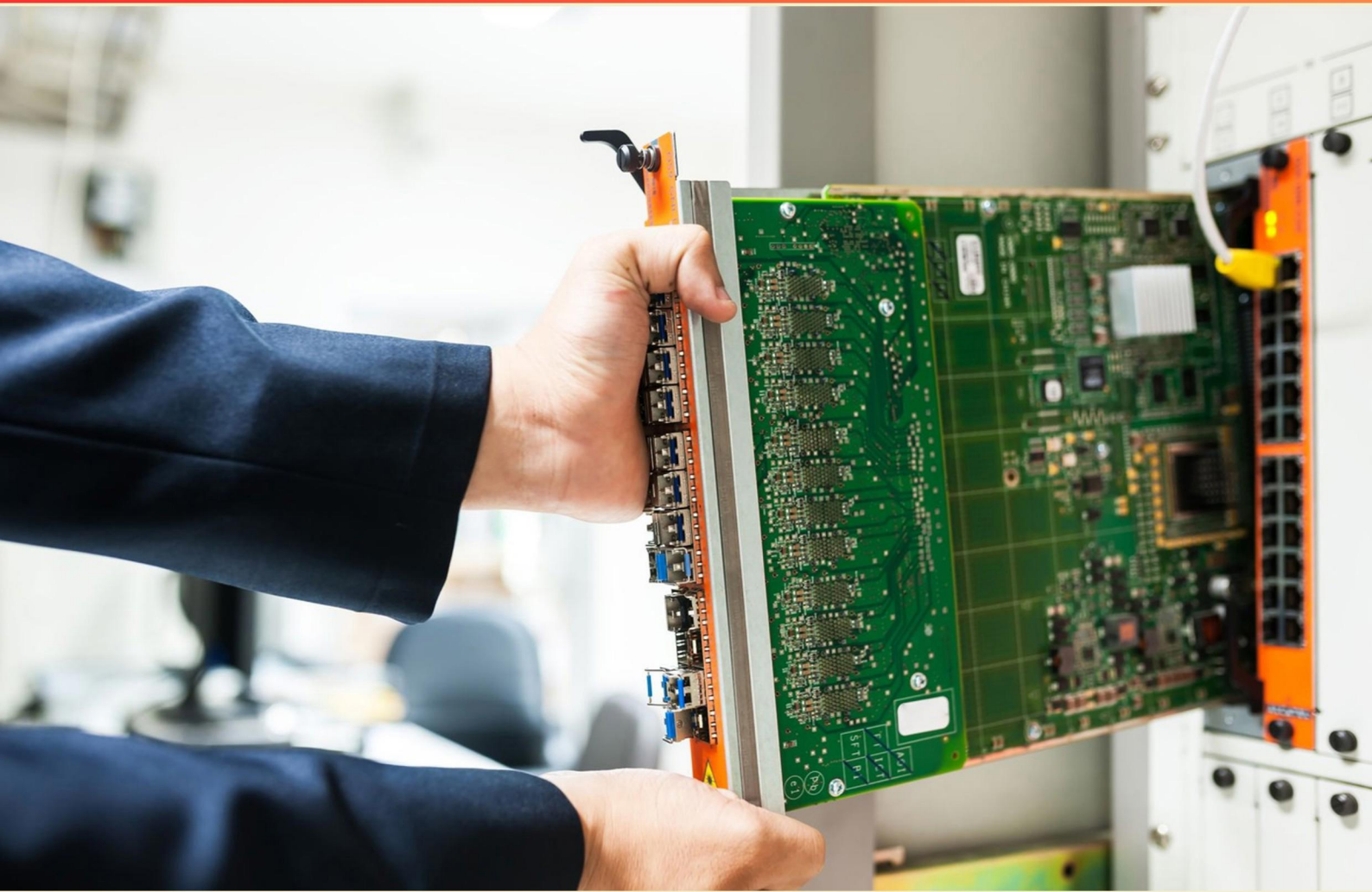


CCDAK APACHE KAFKA PRACTICE TEST (SAMPLE)

STUDY GUIDE



SAMPLE STUDY GUIDE WITH LIMITED QUESTIONS

VISIT US ONLINE FOR
THE FULL VERSION STUDY GUIDE

<https://ccdakapachekafka.examzify.com>



Copyright © 2026 by Examzify - A Kaluba Technologies Inc. product.

ALL RIGHTS RESERVED.

No part of this book may be reproduced or transferred in any form or by any means, graphic, electronic, or mechanical, including photocopying, recording, web distribution, taping, or by any information storage retrieval system, without the written permission of the author.

Notice: Examzify makes every reasonable effort to obtain accurate, complete, and timely information about this product from reliable sources.

SAMPLE

Table of Contents

Copyright	1
Table of Contents	2
Introduction	3
How to Use This Guide	4
Questions	5
Answers	8
Explanations	10
Next Steps	15

SAMPLE

Introduction

Preparing for a certification exam can feel overwhelming, but with the right tools, it becomes an opportunity to build confidence, sharpen your skills, and move one step closer to your goals. At Examzify, we believe that effective exam preparation isn't just about memorization, it's about understanding the material, identifying knowledge gaps, and building the test-taking strategies that lead to success.

This guide was designed to help you do exactly that.

Whether you're preparing for a licensing exam, professional certification, or entry-level qualification, this book offers structured practice to reinforce key concepts. You'll find a wide range of multiple-choice questions, each followed by clear explanations to help you understand not just the right answer, but why it's correct.

The content in this guide is based on real-world exam objectives and aligned with the types of questions and topics commonly found on official tests. It's ideal for learners who want to:

- Practice answering questions under realistic conditions,
- Improve accuracy and speed,
- Review explanations to strengthen weak areas, and
- Approach the exam with greater confidence.

We recommend using this book not as a stand-alone study tool, but alongside other resources like flashcards, textbooks, or hands-on training. For best results, we recommend working through each question, reflecting on the explanation provided, and revisiting the topics that challenge you most.

Remember: successful test preparation isn't about getting every question right the first time, it's about learning from your mistakes and improving over time. Stay focused, trust the process, and know that every page you turn brings you closer to success.

Let's begin.

How to Use This Guide

This guide is designed to help you study more effectively and approach your exam with confidence. Whether you're reviewing for the first time or doing a final refresh, here's how to get the most out of your Examzify study guide:

1. Start with a Diagnostic Review

Skim through the questions to get a sense of what you know and what you need to focus on. Your goal is to identify knowledge gaps early.

2. Study in Short, Focused Sessions

Break your study time into manageable blocks (e.g. 30 – 45 minutes). Review a handful of questions, reflect on the explanations.

3. Learn from the Explanations

After answering a question, always read the explanation, even if you got it right. It reinforces key points, corrects misunderstandings, and teaches subtle distinctions between similar answers.

4. Track Your Progress

Use bookmarks or notes (if reading digitally) to mark difficult questions. Revisit these regularly and track improvements over time.

5. Simulate the Real Exam

Once you're comfortable, try taking a full set of questions without pausing. Set a timer and simulate test-day conditions to build confidence and time management skills.

6. Repeat and Review

Don't just study once, repetition builds retention. Re-attempt questions after a few days and revisit explanations to reinforce learning. Pair this guide with other Examzify tools like flashcards, and digital practice tests to strengthen your preparation across formats.

There's no single right way to study, but consistent, thoughtful effort always wins. Use this guide flexibly, adapt the tips above to fit your pace and learning style. You've got this!

Questions

SAMPLE

- 1. Which command shows partitions without a leader for a topic?**
 - A. `bin/kafka-topics.sh --zookeeper localhost:2181 --describe`
 - B. `bin/kafka-topics.sh --zookeeper localhost:2181 --describe --unavailable-partitions`
 - C. `bin/kafka-topics.sh --bootstrap-server localhost:9092 --describe --unavailable-partitions`
 - D. `bin/kafka-topics.sh --zookeeper localhost:2181 --describe --unavailable-partitions`
- 2. In Avro schema evolution, removing or adding a field that has a default value is classified as which type?**
 - A. Forward
 - B. Full
 - C. Backward
 - D. Partial
- 3. In a ZooKeeper-based Kafka cluster, what type of metadata resides in ZooKeeper, and where is this metadata stored under KRaft?**
 - A. Topic data and messages.
 - B. Cluster membership and some configuration metadata; under KRaft, metadata lives in Kafka's own log and is managed by brokers.
 - C. Consumer offsets.
 - D. User credentials.
- 4. What is a major change with KRaft in terms of ZooKeeper usage?**
 - A. ZooKeeper continues to manage cluster metadata as before.
 - B. KRaft introduces a ZooKeeperless cluster by using an internal controller quorum.
 - C. KRaft stores all metadata in an external database.
 - D. KRaft eliminates ZooKeeper and stores metadata within Kafka via an internal controller quorum.
- 5. Which port is used by followers to connect to the active leader?**
 - A. Client port
 - B. Election port
 - C. Data port
 - D. Leader port

- 6. What is the join type for KTable-to-KTable joins?**
- A. Key-Based
 - B. Key Or Foreign-Key Based
 - C. Window Based
 - D. Foreign Key Based
- 7. In an e-commerce application, which data should be modeled as a KTable?**
- A. Mostly static data like inventory lists, customer lists, and aggregated data like total sales.
 - B. Real-time events such as click streams.
 - C. Binary large objects like images.
 - D. Ephemeral session data.
- 8. By default, how is the target partition chosen for a record with a key in Kafka?**
- A. Round-robin distribution when there is no key
 - B. Hash of the key
 - C. Partition number increments
 - D. Random selection
- 9. What does `fetch.message.max.bytes` configure?**
- A. The maximum size of a message that can be replicated across brokers.
 - B. The default maximum poll interval.
 - C. The maximum size of message that can be fetched by a consumer from the broker.
 - D. The default session timeout
- 10. To publish a 15 MB message, you must configure which broker/topic setting?**
- A. `replica.fetch.max.bytes=15728640`
 - B. `max.poll.interval.ms=300000`
 - C. `fetch.message.max.bytes=15728640`
 - D. `message.max.bytes=15728640` (broker/topic config)

Answers

SAMPLE

1. D
2. B
3. B
4. D
5. D
6. B
7. A
8. B
9. C
10. D

SAMPLE

Explanations

SAMPLE

1. Which command shows partitions without a leader for a topic?

- A. `bin/kafka-topics.sh --zookeeper localhost:2181 --describe`
- B. `bin/kafka-topics.sh --zookeeper localhost:2181 --describe --unavailable-partitions`
- C. `bin/kafka-topics.sh --bootstrap-server localhost:9092 --describe --unavailable-partitions`
- D. `bin/kafka-topics.sh --zookeeper localhost:2181 --describe --unavailable-partitions`

You're being tested on how to identify partitions that currently have no leader. In Kafka, every partition normally has a leader broker that handles all reads and writes for that partition. If that leader broker goes down or is in the middle of a leader election, the partition temporarily has no leader and is considered unavailable. The `kafka-topics.sh` tool can surface this state when you describe a topic and request the partitions that are unavailable. The flag to include those partitions is `--unavailable-partitions`. To get this information, you connect to the cluster state (in this setup via Zookeeper) and run the command with `--zookeeper`, along with `--describe` and `--unavailable-partitions`. This combination will list partitions that currently lack a leader. Using `bootstrap-server` mode can behave differently for this particular flag in some environments, so the Zookeeper-based approach gives the reliable output for showing partitions without a leader.

2. In Avro schema evolution, removing or adding a field that has a default value is classified as which type?

- A. Forward
- B. Full
- C. Backward
- D. Partial

In Avro schema evolution, full compatibility means both backward and forward compatibility hold: data written with the old schema can be read with the new schema, and data written with the new schema can be read with the old one. If you add or remove a field and give that field a default value, you preserve both directions. Adding a field with a default means old data (which doesn't have that field) can be read by the new schema—the missing field is filled in with its default. Removing a field while keeping a sensible default for that field in the context of evolution allows older data to be read by the new schema and keeps newer data readable by older schemas, because differences in fields are either ignored or provided by defaults. This dual-direction compatibility is why the change is considered full compatibility.

3. In a ZooKeeper-based Kafka cluster, what type of metadata resides in ZooKeeper, and where is this metadata stored under KRaft?

- A. Topic data and messages.
- B. Cluster membership and some configuration metadata; under KRaft, metadata lives in Kafka's own log and is managed by brokers.**
- C. Consumer offsets.
- D. User credentials.

The main idea is how metadata is stored and who manages it. In a ZooKeeper-based Kafka cluster, ZooKeeper holds the cluster's basic structural metadata—primarily which brokers are part of the cluster (cluster membership) and certain topic/partition configuration information. It does not hold the actual topic data or the consumer offsets, which are managed by Kafka itself. When moving to KRaft, that metadata responsibility shifts into Kafka's own system. The metadata is stored in Kafka's internal metadata log, replicated across the cluster using the Raft protocol, and managed by the brokers acting as the metadata quorum. There's no separate ZooKeeper involved for cluster metadata in this setup. So, ZooKeeper houses the membership and related configuration data, while under KRaft, the metadata lives in Kafka's own log on the brokers.

4. What is a major change with KRaft in terms of ZooKeeper usage?

- A. ZooKeeper continues to manage cluster metadata as before.
- B. KRaft introduces a ZooKeeperless cluster by using an internal controller quorum.
- C. KRaft stores all metadata in an external database.
- D. KRaft eliminates ZooKeeper and stores metadata within Kafka via an internal controller quorum.**

The change being tested is that KRaft removes the need for ZooKeeper by letting Kafka manage its own cluster metadata inside the Kafka system itself. In KRaft, a built-in controller quorum, powered by the Raft protocol, runs within the Kafka brokers and stores all metadata there, rather than relying on an external ZooKeeper service. This makes the cluster self-contained and eliminates ZooKeeper entirely. That's why the best answer states that ZooKeeper is eliminated and metadata is stored within Kafka via an internal controller quorum. The other options either keep ZooKeeper, suggest an external database, or describe a ZooKeeperless setup without mentioning that the metadata is stored inside Kafka, which is the key distinction of KRaft.

5. Which port is used by followers to connect to the active leader?

- A. Client port
- B. Election port
- C. Data port
- D. Leader port**

In a system that uses a leader-follower replication model, the followers stay in sync by connecting to the current active leader to fetch and apply updates. This replication traffic happens over a dedicated port on the leader that is intended for inter-broker communication. That port is what followers use to establish the replication session and receive new log entries from the leader, keeping the follower replicas up to date. The client port is reserved for end-user clients (producers and consumers) to connect to brokers, not for replication between brokers. The election port is only involved during the process of electing a new leader and isn't used for ongoing replication once leadership is established. The data port isn't a standard term used for this purpose in the typical replication setup. Therefore, the port used by followers to connect to the active leader is the Leader port.

6. What is the join type for KTable-to-KTable joins?

- A. Key-Based
- B. Key Or Foreign-Key Based**
- C. Window Based
- D. Foreign Key Based

The thing being tested is how KTable-to-KTable joins determine which records should be combined. In Kafka Streams, these joins pair up records by their keys. That means you typically join on the key that each KTable carries. If the information you want to join on is not currently the key, you can re-key one or both tables by extracting a foreign key from the value and use that as the new join key. In other words, you can perform a join based on the primary key (the existing key) or based on a foreign key you derive by re-keying, which is why the answer reflects "Key Or Foreign-Key Based." This isn't a windowed join—the join is defined by matching keys, not a time window.

7. In an e-commerce application, which data should be modeled as a KTable?

- A. Mostly static data like inventory lists, customer lists, and aggregated data like total sales.**
- B. Real-time events such as click streams.
- C. Binary large objects like images.
- D. Ephemeral session data.

KTable represents a stateful, updatable view of data keyed by a unique key, storing the latest value for each key. This makes it ideal for data you want to keep current and readily query or join with streams. In an e-commerce app, inventory lists and customer lists are exactly the kind of data you want to reflect as a current snapshot: you update the quantity when stock changes or a customer's information is updated, and downstream processes can rely on the latest state. Aggregated data like total sales per product or category also fits, because you often need a running, current total that other streams can join with or query. Real-time events such as click streams, on the other hand, are best modeled as a KStream since they represent a sequence of events over time rather than a single, latest state per key. Large binary objects aren't well suited for KTable storage because Kafka topics are optimized for key-value, compacted-state data rather than storing big blobs; you'd keep the BLOBs in external storage and store references in the KTable. Ephemeral session data is transient and not meant to become a durable, long-lived view, so it isn't a natural fit for a KTable either. So, data that represents a current, queryable state or derived, up-to-date totals are best modeled as a KTable.

8. By default, how is the target partition chosen for a record with a key in Kafka?

- A. Round-robin distribution when there is no key
- B. Hash of the key**
- C. Partition number increments
- D. Random selection

Kafka's default partitioner uses a hash of the message key to pick the target partition. This hashing is done on the key's bytes and the result is mapped to one of the available partitions (typically by taking the hash modulo the number of partitions). The payoff is that all records with the same key consistently land in the same partition, which preserves per-key ordering within a topic. If there is no key, the partitioner instead distributes messages across partitions in a round-robin fashion to balance load.

9. What does `fetch.message.max.bytes` configure?

- A. The maximum size of a message that can be replicated across brokers.
- B. The default maximum poll interval.
- C. The maximum size of message that can be fetched by a consumer from the broker.**
- D. The default session timeout

Fetching in Kafka happens through fetch requests from a consumer to the broker. `fetch.message.max.bytes` sets a cap on how much data the broker will return in one of those fetch responses. In other words, it limits the total amount of message data that can be pulled in a single fetch, which helps control memory use on the broker and keep network traffic predictable. If a message batch would exceed this limit, the broker returns only the portion that fits, and the consumer may need additional fetches to get the rest. This setting is specifically about the data returned during fetch, not about replication, poll intervals, or session timeouts.

10. To publish a 15 MB message, you must configure which broker/topic setting?

- A. `replica.fetch.max.bytes=15728640`
- B. `max.poll.interval.ms=300000`
- C. `fetch.message.max.bytes=15728640`
- D. `message.max.bytes=15728640` (broker/topic config)**

Kafka restricts how large a single message can be, and that limit is set on the broker as `message.max.bytes`. You can also raise it per topic with `max.message.bytes` if you want a specific topic to allow larger messages. To publish a 15 MB message, you need at least 15728640 bytes configured, either as the broker's `message.max.bytes` or as the topic's `max.message.bytes`. The other options control different behaviors—`replica.fetch.max.bytes` affects how much a follower fetches, `max.poll.interval.ms` is about consumer polling cadence, and `fetch.message.max.bytes` governs how much data a consumer fetches in a single request. They do not set the maximum size for a published message. If you're using a producer, also ensure the producer's `max.request.size` is at least 15728640 so the client request can carry the 15 MB payload.

Next Steps

Congratulations on reaching the final section of this guide. You've taken a meaningful step toward passing your certification exam and advancing your career.

As you continue preparing, remember that consistent practice, review, and self-reflection are key to success. Make time to revisit difficult topics, simulate exam conditions, and track your progress along the way.

If you need help, have suggestions, or want to share feedback, we'd love to hear from you. Reach out to our team at hello@examzify.com.

Or visit your dedicated course page for more study tools and resources:

<https://ccdakapachekafka.examzify.com>

We wish you the very best on your exam journey. You've got this!

SAMPLE