

[By OnlineInterviewQuestions.com](http://OnlineInterviewQuestions.com)

[Apache Kafka Interview Questions](#)

Apache Kafka is buzzing these days, and there are many companies using this technology. When used properly, and using the right cases, Kafka has distinct attributes making Kafka a highly attractive option to integrate data. So, companies these days are looking for aspirants who know Kafka well and can use the right cases. Though you know Kafka very well, sometimes you may not be able to do good in interviews because of missing out on the very basics or minute information. So, to help you out with it, here are a few **Apache Kafka Interview Questions** that will help you land a job.

Read Best Apache Kafka Interview Questions and Answers

Q1. [Please Explain Apache Kafka?](#)

It is a publish-subscribe messaging application and an open source message broker project started by Apache software. Kafka was designed based on transactional logs design.

Q2. [In which language Apache Kafka is written?](#)

Kafka was written in Java and Scala programming languages.

Q3. [What is Kafka message?](#)

Kafka messages can simply be defined as byte arrays that are used by developers to store objects in the format of String, JSON, and Avro.

Q4. [List some use cases of Apache Kafka?](#)

Some of the use cases of Apache Kafka are:

- Message queue
- Event streams
- Tracking and logging

Q5. What is a Kafka cluster?

The Kafka cluster holds all the published records, whether or not they have been consumed using a configurable retention period.

Q6. List few differences between Apache Kafka and RabbitMQ?

Apache Kafka

It is distributed, and data is replicated with durability and availability. The performance rate is high with 100,000 messages per second. It also comes with consumer frameworks that allow reliable log data processing.

RabbitMQ

It has relatively less support for features like replication. The performance rate is 20,000 messages per second. The consumer here is FIFO based that reads from HEAD and process 1 by 1.

Q7. What are some alternatives to Apache Kafka?

Though we cannot find a system with the same concept of Kafka, you can still consider other message brokers such as ActiveMQ, ZeroMQ, RabbitMQ etc.

Q8. List some benefits of Apache Kafka?

Apart from having a traditional messaging technique, Apache Kafka has the following benefits:

- It is fast
- The data is partitioned and streamlined over a cluster for larger scalability.
- It is durable
- Distributed by design

Q9. What are the major APIs available in Apache Kafka?

There are four major APIs available in Apache Kafka:

- Producer API
- Consumer API
- Streams API
- Connector API

Q10. What is the role of a zookeeper in a Kafka cluster?

Zookeeper is used to store and preserve offset related information that is used to consume a particular topic by a specific consumer group, within the Kafka environment.

Q11. What is the message broker?

In Kafka, the message broker is meant to be the message server that holds the capability of storing publisher messages.

Q12. What does serDes mean in Apache Kafka?

SerDes means serializer and de-serializer. It is important for every Kafka stream to provide SerDes for the data types of records and record values to materialize the data when necessary.

Q13. Explain the format of an Apache Kafka message?

Kafka holds the feeds of messages in categories that are called as topics. At a high level, the producers send messages to the Kafka cluster that comprises servers that are called brokers, which in turn serves the messages to the consumers.

Q14. How do I achieve FIFO behavior with Kafka?

To achieve the FIFO behavior with Kafka, follow the steps mentioned below:

Set `enable.auto.commit=false`

- After processing the message, don't make a call to `consumer.commitSync()`;
- Make a call to 'subscribe' and register consumer to a topic.
- Implement a `consumerRebalance Listener` and perform `consumer.seek(topicPartition,offset)`; within the listener.
- Process the messages, hold each message's offset, store the processed message's offset with that of the processed message using atomic-transaction.
- Implement idempotent as a safety net.

Q15. List some difference between flume and Kafka?

Though both of them are used for real-time processing, Kafka is more scalable and ensures message durability.

Q16. How can you send large messages with Kafka (over 15MB)?

By adjusting three or four properties as follows, you can successfully send large messages without encountering any exceptions.

- Consumer side – `fetch.message.max.bytes`
- Broker side – `replica.fetch.max.bytes`
- Broker side – `message.max.bytes`
- Broker side (Per topic) – `max.message.bytes`

Q17. What is a Consumer Group?

It is nothing but an exclusive concept of Kafka. Each consumer groups has one or more consumers who consume subscribed topics.

Q18. What is the difference between Leader and Follower in Kafka?

Each partition in Kafka has one server that plays the role of a leader, while there can be none or more servers that act as followers.

Leader performs the task of all read and write request, while the followers passively replicate the role of a leader. To ensure load balancing, one of the followers takes up the role of a leader in case of Leader failing.

Q19. Explain the role of offset in Kafka?

The offset is a unique id assigned to the partitions, which contains messages. The most important use of offset is that it identifies the messages through the id, which are available in the partitions.

Q20. Can we use Kafka without Zookeeper?

No, it is merely not possible to use Kafka without the Zookeeper, as the user will not be able to connect directly to the Kafka server. And, if for some reason, the Zookeeper is down then the user will not be able to access any of the client requests.

Q21. What is offset in Kafka?

The **offset** is an uncomplicated integer number that is employed by Kafka to preserve the current position of a consumer. The current offset is a director to the latest record that Kafka has sent to a consumer in the most current poll and the offset is a position within a partition for the subsequent message to be sent to a customer. Kafka maintains two sorts of offsets; current offset and committed offset.

Please Visit [OnlineInterviewquestions.com](https://www.onlineinterviewquestions.com) to download more pdfs