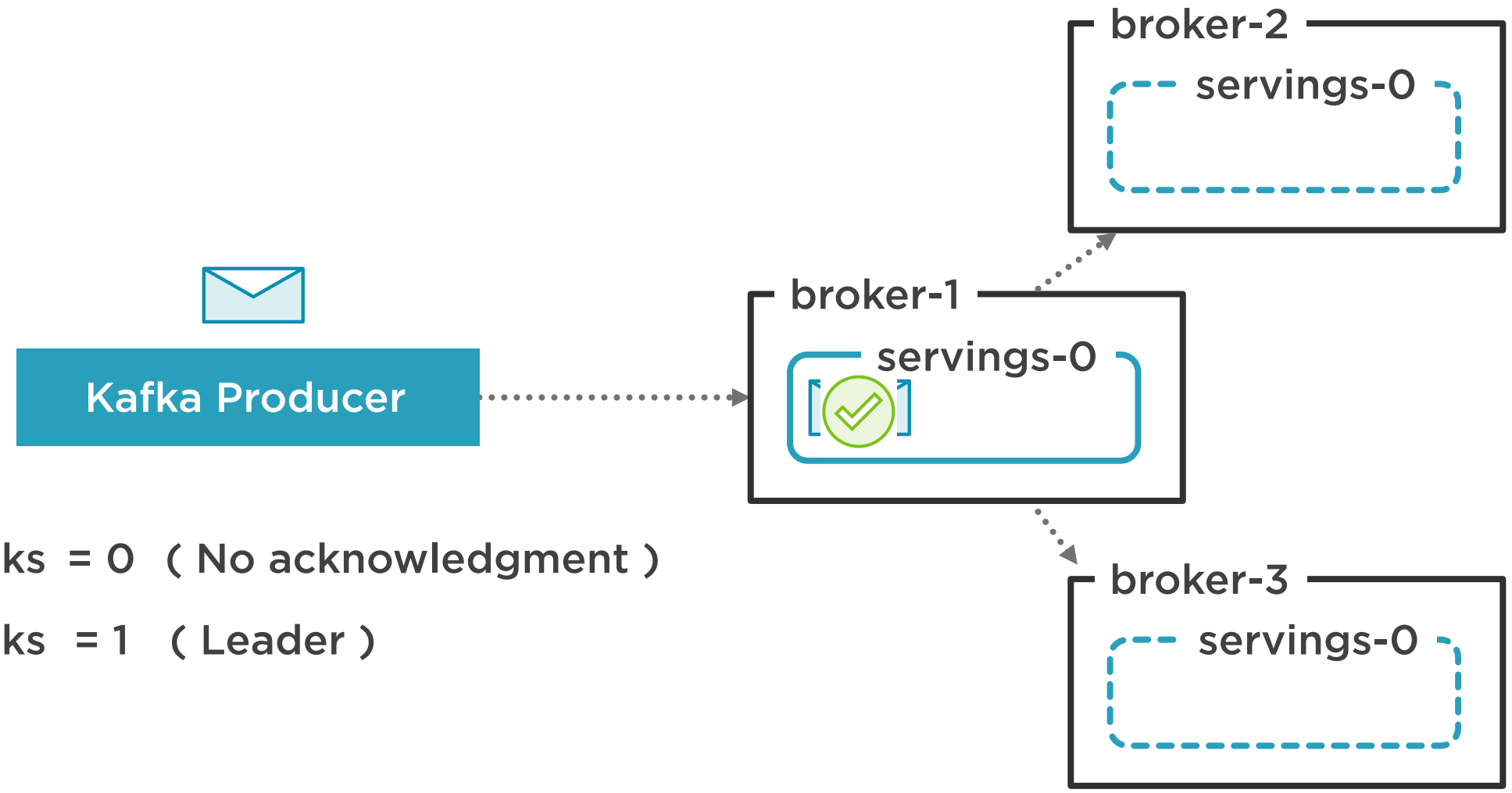

broker-2
servings-0
Kafka Producer
broker-1
servings-0
broker-3
servings-0

Acknowledgment



acks = 0 (No acknowledgment)

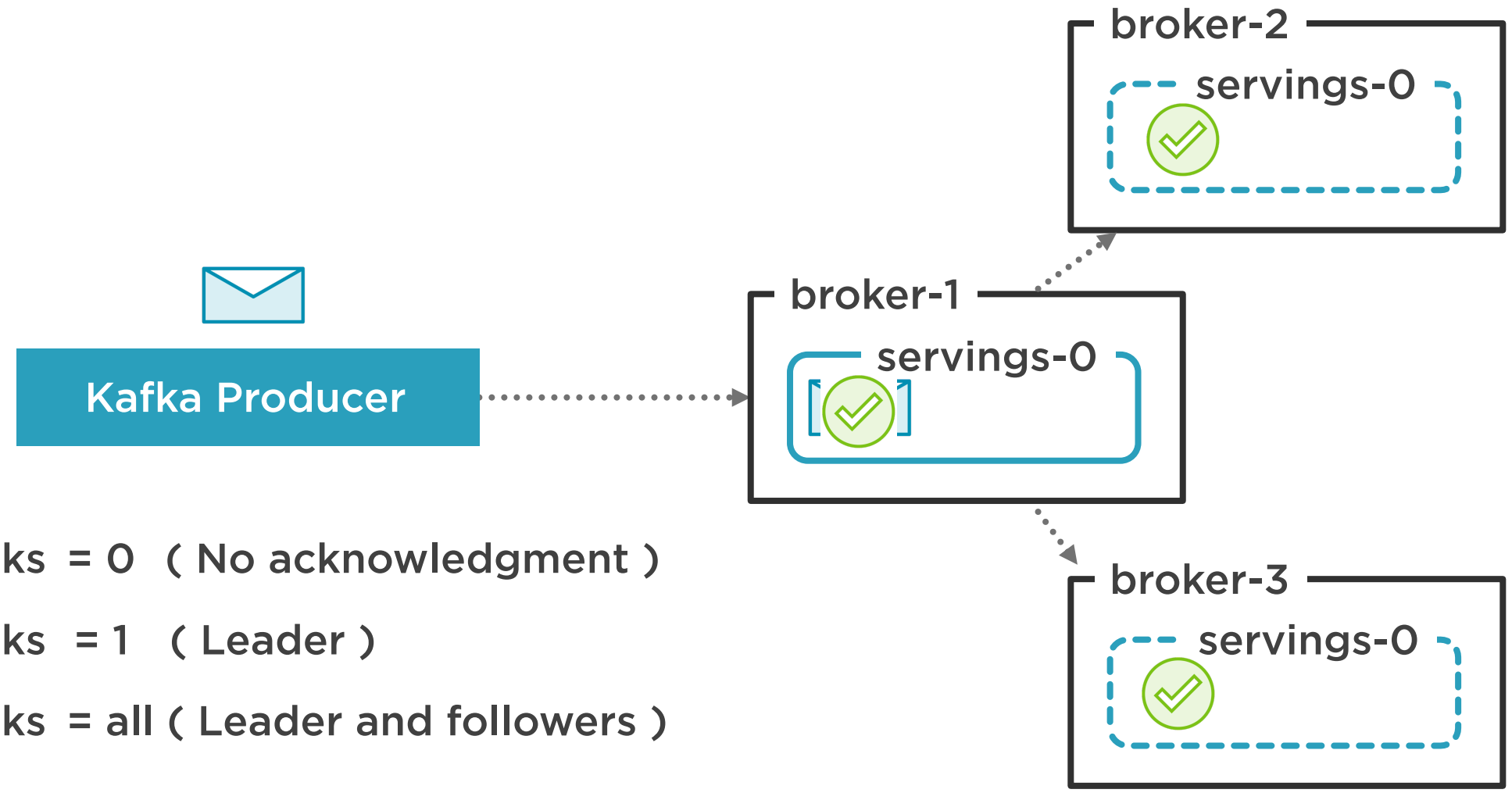
Acknowledgment



acks = 0 (No acknowledgment)

acks = 1 (Leader)

Acknowledgment



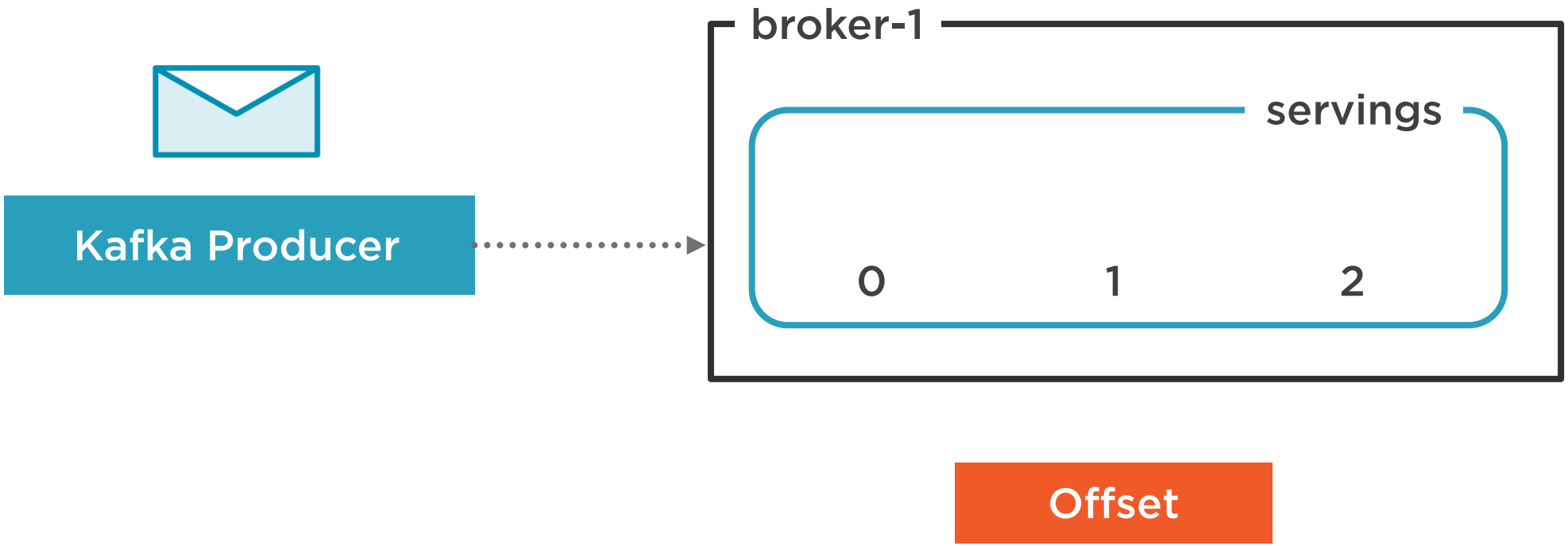
acks = 0 (No acknowledgment)

acks = 1 (Leader)

acks = all (Leader and followers)

Message Ordering

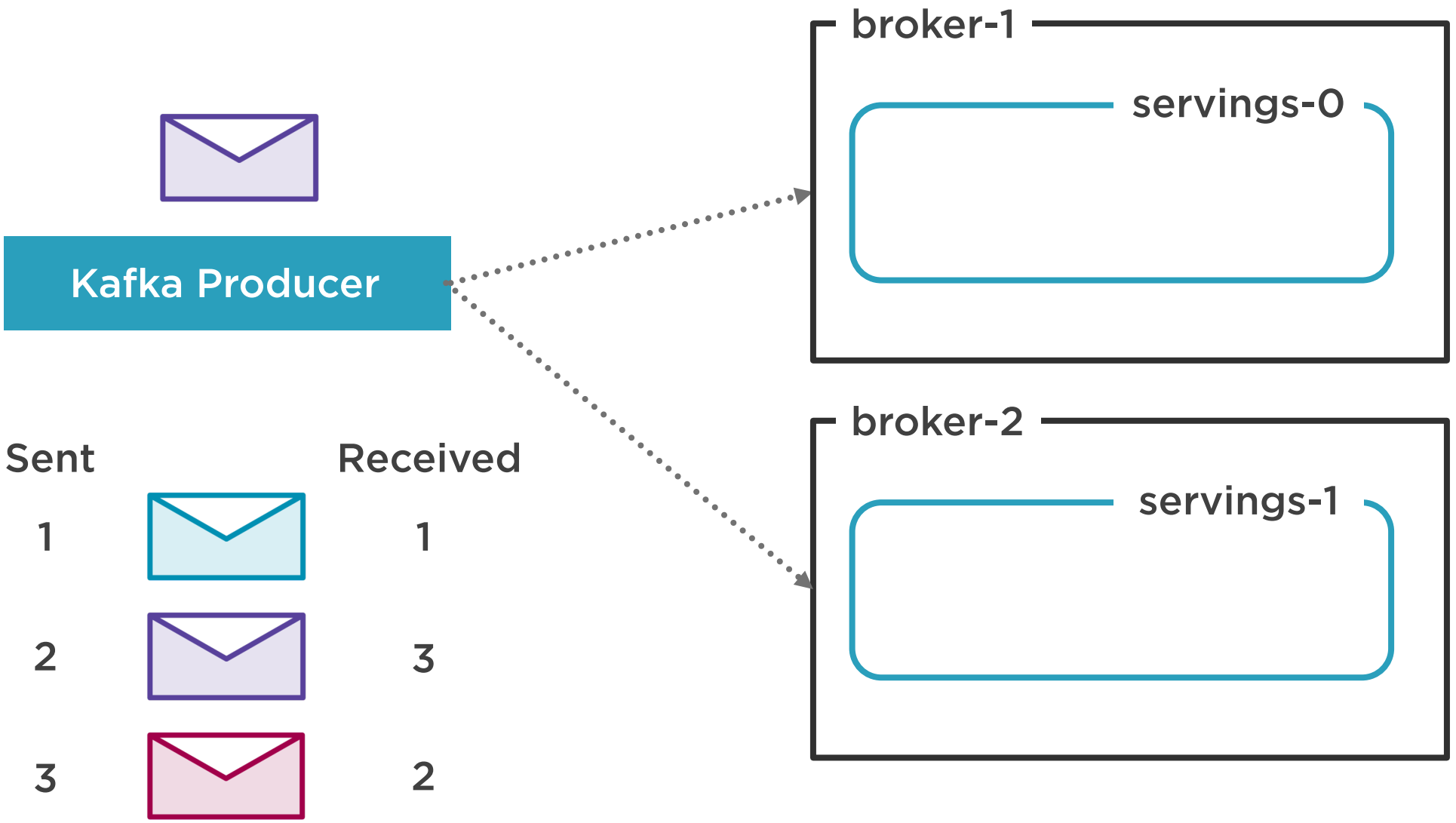
Message Ordering



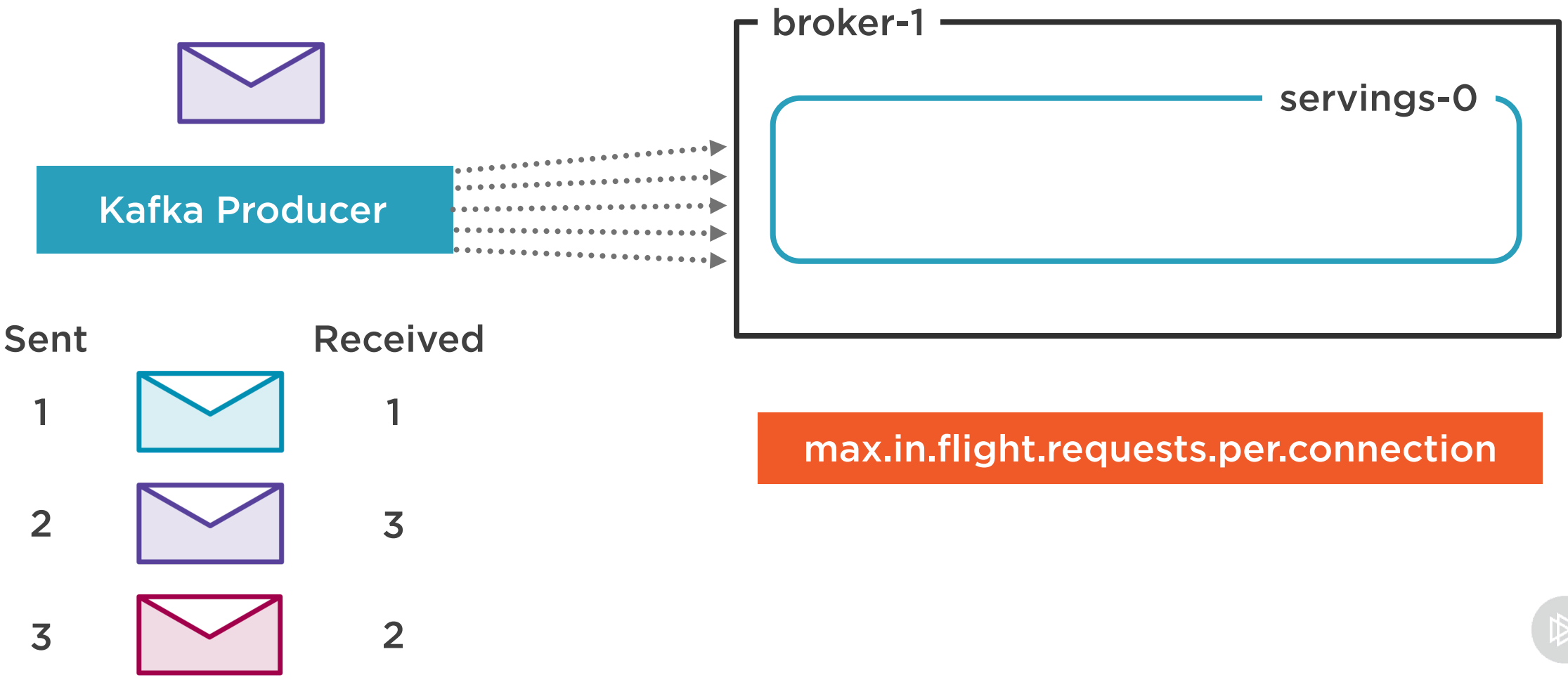
Messages are ordered
per topic

Messages are ordered
per ~~topic~~ partition

Message Ordering

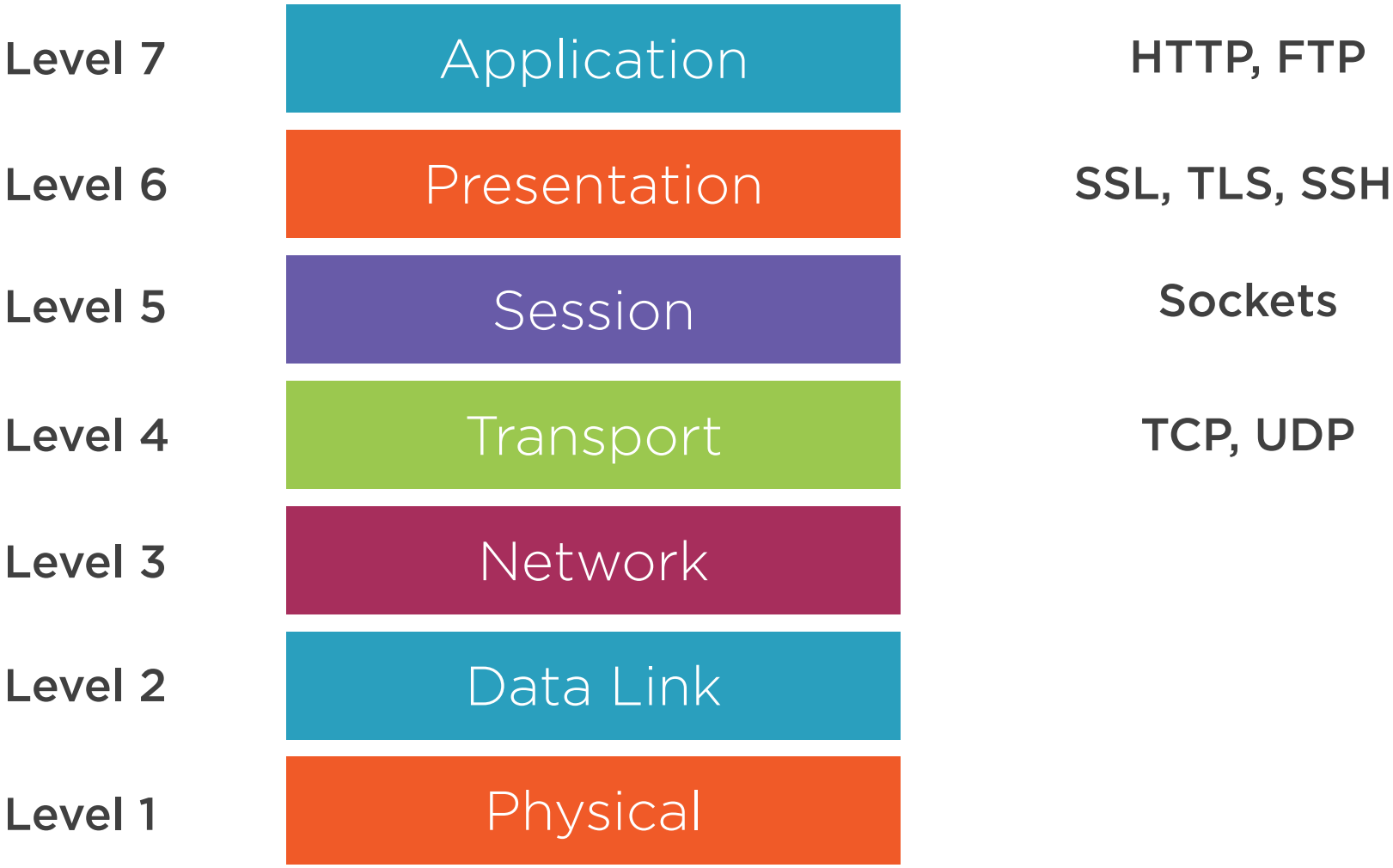


Message Ordering



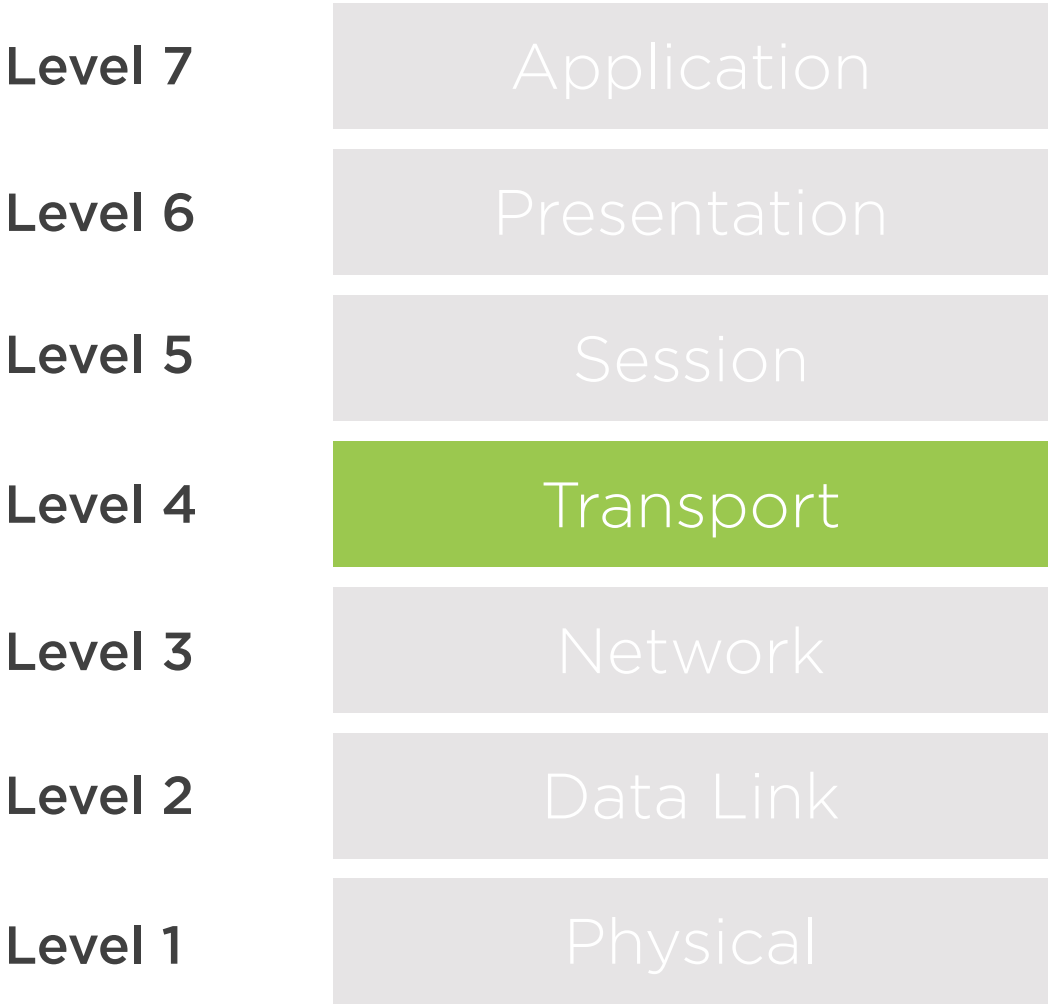
Kafka Producer

Kafka Protocol



OSI Network Model

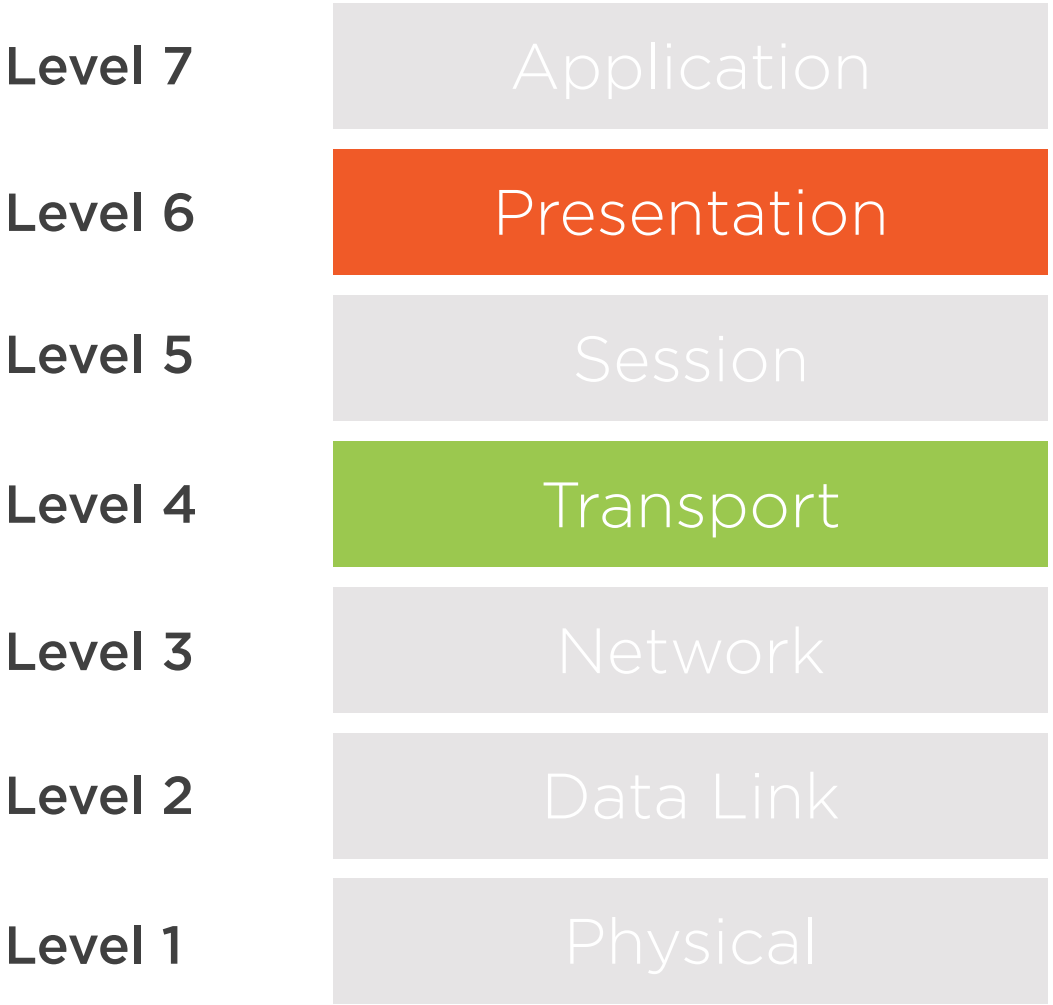
Kafka Protocol



OSI Network Model

TCP – non encrypted communication

Kafka Protocol

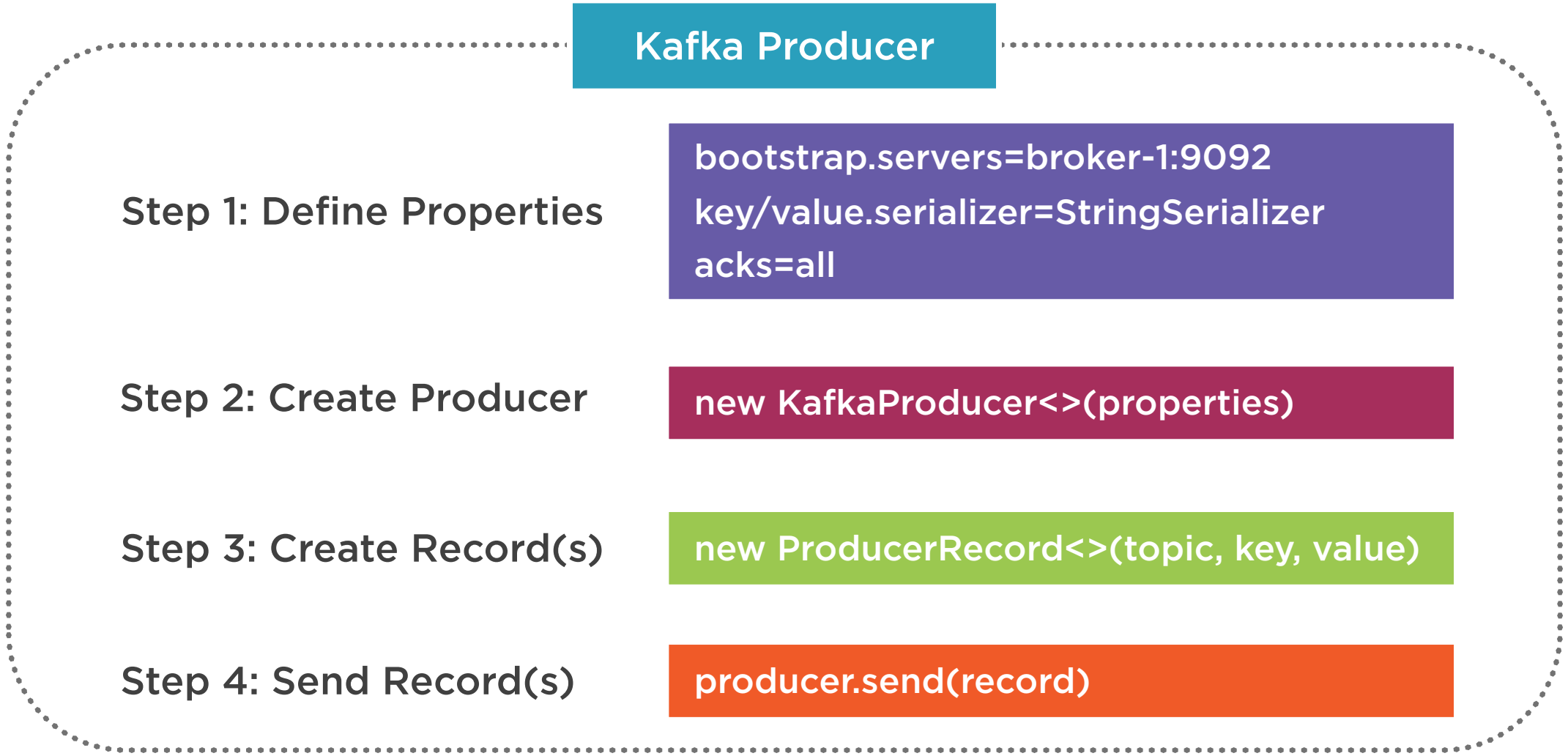


OSI Network Model

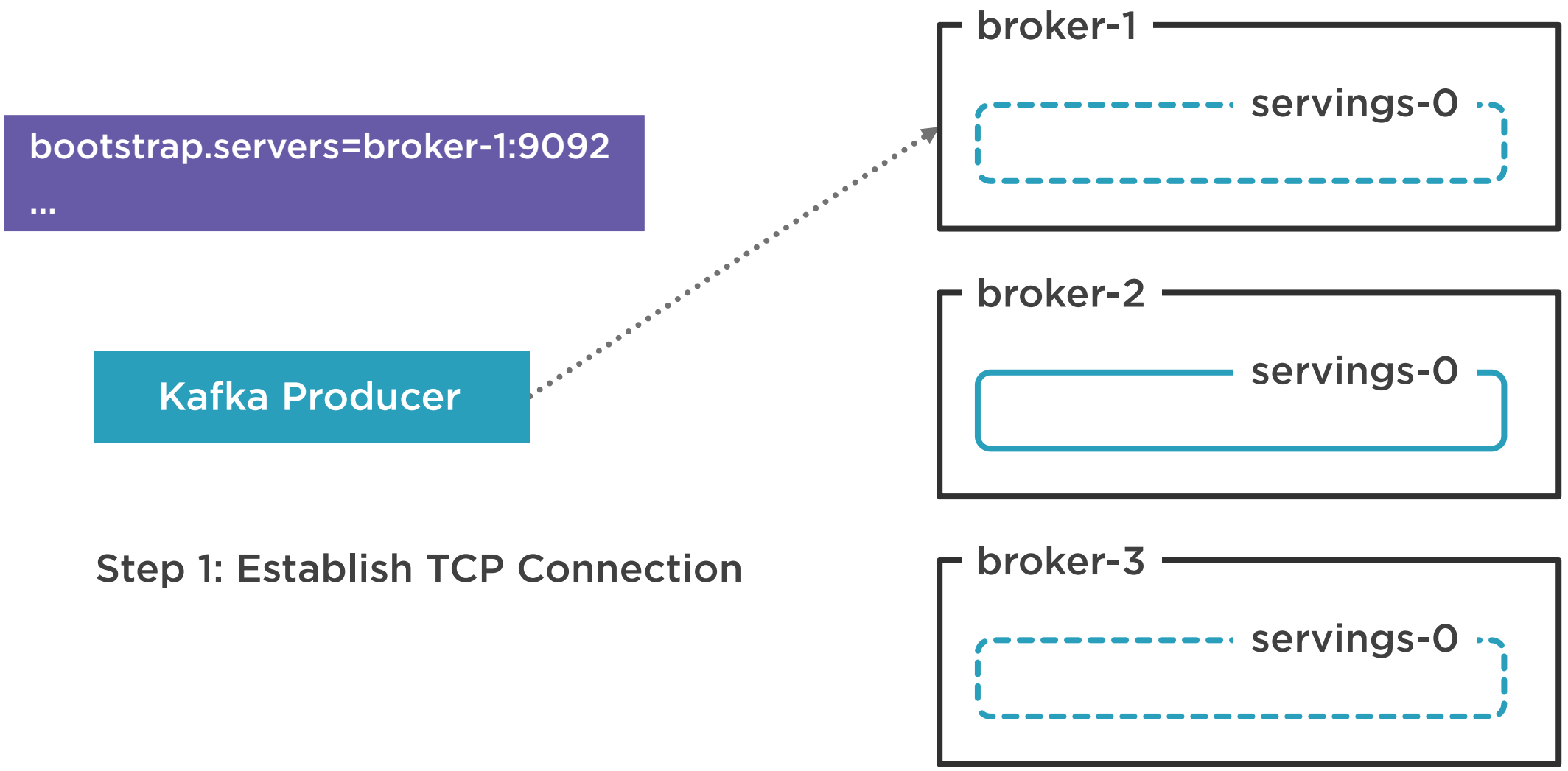
SSL – encrypted communication

TCP – non encrypted communication

Setting up Producer(s)

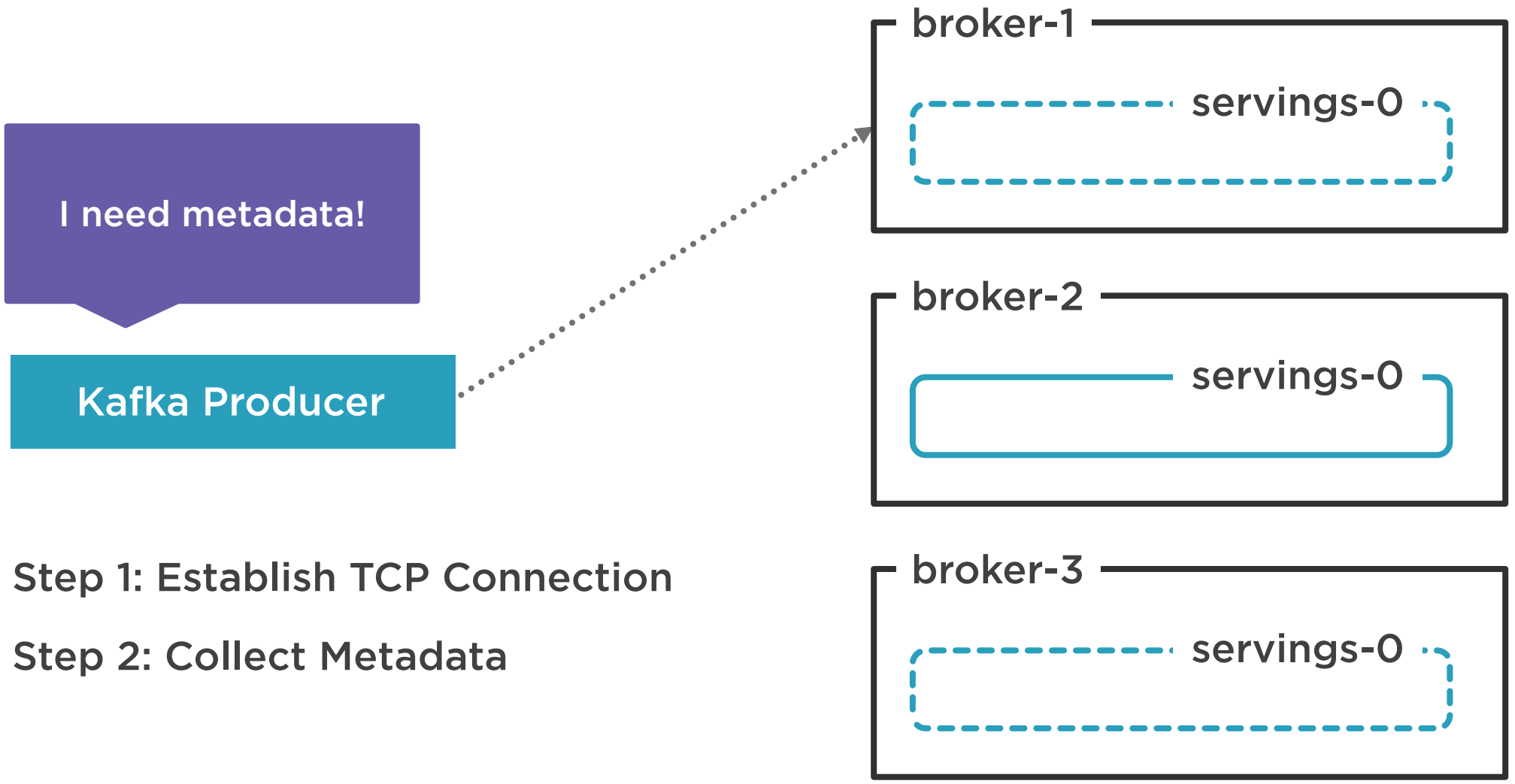


Producing to Kafka



Step 1: Establish TCP Connection

Producing to Kafka



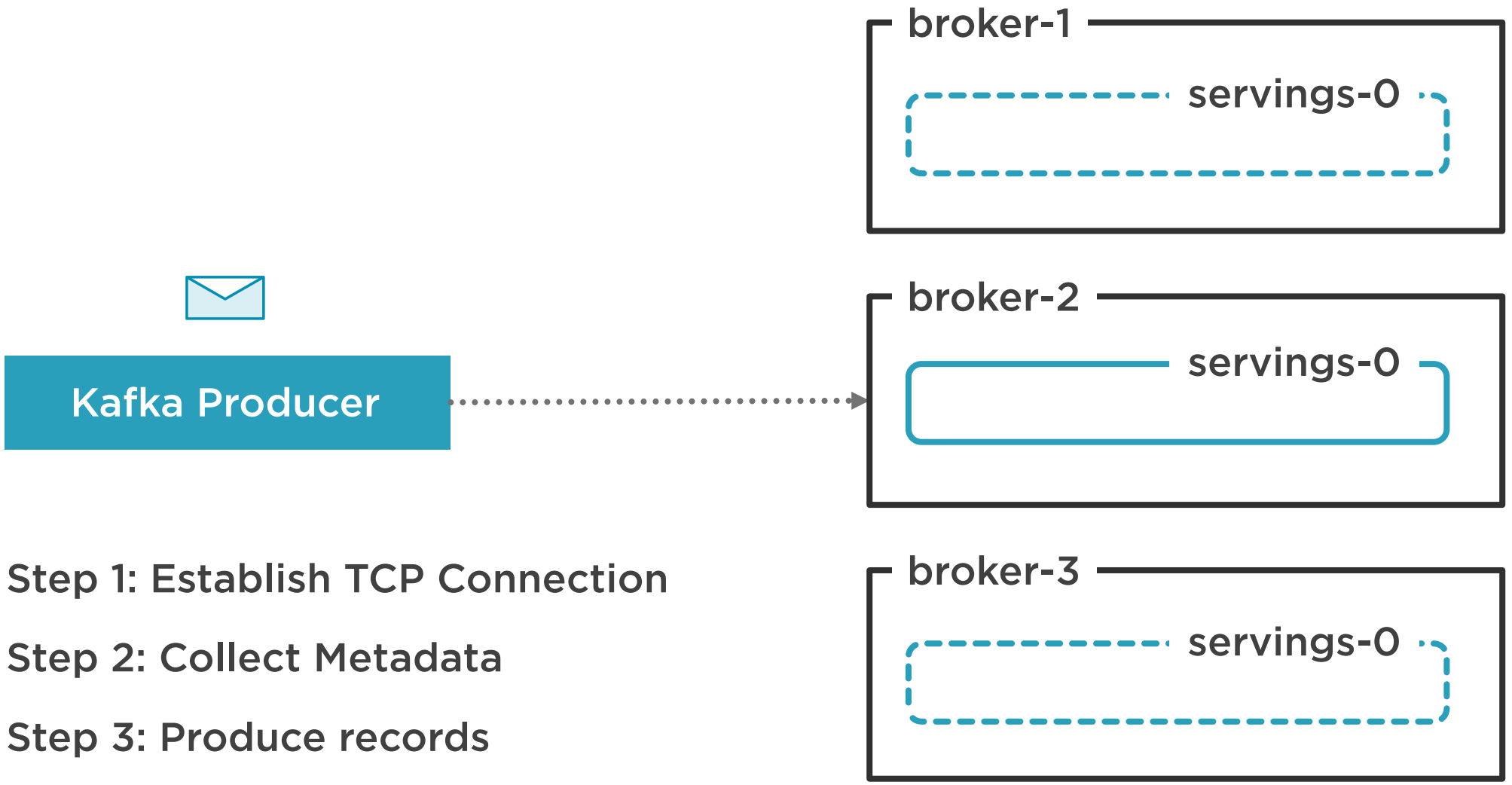
Step 1: Establish TCP Connection

Step 2: Collect Metadata

Brokers Metadata

```
Metadata for all topics (from broker 1: broker-1:9092/1):
3 brokers:
      broker 1  at broker-1:9092
      broker 2 at broker-2:9093
      broker 3 at broker-3:9094
1 topic:
      topic ”servings" with 3 partitions:
            partition 0, leader 2, replicas: 1,2,3
```

Producing to Kafka

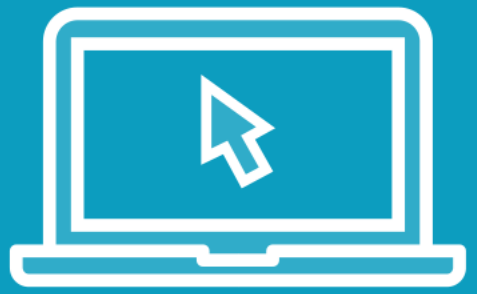


Step 1: Establish TCP Connection

Step 2: Collect Metadata

Step 3: Produce records

Demo

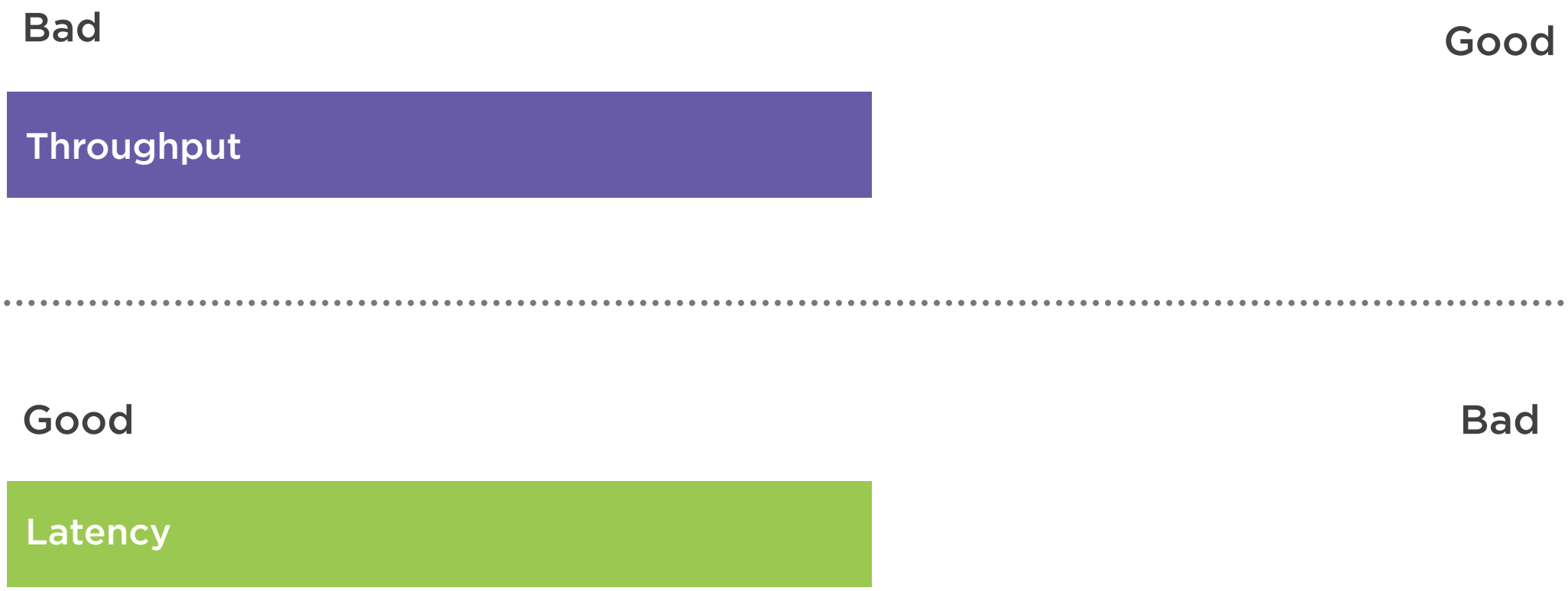


Write a Java Kafka Producer

Servings Data

Optimizing for High Throughput

Throughput vs. Latency



Throughput vs. Latency

Throughput vs. Latency

Bad

Good

Good

Bad

Game of Four

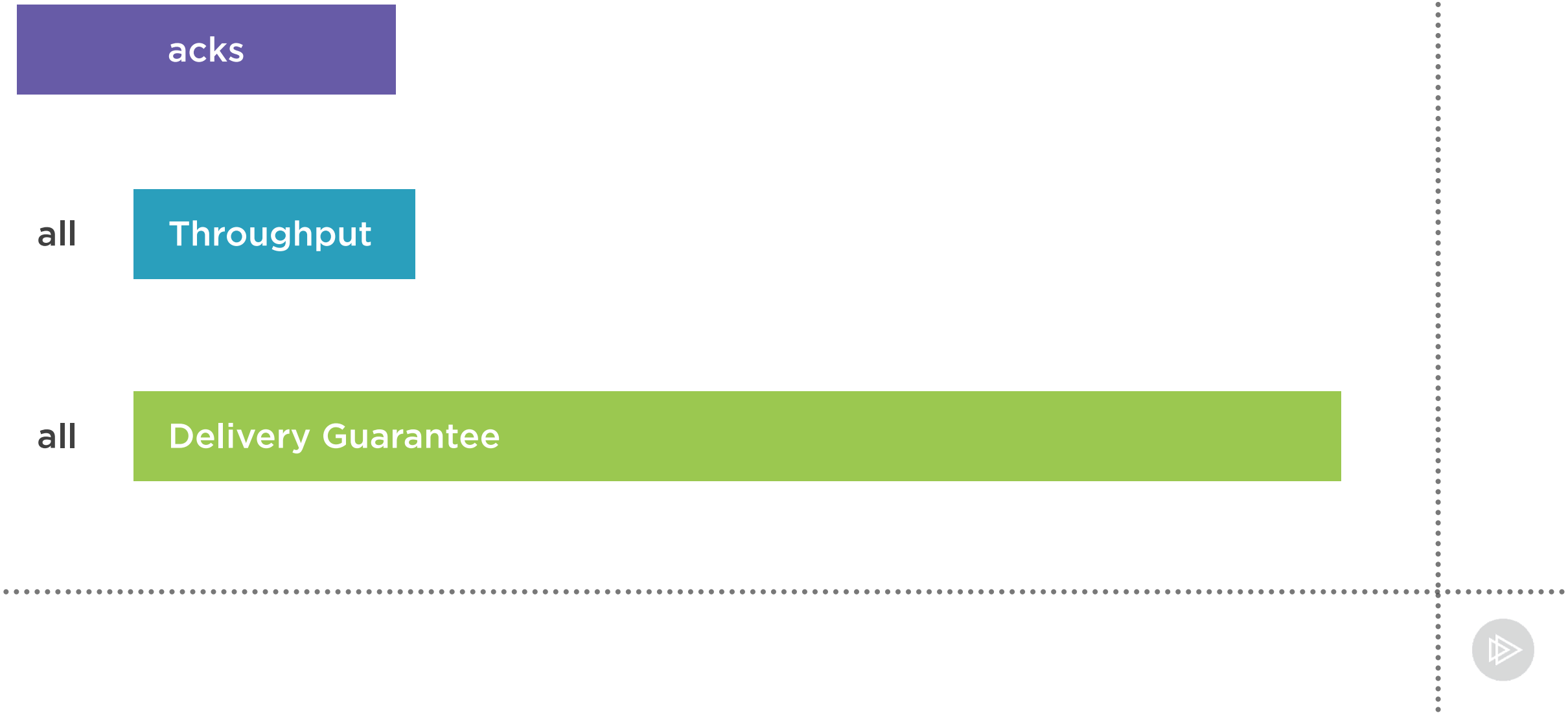
acks

linger.ms

batch.size

max.in.flight.requests.per.connection

Game of Four



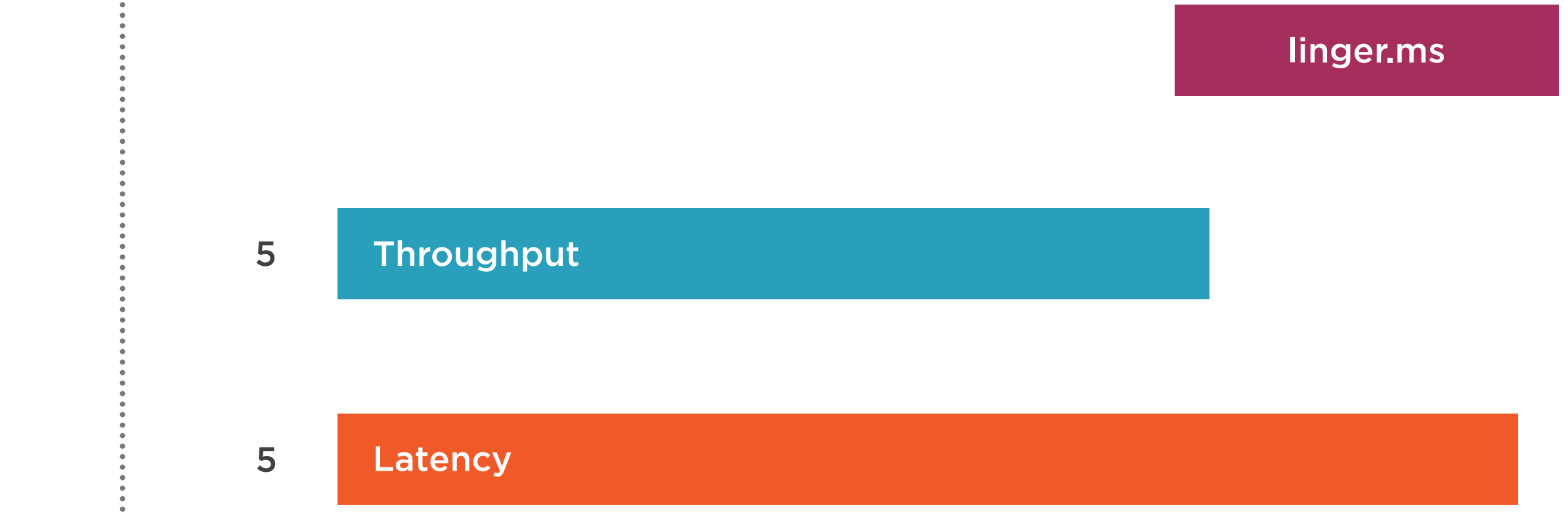
Game of Four

Game of Four

linger.ms

- Throughput
- Latency

Game of Four



Game of Four

max.in.flight.requests.per.connection

1 Throughput

1 Ordering Guarantee

Game of Four

max.in.flight.requests.per.connection

5 Throughput

5 Ordering Guarantee

Game of Four

batch.size

16384 Throughput

16384 Memory Usage

Game of Four

batch.size

1024 Throughput

1024 Memory Usage

Demo

Optimize servings producer for high throughput

Optimizing for Low Latency

Game of Four

acks

linger.ms

batch.size

max.in.flight.requests.per.connection

Game of Four

acks

linger.ms

batch.size

max.in.flight.requests.per.connection

Game of Four

Game of Four

acks

0 Latency

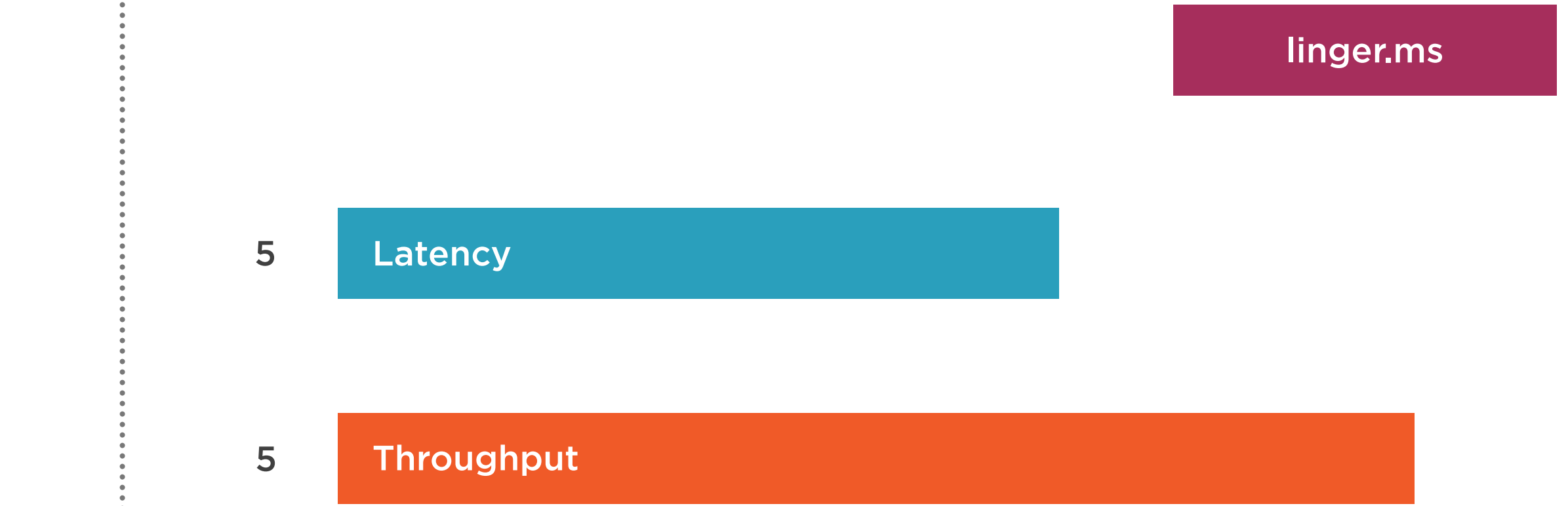
0 Delivery Guarantee

Game of Four

linger.ms

- Latency
- Throughput

Game of Four



Game of Four

batch.size

1024 Latency

1024 Throughput

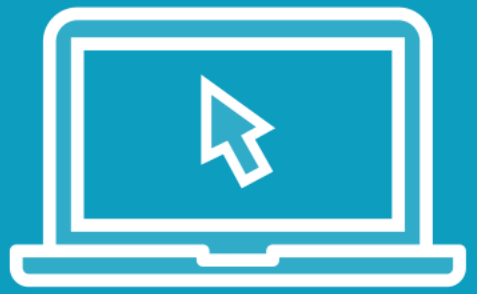
Game of Four

batch.size

16384 Latency

16384 Throughput

Demo



Deliveries Producer

Optimize for Low Latency

Summary

Message Durability

Message Ordering

Kafka Producer

Optimizing for different scenarios