ORACLE

OCI Streaming Service

Level 100

Rohit Rahi

Oracle Cloud Infrastructure

September, 2019

Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Objectives

After completing this lesson, you should be able to:

- Describe the OCI Streaming Service

- Understand Streaming concepts

- Create a stream and publish and consume messages

OCI Streaming Service

- OCI Streaming service provides a fully managed, scalable, durable storage option for continuous, high-volume streams of data that you can consume and process in real-time

- Use cases
    - Log and Event data collection
    - Web/Mobile activity data ingestion
    - IoT Data streaming for processing and alerts
    - Messaging: use streaming to decouple components of large systems

- Oracle managed service with REST APIs (Create, Put, Get, Delete)

- Integrated Monitoring

- 99.95% SLA

Key Concepts

- **Message**: a 64-bit encoded record or array of bytes (think of it as a row or record in a database)

- **Key**: an identifier to group related messages (also a byte array)

- **Stream**: an append-only log of messages (think of it as an ever growing dataset, where new records keep arriving)

- **Topic**: messages are categorized into topics (think of a topic as a database table)

- **Partitions**: Topics are additionally broken down into a number of partitions

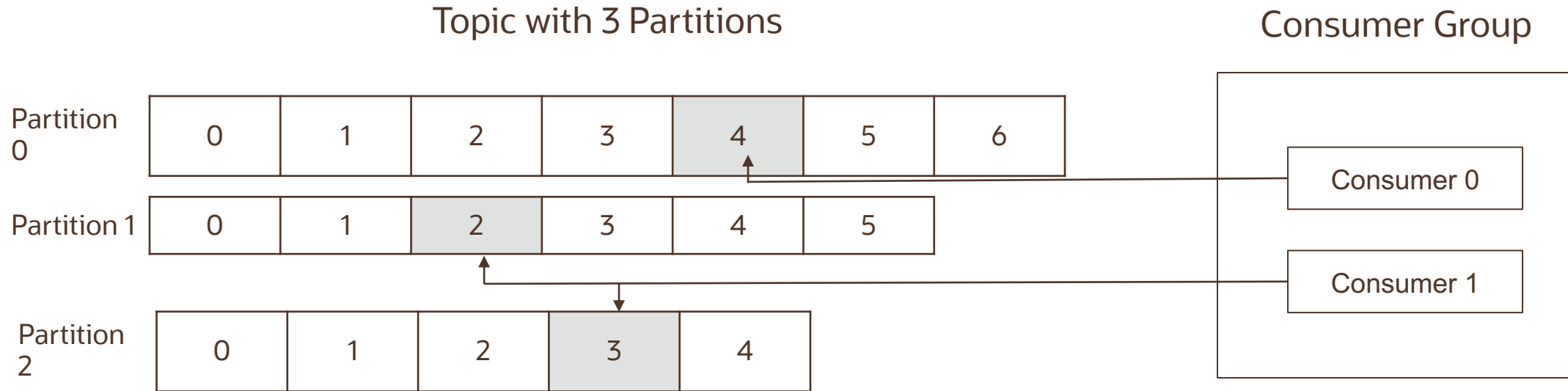| Partition 0	0	1	2	3	4	5	6
Partition 1	0	1	2	3	4	5	
Partition 2	0	1	2	3	4		

Message Write
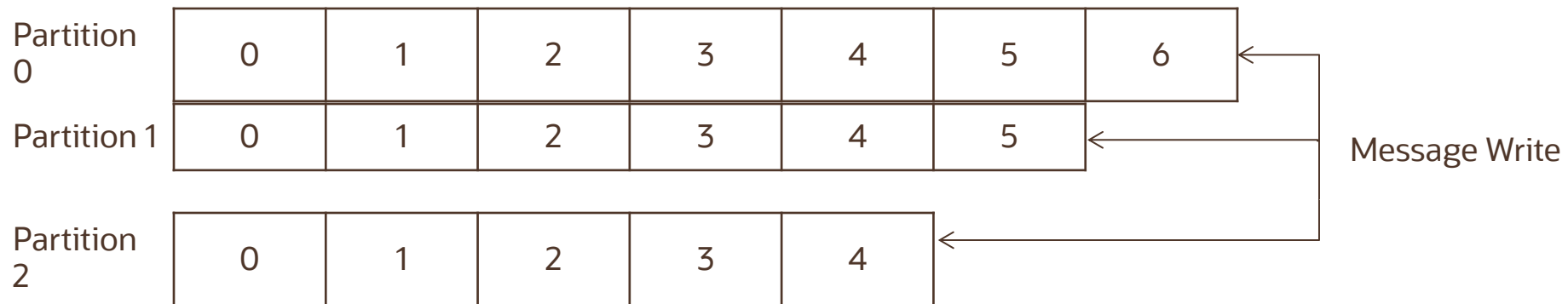
- Each partition can be hosted on a different server (different ADs, within a region), which means that a single topic can be scaled horizontally across multiple servers to provide performance far beyond the ability of a single server

Key Concepts

- **Producer**: create new messages. In general, a message is written to a specific topic
- **Consumer**: read messages
  - In general, consumers subscribes to one or more topics and reads the messages in the order in which they were produced
  - Consumer keeps track of which messages it has already consumed by keeping track of message offset
  - Offset: The location of a message within a stream/partition
  - By storing the offset of the last consumed message, a consumer can stop and restart
- **Consumer Group**: A consumer group (group) is a set of consumers that coordinate to consume messages from all of the partitions in a stream
  - Each partition is only consumed by one member of the group

Consumer Group

Topic with 3 Partitions

Consumer Group

Partition 0	0	1	2	3	4	5	6

Partition 1	0	1	2	3	4	5

Partition 2	0	1	2	3	4

Consumer 0

Consumer 1

Rebalancing – a new consumer joins or a consumer leaves
1. CG stores that offset
2. Rebalancing

Create Stream

COMPARTMENT

OCI-Demo

ociobenablement (root)/OCI-Demo

STREAM NAME

My Stream

Stream Settings

RETENTION (IN HOURS) ⓘ

24

Retention in hours must be between 24 - 168. Default value is 24 hours.

NUMBER OF PARTITIONS ⓘ

**Estimate #
Partitions**

3

The maximum number of partitions is based on the limits of your tenancy

TOTAL WRITE RATE (IN MB PER SECOND) *READ-ONLY*

3 MB/s

TOTAL READ RATE (IN MB PER SECOND) *READ-ONLY*

6 MB/s

TAGS

Tagging is a metadata system that allows you to organize and track resources within your tenancy. Tags are composed of keys and values that can be attached to resources.

Learn more about tagging

TAG NAMESPACE

None (add a free-form tag)

TAG KEY

VALUE

My Stream

Produce Test Message | Move Resource | Add Tags | Delete

Stream Information | Tags

Stream Information

Stream Name: My Stream

OCID: ...6yjh6a Show Copy

Compartment: us_training (root)/TL-SPL-55

Messages Endpoint: https://api.cell-1.us-ashburn-1.streaming.oci.oraclecloud.com

REST API

Settings

Number of partitions: 1

Retention: 24 hours

Read Throughput: 2 MB/s

Write Throughput: 1 MB/s

ACTIVE

Resources

Recent Messages

Produce Monitoring Charts

Consume Monitoring Charts

Recent Messages

Refresh

Key	Value	Offset	Partition	Created
		Refresh to retrieve Recent Messages		

Publishing and consuming messages

- A message is composed of a key (can be null) and a value. The message is published to a partition
- In case of >1 partitions, the partition where the message gets published depends on the message's key
- Messages produced with the same key (not null) will get written to the same partition
- Consuming messages requires use of a cursor
  - TRIM_HORIZON - use if you want to consume an entire stream
  - AT_OFFSET - start consuming at a specified offset
  - AFTER_OFFSET - start consuming after the given offset
  - AT_TIME - start consuming from a given time
  - LATEST - start consuming messages that were published after you created the cursor

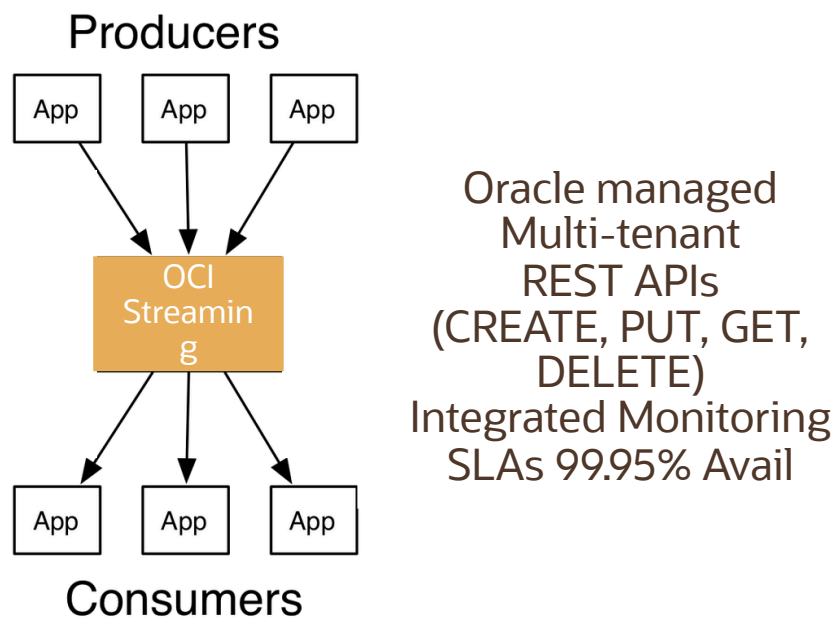 - Ordering guarantee – hash the key

Streaming Demo

Design Considerations

- OCI streaming service supports message retention of up to a maximum 7 days

- Maximum message size supported is 1 MB

- Each partition can handle up to 1000 Emit API call per second and 5 Read API call per second

- Each partition can support up a maximum total data write rate of 1MB per second and a read rate of 2MB per second

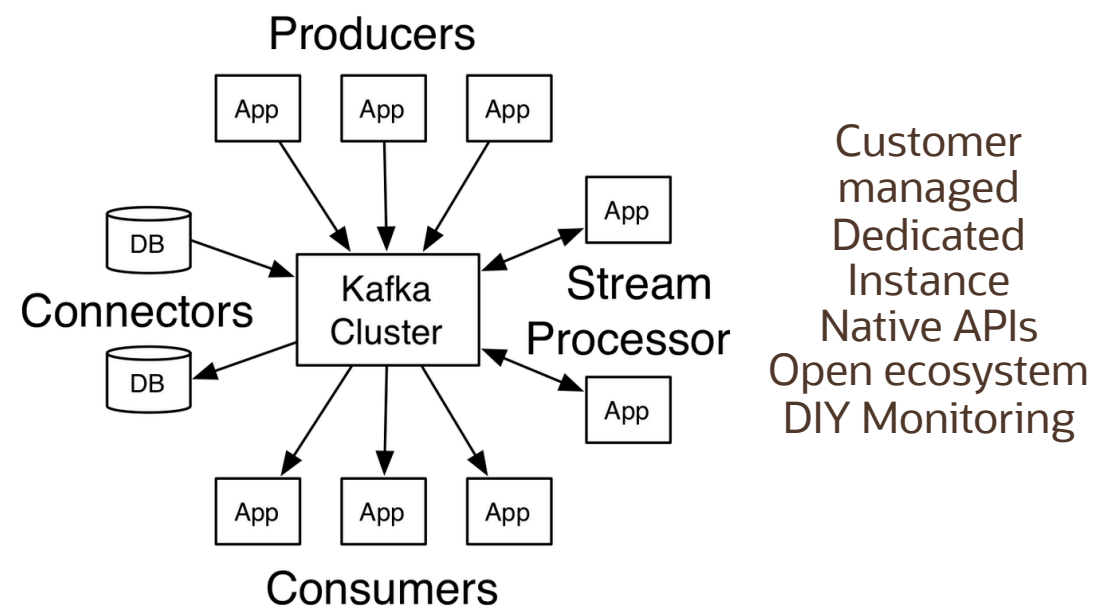- Each tenancy has a limit of 5 partitions (you can request more)

OCI Streaming Vs Apache Kafka

Apache Kafka is an open source pub/sub system;

OCI Streaming Service

Producers

Oracle managed
Multi-tenant
REST APIs
(CREATE, PUT, GET,
DELETE)
Integrated Monitoring
SLAs 99.95% Avail

Consumers

OCI Marketplace Kafka

Producers

Connectors

Stream
Processor

Consumers

Customer
managed
Dedicated
Instance
Native APIs
Open ecosystem
DIY Monitoring
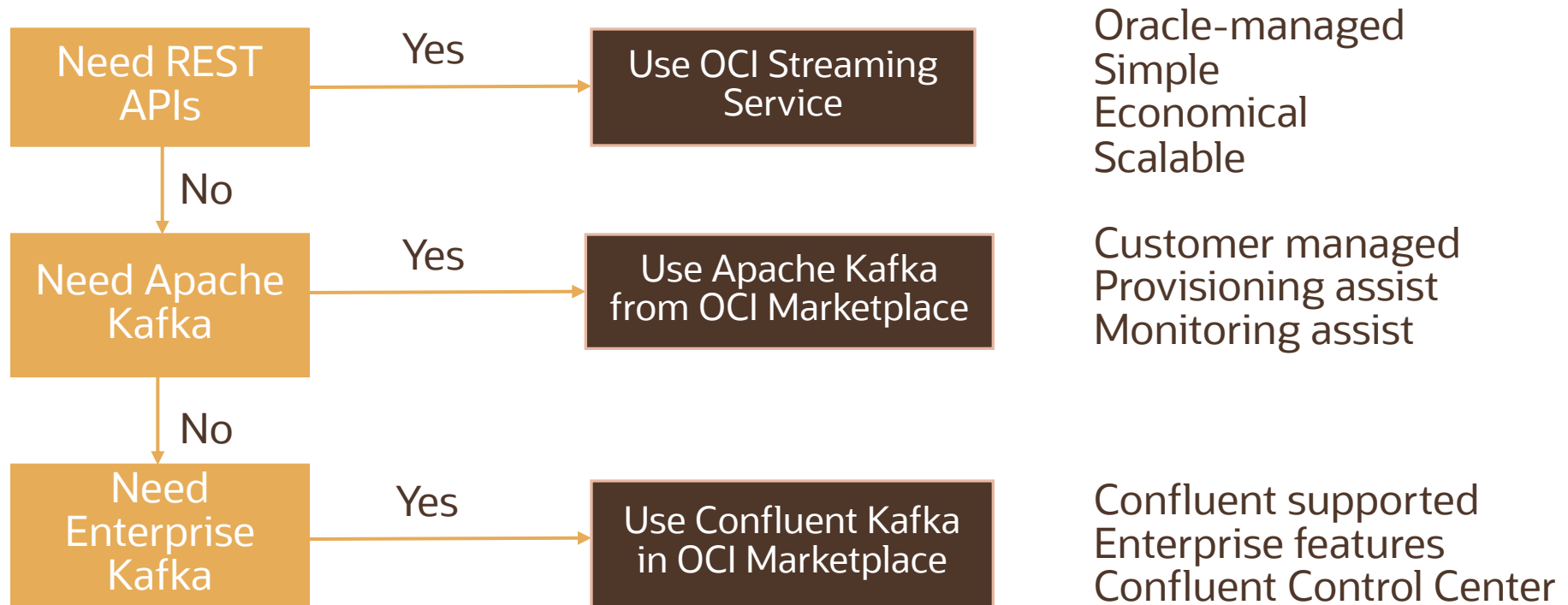
Adding Connectors, Stream Processing, Kafka compatibility in H2 2019

Migration paths from Event Hub Cloud Service

OCI Streaming or Apache Kafka

Need REST APIs	Yes →	**Use OCI Streaming Service**

Oracle-managed
Simple
Economical
Scalable

No ↓

Need Apache Kafka	Yes →	**Use Apache Kafka from OCI Marketplace**

Customer managed
Provisioning assist
Monitoring assist

No ↓

Need Enterprise Kafka	Yes →	**Use Confluent Kafka in OCI Marketplace**

Confluent supported
Enterprise features
Confluent Control Center

Pricing

Pricing model – differentiator – provisioned model (Kinesis)
With AWS, you need to maintain your own offset
Examples of pricing

ORACLE

Oracle Cloud always free tier:
oracle.com/cloud/free/

OCI training and certification:
cloud.oracle.com/en_US/iaas/training
cloud.oracle.com/en_US/iaas/training/certification
education.oracle.com/oracle-certification-path/pFamily_647

OCI hands-on labs:
ocitraining.qloudable.com/provider/oracle

Oracle learning library videos on YouTube:
youtube.com/user/OracleLearning