



Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

Apache Kafka

Hunor Ördög, Norbert-Raymond Pap

Department of Mathematics and Informatics
Babeș–Bolyai University, Cluj-Napoca

2024 May



Contents

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

1 What is Kafka?

2 Kafka components & Internal Architecture

3 Kafka vs RabbitMQ

4 Hotel Reservations example

5 References



What is Kafka?

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References



- Open-source distributed event streaming platform.



What is a “distributed streaming platform”?

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Streams are just infinite data, data that never ends.
- Distributed means Kafka works in a cluster, each node in the cluster is called a **Broker**.
- Kafka is a set of machines working together to be able to handle and process real-time infinite data.



Where does Kafka come from?

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Kafka was originally developed at LinkedIn  in 2010
- Open sourced in early 2011





Kafka Components

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper





Kafka Components

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- **Producer**
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper





Kafka Components

Producer

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- A producer is the source of data who will publish events.





Kafka Components

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper





Kafka Components

Consumer

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

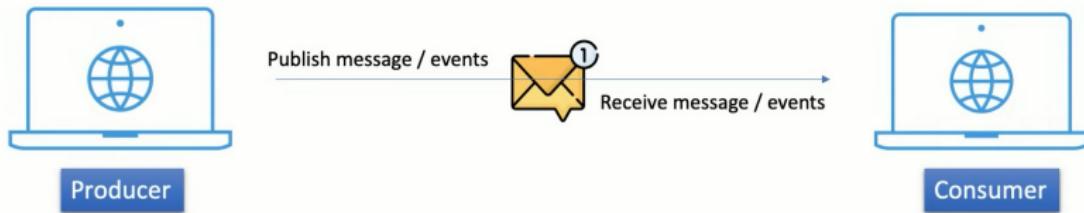
Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- A consumer acts as a receiver, consumes events.





Kafka Components

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper



Kafka Components

Broker

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

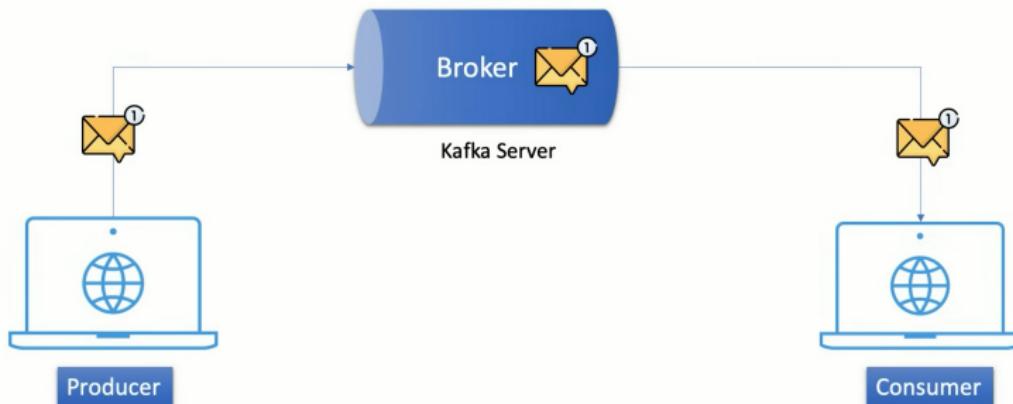
Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- The Kafka Broker is the Kafka server.
- A broker is an intermediate entity that helps in message exchanges between producers and consumers.





Kafka Components

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper



Kafka Components

Cluster

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

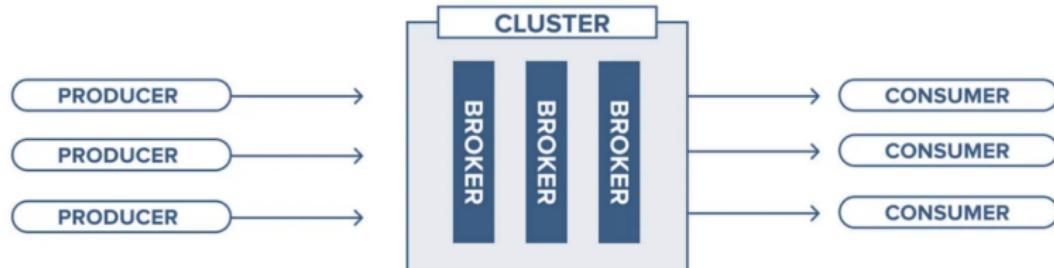
Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- There can be one or more brokers in the Kafka cluster.
- Its distributed architecture is one of the reasons that made Kafka so popular.
- The Brokers is what makes it so resilient, reliable, scalable, and fault-tolerant.





Kafka Components

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper





Kafka Components

Topic

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

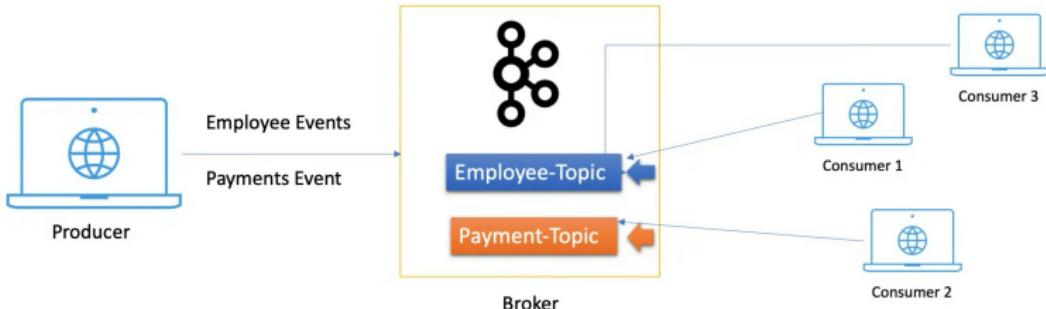
Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Specifies the category of the message or the classification of the message.
- Consumers can then just respond to the messages that belong to the topics they are listening on.





Kafka Components

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper





Kafka Components

Partitions

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

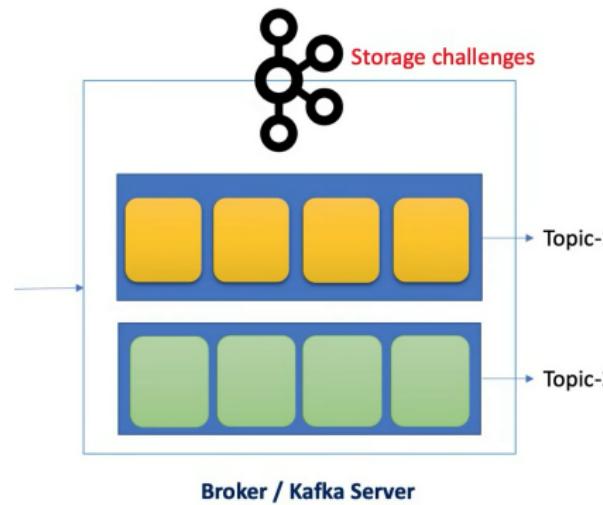
Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- A topic is divided into **partitions**.
- Each topic can have one or more partitions.
- We need to specify the number of partitions when creating the topic.





Kafka Components

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper



Kafka Components

Offset

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

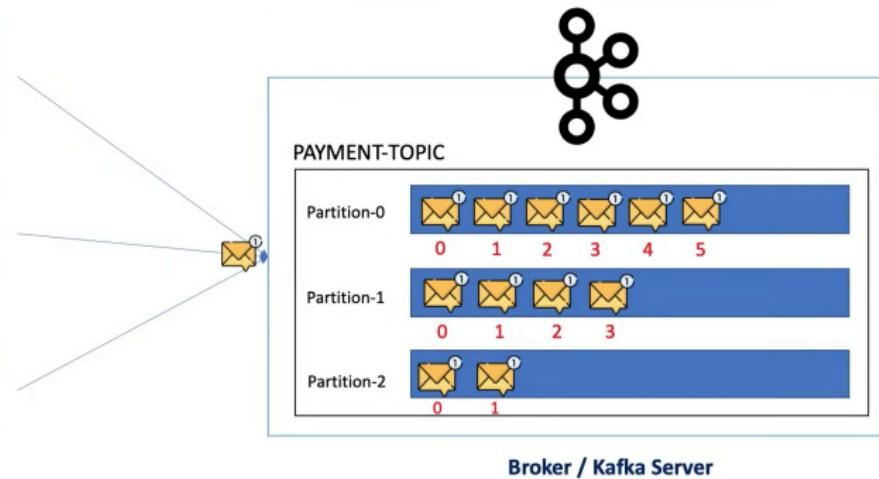
Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Each message will be stored in the broker disk and will receive an **offset**.
- Each partition has its own offsets.
- Kafka stores the messages on the disk (like a database).





Kafka Components

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper





Kafka Components

Consumer Groups

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

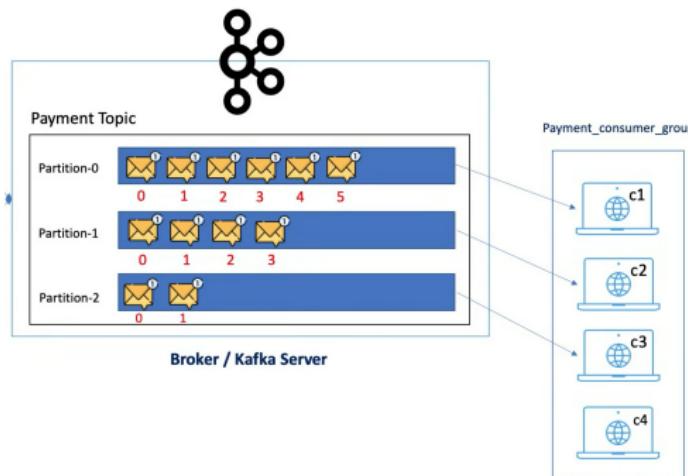
Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- It becomes very costly when a single consumer needs to read from many partitions.
- We need some sort of load-balancing, this is where **consumer groups** enter.
- The ideal is to have the same number of consumers in a group that we have as partitions in a topic, in this way, every consumer read from only one.





Kafka Components

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- Producer
- Consumer
- Broker
- Cluster
- Topic
- Partitions
- Offset
- Consumer Groups
- Zookeeper



Kafka Components

Zookeeper

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

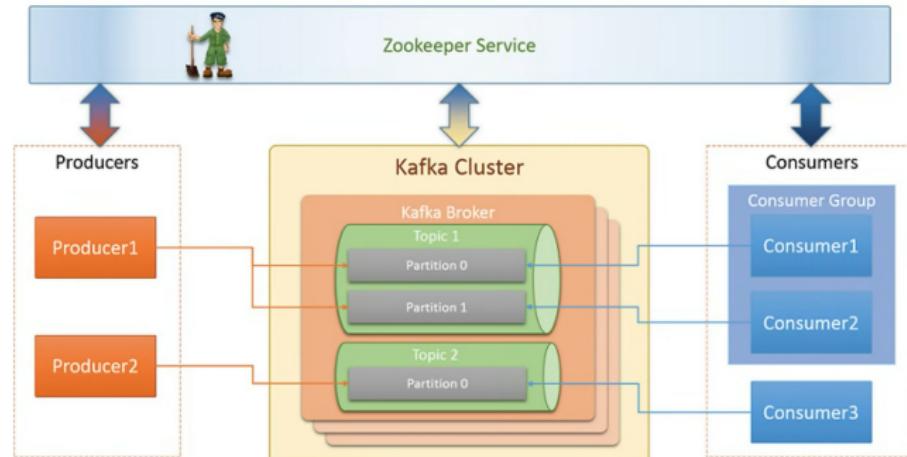
Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- **Zookeeper** is a prerequisite for Kafka.
- Kafka is a distributed system, and it uses **Zookeeper** for coordination and to track the status of Kafka cluster nodes.
- **Zookeeper** also keeps track of Kafka topics, partition, offsets.





Internal Architecture

Partition Leader

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- To ensure the reliability of the cluster, Kafka enters with the concept of the **Partition Leader**.
- Only one leader can exist per partition in all brokers.
- The leader is the only one that receives the messages, their replicas will just sync the data.
- When a leader goes down, a replica will be automatically elected as a new leader by **Zookeeper**.



Internal Architecture

Acknowledgment

Apache Kafka

Honor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- The **Acknowledgment (ack)** is a confirmation that the message was delivered.
- In Kafka we can configure this ack in three different levels:
 - **ack = 0**: We dont want to receive ack from Kafka.
 - **ack = 1**: Default configuration, we want to receive an ack from the partition leader.
 - **ack = all**: We want to receive confirmation from partition leader and their replicas as well.



Kafka vs RabbitMQ

part 1

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

	Kafka	RabbitMQ
Performance	Up to 1 million messages per second (high throughput, low latency)	4,000 - 10,000 messages per second (moderate throughput, low latency)
Message Retention	Policy-based retention (e.g., time-based, size-based)	Acknowledgment-based retention (messages are retained until acknowledged by consumers)
Data Type	Operational (Log-based)	Transactional (Message-oriented)
Consumer Mode	Dumb broker/smart consumer (consumers manage offsets, flexible subscription)	Smart broker/dumb consumer (broker manages acknowledgments, routing, and delivery)



Kafka vs RabbitMQ

part 2

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

	Kafka	RabbitMQ
Topology	Publish/subscribe based	Exchange types: Direct, Fanout, Topic, Headers
Scalability	Natively horizontally scalable	Can be horizontally scalable
Fault Tolerance	Built-in replication and fault tolerance (leader-follower replication)	Replication and high availability (mirrored queues)
Use Cases	Stream processing, real-time analytics, log aggregation, event sourcing	Message queuing, task distribution, RPC, event-driven architectures
Community Support	Large, active community	Large, active community
Ease of Use	Slightly complex setup, particularly for high availability and cluster management	Easy to set up and configure, especially for simple use cases
Language Support	Wide range of client libraries	Client libraries available for most languages

Demo: Hotel Reservations

System Architecture

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

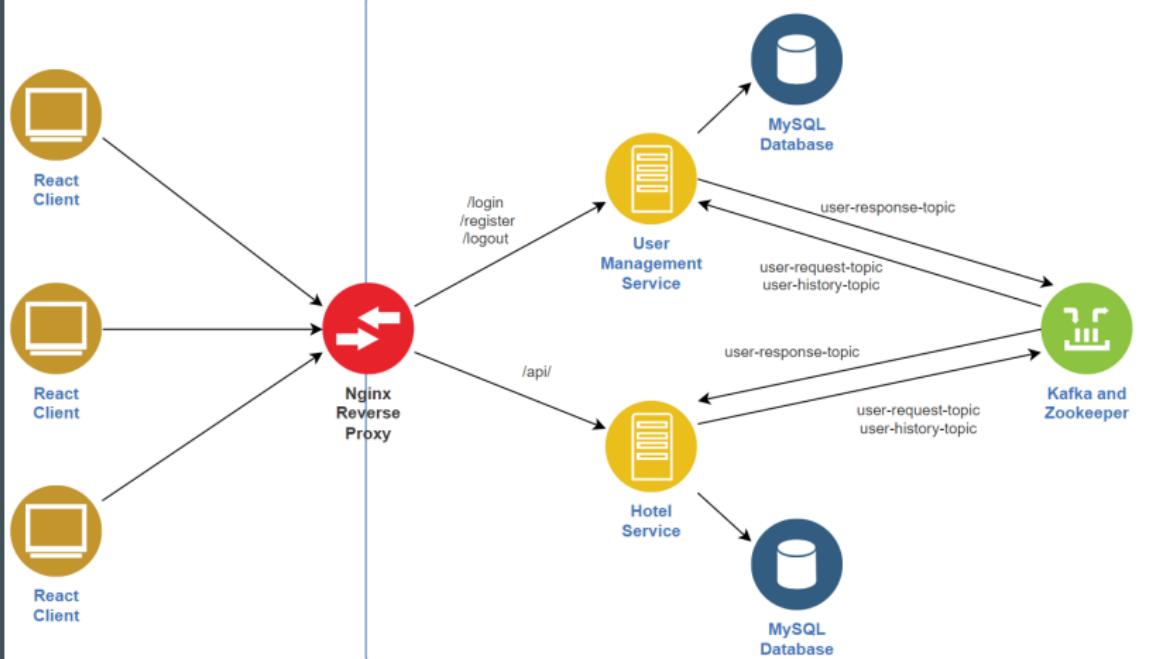
What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References





References

Apache Kafka

Hunor Ördög,
Norbert-
Raymond Pap

What is Kafka?

Kafka
components &
Internal
Architecture

Kafka vs
RabbitMQ

Hotel
Reservations
example

References

- <https://kafka.apache.org/intro>
- <https://www.cloudkarafka.com/blog/2016-11-30-part1-kafka-for-beginners-what-is-apache-kafka.html>
- <https://fullcycle.com.br/apache-kafka-trabalhando-com-mensageria-e-real-time>
- <https://www.educba.com/kafka-replication/>
- <https://www.javatpoint.com/apache-kafka-producer>
- <https://docs.cloudera.com/cdp-private-cloud-base/latest/kafka-developing-applications/topics/kafka-develop-groups-fetching.html>