

RabbitMQ, Kafka, or Pulsar?



Julien Jakubowski

Developer Advocate

<https://bit.ly/48T4Zcv>



StreamNative

- ✓ Founded by the original creators of Apache Pulsar
- ✓ Large contributor to open-source: Apache Pulsar and the ecosystem
- ✓ StreamNative Cloud:
 - Partner with software teams worldwide to accelerate time to production and manage Pulsar at scale.
 - Fully Managed Pulsar service with enterprise features
 - on StreamNative or customer infrastructure





Apache Pulsar is a **Cloud-Native Messaging** and **Data-Streaming** Platform.

Docker/K8S

Elasticity: adapting to workload changes

Data streaming

What is data streaming?

Continuously transmitting & processing an unbounded stream of small data records/events

Processed in order

Events may be:

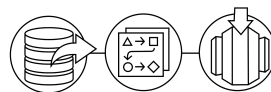
- persisted
- transformed
- aggregated
- replayed

Massive amounts of data, throughput

⇒ Horizontal scalability



What is it for?



Data ingestion



Real-time analytics

Streaming platforms examples



amazon
KINESIS



Message queuing

Goals

Assign time-consuming tasks among a group of *workers*

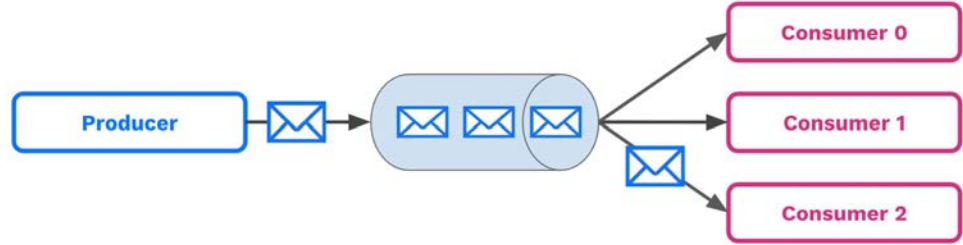
Without impacting the producer:
decoupling

Message Queueing

Each task to be performed is a message placed in a queue

Distribute the tasks among several workers

No strict order required



Essential features

Horizontal scalability & elasticity: add/remove workers dynamically

Do not block the queue when a consumer is busy or fails to consume one message

Redeliver 'failed' messages to retry later

Route undeliverable messages to a dead letter queue

 RabbitMQ

 amazon
SQS

 PULSAR

Message queueing vs data streaming

Data streaming requirements

Events to be processed in order

Ingestion of large amounts of data

Data retention

Handle catch-up reads



Message queueing requirements

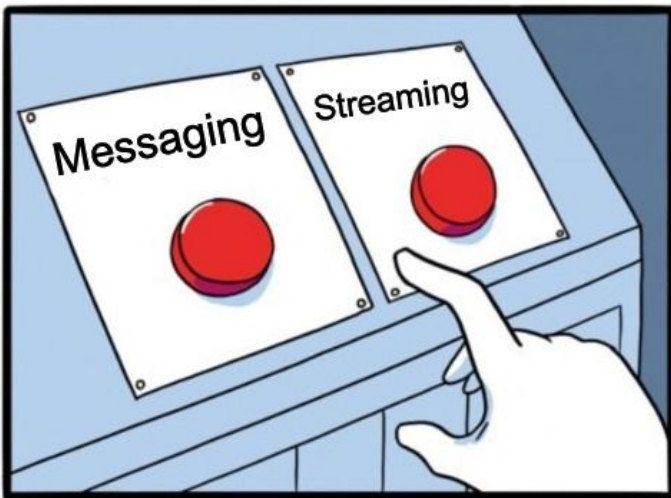
Decoupling tasks executions

Add/remove consumers dynamically

Do not block the queue when a consumer is busy or fails to consume one message

Redeliver 'failed' messages





MESSAGING

**EVENT
STREAMING**

BOTH

**APACHE
PULSAR**



PULSAR DOES BOTH

JAFE-CLARK.TUMBLR

Apache Pulsar key features

1. Unified messaging & streaming platform



2. Three levels of scalability & elasticity

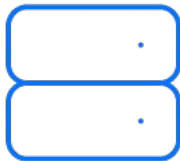


Three levels of scalability and elasticity

The scalability requirements are determined by the bottleneck.



Number of messages
remaining to be
consumed



Number of producers &
consumers



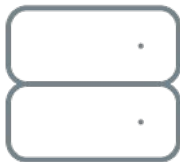
Storage capacity

Three levels of scalability and elasticity

The scalability requirements are determined by the bottleneck.



Number of messages
remaining to be
consumed



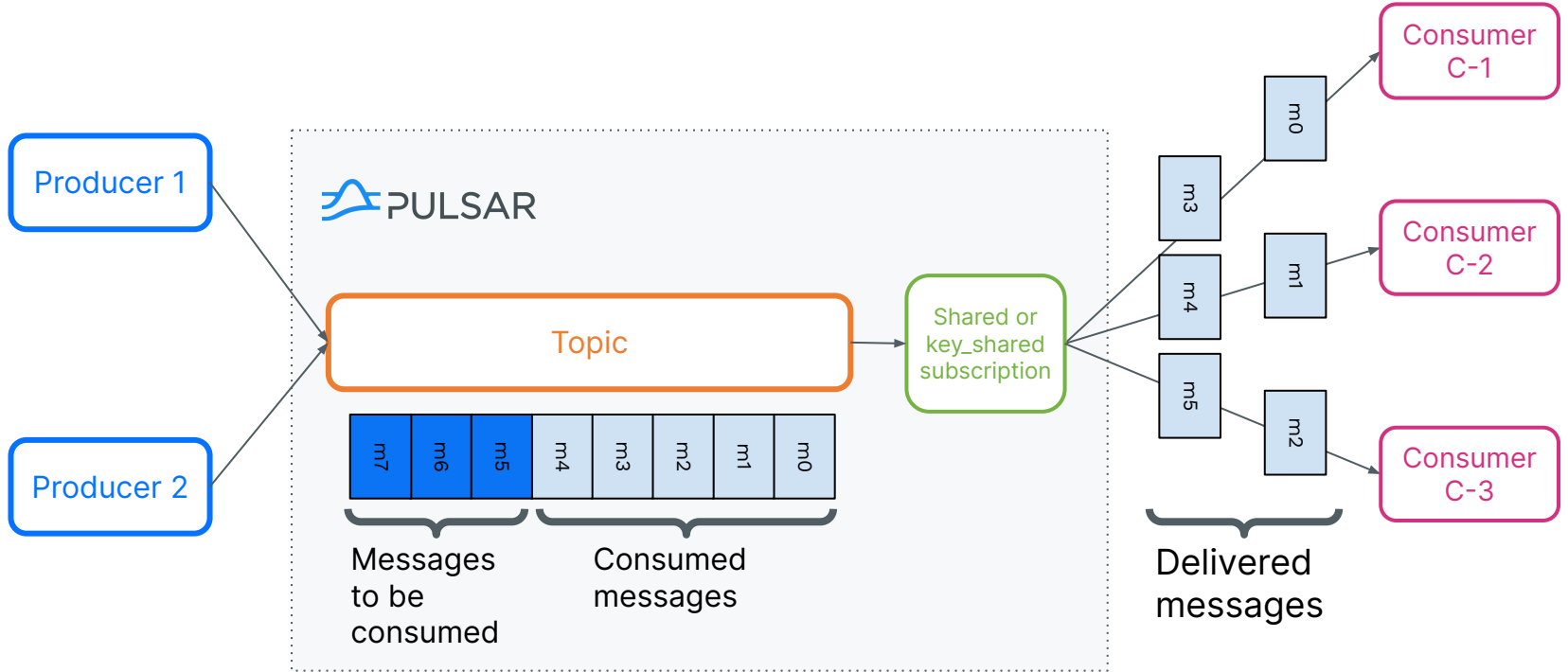
Number of producers &
consumers



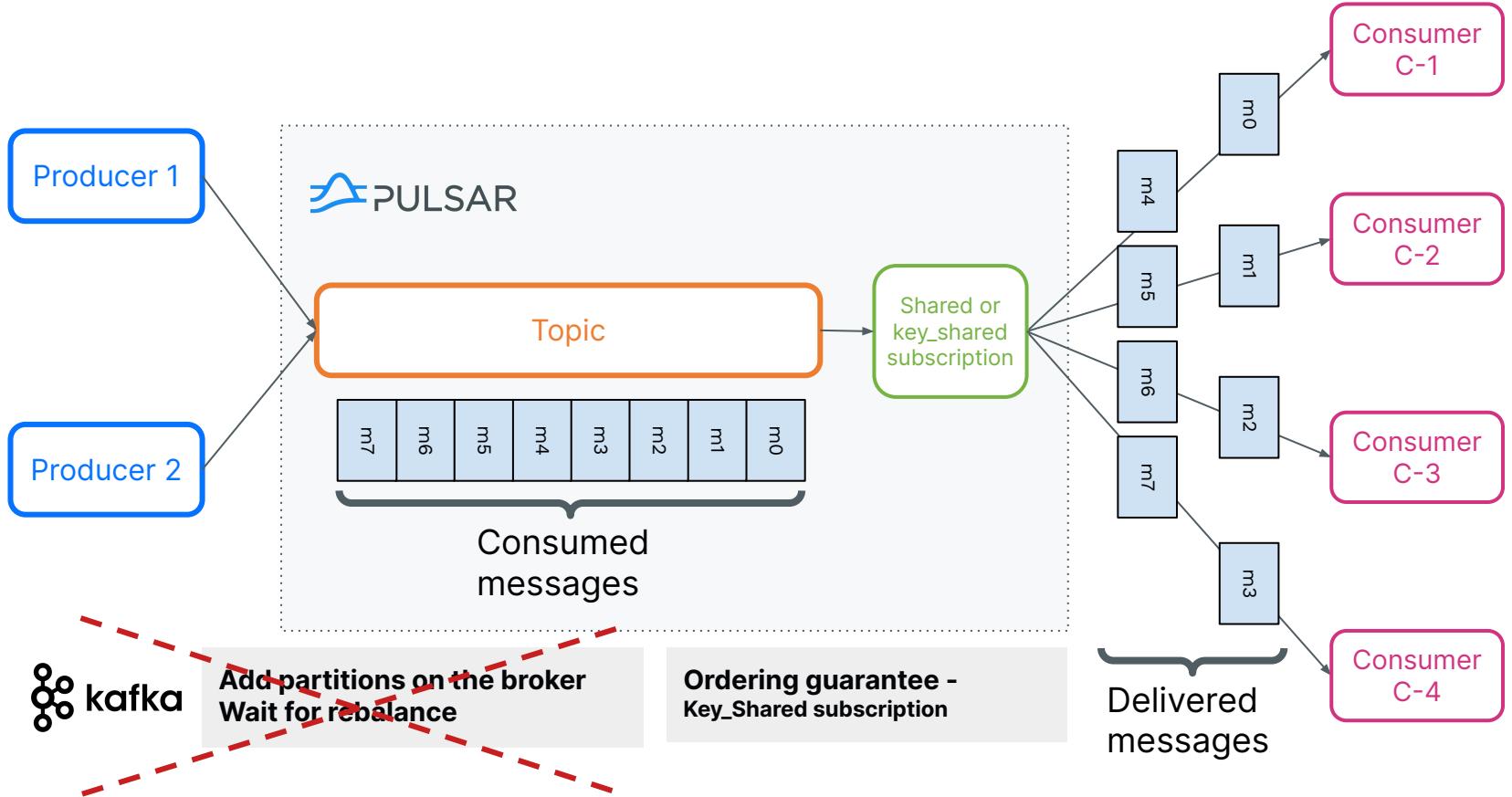
Storage capacity

most common

Adding more consumer instances



Adding more consumer instances

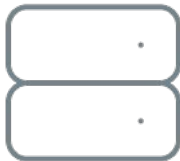


Three levels of elasticity

The scalability requirements are determined by the bottleneck.



Number of messages
remaining to be
consumed



Number of producers &
consumers



Storage capacity

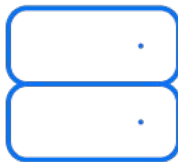
most common

Three levels of elasticity

The scalability requirements are determined by the bottleneck.



Number of messages
remaining to be
consumed



Number of producers &
consumers

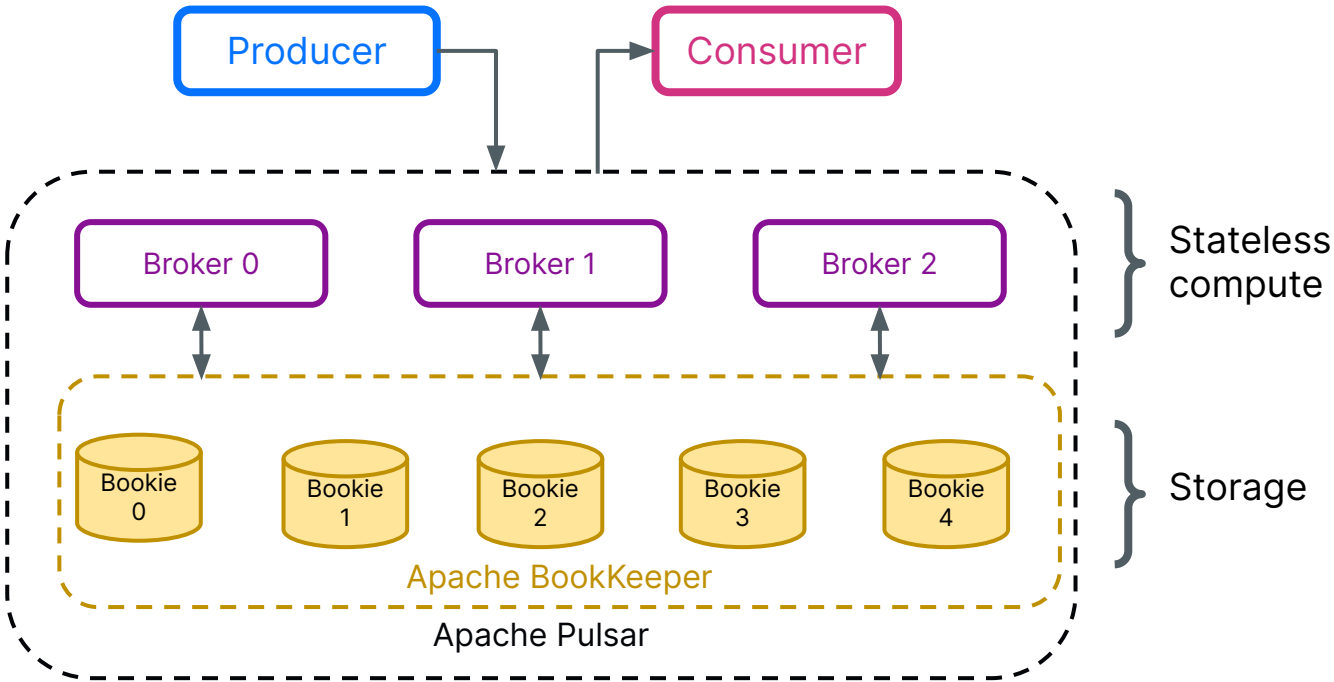


Storage capacity

most common

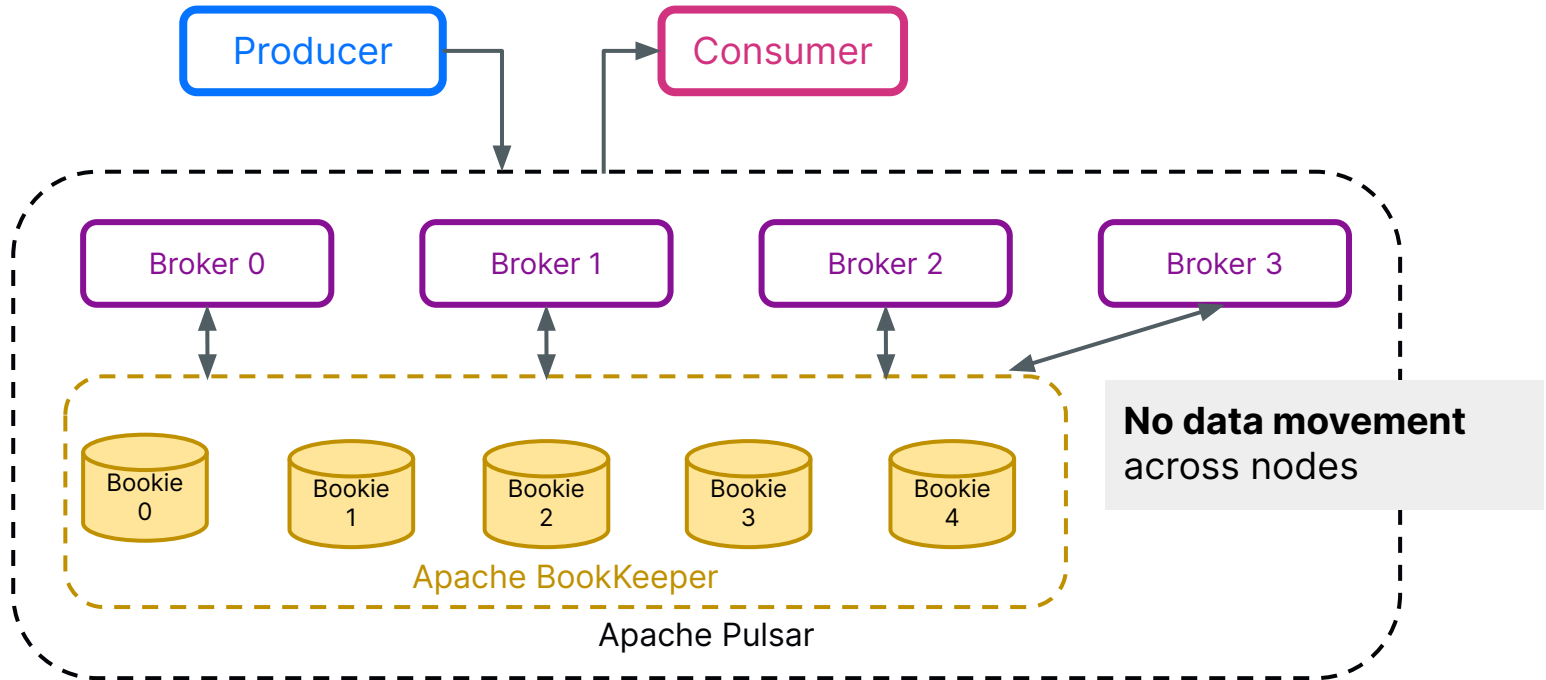


Pulsar elastic multi-layered architecture



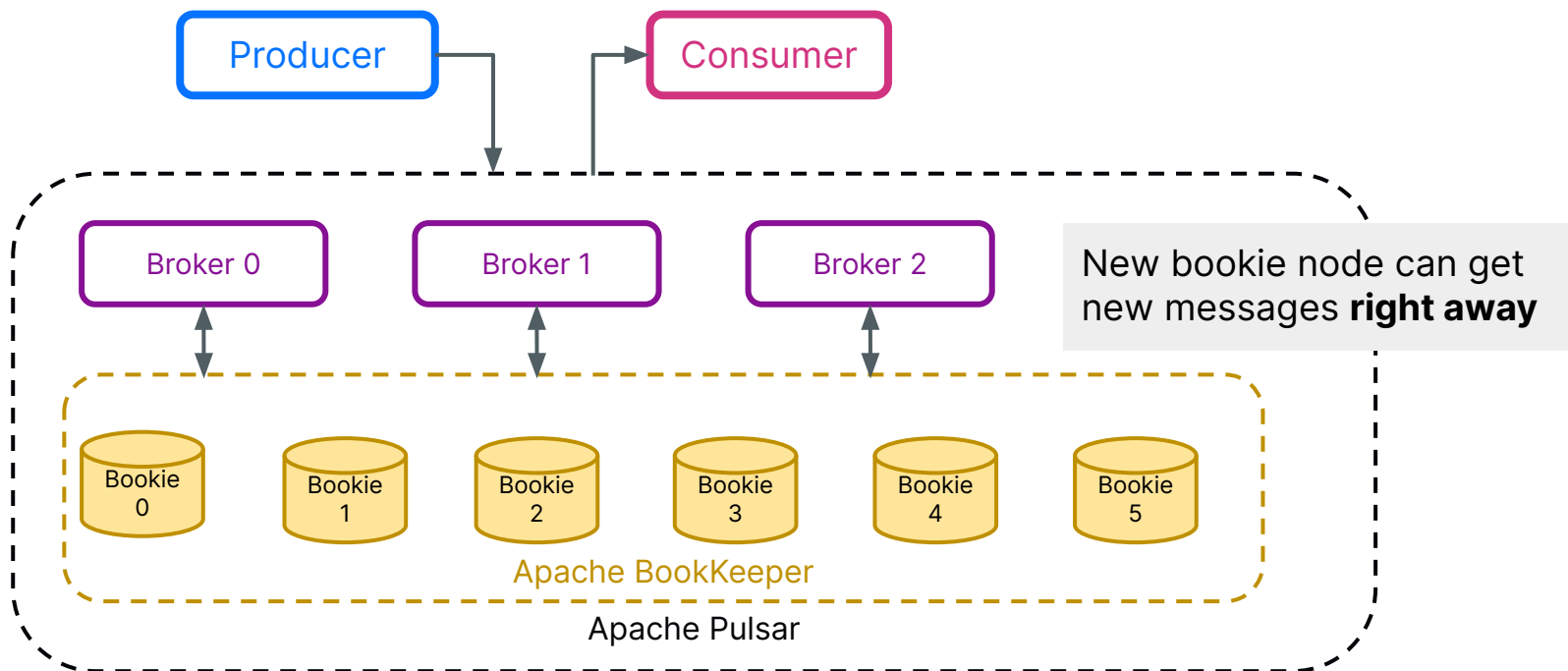


Adding broker nodes to get more processing power





Adding bookies to get more storage capacity



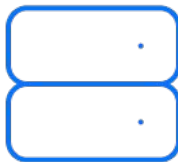
Three levels of elasticity - recap



**Number of messages
to be consumed**

Consumer nodes

No partitions to create on
the broker side



**Number of producers
& consumers**

Broker nodes

Seamless
No performance impact



Storage capacity

Bookie nodes

New nodes get new messages
right away

You can also easily downscale to save on
infrastructure costs

StreamNative Operators for Apache Pulsar



Documentation: <https://docs.streamnative.io/operator>

StreamNative offers these operators under a simple, free-to-use Community License.

Apache Pulsar key features

1. Unified messaging & streaming platform



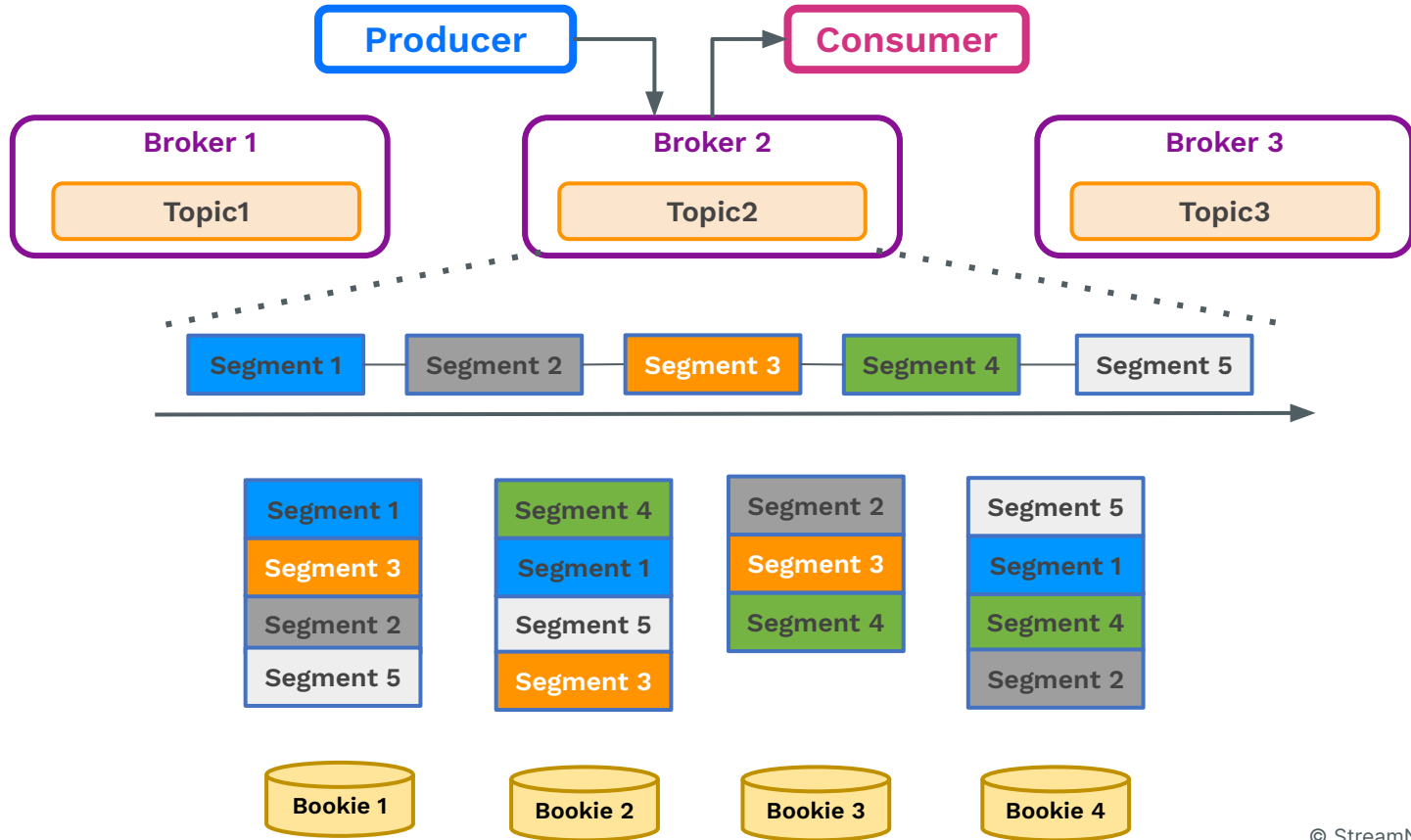
2. Three levels of scalability & elasticity



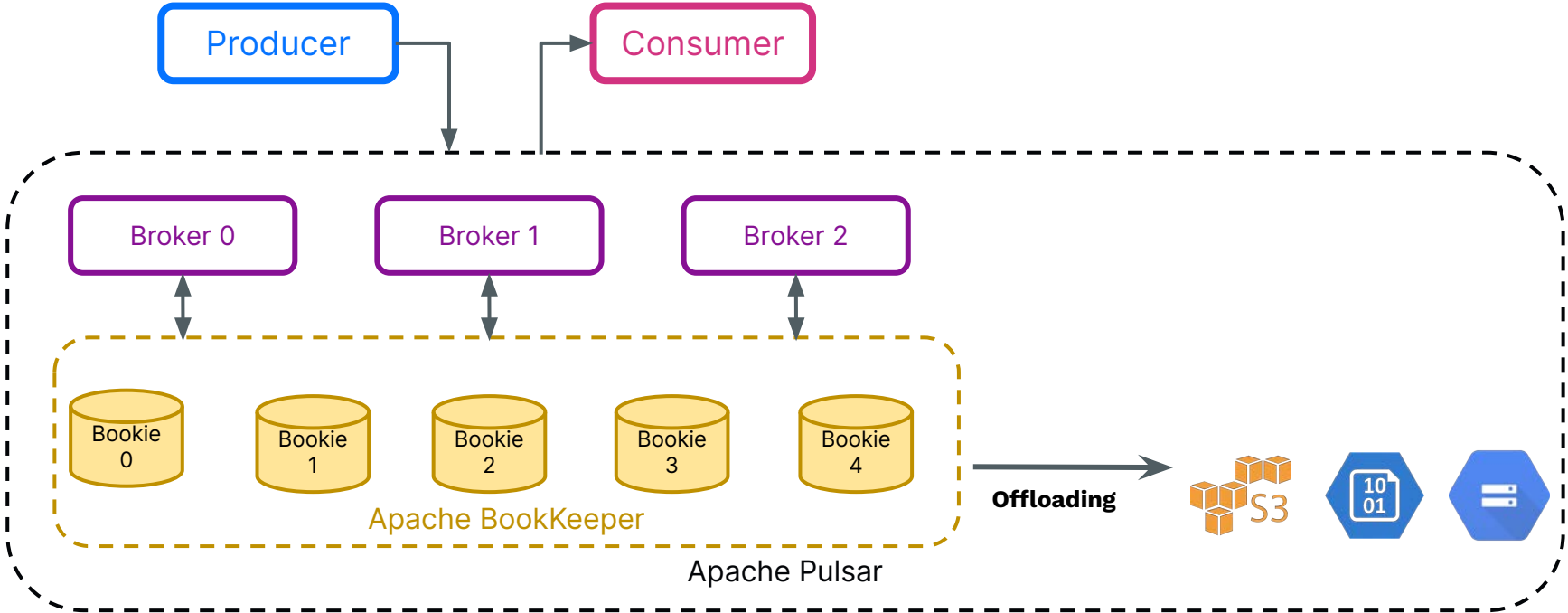
3. Durability



Durability - replication



Durability - Tiered Storage Offload



Apache Pulsar key features

1. Unified messaging & streaming platform

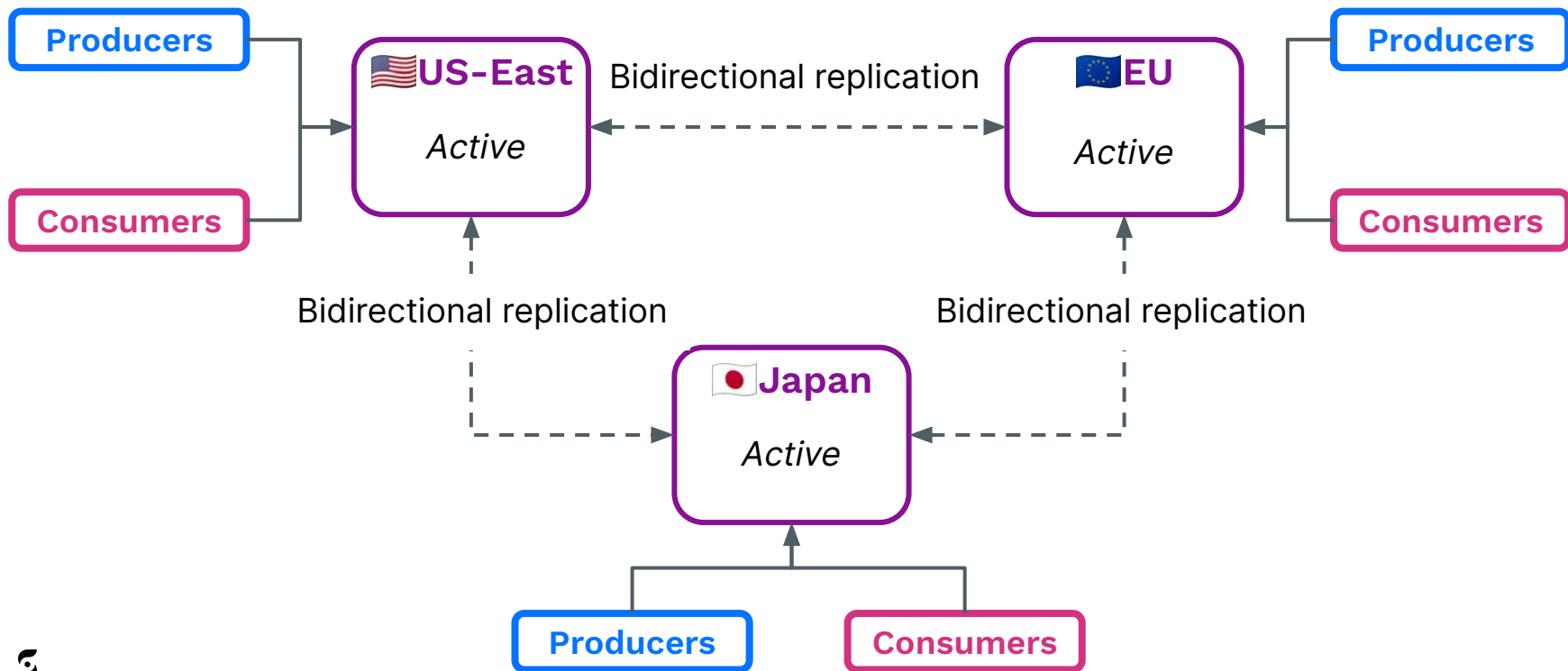
2. Three levels of scalability & elasticity

3. Durability

4. Geo-replication



Built-in geo-replication



Apache Pulsar key features

1. Unified messaging & streaming platform 

2. Three levels of scalability & elasticity 

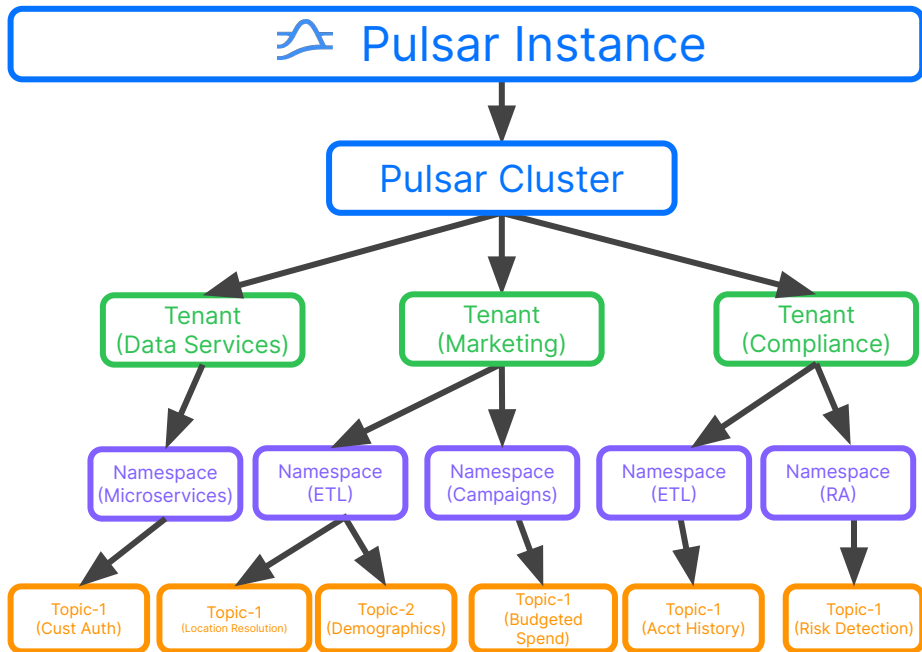
3. Durability 

4. Geo-replication 

5. Multi-tenancy



Native multi-tenancy



What:

- Share a single cluster with multiple teams
- Data isolation

Benefits:

- Easier management
- Resource sharing = costs savings 🤑

Built-in feature

Apache Pulsar key features

1. Unified messaging & streaming platform



2. Three levels of scalability & elasticity



3. Durability



4. Geo-replication



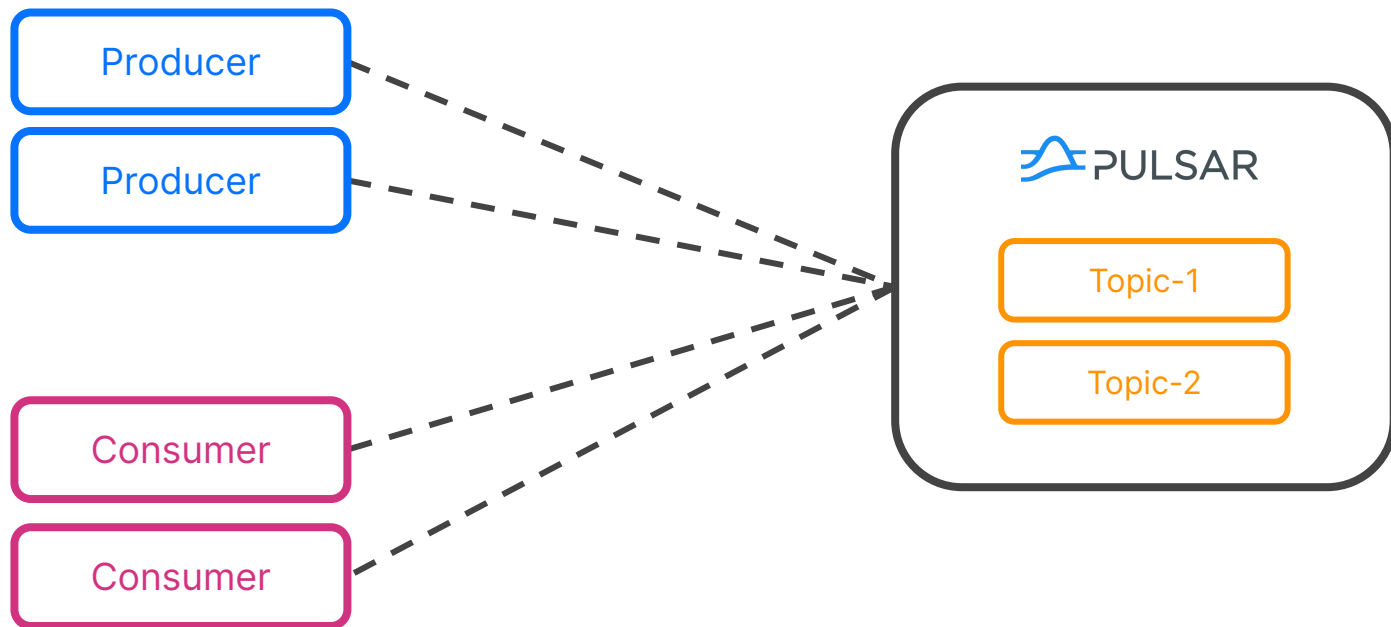
5. Multi-tenancy



6. Compatibility with your software ecosystem



Pulsar provides compatibility for seamless migration path

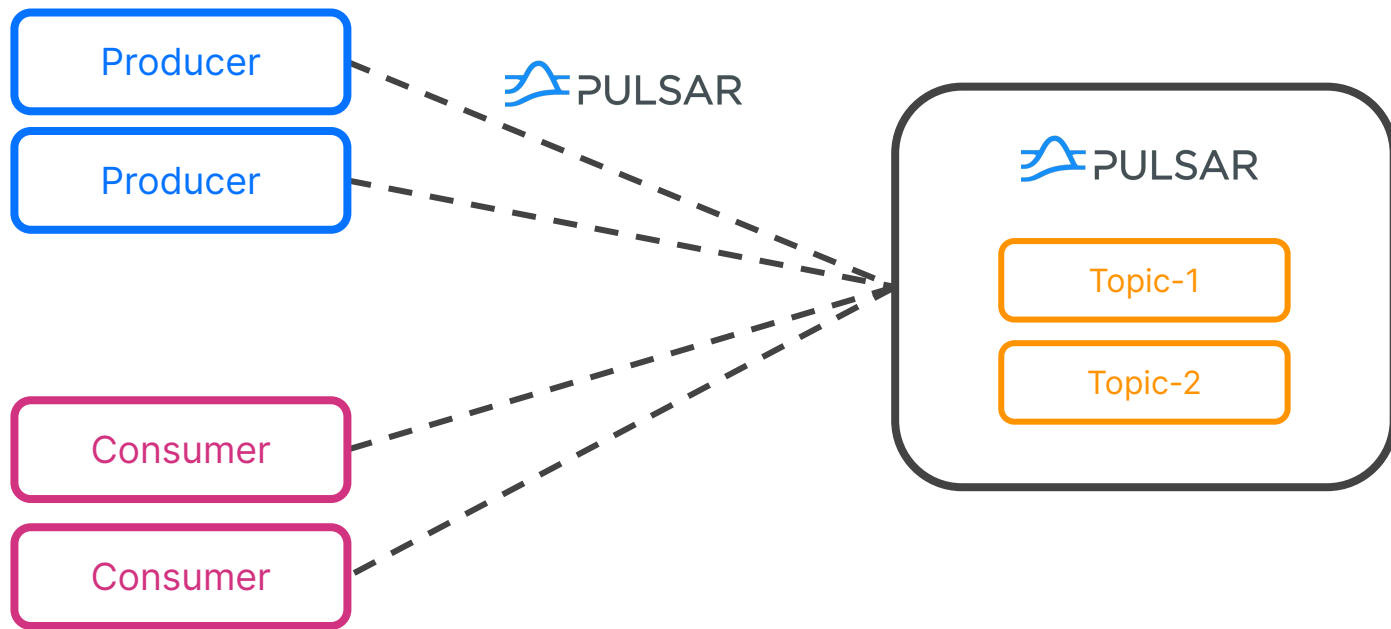


Clients

Protocol

Broker

Pulsar provides compatibility for seamless migration path

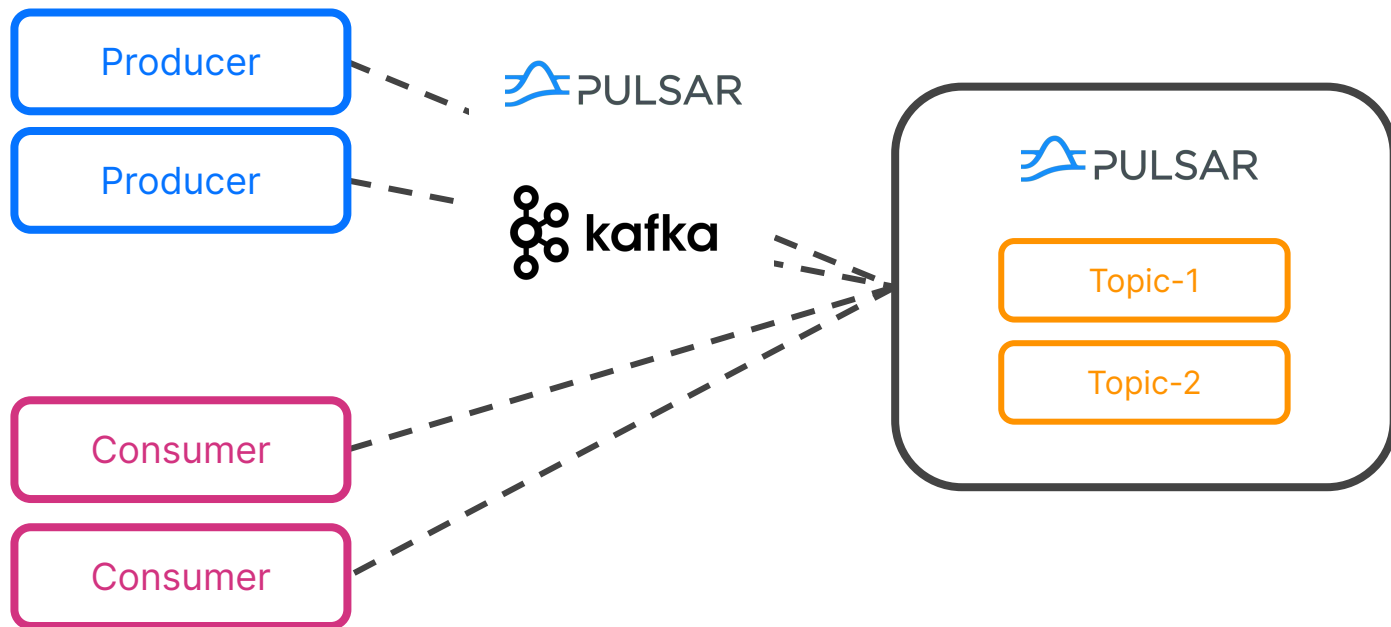


Clients

Protocol

Broker

Pulsar provides compatibility for seamless migration path

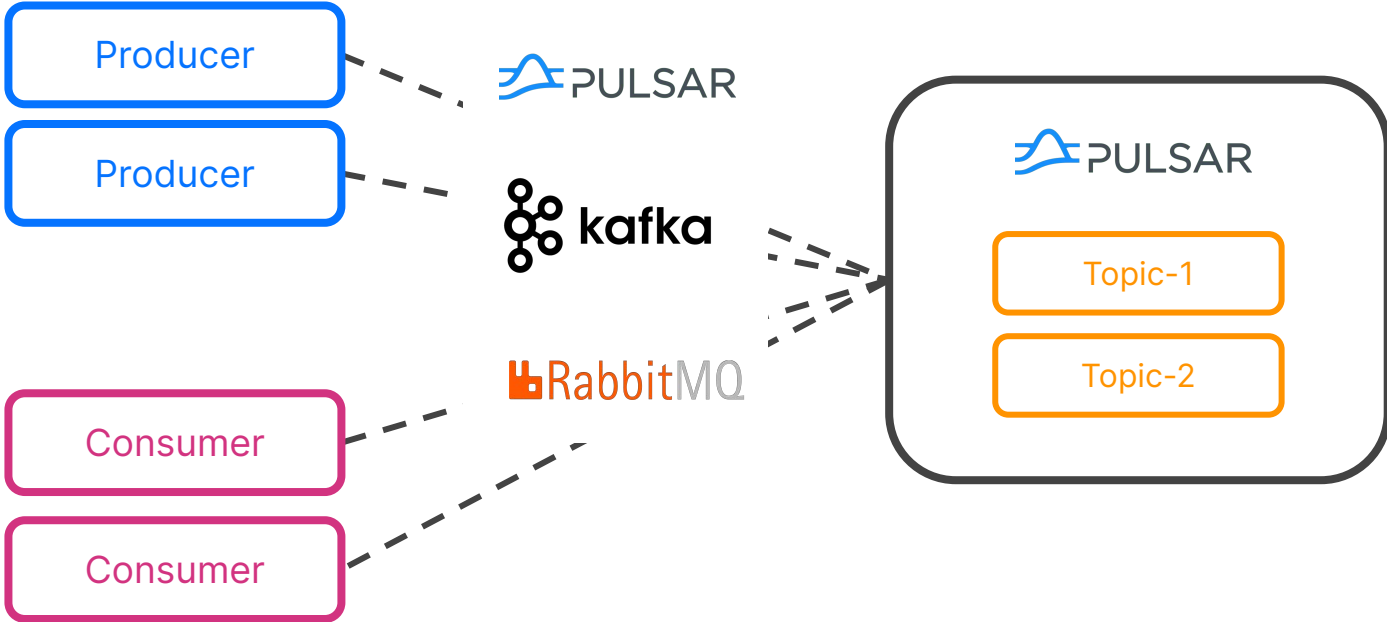


Clients

Protocol

Broker

Pulsar provides compatibility for seamless migration path

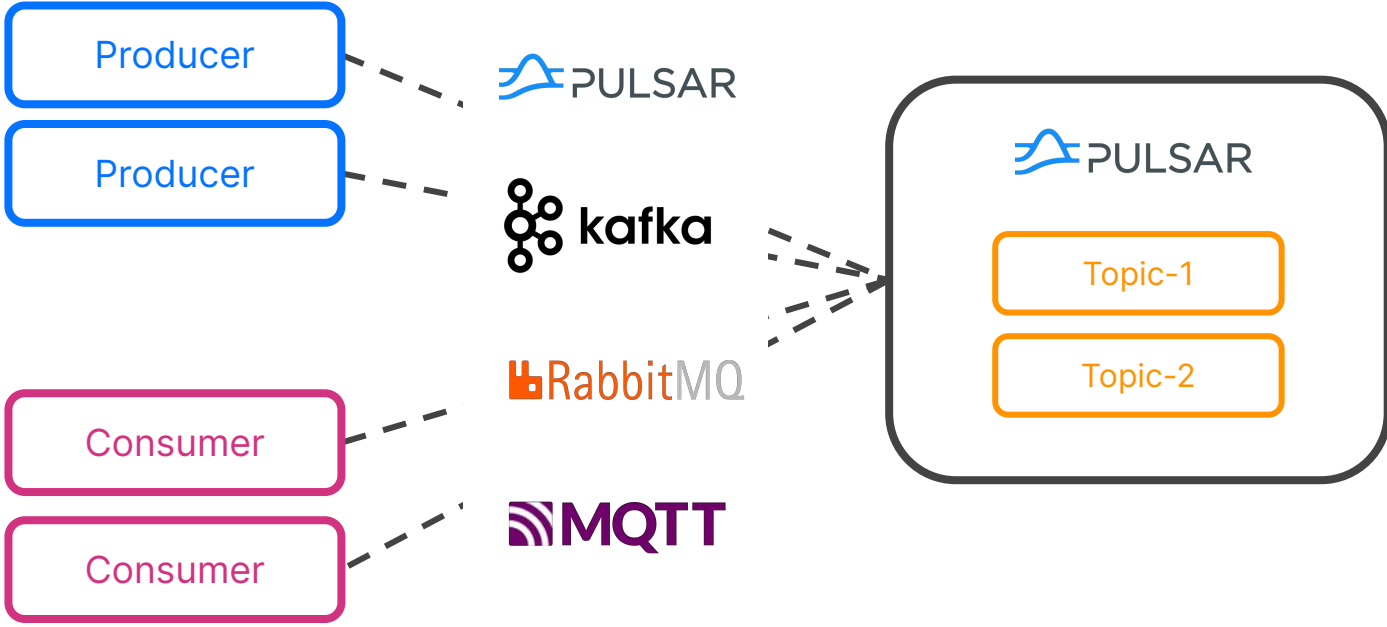


Clients

Protocol

Broker

Pulsar provides compatibility for seamless migration path



Clients

Protocol

Broker

Pulsar ecosystem



Protocol Handlers



Client Libraries



Connectors (Sources & Sinks)



Pulsar Functions (Lightweight Stream Processing)



Processing Engines



Data Offloaders (Tiered Storage)



Apache Pulsar key features

1. Unified messaging & streaming platform



2. Three levels of scalability & elasticity



3. Durability



4. Geo-Replication



5. Multi-tenancy



6. Compatibility with your software ecosystem



7. Open source community



Built-in Super Set of OSS Features

All these features
are Open-Source



Reduced Vendor
Dependency

Vibrant community



600+

Contributors



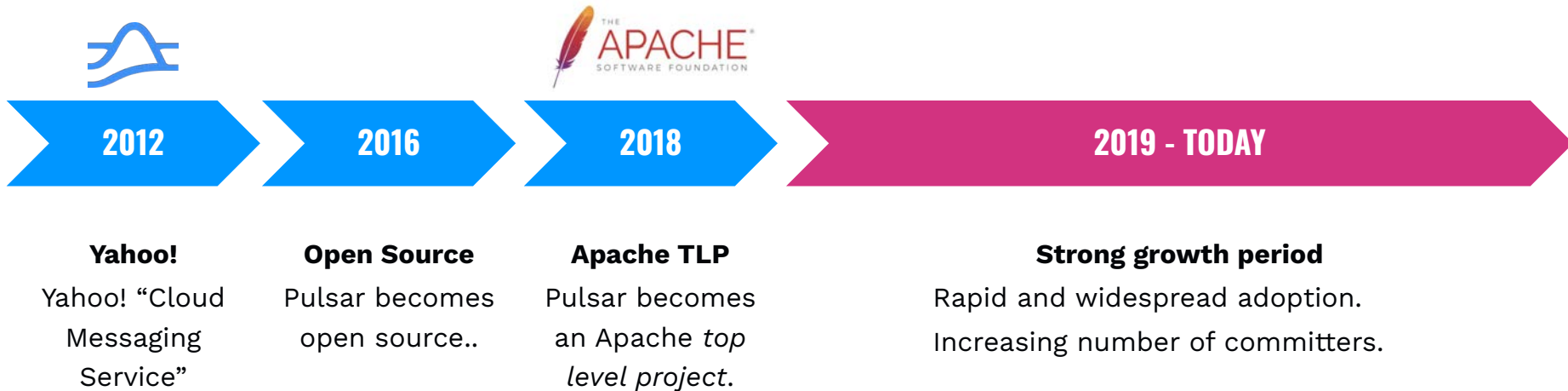
10,000

Slack Members

Apache Pulsar is used by 1000+ orgs



History



Apache Pulsar is battle-tested!



Apache Pulsar key features - recap

1. Unified messaging & streaming platform



2. Three levels of scalability & elasticity



3. Durability, offloading



4. Geo-Replication



5. Multi-tenancy



6. Compatibility with your software ecosystem



7. Open source community



Questions?

Reach out to me 🖱️



<https://streamnative.io/people/julien>

Join the Apache Pulsar Slack 🖱️



<https://bit.ly/3vGE9pJ>

