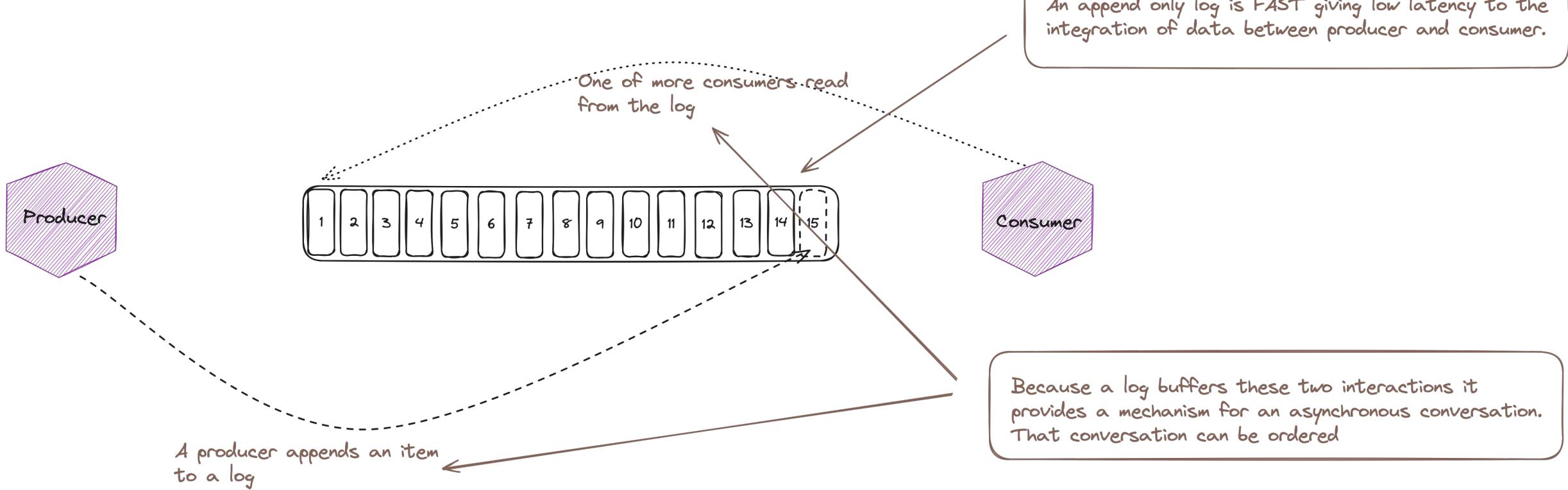


Quick Start Kafka

A Hitchhikers Guide to Kafka

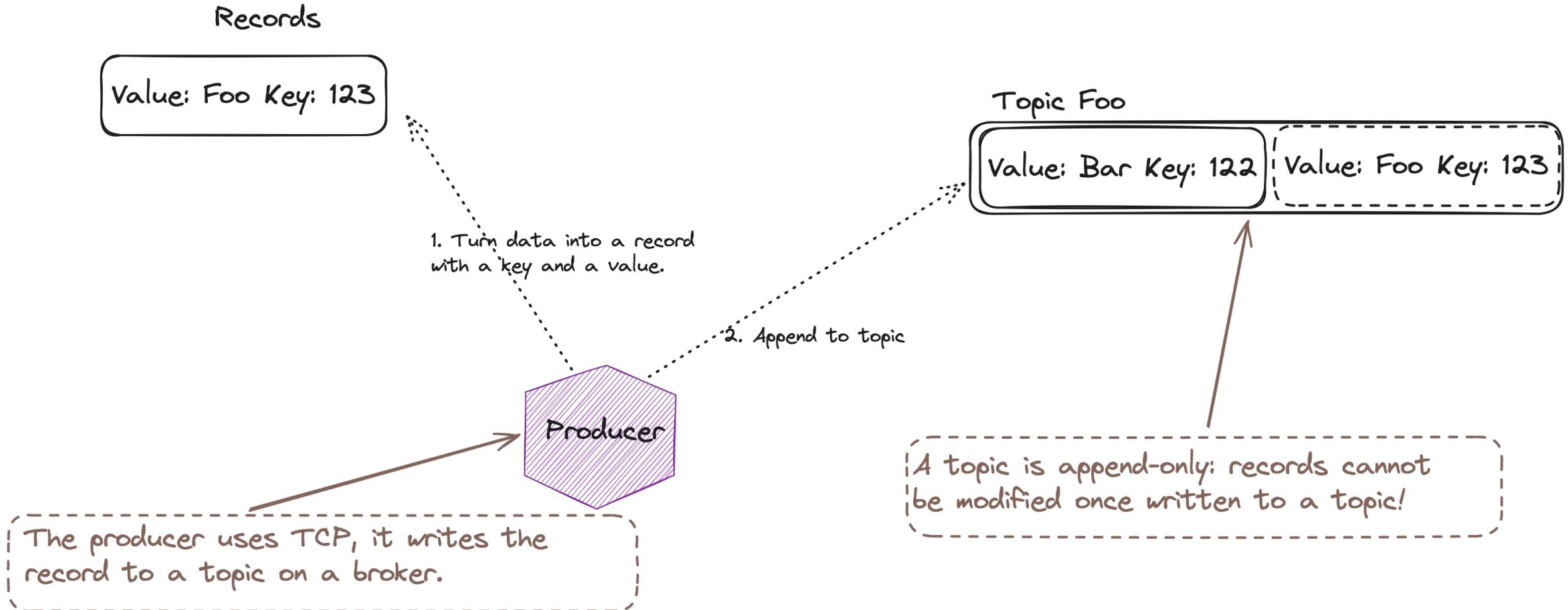
Kafka was created at LinkedIn by Jay Kreps, Neha Narkhede, Jun Rao in 2011.
it became an Apache OSS project in October 2012

Original use case was data engineering, capturing data into LinkedIn's Data Warehouse from diverse sources so that it could be used for later analysis.

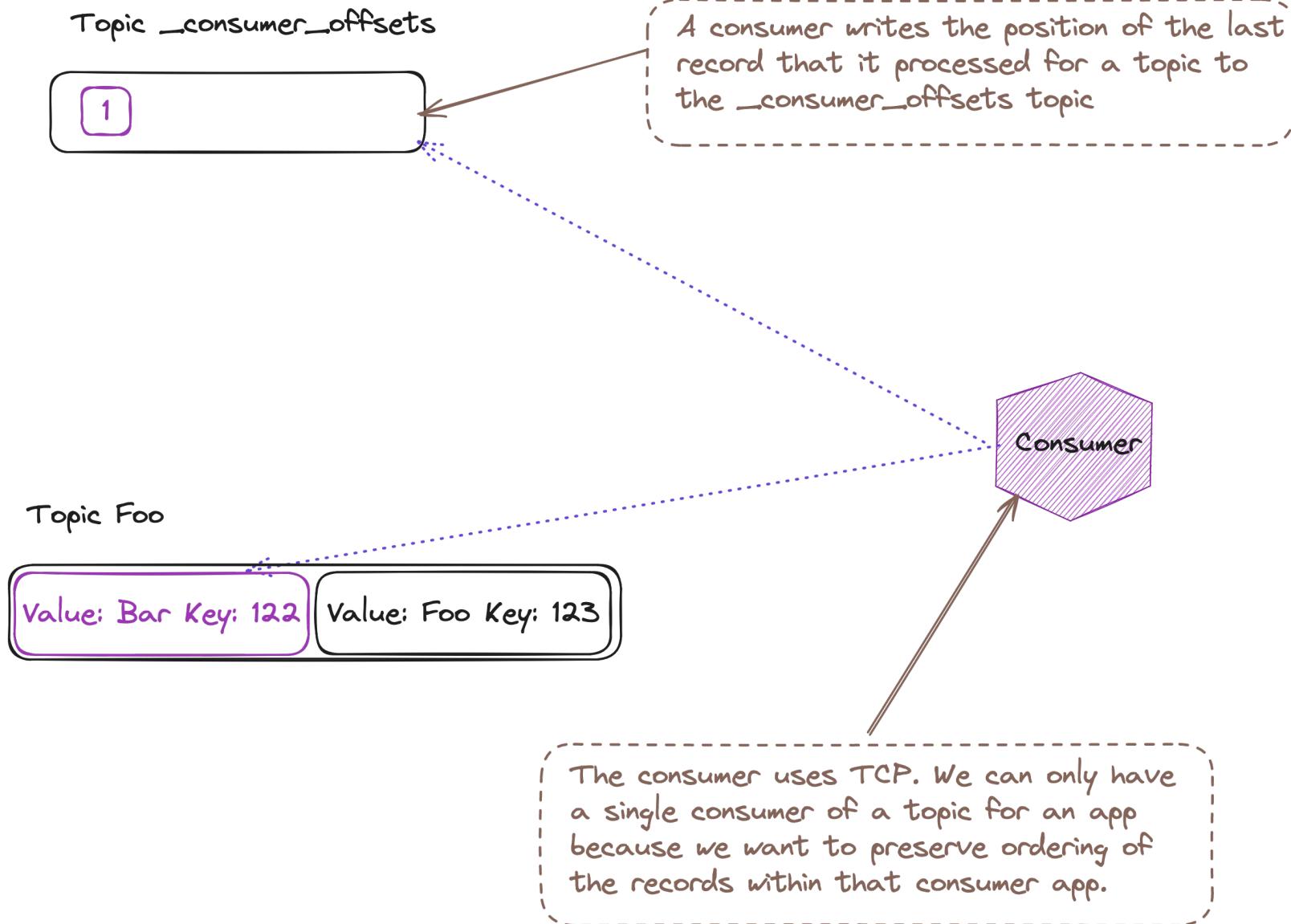


<https://engineering.linkedin.com/distributed-systems/log-what-every-software-engineer-should-know-about-real-time-datas-unifying>

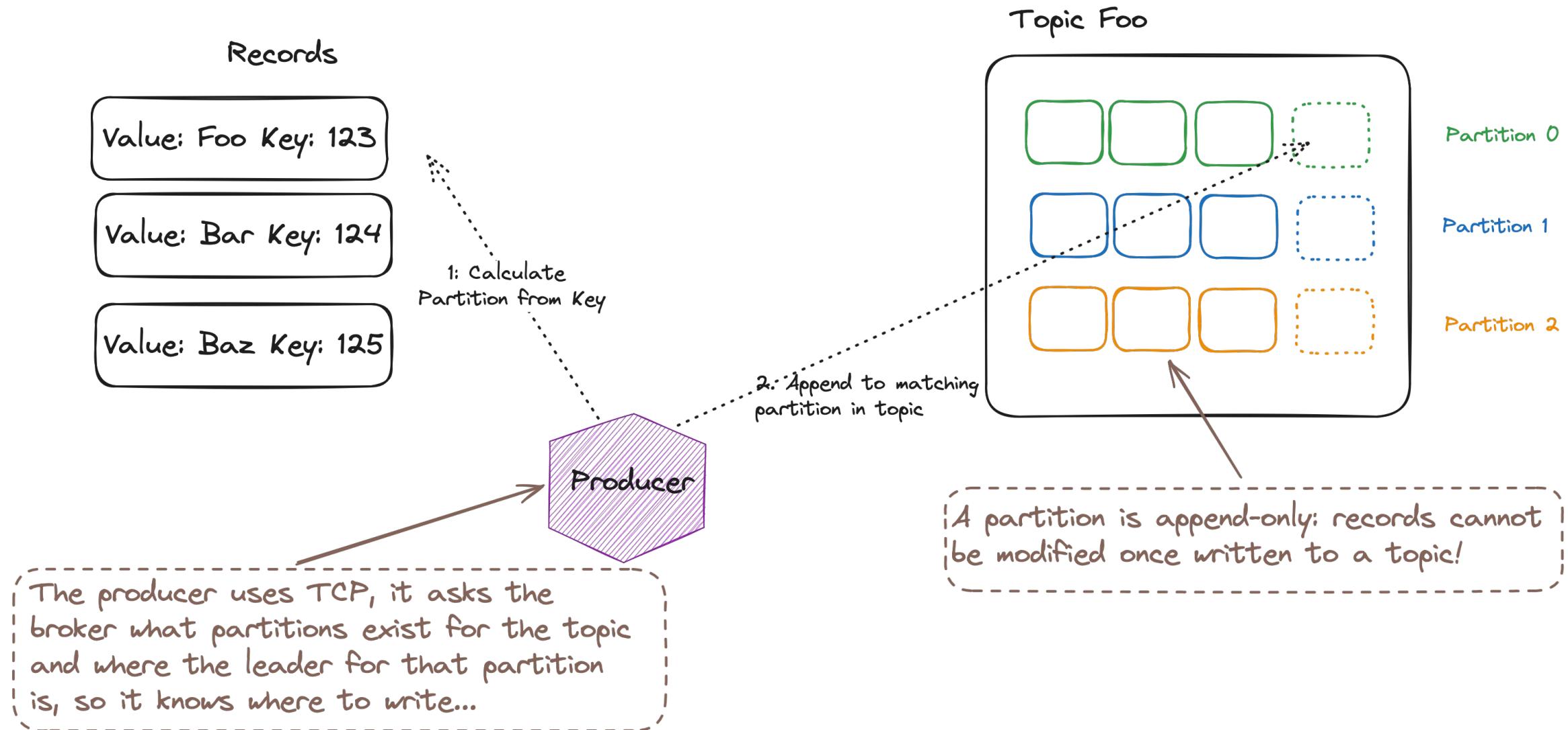
Kafka Topics



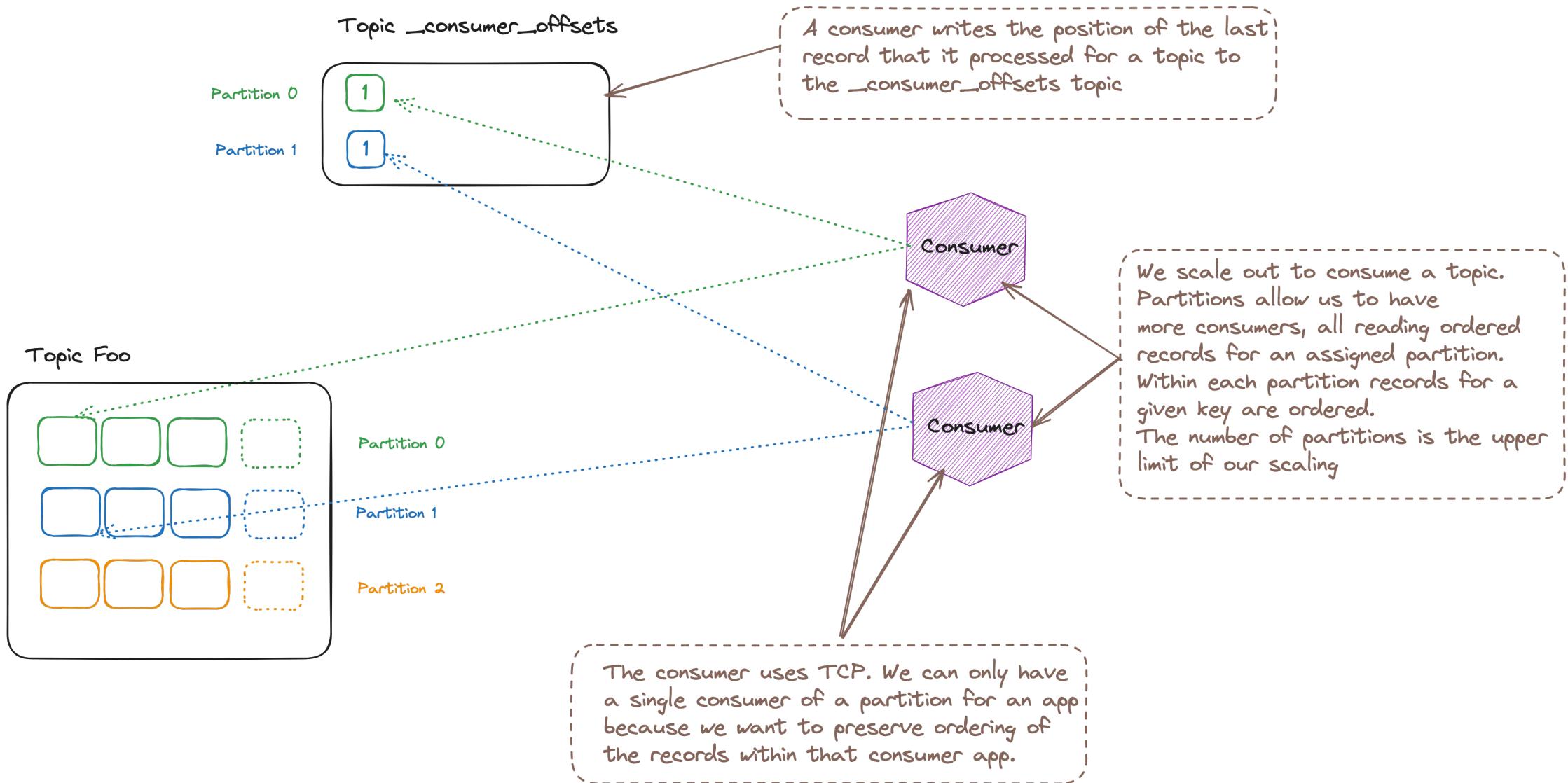
Kafka Topics



Kafka Topics

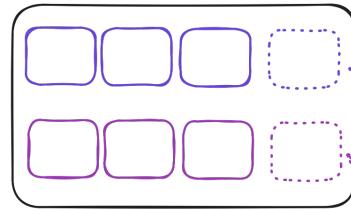


Kafka Topics



Kafka Topics

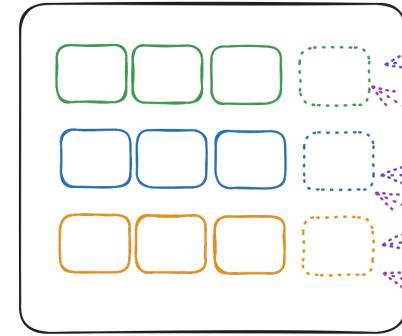
Topic _consumer_offsets



Partition 0

Partition 1

Topic Foo



Partition 0

Partition 1

Partition 2

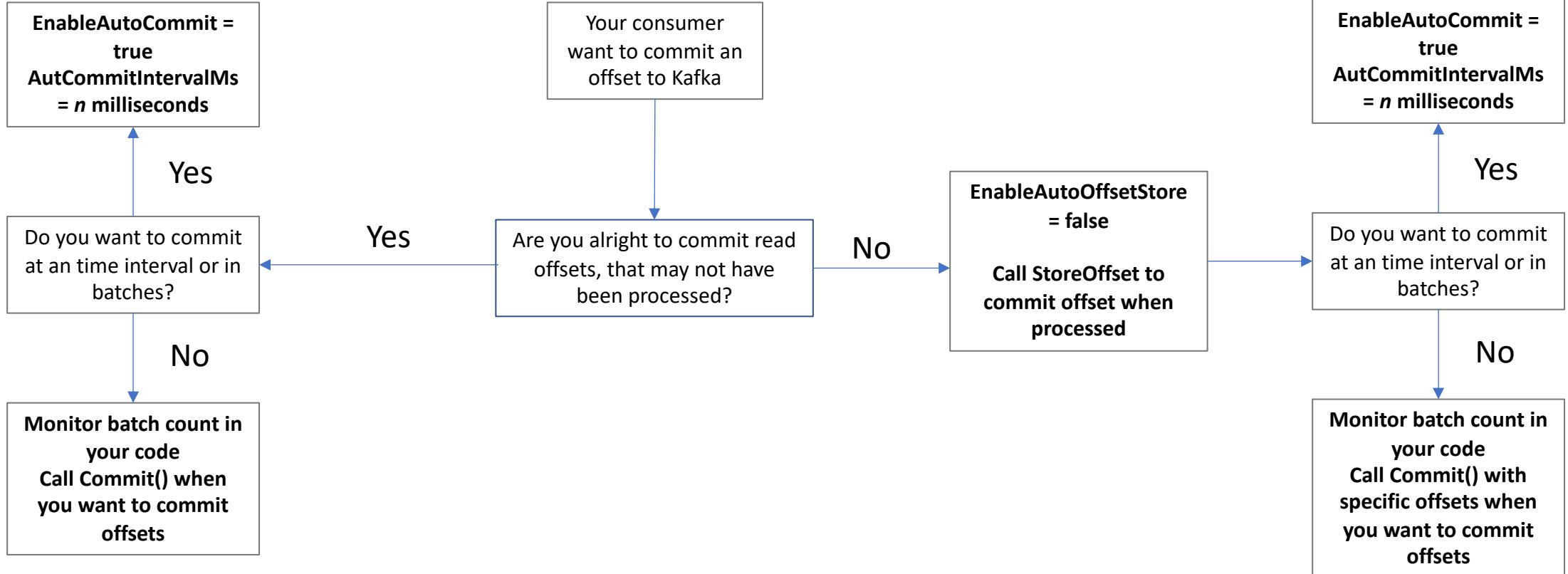
Group: Alpha



Group: Beta



The consumer belongs to a group. Within a group (representing an application) only one consumer can read from a partition on a topic.



Running Kafka

```
1 kafka0:  
2   image: confluentinc/cp-kafka:7.2.1.arm64  
3   hostname: kafka0  
4   container_name: kafka0  
5   ports:  
6     - 9092:9092  
7     - 9997:9997  
8   environment:  
9     KAFKA_BROKER_ID: 1  
10    KAFKA_AUTO_CREATE_TOPICS_ENABLE: "false"  
11    KAFKA_LISTENER_SECURITY_PROTOCOL_MAP:  
12      PLAINTEXT:PLAINTEXT,CONTROLLER:PLAINTEXT,PLAINTEXT_HOST://PLAINTEXT  
13    KAFKA_ADVERTISED_LISTENERS: PLAINTEXT://kafka0:29092,PLAINTEXT_HOST://localhost:9092  
14    KAFKA_INTER_BROKER_LISTENER_NAME: PLAINTEXT  
15    KAFKA_OFFSETS_TOPIC_REPLICATION_FACTOR: 1  
16    KAFKA_GROUP_INITIAL_REBALANCE_DELAY_MS: 0  
17    KAFKA_TRANSACTION_STATE_LOG_MIN_ISR: 1  
18    KAFKA_TRANSACTION_STATE_LOG_REPLICATION_FACTOR: 1  
19    KAFKA_PROCESS_ROLES: 'broker,controller'  
20    KAFKA_NODE_ID: 1  
21    KAFKA_CONTROLLER_QUORUM_VOTERS: '1@kafka0:29093'  
22    KAFKA_LISTENERS: 'PLAINTEXT://kafka0:29092,CONTROLLER://kafka0:29093_PLAINTEXT_HOST://0.0.0.0:9092'  
23    KAFKA_CONTROLLER_LISTENER_NAMES: 'CONTROLLER'  
24    KAFKA_LOG_DIRS: '/tmp/kraft-combined-logs'  
25    KAFKA_JMX_PORT: 9997  
26    KAFKA_JMX_OPTS: -Dcom.sun.management.jmxremote  
     -Dcom.sun.management.jmxremote.authenticate=false -Dcom.sun.management.jmxremote.ssl=false  
     -Djava.rmi.server.hostname=kafka0 -Dcom.sun.management.jmxremote.rmi.port=9997  
27    volumes:  
28      - ./scripts/update_run.sh:/tmp/update_run.sh  
29    command: "bash -c 'if [ ! -f /tmp/update_run.sh ]; then echo \"ERROR: Did you forget to copy the update_run.sh file that came with this docker-compose.yml file?\" && exit 1 ; else /tmp/update_run.sh && /etc/confluent/docker/run ; fi'"
```

If a topic does not exist when a producer writes, what should Kafka do?

The advertised listeners are the endpoints that clients should use to communicate with broker

The roles that this node plays

The listeners are the endpoints that the broker listens on

We need to run a script to run without zookeeper