

PRACTICE QUESTIONS

Google Cloud

Professional Cloud Developer

SHARED BY [u/cleverjaguar1981](#) on aiexamcerts

EDITION	Free Set 1
QUESTIONS	10
SHARED	October 15, 2026
EXAM	Google Cloud Professional Cloud Developer

Shared by [u/cleverjaguar1981](#) on aiexamcerts

10 practice questions for the Google Cloud Professional Cloud Developer exam, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/google-cloud/professional-cloud-developer. Spotted a mistake? Email support@aiexamcerts.com with the exam name and the question number.

1. An order service must publish events without waiting for three downstream services to finish their work. Which Google Cloud service fits best?
 - A Cloud SQL read replicas
 - B A Pub/Sub topic with one subscription per downstream service
 - C A Memorystore for Redis list that each service polls
 - D A single Compute Engine VM running a cron job

Answer: B

Pub/Sub decouples publishers from subscribers: the order service publishes once and returns, and each subscription receives its own copy of every message and consumes it independently. Memorystore is an in-memory cache rather than a durable messaging service, so polling it adds custom code and loses messages on failure. Cloud SQL read replicas serve database reads and are not an event bus.

2. Pub/Sub delivers messages at least once, so consumers must be designed to be _____ so that processing a duplicate message does not change the outcome.

Answer: idempotent (also accepted: idempotency, idempotent processing)

At-least-once delivery means the same message can arrive more than once, especially after a missed acknowledgment. An idempotent handler records what it has already processed, such as by message ID, so a repeat is harmless. Designing for exactly-once processing in the consumer is not something Pub/Sub's delivery guarantee provides on its own.

3. Which Google Cloud service stores Docker images, Maven packages and npm packages in one managed repository service?
 - A Cloud Source Repositories
 - B A Cloud Storage bucket with object lifecycle rules
 - C Artifact Registry
 - D Filestore

Answer: C

Artifact Registry hosts container images as well as language packages such as Maven and npm in one service with IAM controls. Cloud Source Repositories holds source code rather than built artifacts. A bucket can hold files but does not provide native package repository protocols, and Filestore is a network file system.

4. A JSON key file must be checked into the source repository for Application Default Credentials to work in Cloud Run.
- A True
 - B False

Answer: B (False)

Application Default Credentials find credentials automatically on Google Cloud through the attached service account, using the metadata server. Checking a key file into source control is a security risk and is unnecessary there. A key file is only needed for code running outside Google Cloud, and even then it should be kept out of the repository.

5. Which two actions reduce the blast radius of a bad deployment? (Choose two.)
- A Send 100 percent of traffic to the new version at once
 - B Use canary or percentage-based traffic splits
 - C Deploy to every region at the same moment
 - D Roll out region by region with a bake period between stages
 - E Turn off monitoring during the rollout to reduce noise

Answer: B, D

Canary splits expose only a small share of users to a change, and staged regional rollouts with a bake period limit how far a fault can spread before it is caught. Sending everything at once or deploying everywhere together maximizes the impact of a defect. Turning off monitoring removes the signal needed to detect the problem.

6. In a GKE Deployment, which two fields control how many extra pods can be created and how many can be unavailable during a rolling update?
- A replicas and selector
 - B maxSurge and maxUnavailable
 - C readinessProbe and livenessProbe
 - D minReplicas and maxReplicas

Answer: B

The maxSurge and maxUnavailable fields in the rolling update strategy set how many pods can exceed the desired count and how many can be down during the rollout. Replicas sets the desired count, and the probes decide whether a pod receives traffic. The autoscaler bounds control scaling, not the pace of a rolling update.

7. Which Pub/Sub subscription type lets a Cloud Run service receive messages without running a polling loop?
- A A pull subscription that runs a long loop in the service
 - B A push subscription that sends messages to the service's HTTPS endpoint
 - C A snapshot that holds the subscription's backlog for replay
 - D A BigQuery subscription that writes messages into a table

Answer: B

A push subscription posts each message to the service's HTTPS endpoint, so the service does not need to poll. A pull subscription requires the service to keep asking for messages. A snapshot preserves a subscription's position for replay, and a BigQuery subscription writes messages to a table rather than to an application.

8. To send Cloud Logging entries to BigQuery continuously, create a log _____ whose destination is a BigQuery dataset.

Answer: sink (also accepted: log sink)

A log sink routes matching log entries to a destination such as BigQuery, Cloud Storage or Pub/Sub. Routing continues as new entries arrive. A log-based metric instead turns matching entries into a numeric time series, which is a different feature.

9. A team wants an alert when the 99th percentile latency of a Cloud Run service exceeds 500 ms. Which metric type should the policy use?
- A The request count metric
 - B The instance count metric
 - C The request latencies distribution metric
 - D The container CPU utilization metric

Answer: C

The request latencies metric is a distribution, so it supports percentile calculations such as the 99th percentile in a condition. Request count and instance count do not describe latency, and CPU utilization measures load rather than response time. A mean-based latency alert would hide the slow tail the team wants to catch.

10. Which practice gives the earliest warning that a service is on track to miss its availability SLO?
- A Reviewing the monthly invoice for unusual charges
 - B Reading the logs once a week
 - C Running a load test once per quarter
 - D Alerting on the error budget burn rate

Answer: D

Burn rate alerts fire when the error budget is being consumed faster than the SLO allows, often long before the budget is exhausted. Invoices and weekly log reviews come too late, and a quarterly load test cannot detect a live burn. Burn rate alerting is the standard SRE approach for this.