

# Google Cloud

# Professional DevOps Engineer

---

Certification Practice Examination  
Complete Question Bank with Detailed Explanations

*Document Generated: December 28, 2025*

<b>Total Questions</b>	170
<b>Question Format</b>	Multiple Choice (Single & Multiple Answer)
<b>Difficulty Level</b>	Professional Certification Standard
<b>Coverage</b>	Complete DevOps Engineer Exam Topics

## Question 1: GKE Application Performance Monitoring

### Scenario

You support a Node.js application running on Google Kubernetes Engine (GKE) in production. The application makes several HTTP requests to dependent applications. You want to anticipate which dependent applications might cause performance issues.

### Question

What should you do?

- A. Instrument all applications with Stackdriver Profiler.
- B. Instrument all applications with Stackdriver Trace and review inter-service HTTP requests.**
- C. Use Stackdriver Debugger to review the execution of logic within each application to instrument all applications.
- D. Modify the Node.js application to log HTTP request and response times to dependent applications. Use Stackdriver Logging to find dependent applications that are performing poorly.

### ✓ Correct Answer: B

*Explanation: Use Cloud Trace to visualize request latency and identify performance bottlenecks across multiple services.*

---

## Question 2: Sharing Stackdriver Charts Securely

### Scenario

You created a Stackdriver chart for CPU utilization in a dashboard within your workspace project. You want to share the chart with your Site Reliability Engineering (SRE) team only.

### Question

You want to ensure you follow the principle of least privilege. What should you do?

- A. Share the workspace Project ID with the SRE team. Assign the SRE team the Monitoring Viewer IAM role in the workspace project.
- B. Share the workspace Project ID with the SRE team. Assign the SRE team the Dashboard Viewer IAM role in the workspace project.
- C. Click "Share chart by URL" and provide the URL to the SRE team. Assign the SRE team the Monitoring Viewer IAM role in the workspace project.**
- D. Click "Share chart by URL" and provide the URL to the SRE team. Assign the SRE team the Dashboard Viewer IAM role in the workspace project.

### ✓ Correct Answer: C

*Explanation: Share a direct URL to the specific chart and grant the minimal required IAM role ('Monitoring Viewer') to view it.*

## Question 3: SRE Postmortem Culture

### Scenario

Your organization wants to implement Site Reliability Engineering (SRE) culture and principles. Recently, a service that you support had a limited outage. A manager on another team asks you to provide a formal explanation of what happened so they can action remediations.

### Question

What should you do?

- A. Develop a postmortem that includes the root causes, resolution, lessons learned, and a prioritized list of action items. Share it with the manager only.
- B. Develop a postmortem that includes the root causes, resolution, lessons learned, and a prioritized list of action items. Share it on the engineering organization's document portal.**
- C. Develop a postmortem that includes the root causes, resolution, lessons learned, the list of people responsible, and a list of action items for each person. Share it with the manager only.
- D. Develop a postmortem that includes the root causes, resolution, lessons learned, the list of people responsible, and a list of action items for each person. Share it on the engineering organization's document portal.

### ✓ Correct Answer: B

*Explanation: SRE principles advocate for blameless postmortems that are shared widely to promote organizational learning.*

---

## Question 4: Logging from Third-Party Containers in GKE

### Scenario

You have a set of applications running on a GKE cluster with Stackdriver Kubernetes Engine Monitoring. You are deploying a new, unmodifiable third-party application that writes its logs to `/var/log/app\_messages.log`.

### Question

How can you send these log entries to Stackdriver Logging?

- A. Use the default Stackdriver Kubernetes Engine Monitoring agent configuration.
- B. Deploy a Fluentd daemonset to GKE. Then create a customized input and output configuration to tail the log file in the application's pods and write to Stackdriver Logging.
- C. Install Kubernetes on GCE and redeploy your applications. Then customize the built-in Stackdriver Logging configuration to tail the log file.
- D. Write a script to tail the log file within the pod and write entries to standard output. Run the script as a sidecar container with the application's pod.**

### ✓ Correct Answer: D

*Explanation: Use a sidecar container with a shared volume to read the log file and forward its contents to standard output, which is automatically collected by Stackdriver. Configure a shared volume between the containers to allow the script to have read access to /var/log in the application container.*

## Question 5: Troubleshooting Stackdriver Logging Agent

### Scenario

You are running an application in a VM with the Stackdriver Logging agent installed. The application logs via syslog, but the logs are not appearing in the Stackdriver Logs Viewer.

### Question

What is the first thing you should do to troubleshoot?

- A. Look for the agent's test log entry in the Logs Viewer.
- B. Install the most recent version of the Stackdriver agent.
- C. Verify the VM service account access scope includes the `monitoring.write` scope.
- D. SSH to the VM and execute the following command on your VM: `ps ax | grep fluentd`.**

### ✓ Correct Answer: D

*Explanation:* First, verify that the logging agent (fluentd) process is actually running on the VM.

---

## Question 6: Cloud Build Integration with Webhooks

### Scenario

You use a multi-step Cloud Build pipeline and want to send build information to a third-party monitoring platform's webhook.

### Question

What is the most efficient way to achieve this with minimal development effort?

- A. Add logic to each Cloud Build step to HTTP POST the build information to a webhook.
- B. Add a new step at the end of the pipeline in Cloud Build to HTTP POST the build information to a webhook.
- C. Use Stackdriver Logging to create a logs-based metric from the Cloud Build logs. Create an Alert with a Webhook notification type.
- D. Create a Cloud Pub/Sub push subscription to the Cloud Build `cloud-builds` Pub/Sub topic to HTTP POST the build information to a webhook.**

### ✓ Correct Answer: D

*Explanation:* Cloud Build automatically publishes build events to a Pub/Sub topic. Use a push subscription to forward these events to a webhook with minimal development effort.

---

## Question 7: Canary Analysis Configuration

### Scenario

You use Spinnaker for deployments and have a canary stage in your pipeline. Your application has an in-memory cache that loads objects at start time.

### Question

How should you configure the canary analysis to automate the comparison of the canary version against the production version?

- A. Compare the canary with a new deployment of the current production version.
- B. Compare the canary with a new deployment of the previous production version.
- C. Compare the canary with the existing deployment of the current production version.
- D. Compare the canary with the average performance of a sliding window of previous production versions.

### ✓ Correct Answer: A

*Explanation: To ensure a fair comparison, compare the canary against a fresh deployment of the current production version, which also has a "cold" cache like the canary deployment.*

---

## Question 8: Calculating Latency SLI

### Scenario

You support a high-traffic web application and want to implement a Service Level Indicator (SLI) for home page request latency, with an acceptable load time of 100 ms.

### Question

What is the Google-recommended way of calculating this SLI?

- A. Bucketize the request latencies into ranges, and then compute the percentile at 100 ms.
- B. Bucketize the request latencies into ranges, and then compute the median and 90th percentiles.
- C. Count the number of home page requests that load in under 100 ms, and then divide by the total number of home page requests.
- D. Count the number of home page requests that load in under 100 ms, and then divide by the total number of all web application requests.

### ✓ Correct Answer: C

*Explanation: An SLI should be a ratio of "good events" to "total valid events." Here, "good events" are requests faster than 100ms divided by the total number of home page requests.*

---

## Question 9: Reducing Mean Time to Recovery (MTTR)

### Scenario

After a weekend release, a new feature failed in production, causing an extended outage before a rollback was performed.

### Question

What two actions should you take to modify your release process and reduce MTTR?

- A. Before merging new code, require 2 different peers to review the code changes.
- B. Adopt the blue/green deployment strategy when releasing new code via a CD server.**
- C. Integrate a code linting tool to validate coding standards before any code is accepted into the repository.
- D. Require developers to run automated integration tests on their local development environments before release.

### ✓ Correct Answer: B

*Explanation: Blue/green deployments allow for instant rollbacks, and a robust CI pipeline with automated testing catches issues before they reach production.*

---

## Question 10: Secure Log Access for Developers

### Scenario

You have a pool of application servers running on Compute Engine and need to provide developers with a secure and easy way to access application logs for troubleshooting.

### Question

How would you implement this solution on GCP with the least amount of configuration?

- A. Deploy the Stackdriver logging agent to the application servers. Give the developers the IAM Logs Viewer role to access Stackdriver and view logs.**
- B. Deploy the Stackdriver logging agent to the application servers. Give the developers the IAM Logs Private Viewer role.
- C. Deploy the Stackdriver monitoring agent to the application servers. Give the developers the IAM Monitoring Viewer role.
- D. Write a script using `gsutil` to upload logs to a Cloud Storage bucket and give developers IAM Object Viewer access.

### ✓ Correct Answer: A

*Explanation: The Stackdriver logging agent automatically collects logs and sends them to Cloud Logging. The Logs Viewer role provides read-only access to view logs without unnecessary permissions.*

---

## Question 11: Reducing Network Costs in GKE

### Scenario

You support the backend of a mobile game running on a GKE cluster. The application serves HTTP requests from users.

### Question

What should you do to reduce network costs?

- A. Configure the VPC as a Shared VPC Host project.
- B. Configure your network services on the Standard Tier.**
- C. Configure your Kubernetes cluster as a Private Cluster.
- D. Configure a Google Cloud HTTP Load Balancer as Ingress.

### ✓ Correct Answer: B

*Explanation:* Network Service Tiers allow you to optimize for performance (Premium Tier) or cost (Standard Tier). Standard Tier is ideal for reducing egress costs.

---

## Question 12: Post-Incident Communication

### Scenario

You encountered a major service outage that affected all users for multiple hours. The service has now been restored.

### Question

Following SRE recommended practices, what is the first thing you should do to provide an incident summary to relevant stakeholders?

- A. Call individual stakeholders to explain what happened.
- B. Develop a post-mortem to be distributed to stakeholders.**
- C. Send the Incident State Document to all the stakeholders.
- D. Require the engineer responsible to write an apology email to all stakeholders.

### ✓ Correct Answer: B

*Explanation:* Following SRE practices, write a blameless postmortem documenting what happened, why, and what actions will prevent recurrence. Share it broadly for organizational learning.

---

## Question 13: GKE Capacity Planning

### Scenario

You are performing capacity planning for a GKE-based service with a predicted user growth of 10% month-over-month. The service runs on a regional cluster with cluster autoscaler enabled and currently uses 30% of its deployed CPU capacity.

### Question

How should you prepare to handle the predicted growth while ensuring resilience and avoiding unnecessary costs?

- A. Verify the maximum node pool size, enable a horizontal pod autoscaler, and then perform a load test to verify your expected resource needs.**
- B. Because you are deployed on GKE and are using a cluster autoscaler, your GKE cluster will scale automatically, regardless of growth rate.
- C. Because you are at only 30% utilization, you have significant headroom and you won't need to add any additional capacity for this rate of growth.
- D. Proactively add 60% more node capacity to account for six months of 10% growth rate, and then perform a load test.

### ✓ Correct Answer: A

**Explanation:** Proactive planning involves verifying autoscaling limits (both pod and node level) and using load testing to validate that the system can handle future demand.

---

## Question 14: Automated Deployment from GCR

### Scenario

Your application images are built and pushed to Google Container Registry (GCR). You want to build an automated pipeline that deploys the application when the image is updated.

### Question

What should you do to minimize development effort?

- A. Use Cloud Build to trigger a Spinnaker pipeline.
- B. Use Cloud Pub/Sub to trigger a Spinnaker pipeline.**
- C. Use a custom builder in Cloud Build to trigger Jenkins pipeline.
- D. Use Cloud Pub/Sub to trigger a custom deployment service running in GKE.

### ✓ Correct Answer: B

**Explanation:** GCR publishes events to Pub/Sub when images are pushed. Configure a Pub/Sub push subscription to trigger Spinnaker pipeline with minimal development effort.

---

## Question 15: Calculating Reliability Risk

### Scenario

Your product is deployed across three GCP zones. A zone failover causes a 10-minute service disruption. Database failures occur about once per quarter (90 days) and are detected within 5 minutes. A new chat feature requires a new database system that takes twice as long (20 minutes) to fail over.

### Question

What would be the values for the risk of database failover with the new system?

- A. MTTD: 5, MTTR: 10, MTBF: 90, Impact: 33%
- B. MTTD: 5, MTTR: 20, MTBF: 90, Impact: 33%**
- C. MTTD: 5, MTTR: 10, MTBF: 90, Impact: 50%
- D. MTTD: 5, MTTR: 20, MTBF: 90, Impact: 50%

### ✓ Correct Answer: B

*Explanation: MTTR is doubled to 20 minutes, and since one of three zones is impacted, the user impact is 33%. MTTD and MTBF remain unchanged.*

---

## Question 16: Securing GKE Deployments

### Scenario

You are managing the production deployment to a set of GKE clusters and want to ensure only images built by your trusted CI/CD pipeline are deployed.

### Question

What should you do?

- A. Enable Cloud Security Scanner on the clusters.
- B. Enable Vulnerability Analysis on the Container Registry.
- C. Set up the Kubernetes Engine clusters as private clusters.
- D. Set up the Kubernetes Engine clusters with Binary Authorization.**

### ✓ Correct Answer: D

*Explanation: Binary Authorization ensures that only signed and approved container images can be deployed to your GKE clusters, enforcing your CI/CD pipeline's approval process.*

---

## Question 17: Monitoring Container Resource Usage in GKE

### Scenario

You support an e-commerce application with a microservices architecture running on a large GKE cluster.

### Question

How can you identify which containers are using the most CPU and memory?

**A. Use Stackdriver Kubernetes Engine Monitoring.**

- B. Use Prometheus to collect and aggregate logs per container, and then analyze the results in Grafana.
- C. Use the Stackdriver Monitoring API to create custom metrics, and then organize your containers using groups.
- D. Use Stackdriver Logging to export application logs to BigQuery, aggregate logs per container, and then analyze CPU and memory consumption.

**✓ Correct Answer: A**

*Explanation:* Stackdriver Kubernetes Engine Monitoring (now GKE Monitoring) provides built-in visibility into container-level CPU and memory metrics without additional configuration.

## Question 18: Improving Production Stability

### Scenario

Your company experiences frequent bugs and outages because developers and testers use the production environment for development, experiments, and load testing.

### Question

How should you redesign the environment to reduce production issues?

- A. Create an automated testing script in production to detect failures as soon as they occur.
- B. Create a development environment with smaller server capacity and give access only to developers and testers.
- C. Secure the production environment to ensure that developers can't change it and set up one controlled update per year.
- D. Create a development environment for writing code and a test environment for configurations, experiments, and load testing.**

**✓ Correct Answer: D**

*Explanation:* Separate development and test environments from production prevent experimental changes from affecting real users and allow for proper load testing in isolation.

## Question 19: Monitoring App Engine Connections

### Scenario

You support a global application running on App Engine and want to monitor the number of connections using Stackdriver Monitoring.

### Question

What metric should you use?

- A. `flex/connections/current`
- B. `tcp\_ssl\_proxy/new\_connections`
- C. `tcp\_ssl\_proxy/open\_connections`
- D. `flex/instance/connections/current`

### ✓ Correct Answer: A

*Explanation: The flex/connections/current metric tracks the number of concurrent connections for App Engine flexible environment instances.*

---

## Question 20: Troubleshooting Database Timeouts

### Scenario

After an application update, users are reporting database timeout errors. The application runs on Compute Engine and connects to Cloud SQL. The number of concurrent users is stable.

### Question

What should you do to find the most probable cause of the timeouts?

- A. Check the serial port logs of the Compute Engine instance.
- B. Use Stackdriver Profiler to visualize the resource utilization throughout the application.**
- C. Determine whether there is an increased number of connections to the Cloud SQL instance.
- D. Use Cloud Security Scanner to see whether your Cloud SQL is under a DDoS attack.

### ✓ Correct Answer: B

*Explanation: Stackdriver Profiler can help identify inefficient code or resource contention within the application that could be causing the database timeouts.*

---

## Question 21: Versioning Docker Images

### Scenario

Your application images are built using Cloud Build and pushed to GCR. You want to specify a particular application version for deployment based on the release version tagged in source control.

### Question

What should you do when you push the image?

- A. Reference the image digest in the source control tag.
- B. Supply the source control tag as a parameter within the image name.
- C. Use Cloud Build to include the release version tag in the application image.**
- D. Use GCR digest versioning to match the image to the tag in source control.

### ✓ Correct Answer: C

*Explanation: Use Cloud Build to tag images with release version information from source control, making it easy to track which code version is deployed.*

---

## Question 22: SRE Incident Management First Steps

### Scenario

You are on-call for a critical infrastructure service. A major outage is affecting all dependent systems and hundreds of thousands of users. You have declared yourself Incident Commander (IC) and have assigned an Operations Lead (OL) and Communications Lead (CL).

### Question

What should you do next?

- A. Look for ways to mitigate user impact and deploy the mitigations to production.
- B. Contact the affected service owners and update them on the status of the incident.
- C. Establish a communication channel where incident responders and leads can communicate with each other.**
- D. Start a postmortem, add incident information, and circulate the draft internally.

### ✓ Correct Answer: C

*Explanation: The immediate next step after assembling the initial incident team is to establish a clear communication channel (e.g., a chat room, video call) for coordination.*

---

## Question 23: Stackdriver Workspace Strategy

### Scenario

You are developing a monitoring strategy for your GCP projects using Stackdriver Workspaces. You need to quickly identify production issues without false alerts from dev/staging, while adhering to the principle of least privilege.

### Question

What should you do?

- A. Grant relevant team members read access to all GCP production projects. Create Stackdriver workspaces inside each project.
- B. Grant relevant team members the Project Viewer IAM role on all GCP production projects. Create Stackdriver workspaces inside each project.
- C. Choose an existing GCP production project to host the monitoring workspace. Attach the production projects to this workspace. Grant relevant team members read access to the Stackdriver Workspace.
- D. Create a new GCP monitoring project and create a Stackdriver Workspace inside it. Attach the production projects to this workspace. Grant relevant team members read access to the Stackdriver Workspace.**

### ✓ Correct Answer: D

**Explanation:** Create a separate monitoring project to centralize monitoring and apply least privilege by granting access only to the monitoring workspace, not the production projects themselves.

---

## Question 24: Real-time VM Utilization Dashboard

### Scenario

You store VM utilization logs in Stackdriver and need to create an easy-to-share, real-time, interactive dashboard with quarterly aggregated information.

### Question

What GCP solution should you use?

- A. 1. Export VM utilization logs from Stackdriver to BigQuery. 2. Create a dashboard in Data Studio. 3. Share the dashboard with your stakeholders.**
- B. 1. Export logs to Cloud Pub/Sub, then to a SIEM system. 2. Build dashboards in the SIEM.
- C. 1. Export logs to BigQuery, then to a CSV file. 2. Import into Google Sheets and build a dashboard.
- D. 1. Export logs to a Cloud Storage bucket. 2. Build a custom data visualization application to pull and display the logs.

### ✓ Correct Answer: A

**Explanation:** Export logs to BigQuery for powerful querying and aggregation, then use Data Studio for easy-to-share, interactive dashboards with real-time updates.

---

## Question 25: Lowering Compute Engine Costs for Stable Workloads

### Scenario

You need to run a business-critical, stable workload on a fixed set of Compute Engine instances for several months and want to lower costs without impacting performance.

### Question

What should you do?

**A. Purchase Committed Use Discounts.**

- B. Migrate the instances to a Managed Instance Group.
- C. Convert the instances to preemptible virtual machines.
- D. Create an Unmanaged Instance Group for the instances.

**✓ Correct Answer: A**

*Explanation: Committed Use Discounts (CUDs) provide significant discounts in exchange for committing to a certain level of resource usage for a 1 or 3-year term, perfect for stable workloads.*

---

## Question 26: Production Readiness Review (PRR) Outcome

### Scenario

You are part of an SRE team taking over a new service. After a Production Readiness Review (PRR), you determine the service cannot currently meet its Service Level Objectives (SLOs).

### Question

To ensure the service can meet its SLOs in production, what should you do next?

- A. Adjust the SLO targets to be achievable by the service so you can bring it into production.
- B. Notify the development team that they will have to provide production support for the service.
- C. Identify recommended reliability improvements to the service to be completed before handover.**
- D. Bring the service into production with no SLOs and build them when you have collected operational data.

**✓ Correct Answer: C**

*Explanation: The outcome of a failed PRR is a list of actionable items for the development team to address to improve reliability before the SRE team accepts operational ownership.*

---

## Question 27: Responding to a Failing Canary Release

### Scenario

Shortly after deploying a new feature as a canary release, you see a spike in 500 errors and increased latency.

### Question

What should you do first to quickly minimize the negative impact on users?

- A. **Roll back the experimental canary release.**
- B. Start monitoring latency, traffic, errors, and saturation.
- C. Record data for the postmortem document of the incident.
- D. Trace the origin of 500 errors and the root cause of increased latency.

### ✓ Correct Answer: A

*Explanation: When a canary release is causing errors or performance degradation, the first priority is to minimize user impact by rolling back immediately.*

---

## Question 28: Collaborative Terraform Development

### Scenario

You are responsible for Terraform templates, and two new engineers are joining the team. You need a process to prevent conflicting changes and ensure all updates are captured.

### Question

What should you do?

- A. Store code in Git. Allow developers to merge their own changes daily. Package and upload to a versioned Cloud Storage bucket.
- B. Store code in a Git-based version control system. Establish a process with peer code reviews and unit testing before integration. The integrated code in the repository becomes the master version.**
- C. Store code as text files in Google Drive. Manually confirm changes daily and rename the folder with a new version number.
- D. Store code as text files in Google Drive. Manually confirm changes, create a .zip archive, and upload it to a versioned Cloud Storage bucket.

### ✓ Correct Answer: B

*Explanation: Use Git-based version control with peer reviews and automated testing to prevent conflicting changes and ensure code quality before integration.*

---

## Question 29: SLI for Graceful Degradation

### Scenario

You support a high-traffic web application with a microservice architecture. The home page displays widgets from various microservices. When a microservice fails, the page loads with some content missing (degraded mode). Users prefer this to a complete failure.

### Question

What Service Level Indicator (SLI) should you use to set a Service Level Objective (SLO) for this user experience?

**A. A quality SLI: the ratio of non-degraded responses to total responses.**

- B. An availability SLI: the ratio of healthy microservices to the total number of microservices.
- C. A freshness SLI: the proportion of widgets that have been updated within the last 10 minutes.
- D. A latency SLI: the ratio of microservice calls that complete in under 100 ms.

✓ **Correct Answer: A**

**Explanation:** A quality SLI measures the proportion of responses that meet quality standards (non-degraded), directly reflecting the user experience you want to maintain.

---

## Question 30: End-to-End Availability SLI

### Scenario

You support a multi-region web service on GKE behind a Global CLB. User requests first go through a third-party CDN. You have an availability SLI at the CLB level but want to increase coverage for issues like CDN failures or global networking problems.

### Question

Where should you measure this new, more comprehensive SLI? (Choose two.)

- A. Your application servers' logs.
- B. Instrumentation coded directly in the client.**
- C. Metrics exported from the application servers.
- D. GKE health checks for your application servers.

✓ **Correct Answer: B**

**Explanation:** To get a true end-to-end perspective, you must measure from the client-side, either through real user monitoring (instrumentation in the client) or synthetic monitoring (a simulated client).

## Question 31: Monitoring Application Metrics in GKE

### Scenario

Your team is designing a new application for GKE and needs to set up monitoring to collect and aggregate various application-level metrics in a centralized location.

### Question

What should you do to minimize the amount of work required, using GCP services?

- A. Publish various metrics from the application directly to the Stackdriver Monitoring API, and then observe these custom metrics in Stackdriver.
- B. Install the Cloud Pub/Sub client libraries, push metrics to various topics, and then observe the aggregated metrics in Stackdriver.
- C. Install the OpenTelemetry client libraries, configure Stackdriver as the export destination, and then observe the metrics in Stackdriver.**
- D. Emit all metrics as log messages and pass them to the Stackdriver logging collector.

### ✓ Correct Answer: C

*Explanation: OpenTelemetry is the industry-standard, vendor-neutral way to instrument applications for metrics and traces, with native support for exporting data to Stackdriver/Cloud Monitoring.*

---

## Question 32: Automating Service Recovery

### Scenario

You support a production service on a single Compute Engine instance that crashes regularly, requiring you to manually delete the instance and create a new one.

### Question

How can you reduce manual operations while following SRE principles?

- A. File a bug with the development team so they can find the root cause of the crashing instance.
- B. Create a Managed Instance Group with a single instance and use health checks to determine the system status.**
- C. Add a Load Balancer in front of the Compute Engine instance and use health checks.
- D. Create a Stackdriver Monitoring dashboard with SMS alerts to be able to start recreating the crashed instance promptly.

### ✓ Correct Answer: B

*Explanation: A Managed Instance Group (MIG) with a health check will automatically recreate the instance if it becomes unhealthy, automating the recovery process.*

---

## Question 33: Securely Managing Secrets in a CI/CD Pipeline

### Scenario

Your application artifacts are built and deployed via a CI/CD pipeline. You want the pipeline to securely access application secrets and also be able to rotate them easily.

### Question

What should you do?

- A. Prompt developers for secrets at build time.
- B. Store secrets in a separate configuration file on Git.
- C. Store secrets in Cloud Storage encrypted with a key from Cloud KMS. Provide the CI/CD pipeline with access to Cloud KMS via IAM.**
- D. Encrypt the secrets and store them in the source code repository. Store a decryption key in a separate repository.

### ✓ Correct Answer: C

*Explanation: Cloud KMS provides secure key management and encryption, and granting CI/CD pipeline IAM access to decrypt secrets is the recommended secure approach for managing secrets.*

---

## Question 34: Incident Communication Strategy

### Scenario

You are the Communications Lead for a large, ongoing incident affecting customer-facing applications. There is no ETA for resolution. You are receiving emails from both internal stakeholders and customers.

### Question

How can you efficiently provide updates to everyone affected?

- A. Focus on responding to internal stakeholders at least every 30 minutes. Commit to "next update" times.
- B. Provide periodic updates to all stakeholders in a timely manner. Commit to a "next update" time in all communications.**
- C. Delegate internal emails to another team member and focus on responding directly to customers.
- D. Provide all internal emails to the Incident Commander and focus on responding directly to customers.

### ✓ Correct Answer: B

*Explanation: Effective incident communication involves providing regular, consistent updates to \*all\* affected parties and managing expectations by committing to a time for the next update.*

---

## Question 35: Authenticating Cloud Build to GKE

### Scenario

Your team uses Cloud Build for all CI/CD pipelines. You want to use the `kubectl` builder to deploy new images to GKE.

### Question

What is the most straightforward way to authenticate to GKE?

- A. Assign the Container Developer role to the Cloud Build service account.**
- B. Specify the Container Developer role for Cloud Build in the `cloudbuild.yaml` file.
- C. Create a new service account with the Container Developer role and use it to run Cloud Build.
- D. Create a separate step in Cloud Build to retrieve service account credentials and pass these to `kubectl`.

### ✓ Correct Answer: A

*Explanation: Cloud Build runs as a service account. Granting this service account the necessary IAM roles (like 'Kubernetes Engine Developer') is the standard way to give it permissions to access other GCP services.*

---

## Question 36: Visualizing Cache Misses

### Scenario

You support an application that logs an entry to Stackdriver Logging for every cache miss.

### Question

How can you visualize the frequency of cache misses over time?

- A. Link Stackdriver Logging as a source in Google Data Studio. Filter the logs on the cache misses.
- B. Configure Stackdriver Profiler to identify and visualize when the cache misses occur based on the logs.
- C. Create a logs-based metric in Stackdriver Logging and a dashboard for that metric in Stackdriver Monitoring.**
- D. Configure BigQuery as a sink for Stackdriver Logging. Create a scheduled query to filter the cache miss logs.

### ✓ Correct Answer: C

*Explanation: Logs-based metrics allow you to create metrics from log entries, which can then be visualized in Monitoring dashboards to track trends over time.*

---

## Question 37: Capacity Planning for a Multi-Region Service

### Scenario

You need to deploy a new, resource-intensive service to production. The service will use a multi-region Managed Instance Group (MIG) for auto-scaling.

### Question

What should you do to plan for capacity?

- A. Use the `n1-highcpu-96` machine type in the MIG configuration.
- B. Monitor results of Stackdriver Trace to determine the required amount of resources.
- C. Validate that the resource requirements are within the available quota limits of each region.**
- D. Deploy the service in one region and use a global load balancer to route traffic.

### ✓ Correct Answer: C

*Explanation: Before deploying resource-intensive services, verify that you have sufficient quota in each region to avoid deployment failures due to resource constraints.*

---

## Question 38: Handling PII in Logs

### Scenario

You are collecting logs with Stackdriver from an application on Compute Engine. You discover that personally identifiable information (PII) is leaking into log fields. All PII entries begin with the text `userinfo`.

### Question

How can you capture these log entries for secure review while preventing them from leaking to the main Stackdriver Logging view?

- A. Create a basic log filter matching `userinfo`, and then configure a log export to Cloud Storage.
- B. Use a Fluentd filter plugin with the Stackdriver Agent to remove log entries containing `userinfo`, and then copy the entries to a Cloud Storage bucket.
- C. Create an advanced log filter matching `userinfo`, configure a log export to Cloud Storage, and then configure a log exclusion with `userinfo` as a filter.**
- D. Use a Fluentd filter plugin to remove entries, create an advanced log filter, and configure a log export.

### ✓ Correct Answer: C

*Explanation: The correct approach is to use a combination of a sink (export) to capture the sensitive logs for secure review and an exclusion filter to prevent them from being ingested into the general Logs Explorer.*

---

## Question 39: Troubleshooting a Failing Cloud Build Pipeline

### Scenario

You have a CI/CD pipeline that uses Cloud Build. After a change in the Cloud Build YAML configuration, the pipeline is no longer building new artifacts.

### Question

Following SRE practices, what should you do to resolve the issue?

- A. Disable the CI pipeline and revert to manually building and pushing the artifacts.
- B. Change the CI pipeline to push the artifacts to Container Registry instead of Docker Hub.
- C. Upload the configuration YAML file to Cloud Storage and use Error Reporting to identify and fix the issue.
- D. Run a Git compare between the previous and current Cloud Build Configuration files to find and fix the bug.**

### ✓ Correct Answer: D

*Explanation: The most direct way to find the cause of a problem introduced by a configuration change is to diff the current version against the last known good version.*

---

## Question 40: Preventing Future Incidents

### Scenario

Your company follows SRE principles. You are writing a postmortem for a severe, user-affecting incident that was triggered by a software change.

### Question

What should you do to prevent similar severe incidents in the future?

- A. Identify engineers responsible for the incident and escalate to their senior management.
- B. Ensure that test cases that catch errors of this type are run successfully before new software releases.**
- C. Follow up with the employees who reviewed the changes and prescribe practices they should follow in the future.
- D. Design a policy that will require on-call teams to immediately call engineers and management if an incident occurs.

### ✓ Correct Answer: B

*Explanation: The best way to prevent similar incidents is to add automated test coverage that would have caught the issue before it reached production.*

---

## Question 41: Measuring Application Reliability from a User Perspective

### Scenario

You support a high-traffic web application on GCP and need to measure its reliability from a user's perspective without making any engineering changes to the application itself.

### Question

What should you do? (Choose two.)

- A. Review current application metrics and add new ones as needed.
- B. Modify the code to capture additional information for user interaction.
- C. Analyze the web proxy logs only and capture response time of each request.
- D. Create new synthetic clients to simulate a user journey using the application.**

### ✓ Correct Answer: D

*Explanation: To measure reliability without code changes, you can either analyze existing server-side logs (passive) or simulate user behavior with synthetic clients (active).*

---

## Question 42: Granting Log Export Permissions

### Scenario

You manage an application that writes logs to Stackdriver Logging. You need to give some team members the ability to export logs.

### Question

What should you do?

- A. Grant the team members the IAM role of `logging.configWriter` on Cloud IAM.**
- B. Configure Access Context Manager to allow only these members to export logs.
- C. Create and grant a custom IAM role with the permissions `logging.sinks.list` and `logging.sink.get` .
- D. Create an Organizational Policy in Cloud IAM to allow only these members to create log exports.

### ✓ Correct Answer: A

*Explanation: The logging.configWriter role provides permissions to create and manage log exports (sinks) in Cloud Logging.*

---

## Question 43: Restricting GKE Image Sources

### Scenario

Your application services run in GKE. You want to ensure that only images from your centrally managed GCR registry in the `altostrat-images` project can be deployed to the cluster.

### Question

What should you do to achieve this with minimal development time?

- A. Create a custom builder for Cloud Build that will only push images to `gcr.io/altostrat-images`.
- B. Use a Binary Authorization policy that includes the whitelist name pattern `gcr.io/altostrat-images`.**
- C. Add logic to the deployment pipeline to check that all manifests contain only images from `gcr.io/altostrat-images`.
- D. Add a tag to each image in `gcr.io/altostrat-images` and check for this tag at deployment.

### ✓ Correct Answer: B

*Explanation: Binary Authorization is the GKE-native feature designed specifically for creating policies that restrict deployments to trusted image registries.*

---

## Question 44: Scaling a GKE Application Based on Traffic

### Scenario

Your team has deployed an NGINX-based application to GKE, exposed via an HTTP GCLB ingress. You want to scale the frontend deployment based on an appropriate SLI.

### Question

What should you do?

- A. Configure the horizontal pod autoscaler (HPA) to use the average response time from the Liveness and Readiness probes.
- B. Configure the vertical pod autoscaler (VPA) and enable the cluster autoscaler.
- C. Install the Stackdriver custom metrics adapter and configure an HPA to use the number of requests provided by the GCLB.**
- D. Expose the NGINX stats endpoint and configure the HPA to use the request metrics exposed by the NGINX deployment.

### ✓ Correct Answer: C

*Explanation: To scale based on external metrics like load balancer requests, you need to install the custom metrics adapter, which makes these metrics available to the HPA.*

---

## Question 45: Initial Incident Management Roles

### Scenario

Your company follows SRE practices. You are the Incident Commander for a new, customer-impacting incident.

### Question

What two incident management roles should you assign immediately to assist in an effective response?

- A. Operations Lead**
- B. Engineering Lead
- C. Communications Lead**
- D. Customer Impact Assessor

### ✓ Correct Answer: A & C

***Explanation:** The two most critical roles to fill immediately after the Incident Commander are the Operations Lead (to manage the technical response) and the Communications Lead (to manage all communications).*

---

## Question 46: Configuring SMS Alerts in Stackdriver

### Scenario

You support an application on GCP and want to configure SMS notifications for the most critical alerts in Stackdriver Monitoring.

### Question

What should you do?

- A. Download and configure a third-party integration between Stackdriver Monitoring and an SMS gateway.
- B. Select the Webhook notifications option and configure it to use a third-party integration tool.
- C. Ensure your team members set their SMS/phone numbers in their Stackdriver Profile. Select the SMS notification option for each alerting policy and select the appropriate numbers.**
- D. Configure a Slack notification and use a Slack-to-SMS integration.

### ✓ Correct Answer: C

***Explanation:** Stackdriver Monitoring supports SMS notifications when team members configure their phone numbers in their profiles and SMS is selected as a notification channel.*

---

## Question 47: Visualizing HTTP Latency Distribution

### Scenario

You are managing an application that exposes an HTTP endpoint without a load balancer. The latency of HTTP responses is critical, and you want to understand the full distribution of latencies experienced by your users.

### Question

Using Stackdriver Monitoring, what should you do?

- A. Create a `DELTA` metric with `DOUBLE` valueType and use a Stacked Bar graph.
- B. Create a `CUMULATIVE` metric with `DOUBLE` valueType and use a Line graph.
- C. In your application, create a metric with a `metricKind` set to `GAUGE` and a `valueType` set to `DISTRIBUTION`. In Stackdriver's Metrics Explorer, use a Heatmap graph to visualize the metric.**
- D. Create a `METRIC\_KIND\_UNSPECIFIED` metric with `INT64` valueType and use a Stacked Area graph.

### ✓ Correct Answer: C

*Explanation: DISTRIBUTION metrics are designed for tracking latency distributions, and Heatmap charts are ideal for visualizing how latency values are distributed over time.*

---

## Question 48: Collecting Detailed Metrics Inside and Outside GCP

### Scenario

Your team is designing a new application for deployment both inside and outside of GCP. You need to collect detailed metrics like system resource utilization and want to use centralized GCP services with minimal setup.

### Question

What should you do?

- A. Import the Stackdriver Profiler package and configure it to relay function timing data to Stackdriver.**
- B. Import the Stackdriver Debugger package and configure the application to emit debug messages with timing information.
- C. Instrument the code using a timing library and publish the metrics via a health check endpoint that is scraped by Stackdriver.
- D. Install an Application Performance Monitoring (APM) tool in both locations and configure an export to a central data storage location.

### ✓ Correct Answer: A

*Explanation: Stackdriver Profiler is designed to work in hybrid and multi-cloud environments, allowing you to collect performance data from applications running anywhere.*

---

## Question 49: Suitable Workloads for Preemptible VMs

### Scenario

You need to reduce the cost of VMs for your organization and decide to use preemptible VM instances.

### Question

Which applications are suitable for preemptible VMs? (Choose two.)

- A. A scalable in-memory caching system.
- B. The organization's public-facing website.
- C. A distributed, eventually consistent NoSQL database cluster with sufficient quorum.**
- D. A GPU-accelerated video rendering platform that retrieves and stores videos in a storage bucket.**

### ✓ Correct Answer: C & D

**Explanation:** Preemptible VMs are ideal for fault-tolerant and stateless batch processing workloads, like rendering jobs or distributed databases with sufficient redundancy where instances can be replaced without major impact.

---

## Question 50: Enforcing Deployment Approvals in a Container Workflow

### Scenario

Your organization uses a container-based workflow with a continuous deployment pipeline to a production Kubernetes cluster. A security auditor is concerned that developers could push unapproved changes to production.

### Question

What should you do to enforce approvals?

- A. Configure the build system with protected branches that require pull request approval.
- B. Use an Admission Controller to verify that incoming requests originate from approved sources.
- C. Leverage Kubernetes Role-Based Access Control (RBAC) to restrict access to only approved users.
- D. Enable binary authorization inside the Kubernetes cluster and configure the build pipeline as an attester.**

### ✓ Correct Answer: D

**Explanation:** Binary Authorization with attestations provides a cryptographically verifiable way to ensure that only images that have passed specific stages (like QA approval) in your pipeline can be deployed.

---

## Question 51: Question Removed or Missing

### Scenario

You are running a real-time gaming application on Compute Engine with separate production and testing VPCs. You suspect a malicious process is communicating intermittently from your production frontend servers.

### Question

How can you ensure that network traffic is captured for analysis?

- A. Enable VPC Flow Logs on the production VPC network frontend and backend subnets only with a sample volume scale of 0.5.
- B. Enable VPC Flow Logs on the production VPC network frontend and backend subnets only with a sample volume scale of 1.0.**
- C. Enable VPC Flow Logs on the testing and production VPCs with a volume scale of 0.5.
- D. Enable VPC Flow Logs on the testing and production VPCs with a volume scale of 1.0.

### ✓ Correct Answer: B

**Explanation:** To capture intermittent traffic reliably, enable VPC Flow Logs with a sampling rate of 1.0 (100%) on the production frontend and backend subnets where the suspicious activity occurs.

---

## Question 52: 53: Collaborative Terraform Workflow

### Scenario

Your team of Infrastructure DevOps Engineers is growing, and you are starting to use Terraform to manage infrastructure.

### Question

What is the proper way to implement code versioning and sharing?

- A. Store the Terraform code in a version-control system. Establish procedures for pushing new versions and merging with the master.**
- B. Store the Terraform code in a network shared folder with child folders for each version release.
- C. Store the Terraform code in a Cloud Storage bucket using object versioning.
- D. Store the Terraform code in a shared Google Drive folder.

### ✓ Correct Answer: A

**Explanation:** Use Git-based version control for Infrastructure as Code with proper branching, pull requests, and merge procedures to enable collaboration and maintain code history.

---

## Question 53: 54: Troubleshooting Missing Logs in Stackdriver

### Scenario

You are using Stackdriver to monitor applications on GCP. After deploying a new application, its logs are not appearing on the Stackdriver dashboard.

### Question

What should you do to troubleshoot the issue?

- A. **Confirm that the Stackdriver agent has been installed in the hosting virtual machine.**
- B. Confirm that your account has the proper permissions to use the Stackdriver dashboard.
- C. Confirm that port 25 has been opened in the firewall.
- D. Confirm that the application is using the required client library and the service account key has proper permissions.

### ✓ Correct Answer: A

*Explanation: The most fundamental requirement for collecting logs from a VM is that the Stackdriver Logging agent must be installed and running correctly.*

---

## Question 54: 55: Container Vulnerability Scanning

### Scenario

Your organization has adopted a container-based workflow with a CI/CD pipeline. A security audit raised concerns about vulnerabilities in the code being pushed to production.

### Question

What should you do to ensure the security and patch level of all code running through the pipeline?

- A. **Set up Container Analysis to scan and report Common Vulnerabilities and Exposures (CVEs).**
- B. Configure the containers in the build pipeline to always update themselves before release.
- C. Reconfigure the existing OS vulnerability software to exist inside the container.
- D. Implement static code analysis tooling against the Dockerfiles.

### ✓ Correct Answer: A

*Explanation: Container Analysis (now part of Artifact Analysis) automatically scans container images for known vulnerabilities (CVEs) when images are pushed to Container Registry.*

---

## Question 55: 56: Reducing Cloud Build Time

### Scenario

You use Cloud Build to build your application and want to reduce the build time while minimizing cost and development effort.

### Question

What should you do?

- A. Use Cloud Storage to cache intermediate artifacts.**
- B. Run multiple Jenkins agents to parallelize the build.
- C. Use multiple smaller build steps to minimize execution time.
- D. Use larger Cloud Build virtual machines (VMs) by using the `machine-type` option.

✓ Correct Answer: A

---

## Question 56: 57: Mitigating an Incident Caused by a New Release

### Scenario

You support a web application on Compute Engine. Shortly after a new feature release, your monitoring shows that all users are experiencing latency at login.

### Question

What should you do first to mitigate the impact on users?

- A. Roll back the recent release.**
- B. Review the Stackdriver monitoring.
- C. Upsize the virtual machines running the login services.
- D. Deploy a new release to see whether it fixes the problem.

✓ Correct Answer: A

*Explanation: When a new release causes immediate user-facing issues, the first priority is to mitigate impact by rolling back to the last known good version.*

---

## Question 57: 58: Managing Sensitive Information

### Scenario

You are deploying an application that needs to access sensitive information. You need to ensure this information is encrypted and the risk of exposure is minimal.

### Question

What should you do?

- A. Store the encryption keys in Cloud Key Management Service (KMS) and rotate the keys frequently.
- B. Inject the secret at the time of instance creation via an encrypted configuration management system.
- C. Integrate the application with a Single Sign-On (SSO) system.
- D. Leverage a continuous build pipeline that produces multiple versions of the secret for each instance.

✓ Correct Answer: A

---

## Question 58: 59: Reducing Alert Fatigue

### Scenario

You encounter a large number of outages in your production systems. You receive alerts for all of them, even though the systems are automatically restarted within a minute. This is causing staff burnout.

### Question

Following SRE practices, what should you do to prevent burnout?

- A. Eliminate unactionable alerts.
- B. Create an incident report for each of the alerts.
- C. Distribute the alerts to engineers in different time zones.
- D. Redefine the related SLO so that the error budget is not exhausted.

✓ Correct Answer: A

---

*Explanation: Every alert should be actionable. If an issue is resolved automatically and requires no human intervention, it should not trigger a page or alert.*

## Question 59: 60: Preparing for a Busy Season

### Scenario

You have migrated an e-commerce application to GCP and want to prepare it for the upcoming busy season.

### Question

What should you do first?

- A. Load test the application to profile its performance for scaling.**
- B. Enable AutoScaling on the production clusters, in case there is growth.
- C. Pre-provision double the compute power used last season, expecting growth.
- D. Create a runbook on inflating the disaster recovery (DR) environment if there is growth.

### ✓ Correct Answer: A

*Explanation: Before a busy season, perform load testing to understand your application's performance characteristics and identify bottlenecks or scaling limits.*

---

## Question 60: 61: Troubleshooting App Engine Latency

### Scenario

You support a web application on App Engine using Cloud SQL and Cloud Storage. After a traffic spike, latency remains high for all requests even though traffic levels have returned to normal. CPU use and the number of running processes are also high.

### Question

You expect another traffic spike soon. What should you do to prevent latency?

- A. Upgrade the GCS buckets to Multi-Regional.
- B. Enable high availability on the Cloud SQL instances.
- C. Move the application from App Engine to Compute Engine.
- D. Modify the App Engine configuration to have additional idle instances.**

### ✓ Correct Answer: D

*Explanation: The symptoms point to a "cold start" problem where new instances are being created too slowly. Increasing the number of idle instances ensures that pre-warmed instances are ready to handle sudden traffic spikes.*

---

## **Question 61: 57**

### **Question**

You support a web application that is hosted on Compute Engine. The application provides a booking service for thousands of users. Shortly after the release of a new feature, your monitoring dashboard shows that all users are experiencing latency at login. You want to mitigate the impact of the incident on the users of your service. What should you do first?

- A. Roll back the recent release. Most Voted
- 

## **Question 62: Implementing Jenkins on GCP**

### **Scenario**

Your application runs on GCP. You need to implement Jenkins for deploying application releases to GCP.

### **Question**

To streamline the release process, lower operational toil, and keep user data secure, what should you do?

- A. Implement Jenkins on local workstations.
- B. Implement Jenkins on Kubernetes on-premises.
- C. Implement Jenkins on Google Cloud Functions.
- D. Implement Jenkins on Compute Engine virtual machines.**

### **✓ Correct Answer: D**

**Explanation:** Running Jenkins on Compute Engine VMs provides the flexibility, security, and integration with GCP services needed for a production CI/CD system while keeping operational toil manageable.

---

## Question 63: Long-Term Log Archiving

### Scenario

You are working with a government agency that requires you to archive application logs for seven years. You need to configure Stackdriver to export and store these logs while minimizing storage costs.

### Question

What should you do?

- A. Create a Cloud Storage bucket and develop your application to send logs directly to it.
- B. Develop an App Engine application that pulls logs from Stackdriver and saves them in BigQuery.
- C. Create an export in Stackdriver and configure Cloud Pub/Sub to store logs in permanent storage.
- D. Create a sink in Stackdriver, create a Cloud Storage bucket for archiving, and select the bucket as the log export destination.**

✓ Correct Answer: D

---

## Question 64: Customizing Error Reporting in App Engine

### Scenario

You support a Python trading application on App Engine flexible environment and want to customize the error information being sent to Stackdriver Error Reporting.

### Question

What should you do?

- A. Install the Stackdriver Error Reporting library for Python and run your code on a Compute Engine VM.
- B. Install the Stackdriver Error Reporting library for Python and run your code on GKE.
- C. Install the Stackdriver Error Reporting library for Python and run your code on App Engine flexible environment.
- D. Use the Stackdriver Error Reporting API to write errors from your application to `ReportedErrorEvent`, and then generate log entries with properly formatted error messages in Stackdriver Logging.**

✓ Correct Answer: D

**Explanation:** For maximum control over the error data sent, you should use the Error Reporting API directly to construct a `ReportedErrorEvent` and log it in the specific JSON format that Error Reporting expects.

---

## Question 65: Setting a Latency SLO

### Scenario

You need to define latency SLOs for a high-traffic, multi-region web application. Customers are currently happy. Current measurements over a 28-day window show the 90th percentile of latency is 120ms and the 95th percentile is 275ms.

### Question

What latency SLO would you recommend publishing?

- A. 90th percentile - 100ms; 95th percentile - 250ms
- B. 90th percentile - 120ms; 95th percentile - 275ms
- C. 90th percentile - 150ms; 95th percentile - 300ms**
- D. 90th percentile - 250ms; 95th percentile - 400ms

### ✓ Correct Answer: C

*Explanation: Set your SLO slightly looser than your current performance to create an "error budget" that allows for some performance degradation without breaching the SLO. This provides a buffer for deployments and minor issues.*

---

## Question 66: Question Removed

### Scenario

Your company is developing applications on GKE. Each team manages a different application and needs separate development and production environments. Teams should not be able to access other teams' environments.

### Question

How should you set this up to minimize costs?

- A. Create one GCP Project per team. In each project, create a cluster for Development and one for Production.
- B. Create one GCP Project per team. In each project, create a single cluster with a Kubernetes namespace for Development and one for Production.
- C. Create a Development and a Production GKE cluster in separate projects. In each cluster, create a Kubernetes namespace per team, and then configure Identity-Aware Proxy.
- D. Create a Development and a Production GKE cluster in separate projects. In each cluster, create a Kubernetes namespace per team, and then configure Kubernetes Role-based access control (RBAC).**

### ✓ Correct Answer: D

*Explanation: The most cost-effective and standard way to provide multi-tenancy within a GKE cluster is to use Kubernetes namespaces for logical separation and RBAC for access control.*

---

## Question 67: 68: Optimizing GCR Image Push Bandwidth

### Scenario

Some of your production services run in GKE in the `eu-west-1` region. Your build system runs in the `us-west-1` region. You want to push container images from your build system to a scalable registry to maximize the bandwidth for transferring the images to the cluster.

### Question

What should you do?

- A. Push the images to Google Container Registry (GCR) using the `gcr.io` hostname.
- B. Push the images to GCR using the `us.gcr.io` hostname.
- C. Push the images to GCR using the `eu.gcr.io` hostname.**
- D. Push the images to a private image registry running on a Compute Engine instance in `eu-west-1`.

### ✓ Correct Answer: C

*Explanation: Use the regional GCR endpoint (eu.gcr.io) closest to your cluster to minimize latency and maximize bandwidth when pulling images.*

---

## Question 68: 69: Cost Breakdown by System

### Scenario

You manage several production systems on Compute Engine in the same GCP project. Each system has its own dedicated set of instances.

### Question

How can you determine how much it costs to run each of the systems?

- A. In the GCP Console, use the Cost Breakdown section to visualize the costs per system.
- B. Assign all instances a label specific to the system they run. Configure BigQuery billing export and query costs per label.**
- C. Enrich all instances with metadata specific to the system they run. Configure Stackdriver Logging to export to BigQuery and query costs based on metadata.
- D. Name each VM after the system it runs. Set up a usage report export to a Cloud Storage bucket and query costs based on VM name.

### ✓ Correct Answer: B

*Explanation: Labels are the standard way to organize and track GCP resources. Export billing data to BigQuery and query by label to get cost breakdown per system.*

---

## Question 69: 70: Securely Using Secrets in Cloud Build

### Scenario

You use Cloud Build to build and deploy your application. You want to securely incorporate database credentials and other secrets into the build pipeline with minimal effort.

### Question

What should you do?

- A. Create a Cloud Storage bucket with encryption at rest. Store secrets in the bucket and grant Cloud Build access.
- B. Encrypt secrets and store them in the application repository. Store a decryption key in a separate repository.
- C. Use client-side encryption for secrets in a Cloud Storage bucket. Store a decryption key in the bucket.
- D. Use Cloud Key Management Service (Cloud KMS) to encrypt the secrets and include them in your Cloud Build deployment configuration. Grant Cloud Build access to the KeyRing.**
- C. Install the Stackdriver Error Reporting library for Python, and then run your code on App Engine flexible environment.
- D. Use the Stackdriver Error Reporting API to write errors from your application to ReportedErrorEvent, and then generate log entries with properly formatted error messages in Stackdriver Logging. Most Voted**

### ✓ Correct Answer: D

**Explanation:** Cloud Build has native integration with Cloud KMS for decrypting encrypted variables at build time. This is the most secure and straightforward method. **D. Use the Stackdriver Error Reporting API to write errors from your application to ReportedErrorEvent, and then generate log entries with properly formatted error messages in Stackdriver Logging. Most Voted**

---

## Question 70: 65

### Question

You need to define Service Level Objectives (SLOs) for a high-traffic multi-region web application. Customers expect the application to always be available and have fast response times. Customers are currently happy with the application performance and availability. Based on current measurement, you observe that the 90 percentile of latency is 120ms and the 95 percentile of latency is 275ms over a 28-day window. What latency SLO would you recommend to the team to th th publish?

- A. 90 percentile => 100ms th 95 percentile => 250ms th
- B. 90 percentile => 120ms th 95 percentile => 275ms th
- C. 90 percentile => 150ms th 95 percentile => 300ms th Most Voted**
- D. 90 percentile => 250ms th 95 percentile => 400ms th

### ✓ Correct Answer: C

---

## Question 71: 67

### Question

Your company is developing applications that are deployed on Google Kubernetes Engine (GKE). Each team manages a different application. You need to create the development and production environments for each team, while minimizing costs. Different teams should not be able to access other teams' environments. What should you do?

- A. Create one GCP Project per team. In each project, create a cluster for Development and one for Production. Grant the teams IAM access to their respective clusters.
- B. Create one GCP Project per team. In each project, create a cluster with a Kubernetes namespace for Development and one for Production. Grant the teams IAM access to their respective clusters.
- C. Create a Development and a Production GKE cluster in separate projects. In each cluster, create a Kubernetes namespace per team, and then configure Identity Aware Proxy so that each team can only access its own namespace.
- D. Create a Development and a Production GKE cluster in separate projects. In each cluster, create a Kubernetes namespace per team, and then configure Kubernetes Role-based access control (RBAC) so that each team can only access its own namespace. Most Voted**

✓ Correct Answer: D

---

## Question 72: 68

### Question

Some of your production services are running in Google Kubernetes Engine (GKE) in the eu-west-1 region. Your build system runs in the us-west-1 region. You want to push the container images from your build system to a scalable registry to maximize the bandwidth for transferring the images to the cluster. What should you do?

- A. Push the images to Google Container Registry (GCR) using the gcr.io hostname.
- B. Push the images to Google Container Registry (GCR) using the us.gcr.io hostname.
- C. Push the images to Google Container Registry (GCR) using the eu.gcr.io hostname. Most Voted**
- D. Push the images to a private image registry running on a Compute Engine instance in the eu-west-1 region.

✓ Correct Answer: C

---

## Question 73: 69

### Question

You manage several production systems that run on Compute Engine in the same Google Cloud Platform (GCP) project. Each system has its own set of dedicated Compute Engine instances. You want to know how much it costs to run each of the systems. What should you do?

- A. In the Google Cloud Platform Console, use the Cost Breakdown section to visualize the costs per system.
- B. Assign all instances a label specific to the system they run. Configure BigQuery billing export and query costs per label. Most Voted**
- C. Enrich all instances with metadata specific to the system they run. Configure Stackdriver Logging to export to BigQuery, and query costs based on the metadata.
- D. Name each virtual machine (VM) after the system it runs. Set up a usage report export to a Cloud Storage bucket. Configure the bucket as a source in BigQuery to query costs based on VM name.

✓ Correct Answer: B

---

## Question 74: 70

### Question

You use Cloud Build to build and deploy your application. You want to securely incorporate database credentials and other application secrets into the build pipeline. You also want to minimize the development effort. What should you do?

- A. Create a Cloud Storage bucket and use the built-in encryption at rest. Store the secrets in the bucket and grant Cloud Build access to the bucket.
- B. Encrypt the secrets and store them in the application repository. Store a decryption key in a separate repository and grant Cloud Build access to the repository.
- C. Use client-side encryption to encrypt the secrets and store them in a Cloud Storage bucket. Store a decryption key in the bucket and grant Cloud Build access to the bucket.
- D. Use Cloud Key Management Service (Cloud KMS) to encrypt the secrets and include them in your Cloud Build deployment configuration. Grant Cloud Build access to the KeyRing. Most Voted**

✓ Correct Answer: D

**Explanation:** Cloud KMS integration with Cloud Build allows secure secret management. Encrypt secrets with KMS, store them in your build config, and grant Cloud Build decrypt permissions for seamless, secure access.

---

## Question 75: 71

### Question

You support a popular mobile game application deployed on Google Kubernetes Engine (GKE) across several Google Cloud regions. Each region has multiple Kubernetes clusters. You receive a report that none of the users in a specific region can connect to the application. You want to resolve the incident while following Site Reliability Engineering practices. What should you do first?

**A. Reroute the user traffic from the affected region to other regions that don't report issues. Most Voted**

- B. Use Stackdriver Monitoring to check for a spike in CPU or memory usage for the affected region.
- C. Add an extra node pool that consists of high memory and high CPU machine type instances to the cluster.
- D. Use Stackdriver Logging to filter on the clusters in the affected region, and inspect error messages in the logs.

**✓ Correct Answer: A**

*Explanation: Following SRE incident management practices, the first priority is to mitigate user impact. Reroute traffic away from the affected region immediately, then investigate the root cause.*

---

## Question 76: 72

### Question

You are writing a postmortem for an incident that severely affected users. You want to prevent similar incidents in the future. Which two of the following sections should you include in the postmortem? (Choose two.)

**A. An explanation of the root cause of the incident. Most Voted**

- B. A list of employees responsible for causing the incident

**C. A list of action items to prevent a recurrence of the incident Most Voted**

- D. Your opinion of the incident's severity compared to past incidents
- E. Copies of the design documents for all the services impacted by the incident

**✓ Correct Answer: A & C**

*Explanation: Blameless postmortems should include root cause analysis and actionable remediation items. Never include blame or opinions - focus on system improvements and preventing recurrence.*

---

## Question 77: 73

### Question

You are ready to deploy a new feature of a web-based application to production. You want to use Google Kubernetes Engine (GKE) to perform a phased rollout to half of the web server pods. What should you do?

- A. Use a partitioned rolling update. **Most Voted**
- B. Use Node taints with NoExecute.
- C. Use a replica set in the deployment specification.
- D. Use a stateful set with parallel pod management policy.

### ✓ Correct Answer: A

*Explanation:* A partitioned rolling update in GKE allows you to control the rollout by updating a specific number or percentage of pods at a time, perfect for phased deployments.

---

## Question 78: 74

### Question

You are responsible for the reliability of a high-volume enterprise application. A large number of users report that an important subset of the application's functionality " a data intensive reporting feature " is consistently failing with an HTTP 500 error. When you investigate your application's dashboards, you notice a strong correlation between the failures and a metric that represents the size of an internal queue used for generating reports. You trace the failures to a reporting backend that is experiencing high I/O wait times. You quickly fix the issue by resizing the backend's persistent disk (PD). How you need to create an availability Service Level Indicator (SLI) for the report generation feature. How would you define it?

- A. As the I/O wait times aggregated across all report generation backends
- B. As the proportion of report generation requests that result in a successful response** **Most Voted**
- C. As the application's report generation queue size compared to a known-good threshold
- D. As the reporting backend PD throughout capacity compared to a known-good threshold

### ✓ Correct Answer: B

*Explanation:* An availability SLI should measure user-facing success. The proportion of successful responses directly reflects the user experience, not internal metrics like queue size or I/O wait.

---

## Question 79: 75

### Question

You have an application running in Google Kubernetes Engine. The application invokes multiple services per request but responds too slowly. You need to identify which downstream service or services are causing the delay. What should you do?

- A. Analyze VPC flow logs along the path of the request.
- B. Investigate the Liveness and Readiness probes for each service.
- C. Create a Dataflow pipeline to analyze service metrics in real time.
- D. Use a distributed tracing framework such as OpenTelemetry or Stackdriver Trace. Most Voted**

### ✓ Correct Answer: D

**Explanation:** Distributed tracing tools like OpenTelemetry or Cloud Trace track requests across multiple services, showing exactly where time is spent and which service causes delays.

---

## Question 80: 76

### Question

You are creating and assigning action items in a postmodern for an outage. The outage is over, but you need to address the root causes. You want to ensure that your team handles the action items quickly and efficiently. How should you assign owners and collaborators to action items?

- A. Assign one owner for each action item and any necessary collaborators. Most Voted**
- B. Assign multiple owners for each item to guarantee that the team addresses items quickly.
- C. Assign collaborators but no individual owners to the items to keep the postmortem blameless.
- D. Assign the team lead as the owner for all action items because they are in charge of the SRE team.

### ✓ Correct Answer: A

**Explanation:** Each action item needs one clear owner accountable for completion, with collaborators as needed. Multiple owners creates confusion, and no owners means no accountability.

---

## Question 81: 77

### Question

Your development team has created a new version of their service's API. You need to deploy the new versions of the API with the least disruption to third-party developers and end users of third-party installed applications. What should you do?

- A. Introduce the new version of the API. Announce deprecation of the old version of the API. Deprecate the old version of the API. Contact remaining users of the old API. Provide best effort support to users of the old API. Turn down the old version of the API. Most Voted**
- B. Announce deprecation of the old version of the API. Introduce the new version of the API. Contact remaining users on the old API. Deprecate the old version of the API. Turn down the old version of the API. Provide best effort support to users of the old API.
- C. Announce deprecation of the old version of the API. Contact remaining users on the old API. Introduce the new version of the API. Deprecate the old version of the API. Provide best effort support to users of the old API. Turn down the old version of the API.
- D. Introduce the new version of the API. Contact remaining users of the old API. Announce deprecation of the old version of the API. Deprecate the old version of the API. Turn down the old version of the API. Provide best effort support to users of the old API.

### ✓ Correct Answer: A

*Explanation: Introduce the new API first (giving users time to migrate), announce deprecation, then deprecate after notice, contact remaining users, provide support, and finally turn down the old version.*

---

## Question 82: 78

### Question

You are running an application on Compute Engine and collecting logs through Stackdriver. You discover that some personally identifiable information (PII) is leaking into certain log entry fields. You want to prevent these fields from being written in new log entries as quickly as possible. What should you do?

- A. Use the filter-record-transformer Fluentd filter plugin to remove the fields from the log entries in flight. Most Voted**
- B. Use the fluent-plugin-record-reformer Fluentd output plugin to remove the fields from the log entries in flight.
- C. Wait for the application developers to patch the application, and then verify that the log entries are no longer exposing PII.
- D. Stage log entries to Cloud Storage, and then trigger a Cloud Function to remove the fields and write the entries to Stackdriver via the Stackdriver Logging API.

### ✓ Correct Answer: A

*Explanation: The filter-record-transformer Fluentd plugin can modify log entries in real-time before they're sent to Cloud Logging, removing PII fields immediately.*

---

## Question 83: 79

### Question

You support a service that recently had an outage. The outage was caused by a new release that exhausted the service memory resources. You rolled back the release successfully to mitigate the impact on users. You are now in charge of the post-mortem for the outage. You want to follow Site Reliability Engineering practices when developing the post-mortem. What should you do?

- A. Focus on developing new features rather than avoiding the outages from recurring.
- B. Focus on identifying the contributing causes of the incident rather than the individual responsible for the cause. Most Voted**
- C. Plan individual meetings with all the engineers involved. Determine who approved and pushed the new release to production.
- D. Use the Git history to find the related code commit. Prevent the engineer who made that commit from working on production services.

### ✓ Correct Answer: B

**Explanation:** Blameless postmortems focus on system improvements, not individual blame. Identify contributing causes and systemic issues to prevent future incidents.

---

## Question 84: 80

### Question

You support a user-facing web application. When analyzing the application's error budget over the previous six months, you notice that the application has never consumed more than 5% of its error budget in any given time window. You hold a Service Level Objective (SLO) review with business stakeholders and confirm that the SLO is set appropriately. You want your application's SLO to more closely reflect its observed reliability. What steps can you take to further that goal while balancing velocity, reliability, and business needs? (Choose two.)

- A. Add more serving capacity to all of your application's zones.
- B. Have more frequent or potentially risky application releases.
- C. Tighten the SLO match the application's observed reliability.
- D. Implement and measure additional Service Level Indicators (SLIs) fro the application. Most Voted**
- E. Announce planned downtime to consume more error budget and ensure that users are not depending on a tighter SLO. Most Voted**

### ✓ Correct Answer: D

**Explanation:** If error budget is underutilized, you can increase velocity with riskier releases, or add SLIs for completeness. Consider if users depend on tighter reliability than the SLO specifies.

---

## Question 85: 81

### Question

You support a service with a well-defined Service Level Objective (SLO). Over the previous 6 months, your service has consistently met its SLO and customer satisfaction has been consistently high. Most of your service's operations tasks are automated and few repetitive tasks occur frequently. You want to optimize the balance between reliability and deployment velocity while following site reliability engineering best practices. What should you do? (Choose two.)

- A. Make the service's SLO more strict.
- B. Increase the service's deployment velocity and/or risk. Most Voted**
- C. Shift engineering time to other services that need more reliability. Most Voted**
- D. Get the product team to prioritize reliability work over new features.
- E. Change the implementation of your Service Level Indicators (SLIs) to increase coverage.

### ✓ Correct Answer: B & C

*Explanation: When a service consistently meets its SLO with low operational toil, increase deployment velocity to deliver more features, and redirect engineering effort to services needing reliability improvements.*

---

## Question 86: 82

### Question

Your company follows Site Reliability Engineering principles. You are writing a postmortem for an incident, triggered by a software change that severely affected users. You want to prevent severe incident from happening in the future. What should you do?

- A. Identify engineers responsible for the incident and escalate to the senior management.
- B. Ensure that test cases that catch errors of this type are run successfully before new software releases. Most Voted**
- C. Follow up with the employees who reviewed the changes and prescribe practices they should follow in the future.
- D. Design a policy that will require on-call teams to immediately call engineers and management to discuss a plan of action if an incident occurs.

### ✓ Correct Answer: B

*Explanation: Prevent similar incidents by ensuring comprehensive automated testing catches issues before deployment. Tests should cover the failure mode that caused the incident.*

---

## Question 87: 83

### Question

Your organization uses a change advisory board (CAB) to approve all changes to an existing service. You want to revise this process to eliminate any negative impact on the software delivery performance. What should you do? (Choose two.)

- A. Replace the CAB with a senior manager to ensure continuous oversight from development to deployment.
- B. Let developers merge their own changes, but ensure that the team's deployment platform can roll back changes if any issues are discovered.
- C. Move to a peer-review based process for individual changes that is enforced at code check-in time and supported by automated tests. Most Voted**
- D. Batch changes into larger but less frequent software releases.
- E. Ensure that the team's development platform enables developers to get fast feedback on the impact of their changes. **Most Voted**

### ✓ Correct Answer: C

*Explanation: Replace heavyweight CABs with peer reviews at code check-in, automated testing, and fast feedback loops. This maintains quality while improving delivery velocity.*

---

## Question 88: 84

### Question

Your organization has a containerized web application that runs on-premises. As part of the migration plan to Google Cloud, you need to select a deployment strategy and platform that meets the following acceptance criteria: 1. The platform must be able to direct traffic from Android devices to an Android-specific microservice. 2. The platform must allow for arbitrary percentage-based traffic splitting 3. The deployment strategy must allow for continuous testing of multiple versions of any microservice. What should you do?

- A. Deploy the canary release of the application to Cloud Run. Use traffic splitting to direct 10% of user traffic to the canary release based on the revision tag.
- B. Deploy the canary release of the application to App Engine. Use traffic splitting to direct a subset of user traffic to the new version based on the IP address.
- C. Deploy the canary release of the application to Compute Engine. Use Anthos Service Mesh with Compute Engine to direct 10% of user traffic to the canary release by configuring the virtual service.
- D. Deploy the canary release to Google Kubernetes Engine with Anthos Service Mesh. Use traffic splitting to direct 10% of user traffic to the new version based on the user-agent header configured in the virtual service. Most Voted**

### ✓ Correct Answer: D

*Explanation: GKE with Anthos Service Mesh provides advanced traffic routing based on headers (like user-agent for device detection), percentage-based splitting, and multi-version testing capabilities.*

---

## Question 89: 85

### Question

Your team is running microservices in Google Kubernetes Engine (GKE). You want to detect consumption of an error budget to protect customers and define release policies. What should you do?

- A. Create SLIs from metrics. Enable Alert Policies if the services do not pass.
- B. Use the metrics from Anthos Service Mesh to measure the health of the microservices.
- C. Create a SLO. Create an Alert Policy on select\_slo\_burn\_rate. Most Voted**
- D. Create a SLO and configure uptime checks for your services. Enable Alert Policies if the services do not pass.

### ✓ Correct Answer: C

*Explanation: Create an SLO and alert on the burn rate metric (select\_slo\_burn\_rate). This proactively detects when error budget is being consumed faster than expected.*

---

## Question 90: 86

### Question

Your organization wants to collect system logs that will be used to generate dashboards in Cloud Operations for their Google Cloud project. You need to configure all current and future Compute Engine instances to collect the system logs, and you must ensure that the Ops Agent remains up to date. What should you do?

- A. Use the gcloud CLI to install the Ops Agent on each VM listed in the Cloud Asset Inventory,
- B. Select all VMs with an Agent status of Not detected on the Cloud Operations VMs dashboard. Then select Install agents.
- C. Use the gcloud CLI to create an Agent Policy. Most Voted**
- D. Install the Ops Agent on the Compute Engine image by using a startup script

### ✓ Correct Answer: C

*Explanation: Agent Policies automatically install and update the Ops Agent on all matching VMs (current and future), ensuring consistent log collection without manual intervention.*

---

## Question 91: 87

### Question

Your company has a Google Cloud resource hierarchy with folders for production, test, and development. Your cyber security team needs to review your company's Google Cloud security posture to accelerate security issue identification and resolution. You need to centralize the logs generated by Google Cloud services from all projects only inside your production folder to allow for alerting and near-real time analysis. What should you do?

- A. Enable the Workflows API and route all the logs to Cloud Logging.
- B. Create a central Cloud Monitoring workspace and attach all related projects.
- C. Create an aggregated log sink associated with the production folder that uses a Pub/Sub topic as the destination. Most Voted**
- D. Create an aggregated log sink associated with the production folder that uses a Cloud Logging bucket as the destination.

### ✓ Correct Answer: C

*Explanation:* Aggregated log sinks at the folder level collect logs from all projects within that folder. Pub/Sub enables near real-time streaming for alerting and analysis.

---

## Question 92: 88

### Question

You are configuring the frontend tier of an application deployed in Google Cloud. The frontend tier is hosted in nginx and deployed using a managed instance group with an Envoy-based external HTTP(S) load balancer in front. The application is deployed entirely within the europe-west2 region, and only serves users based in the United Kingdom. You need to choose the most cost-effective network tier and load balancing configuration. What should you use?

- A. Premium Tier with a global load balancer
- B. Premium Tier with a regional load balancer
- C. Standard Tier with a global load balancer
- D. Standard Tier with a regional load balancer Most Voted**

### ✓ Correct Answer: D

*Explanation:* For a single-region deployment serving users in one geographic area, Standard Tier with regional load balancing is the most cost-effective option.

---

## Question 93: 89

### Question

You recently deployed your application in Google Kubernetes Engine (GKE) and now need to release a new version of the application. You need the ability to instantly roll back to the previous version of the application in case there are issues with the new version. Which deployment model should you use?

- A. Perform a rolling deployment, and test your new application after the deployment is complete.
- B. Perform A/B testing, and test your application periodically after the deployment is complete.
- C. Perform a canary deployment, and test your new application periodically after the new version is deployed.
- D. Perform a blue/green deployment, and test your new application after the deployment is complete. Most Voted**

### ✓ Correct Answer: D

*Explanation: Blue/green deployment runs both versions simultaneously and allows instant rollback by switching traffic back to the blue (old) version if issues arise with green (new) version.*

---

## Question 94: 90

### Question

You are building and deploying a microservice on Cloud Run for your organization. Your service is used by many applications internally. You are deploying a new release, and you need to test the new version extensively in the staging and production environments. You must minimize user and developer impact. What should you do?

- A. Deploy the new version of the service to the staging environment. Split the traffic, and allow 1% of traffic through to the latest version. Test the latest version. If the test passes, gradually roll out the latest version to the staging and production environments. Most Voted**
- B. Deploy the new version of the service to the staging environment. Split the traffic, and allow 50% of traffic through to the latest version. Test the latest version. If the test passes, send all traffic to the latest version. Repeat for the production environment.
- C. Deploy the new version of the service to the staging environment with a new-release tag without serving traffic. Test the new-release version. If the test passes, gradually roll out this tagged version. Repeat for the production environment. Most Voted**
- D. Deploy a new environment with the green tag to use as the staging environment. Deploy the new version of the service to the green environment and test the new version. If the tests pass, send all traffic to the green environment and delete the existing staging environment. Repeat for the production environment.

### ✓ Correct Answer: A & C

---

## Question 95: 91

### Question

You work for a global organization and run a service with an availability target of 99% with limited engineering resources. For the current calendar month, you noticed that the service has 99.5% availability. You must ensure that your service meets the defined availability goals and can react to business changes, including the upcoming launch of new features. You also need to reduce technical debt while minimizing operational costs. You want to follow Google-recommended practices. What should you do?

- A. Add N+1 redundancy to your service by adding additional compute resources to the service.
- B. Identify, measure, and eliminate toil by automating repetitive tasks. Most Voted**
- C. Define an error budget for your service level availability and minimize the remaining error budget.
- D. Allocate available engineers to the feature backlog while you ensure that the service remains within the availability target.

### ✓ Correct Answer: B

*Explanation: With availability above target and limited resources, reduce toil by automating repetitive tasks. This frees engineers for feature development while maintaining reliability.*

---

## Question 96: 92

### Question

You are developing the deployment and testing strategies for your CI/CD pipeline in Google Cloud. You must be able to:

- Reduce the complexity of release deployments and minimize the duration of deployment rollbacks.
- Test real production traffic with a gradual increase in the number of affected users. You want to select a deployment and testing strategy that meets your requirements. What should you do?

- A. Recreate deployment and canary testing
- B. Blue/green deployment and canary testing Most Voted**
- C. Rolling update deployment and A/B testing
- D. Rolling update deployment and shadow testing

### ✓ Correct Answer: B

*Explanation: Blue/green deployment enables instant rollback (just switch traffic back), while canary testing gradually increases production traffic exposure to new version.*

---

## Question 97: 93

### Question

You are creating a CI/CD pipeline to perform Terraform deployments of Google Cloud resources. Your CI/CD tooling is running in Google Kubernetes Engine (GKE) and uses an ephemeral Pod for each pipeline run. You must ensure that the pipelines that run in the Pods have the appropriate Identity and Access Management (IAM) permissions to perform the Terraform deployments. You want to follow Google-recommended practices for identity management. What should you do? (Choose two.)

- A. Create a new Kubernetes service account, and assign the service account to the Pods. Use Workload Identity to authenticate as the Google service account. **Most Voted**
- B. Create a new JSON service account key for the Google service account, store the key as a Kubernetes secret, inject the key into the Pods, and set the GOOGLE\_APPLICATION\_CREDENTIALS environment variable.
- C. Create a new Google service account, and assign the appropriate IAM permissions. **Most Voted**
- D. Create a new JSON service account key for the Google service account, store the key in the secret management store for the CI/CD tool, and configure Terraform to use this key for authentication.
- E. Assign the appropriate IAM permissions to the Google service account associated with the Compute Engine VM instances that run the Pods.

### ✓ Correct Answer: A & C

*Explanation: Create a Google service account with necessary IAM permissions, then use Workload Identity to bind a Kubernetes service account to it. This avoids managing JSON keys.*

---

## Question 98: 94

### Question

You are the on-call Site Reliability Engineer for a microservice that is deployed to a Google Kubernetes Engine (GKE) Autopilot cluster. Your company runs an online store that publishes order messages to Pub/Sub, and a microservice receives these messages and updates stock information in the warehousing system. A sales event caused an increase in orders, and the stock information is not being updated quickly enough. This is causing a large number of orders to be accepted for products that are out of stock. You check the metrics for the microservice and compare them to typical levels: You need to ensure that the warehouse system accurately reflects product inventory at the time orders are placed and minimize the impact on customers. What should you do?

- A. Decrease the acknowledgment deadline on the subscription.
- B. Add a virtual queue to the online store that allows typical traffic levels.
- C. Increase the number of Pod replicas. **Most Voted**
- D. Increase the Pod CPU and memory limits.

### ✓ Correct Answer: C

*Explanation: The high unacknowledged message count indicates the service can't process messages fast enough. Increase Pod replicas to scale horizontally and process more messages concurrently.*

---

## Question 99: 95

### Question

Your team deploys applications to three Google Kubernetes Engine (GKE) environments: development, staging, and production. You use GitHub repositories as your source of truth. You need to ensure that the three environments are consistent. You want to follow Google-recommended practices to enforce and install network policies and a logging DaemonSet on all the GKE clusters in those environments. What should you do?

- A. Use Google Cloud Deploy to deploy the network policies and the DaemonSet. Use Cloud Monitoring to trigger an alert if the network policies and DaemonSet drift from your source in the repository.
- B. Use Google Cloud Deploy to deploy the DaemonSet and use Policy Controller to configure the network policies. Use Cloud Monitoring to detect drifts from the source in the repository and Cloud Functions to correct the drifts.
- C. Use Cloud Build to render and deploy the network policies and the DaemonSet. Set up Config Sync to sync the configurations for the three environments.

**D. Use Cloud Build to render and deploy the network policies and the DaemonSet. Set up a Policy Controller to enforce the configurations for the three environments. Most Voted**

### ✓ Correct Answer: D

*Explanation: Cloud Build renders and deploys the resources, while Policy Controller enforces the desired state across all clusters. This ensures consistency and prevents configuration drift.*

---

## Question 100: 96

### Question

You are using Terraform to manage infrastructure as code within a CI/CD pipeline. You notice that multiple copies of the entire infrastructure stack exist in your Google Cloud project, and a new copy is created each time a change to the existing infrastructure is made. You need to optimize your cloud spend by ensuring that only a single instance of your infrastructure stack exists at a time. You want to follow Google-recommended practices. What should you do?

- A. Create a new pipeline to delete old infrastructure stacks when they are no longer needed.
- B. Confirm that the pipeline is storing and retrieving the terraform.tfstate file from Cloud Storage with the Terraform gcs backend. Most Voted**
- C. Verify that the pipeline is storing and retrieving the terraform.tfstate file from a source control.
- D. Update the pipeline to remove any existing infrastructure before you apply the latest configuration.

### ✓ Correct Answer: B

*Explanation: Terraform state file tracks infrastructure. Store it in Cloud Storage using gcs backend so all pipeline runs reference the same state, preventing duplicate infrastructure creation.*

---

## Question 101: 97

### Question

You are creating Cloud Logging sinks to export log entries from Cloud Logging to BigQuery for future analysis. Your organization has a Google Cloud folder named Dev that contains development projects and a folder named Prod that contains production projects. Log entries for development projects must be exported to dev\_dataset, and log entries for production projects must be exported to prod\_dataset. You need to minimize the number of log sinks created, and you want to ensure that the log sinks apply to future projects. What should you do?

- A. Create a single aggregated log sink at the organization level.
- B. Create a log sink in each project.
- C. Create two aggregated log sinks at the organization level, and filter by project ID.

**D. Create an aggregated log sink in the Dev and Prod folders. Most Voted**

### ✓ Correct Answer: D

*Explanation: Aggregated log sinks at folder level automatically apply to all current and future projects in that folder, minimizing sink creation and management.*

---

## Question 102: 98

### Question

Your company runs services by using multiple globally distributed Google Kubernetes Engine (GKE) clusters. Your operations team has set up workload monitoring that uses Prometheus-based tooling for metrics, alerts, and generating dashboards. This setup does not provide a method to view metrics globally across all clusters. You need to implement a scalable solution to support global Prometheus querying and minimize management overhead. What should you do?

- A. Configure Prometheus cross-service federation for centralized data access.
- B. Configure workload metrics within Cloud Operations for GKE.
- C. Configure Prometheus hierarchical federation for centralized data access.

**D. Configure Google Cloud Managed Service for Prometheus. Most Voted**

### ✓ Correct Answer: D

*Explanation: Google Cloud Managed Service for Prometheus provides a scalable, global solution for Prometheus metrics across multiple GKE clusters with minimal management overhead.*

---

## Question 103: 99

### Question

You need to build a CI/CD pipeline for a containerized application in Google Cloud. Your development team uses a central Git repository for trunk-based development. You want to run all your tests in the pipeline for any new versions of the application to improve the quality. What should you do?

- A. 1. Install a Git hook to require developers to run unit tests before pushing the code to a central repository. 2. Trigger Cloud Build to build the application container. Deploy the application container to a testing environment, and run integration tests. 3. If the integration tests are successful, deploy the application container to your production environment, and run acceptance tests.
- B. 1. Install a Git hook to require developers to run unit tests before pushing the code to a central repository. If all tests are successful, build a container. 2. Trigger Cloud Build to deploy the application container to a testing environment, and run integration tests and acceptance tests. 3. If all tests are successful, tag the code as production ready. Trigger Cloud Build to build and deploy the application container to the production environment.
- C. 1. Trigger Cloud Build to build the application container, and run unit tests with the container. 2. If unit tests are successful, deploy the application container to a testing environment, and run integration tests. 3. If the integration tests are successful, the pipeline deploys the application container to the production environment. After that, run acceptance tests.
- D. 1. Trigger Cloud Build to run unit tests when the code is pushed. If all unit tests are successful, build and push the application container to a central registry. 2. Trigger Cloud Build to deploy the container to a testing environment, and run integration tests and acceptance tests. 3. If all tests are successful, the pipeline deploys the application to the production environment and runs smoke tests Most Voted**

✓ Correct Answer: D

**Explanation:** Cloud Build triggers automate testing (unit → integration → acceptance) and deployment. Smoke tests in production validate the deployment succeeded without impacting quality.

---

## Question 104: 100

### Question

The new version of your containerized application has been tested and is ready to be deployed to production on Google Kubernetes Engine (GKE). You could not fully load-test the new version in your pre-production environment, and you need to ensure that the application does not have performance problems after deployment. Your deployment must be automated. What should you do?

- A. Deploy the application through a continuous delivery pipeline by using canary deployments. Use Cloud Monitoring to look for performance issues, and ramp up traffic as supported by the metrics. Most Voted**
- B. Deploy the application through a continuous delivery pipeline by using blue/green deployments. Migrate traffic to the new version of the application and use Cloud Monitoring to look for performance issues.
- C. Deploy the application by using kubectl and use Config Connector to slowly ramp up traffic between versions. Use Cloud Monitoring to look for performance issues.
- D. Deploy the application by using kubectl and set the spec.updateStrategy.type field to RollingUpdate. Use Cloud Monitoring to look for performance issues, and run the kubectl rollback command if there are any issues.

### ✓ Correct Answer: A

**Explanation:** Canary deployment gradually increases production traffic to the new version while monitoring performance. If issues arise, rollback impacts fewer users than blue/green.

---

## Question 105: 101

### Question

You are managing an application that runs in Compute Engine. The application uses a custom HTTP server to expose an API that is accessed by other applications through an internal TCP/UDP load balancer. A firewall rule allows access to the API port from 0.0.0.0/0. You need to configure Cloud Logging to log each IP address that accesses the API by using the fewest number of steps. What should you do first?

- A. Enable Packet Mirroring on the VPC.
- B. Install the Ops Agent on the Compute Engine instances.
- C. Enable logging on the firewall rule.

**D. Enable VPC Flow Logs on the subnet. Most Voted**

### ✓ Correct Answer: D

*Explanation: VPC Flow Logs capture connection data including source IP addresses for all traffic on the subnet, providing the requested IP logging with minimal configuration.*

---

## Question 106: 102

### Question

Your company runs an ecommerce website built with JVM-based applications and microservice architecture in Google Kubernetes Engine (GKE). The application load increases during the day and decreases during the night. Your operations team has configured the application to run enough Pods to handle the evening peak load. You want to automate scaling by only running enough Pods and nodes for the load. What should you do?

- A. Configure the Vertical Pod Autoscaler, but keep the node pool size static.
- B. Configure the Vertical Pod Autoscaler, and enable the cluster autoscaler.
- C. Configure the Horizontal Pod Autoscaler, but keep the node pool size static.

**D. Configure the Horizontal Pod Autoscaler, and enable the cluster autoscaler. Most Voted**

### ✓ Correct Answer: D

*Explanation: Horizontal Pod Autoscaler scales Pods based on load (CPU/memory/custom metrics), while cluster autoscaler scales nodes. Together they provide end-to-end automatic scaling.*

---

## Question 107: 103

### Question

Your organization wants to increase the availability target of an application from 99.9% to 99.99% for an investment of \$2,000. The application's current revenue is \$1,000,000. You need to determine whether the increase in availability is worth the investment for a single year of usage. What should you do?

**A. Calculate the value of improved availability to be \$900, and determine that the increase in availability is not worth the investment. Most Voted**

- B. Calculate the value of improved availability to be \$1,000, and determine that the increase in availability is not worth the investment.
- C. Calculate the value of improved availability to be \$1,000, and determine that the increase in availability is worth the investment.
- D. Calculate the value of improved availability to be \$9,000, and determine that the increase in availability is worth the investment.

✓ **Correct Answer: A**

**Explanation:** Value = Revenue × (New availability - Old availability) = \$1M × (0.9999 - 0.999) = \$1M × 0.0001 = \$100 expected loss reduction (not \$900). The investment exceeds the value.

---

## Question 108: 104

### Question

A third-party application needs to have a service account key to work properly. When you try to export the key from your cloud project, you receive an error: "The organization policy constraint iam.disableServiceAccountKeyCreation is enforced." You need to make the third-party application work while following Google-recommended security practices. What should you do?

- A. Enable the default service account key, and download the key.
- B. Remove the iam.disableServiceAccountKeyCreation policy at the organization level, and create a key.
- C. Disable the service account key creation policy at the project's folder, and download the default key.
- D. Add a rule to set the iam.disableServiceAccountKeyCreation policy to off in your project and create a key. Most Voted**

✓ **Correct Answer: D**

**Explanation:** Override the org policy at the project level using `iam.disableServiceAccountKeyCreation=false` to allow key creation where specifically needed while maintaining org-wide security.

---

## Question 109: 105

### Question

Your team is writing a postmortem after an incident on your external facing application. Your team wants to improve the postmortem policy to include triggers that indicate whether an incident requires a postmortem. Based on Site Reliability Engineering (SRE) practices, what triggers should be defined in the postmortem policy? (Choose two.)

- A. An external stakeholder asks for a postmortem
- B. Data is lost due to an incident. Most Voted**
- C. An internal stakeholder requests a postmortem.
- D. The monitoring system detects that one of the instances for your application has failed. E. The CD pipeline detects an issue and rolls back a problematic release. Most Voted

### ✓ Correct Answer: B

*Explanation: SRE practices require postmortems for significant incidents with customer impact: data loss directly affects users, and rollbacks indicate a serious issue that reached deployment.*

---

## Question 110: 106

### Question

You are implementing a CI/CD pipeline for your application in your company's multi-cloud environment. Your application is deployed by using custom Compute Engine images and the equivalent in other cloud providers. You need to implement a solution that will enable you to build and deploy the images to your current environment and is adaptable to future changes. Which solution stack should you use?

- A. Cloud Build with Packer Most Voted**
- B. Cloud Build with Google Cloud Deploy
- C. Google Kubernetes Engine with Google Cloud Deploy
- D. Cloud Build with kpt

### ✓ Correct Answer: A

*Explanation: Packer builds machine images across multiple cloud providers using the same configuration template. Cloud Build provides the CI/CD automation for multi-cloud deployments.*

---

## Question 111: 107

### Question

Your application's performance in Google Cloud has degraded since the last release. You suspect that downstream dependencies might be causing some requests to take longer to complete. You need to investigate the issue with your application to determine the cause. What should you do?

- A. Configure Error Reporting in your application.
- B. Configure Google Cloud Managed Service for Prometheus in your application.
- C. Configure Cloud Profiler in your application.
- D. Configure Cloud Trace in your application. Most Voted**

### ✓ Correct Answer: D

*Explanation: Cloud Trace provides distributed tracing to track requests across services and identify latency issues in downstream dependencies. It shows the complete request flow.*

---

## Question 112: 108

### Question

You are creating a CI/CD pipeline in Cloud Build to build an application container image. The application code is stored in GitHub. Your company requires that production image builds are only run against the main branch and that the change control team approves all pushes to the main branch. You want the image build to be as automated as possible. What should you do? (Choose two.)

- A. Create a trigger on the Cloud Build job. Set the repository event setting to 'Pull request'.
- B. Add the OWNERS file to the Included files filter on the trigger.
- C. Create a trigger on the Cloud Build job. Set the repository event setting to 'Push to a branch' Most Voted**
- D. Configure a branch protection rule for the main branch on the repository. Most Voted E. Enable the Approval option on the trigger.**

### ✓ Correct Answer: C & D

*Explanation: Trigger Cloud Build on pushes to main branch, and use GitHub branch protection rules to require approvals before merging to main. This ensures change control.*

---

## Question 113: 109

### Question

You built a serverless application by using Cloud Run and deployed the application to your production environment. You want to identify the resource utilization of the application for cost optimization. What should you do?

- A. Use Cloud Trace with distributed tracing to monitor the resource utilization of the application.
- B. Use Cloud Profiler with Ops Agent to monitor the CPU and memory utilization of the application. Most Voted**
- C. Use Cloud Monitoring to monitor the container CPU and memory utilization of the application. Most Voted**
- D. Use Cloud Ops to create logs-based metrics to monitor the resource utilization of the application.

### ✓ Correct Answer: B & C

*Explanation: Cloud Profiler identifies CPU and memory hotspots in code, while Cloud Monitoring provides container-level resource utilization metrics. Both help optimize costs.*

---

## Question 114: 110

### Question

Your company is using HTTPS requests to trigger a public Cloud Run-hosted service accessible at the `https://booking-engine-abcdef.a.run.app` URL. You need to give developers the ability to test the latest revisions of the service before the service is exposed to customers. What should you do?

- A. Run the `gcloud run deploy booking-engine --no-traffic --tag dev` command. Use the `https://dev--booking-engine-abcdef.a.run.app` URL for testing. Most Voted**
- B. Run the `gcloud run services update-traffic booking-engine --to-revisions LATEST=1` command. Use the `https://booking-engine-abcdef.a.run.app` URL for testing.
- C. Pass the `curl -H "Authorization:Bearer $(gcloud auth print-identity-token)"` auth token. Use the `https://booking-engine-abcdef.a.run.app` URL to test privately.
- D. Grant the `roles/run.invoker` role to the developers testing the `booking-engine` service. Use the `https://booking-engine-abcdef.private.run.app` URL for testing.

### ✓ Correct Answer: A

*Explanation: The `--no-traffic` flag deploys a revision without receiving production traffic, while `--tag dev` creates a unique URL (`dev--booking-engine-*`) for testing new revisions.*

---

## Question 115: 111

### Question

You are configuring connectivity across Google Kubernetes Engine (GKE) clusters in different VPCs. You notice that the nodes in Cluster A are unable to access the nodes in Cluster B. You suspect that the workload access issue is due to the network configuration. You need to troubleshoot the issue but do not have execute access to workloads and nodes. You want to identify the layer at which the network connectivity is broken. What should you do?

- A. Install a toolbox container on the node in Cluster Confirm that the routes to Cluster B are configured appropriately.

**B. Use Network Connectivity Center to perform a Connectivity Test from Cluster A to Cluster B. Most Voted**

- C. Use a debug container to run the traceroute command from Cluster A to Cluster B and from Cluster B to Cluster A. Identify the common failure point.
- D. Enable VPC Flow Logs in both VPCs, and monitor packet drops.

**✓ Correct Answer: B**

*Explanation: Network Connectivity Center's Connectivity Test simulates network paths and identifies configuration issues (firewall rules, routes) without needing exec access to workloads.*

---

## Question 116: 112

### Question

You manage an application that runs in Google Kubernetes Engine (GKE) and uses the blue/green deployment methodology. Extracts of the Kubernetes manifests are shown below: The Deployment app-green was updated to use the new version of the application. During post-deployment monitoring, you notice that the majority of user requests are failing. You did not observe this behavior in the testing environment. You need to mitigate the incident impact on users and enable the developers to troubleshoot the issue. What should you do?

- A. Update the Deployment app-blue to use the new version of the application.
- B. Update the Deployment app-green to use the previous version of the application.
- C. Change the selector on the Service app-svc to app: my-app.

**D. Change the selector on the Service app-svc to app: my-app, version: blue. Most Voted**

**✓ Correct Answer: D**

*Explanation: In blue/green deployments, instantly rollback by changing the Service selector to point back to the blue (old, working) Deployment. This immediately restores service.*

---

## Question 117: 113

### Question

You are running a web application deployed to a Compute Engine managed instance group. Ops Agent is installed on all instances. You recently noticed suspicious activity from a specific IP address. You need to configure Cloud Monitoring to view the number of requests from that specific IP address with minimal operational overhead. What should you do?

- A. Configure the Ops Agent with a logging receiver. Create a logs-based metric.
- B. Create a script to scrape the web server log. Export the IP address request metrics to the Cloud Monitoring API. **Most Voted**
- C. Update the application to export the IP address request metrics to the Cloud Monitoring API.
- D. Configure the Ops Agent with a metrics receiver.

### ✓ Correct Answer: B

**Explanation:** Ops Agent collects web server logs. Create a script to parse logs for the specific IP address and export as custom metrics to Cloud Monitoring with minimal overhead.

---

## Question 118: 114

### Question

Your organization is using Helm to package containerized applications. Your applications reference both public and private charts. Your security team flagged that using a public Helm repository as a dependency is a risk. You want to manage all charts uniformly, with native access control and VPC Service Controls. What should you do?

- A. Store public and private charts in OCI format by using Artifact Registry. **Most Voted****
- B. Store public and private charts by using GitHub Enterprise with Google Workspace as the identity provider.
- C. Store public and private charts by using Git repository. Configure Cloud Build to synchronize contents of the repository into a Cloud Storage bucket. Connect Helm to the bucket by using [https://\[bucket\].storage.googleapis.com/\[helmchart\]](https://[bucket].storage.googleapis.com/[helmchart]) as the Helm repository.
- D. Configure a Helm chart repository server to run in Google Kubernetes Engine (GKE) with Cloud Storage bucket as the storage backend.

### ✓ Correct Answer: A

**Explanation:** Artifact Registry supports OCI-format Helm charts with native IAM access control and VPC Service Controls. Store both public (mirrored) and private charts uniformly.

---

## Question 119: 115

### Question

You use Terraform to manage an application deployed to a Google Cloud environment. The application runs on instances deployed by a managed instance group. The Terraform code is deployed by using a CI/CD pipeline. When you change the machine type on the instance template used by the managed instance group, the pipeline fails at the terraform apply stage with the following error message: You need to update the instance template and minimize disruption to the application and the number of pipelines runs. What should you do?

- A. Delete the managed instance group and recreate it after updating the instance template.
- B. Add a new instance template, update the managed instance group to use the new instance template, and delete the old instance template.
- C. Remove the managed instance group from the Terraform state file, update the instance template, and reimport the managed instance group.
- D. Set the `create_before_destroy` meta-argument to true in the lifecycle block on the instance template. Most Voted**

### ✓ Correct Answer: D

*Explanation: `create_before_destroy=true` tells Terraform to create the new instance template before destroying the old one, avoiding the dependency conflict with the managed instance group.*

---

## Question 120: 116

### Question

Your company operates in a highly regulated domain that requires you to store all organization logs for seven years. You want to minimize logging infrastructure complexity by using managed services. You need to avoid any future loss of log capture or stored logs due to misconfiguration or human error. What should you do?

- A. Use Cloud Logging to configure an aggregated sink at the organization level to export all logs into a BigQuery dataset.
- B. Use Cloud Logging to configure an aggregated sink at the organization level to export all logs into Cloud Storage with a seven-year retention policy and Bucket Lock. Most Voted**
- C. Use Cloud Logging to configure an export sink at each project level to export all logs into a BigQuery dataset
- D. Use Cloud Logging to configure an export sink at each project level to export all logs into Cloud Storage with a seven-year retention policy and Bucket Lock.

### ✓ Correct Answer: B

*Explanation: Aggregated sink at org level captures all logs. Cloud Storage with retention policy and Bucket Lock ensures logs are immutable and retained for 7 years.*

---

## Question 121: 117

### Question

You are building the CI/CD pipeline for an application deployed to Google Kubernetes Engine (GKE). The application is deployed by using a Kubernetes Deployment, Service, and Ingress. The application team asked you to deploy the application by using the blue/green deployment methodology. You need to implement the rollback actions. What should you do?

- A. Run the kubectl rollout undo command.
- B. Delete the new container image, and delete the running Pods.
- C. Update the Kubernetes Service to point to the previous Kubernetes Deployment. Most Voted**
- D. Scale the new Kubernetes Deployment to zero

### ✓ Correct Answer: C

*Explanation: Blue/green deployment maintains both old and new versions. Rollback is accomplished by updating the Service's selector to point back to the previous Deployment.*

---

## Question 122: 118

### Question

You are building and running client applications in Cloud Run and Cloud Functions. Your client requires that all logs must be available for one year so that the client can import the logs into their logging service. You must minimize required code changes. What should you do?

- A. Update all images in Cloud Run and all functions in Cloud Functions to send logs to both Cloud Logging and the client's logging service. Ensure that all the ports required to send logs are open in the VPC firewall.
- B. Create a Pub/Sub topic, subscription, and logging sink. Configure the logging sink to send all logs into the topic. Give your client access to the topic to retrieve the logs.
- C. Create a storage bucket and appropriate VPC firewall rules. Update all images in Cloud Run and all functions in Cloud Functions to send logs to a file within the storage bucket.
- D. Create a logs bucket and logging sink. Set the retention on the logs bucket to 365 days. Configure the logging sink to send logs to the bucket. Give your client access to the bucket to retrieve the logs. Most Voted**

### ✓ Correct Answer: D

*Explanation: Log buckets with retention policies and logging sinks automate log storage and retention. Grant client access to bucket for retrieval without code changes.*

---

## Question 123: 119

### Question

You are building and running client applications in Cloud Run and Cloud Functions. Your client requires that all logs must be available for one year so that the client can import the logs into their logging service. You must minimize required code changes. What should you do?

- A. Deploy Falco or Twistlock on GKE to monitor for vulnerabilities on your running Pods.
- B. Configure Identity and Access Management (IAM) policies to create a least privilege model on your GKE clusters.
- C. Use Binary Authorization to attest images during your CI/CD pipeline. Most Voted**
- D. Enable Container Analysis in Artifact Registry, and check for common vulnerabilities and exposures (CVEs) in your container images.

### ✓ Correct Answer: C

*Explanation: Binary Authorization enforces deployment policies by requiring cryptographic attestations during CI/CD. Only attested (verified) images can be deployed to GKE clusters.*

---

## Question 124: 120

### Question

You have an application that runs in Google Kubernetes Engine (GKE). The application consists of several microservices that are deployed to GKE by using Deployments and Services. One of the microservices is experiencing an issue where a Pod returns 403 errors after the Pod has been running for more than five hours. Your development team is working on a solution, but the issue will not be resolved for a month. You need to ensure continued operations until the microservice is fixed. You want to follow Google-recommended practices and use the fewest number of steps. What should you do?

- A. Create a cron job to terminate any Pods that have been running for more than five hours.
- B. Add a HTTP liveness probe to the microservice's deployment. Most Voted**
- C. Monitor the Pods, and terminate any Pods that have been running for more than five hours.
- D. Configure an alert to notify you whenever a Pod returns 403 errors

### ✓ Correct Answer: B

*Explanation: HTTP liveness probe checks Pod health. When a Pod fails the probe (e.g., returns 403 errors), Kubernetes automatically restarts it, temporarily resolving the issue.*

---

## Question 125: 121

### Question

You want to share a Cloud Monitoring custom dashboard with a partner team. What should you do?

- A. Provide the partner team with the dashboard URL to enable the partner team to create a copy of the dashboard. **Most Voted**
- B. Export the metrics to BigQuery. Use Looker Studio to create a dashboard, and share the dashboard with the partner team.
- C. Copy the Monitoring Query Language (MQL) query from the dashboard, and send the ML query to the partner team.
- D. Download the JSON definition of the dashboard, and send the JSON file to the partner team.

### ✓ Correct Answer: A

*Explanation: Cloud Monitoring dashboards have shareable URLs. The partner team can copy the dashboard using the URL without needing complex exports or access to underlying data.*

---

## Question 126: 122

### Question

You are building an application that runs on Cloud Run. The application needs to access a third-party API by using an API key. You need to determine a secure way to store and use the API key in your application by following Google-recommended practices. What should you do?

- A. Save the API key in Secret Manager as a secret. Reference the secret as an environment variable in the Cloud Run application. **Most Voted**
- B. Save the API key in Secret Manager as a secret key. Mount the secret key under the /sys/api\_key directory, and decrypt the key in the Cloud Run application.
- C. Save the API key in Cloud Key Management Service (Cloud KMS) as a key. Reference the key as an environment variable in the Cloud Run application.
- D. Encrypt the API key by using Cloud Key Management Service (Cloud KMS), and pass the key to Cloud Run as an environment variable. Decrypt and use the key in Cloud Run.

### ✓ Correct Answer: A

*Explanation: Secret Manager securely stores sensitive data like API keys. Reference secrets as environment variables in Cloud Run without embedding credentials in code or configuration.*

---

## Question 127: 123

### Question

You are currently planning how to display Cloud Monitoring metrics for your organization's Google Cloud projects. Your organization has three folders and six projects: You want to configure Cloud Monitoring dashboards to only display metrics from the projects within one folder. You need to ensure that the dashboards do not display metrics from projects in the other folders. You want to follow Google-recommended practices. What should you do?

- A. Create a single new scoping project.
- B. Create new scoping projects for each folder. Most Voted**
- C. Use the current app-one-prod project as the scoping project.
- D. Use the current app-one-dev, app-one-staging, and app-one-prod projects as the scoping project for each folder.

### ✓ Correct Answer: B

*Explanation: Scoping projects define metrics visibility boundaries. Create one scoping project per folder to ensure dashboards only display metrics from projects within that folder.*

---

## Question 128: 124

### Question

Your company's security team needs to have read-only access to Data Access audit logs in the \_Required bucket. You want to provide your security team with the necessary permissions following the principle of least privilege and Google-recommended practices. What should you do?

- A. Assign the roles/logging.viewer role to each member of the security team.
- B. Assign the roles/logging.viewer role to a group with all the security team members.
- C. Assign the roles/logging.privateLogViewer role to each member of the security team.
- D. Assign the roles/logging.privateLogViewer role to a group with all the security team members. Most Voted**

### ✓ Correct Answer: D

*Explanation: Data Access audit logs require roles/logging.privateLogViewer role. Assign to a group (not individuals) following least privilege and Google-recommended practices.*

---

## Question 129: 125

### Question

Your team is building a service that performs compute-heavy processing on batches of data. The data is processed faster based on the speed and number of CPUs on the machine. These batches of data vary in size and may arrive at any time from multiple third-party sources. You need to ensure that third parties are able to upload their data securely. You want to minimize costs, while ensuring that the data is processed as quickly as possible. What should you do?

- A. Provide a secure file transfer protocol (SFTP) server on a Compute Engine instance so that third parties can upload batches of data, and provide appropriate credentials to the server. Create a Cloud Function with a `google.storage.object.finalize` Cloud Storage trigger. Write code so that the function can scale up a Compute Engine autoscaling managed instance group. Use an image pre-loaded with the data processing software that terminates the instances when processing completes.
- B. Provide a Cloud Storage bucket so that third parties can upload batches of data, and provide appropriate Identity and Access Management (IAM) access to the bucket. Use a standard Google Kubernetes Engine (GKE) cluster and maintain two services: one that processes the batches of data, and one that monitors Cloud Storage for new batches of data. Stop the processing service when there are no batches of data to process.
- C. Provide a Cloud Storage bucket so that third parties can upload batches of data, and provide appropriate Identity and Access Management (IAM) access to the bucket. Create a Cloud Function with a `google.storage.object.finalize` Cloud Storage trigger. Write code so that the function can scale up a Compute Engine autoscaling managed instance group. Use an image pre-loaded with the data processing software that terminates the instances when processing completes. Most Voted**
- D. Provide a Cloud Storage bucket so that third parties can upload batches of data, and provide appropriate Identity and Access Management (IAM) access to the bucket. Use Cloud Monitoring to detect new batches of data in the bucket and trigger a Cloud Function that processes the data. Set a Cloud Function to use the largest CPU possible to minimize the runtime of the processing.

### ✓ Correct Answer: C

**Explanation:** Cloud Storage with IAM provides secure uploads. Cloud Function triggers on object finalization scale up autoscaling MIG with pre-loaded software, optimizing cost and speed.

---

## Question 130: 126

### Question

You are reviewing your deployment pipeline in Google Cloud Deploy. You must reduce toil in the pipeline, and you want to minimize the amount of time it takes to complete an end-to-end deployment. What should you do? (Choose two.)

- A. Create a trigger to notify the required team to complete the next step when manual intervention is required.

- B. Divide the automation steps into smaller tasks. Most Voted**

- C. Use a script to automate the creation of the deployment pipeline in Google Cloud Deploy.

- D. Add more engineers to finish the manual steps. E. Automate promotion approvals from the development environment to the test environment. Most Voted

### ✓ Correct Answer: B

**Explanation:** Reduce toil by breaking large tasks into smaller automated steps and automating approvals for lower environments (dev→test), minimizing manual intervention.

---

## Question 131: 127

### Question

You work for a global organization and are running a monolithic application on Compute Engine. You need to select the machine type for the application to use that optimizes CPU utilization by using the fewest number of steps. You want to use historical system metrics to identify the machine type for the application to use. You want to follow Google-recommended practices. What should you do?

- A. Use the Recommender API and apply the suggested recommendations. Most Voted**
- B. Create an Agent Policy to automatically install Ops Agent in all VMs.
  - C. Install the Ops Agent in a fleet of VMs by using the gcloud CLI.
  - D. Review the Cloud Monitoring dashboard for the VM and choose the machine type with the lowest CPU utilization.

### ✓ Correct Answer: A

*Explanation: Recommender API analyzes historical metrics and provides machine type recommendations to optimize CPU utilization. Applying recommendations is a one-step solution.*

---

## Question 132: 128

### Question

You deployed an application into a large Standard Google Kubernetes Engine (GKE) cluster. The application is stateless and multiple pods run at the same time. Your application receives inconsistent traffic. You need to ensure that the user experience remains consistent regardless of changes in traffic and that the resource usage of the cluster is optimized. What should you do?

- A. Configure a cron job to scale the deployment on a schedule

**B. Configure a Horizontal Pod Autoscaler. Most Voted**

- C. Configure a Vertical Pod Autoscaler
- D. Configure cluster autoscaling on the node pool.

### ✓ Correct Answer: B

*Explanation: Horizontal Pod Autoscaler scales the number of Pods based on CPU/memory utilization or custom metrics, automatically handling inconsistent traffic patterns.*

---

## Question 133: 129

### Question

You need to deploy a new service to production. The service needs to automatically scale using a managed instance group and should be deployed across multiple regions. The service needs a large number of resources for each instance and you need to plan for capacity. What should you do?

- A. Monitor results of Cloud Trace to determine the optimal sizing.
- B. Use the n2-highcpu-96 machine type in the configuration of the managed instance group.
- C. Deploy the service in multiple regions and use an internal load balancer to route traffic.
- D. Validate that the resource requirements are within the available project quota limits of each region. Most Voted**

### ✓ Correct Answer: D

**Explanation:** Before deploying resource-intensive services across multiple regions, verify that project quotas in each region can accommodate the required resources to avoid deployment failures.

---

## Question 134: 130

### Question

You are analyzing Java applications in production. All applications have Cloud Profiler and Cloud Trace installed and configured by default. You want to determine which applications need performance tuning. What should you do? (Choose two.)

- A. Examine the wall-clock time and the CPU time of the application. If the difference is substantial increase the CPU resource allocation.
- B. Examine the wall-clock time and the CPU time of the application. If the difference is substantial, increase the memory resource allocation.
- C. Examine the wall-clock time and the CPU time of the application. If the difference is substantial, increase the local disk storage allocation.
- D. Examine the latency time the wall-clock time and the CPU time of the application. If the latency time is slowly burning down the error budget, and the difference between wall-clock time and CPU time is minimal mark the application for optimization. Most Voted E. Examine the heap usage of the application. If the usage is low, mark the application for optimization. Most Voted**

### ✓ Correct Answer: D

**Explanation:** If latency burns error budget with minimal CPU difference, the app is I/O bound (needs optimization). Low heap usage indicates inefficient memory allocation patterns.

---

## Question 135: 131

### Question

Your organization stores all application logs from multiple Google Cloud projects in a central Cloud Logging project. Your security team wants to enforce a rule that each project team can only view their respective logs and only the operations team can view all the logs. You need to design a solution that meets the security team's requirements while minimizing costs. What should you do?

- A. Grant each project team access to the project \_Default view in the central logging project. Grant viewer access to the operations team in the central logging project.
- B. Create Identity and Access Management (IAM) roles for each project team and restrict access to the \_Default log view in their individual Google Cloud project. Grant viewer access to the operations team in the central logging project.
- C. Create log views for each project team and only show each project team their application logs. Grant the operations team access to the \_AllLogs view in the central logging project. Most Voted**
- D. Export logs to BigQuery tables for each project team. Grant project teams access to their tables. Grant logs writer access to the operations team in the central logging project.

### ✓ Correct Answer: C

*Explanation: Create separate log views per project team showing only their logs. Operations team gets \_AllLogs view access. This isolates access without duplicating logs to BigQuery.*

---

## Question 136: 132

### Question

Your company uses Jenkins running on Google Cloud VM instances for CI/CD. You need to extend the functionality to use infrastructure as code automation by using Terraform. You must ensure that the Terraform Jenkins instance is authorized to create Google Cloud resources. You want to follow Google-recommended practices. What should you do?

- A. Confirm that the Jenkins VM instance has an attached service account with the appropriate Identity and Access Management (IAM) permissions. Most Voted**
- B. Use the Terraform module so that Secret Manager can retrieve credentials.
- C. Create a dedicated service account for the Terraform instance. Download and copy the secret key value to the GOOGLE\_CREDENTIALS environment variable on the Jenkins server. Most Voted**
- D. Add the gcloud auth application-default login command as a step in Jenkins before running the Terraform commands.

### ✓ Correct Answer: A & C

*Explanation: Best practice is using VM's attached service account with IAM permissions. Alternative is dedicated service account with JSON key stored in GOOGLE\_CREDENTIALS environment variable.*

---

## Question 137: 133

### Question

You encounter a large number of outages in the production systems you support. You receive alerts for all the outages, the alerts are due to unhealthy systems that are automatically restarted within a minute. You want to set up a process that would prevent staff burnout while following Site Reliability Engineering (SRE) practices. What should you do?

#### A. Eliminate alerts that are not actionable Most Voted

- B. Redefine the related SLO so that the error budget is not exhausted
- C. Distribute the alerts to engineers in different time zones
- D. Create an incident report for each of the alerts

#### ✓ Correct Answer: A

*Explanation: Following SRE practices, eliminate non-actionable alerts (systems auto-recover within a minute). Only alert on issues requiring human intervention to prevent alert fatigue.*

---

## Question 138: 134

### Question

As part of your company's initiative to shift left on security, the InfoSec team is asking all teams to implement guard rails on all the Google Kubernetes Engine (GKE) clusters to only allow the deployment of trusted and approved images. You need to determine how to satisfy the InfoSec team's goal of shifting left on security. What should you do?

- A. Enable Container Analysis in Artifact Registry, and check for common vulnerabilities and exposures (CVEs) in your container images

#### B. Use Binary Authorization to attest images during your CI/CD pipeline Most Voted

- C. Configure Identity and Access Management (IAM) policies to create a least privilege model on your GKE clusters.
- D. Deploy Falco or Twistlock on GKE to monitor for vulnerabilities on your running Pods

#### ✓ Correct Answer: B

*Explanation: Binary Authorization with attestation in CI/CD pipeline ensures only trusted, approved images (verified during build) can be deployed to GKE, implementing shift-left security.*

---

## Question 139: 135

### Question

Your company operates in a highly regulated domain. Your security team requires that only trusted container images can be deployed to Google Kubernetes Engine (GKE). You need to implement a solution that meets the requirements of the security team while minimizing management overhead. What should you do?

**A. Configure Binary Authorization in your GKE clusters to enforce deploy-time security policies.**

**Most Voted**

- B. Grant the roles/artifactregistry.writer role to the Cloud Build service account. Confirm that no employee has Artifact Registry write permission.
- C. Use Cloud Run to write and deploy a custom validator. Enable an Eventarc trigger to perform validations when new images are uploaded.
- D. Configure Kritis to run in your GKE clusters to enforce deploy-time security policies.

**✓ Correct Answer: A**

*Explanation: Binary Authorization is Google's managed solution for deploy-time security policies on GKE. It enforces image trust with minimal overhead compared to third-party solutions.*

---

## Question 140: 136

### Question

Your CTO has asked you to implement a postmortem policy on every incident for internal use. You want to define what a good postmortem is to ensure that the policy is successful at your company. What should you do? (Choose two.)

A. Ensure that all postmortems include what caused the incident, identify the person or team responsible for causing the incident, and how to prevent a future occurrence of the incident.

B. Ensure that all postmortems include what caused the incident, how the incident could have been worse, and how to prevent a future occurrence of the incident.

**C. Ensure that all postmortems include the severity of the incident, how to prevent a future occurrence of the incident, and what caused the incident without naming internal system components. Most Voted**

D. Ensure that all postmortems include how the incident was resolved and what caused the incident without naming customer information. E. Ensure that all postmortems include all incident participants in postmortem authoring and share postmortems as widely as possible. Most Voted

**✓ Correct Answer: C**

*Explanation: Good postmortems are blameless, include all participants in authoring, are widely shared, document severity/cause/prevention, and focus on systems not individuals.*

---

## Question 141: 137

### Question

You are developing reusable infrastructure as code modules. Each module contains integration tests that launch the module in a test project. You are using GitHub for source control. You need to continuously test your feature branch and ensure that all code is tested before changes are accepted. You need to implement a solution to automate the integration tests. What should you do?

- A. Use a Jenkins server for CI/CD pipelines. Periodically run all tests in the feature branch.
  - B. Ask the pull request reviewers to run the integration tests before approving the code.
  - C. Use Cloud Build to run the tests. Trigger all tests to run after a pull request is merged.
- D. Use Cloud Build to run tests in a specific folder. Trigger Cloud Build for every GitHub pull request. Most Voted**

### ✓ Correct Answer: D

*Explanation: Cloud Build triggers on every GitHub pull request, running tests before code is merged. This ensures all feature branch changes are tested continuously.*

---

## Question 142: 138

### Question

Your company processes IoT data at scale by using Pub/Sub, App Engine standard environment, and an application written in Go. You noticed that the performance inconsistently degrades at peak load. You could not reproduce this issue on your workstation. You need to continuously monitor the application in production to identify slow paths in the code. You want to minimize performance impact and management overhead. What should you do?

- A. Use Cloud Monitoring to assess the App Engine CPU utilization metric.
  - B. Install a continuous profiling tool into Compute Engine. Configure the application to send profiling data to the tool.
  - C. Periodically run the go tool pprof command against the application instance. Analyze the results by using flame graphs.
- D. Configure Cloud Profiler, and initialize the [cloud.google.com/go/profiler](https://cloud.google.com/go/profiler) library in the application. Most Voted**

### ✓ Correct Answer: D

*Explanation: Cloud Profiler is a managed continuous profiling service. Initializing the library in Go applications provides low-overhead performance monitoring to identify slow code paths.*

---

## Question 143: 139

### Question

Your company runs services by using Google Kubernetes Engine (GKE). The GKE clusters in the development environment run applications with verbose logging enabled. Developers view logs by using the kubectl logs command and do not use Cloud Logging. Applications do not have a uniform logging structure defined. You need to minimize the costs associated with application logging while still collecting GKE operational logs. What should you do?

- A. Run the gcloud container clusters update --logging=SYSTEM command for the development cluster.
- B. Run the gcloud container clusters update --logging=WORKLOAD command for the development cluster.
- C. Run the gcloud logging sinks update \_Default --disabled command in the project associated with the development environment.
- D. Add the severity >= DEBUG resource.type = "k8s\_container" exclusion filter to the \_Default logging sink in the project associated with the development environment. Most Voted**

### ✓ Correct Answer: D

*Explanation:* Add exclusion filter to \_Default sink to exclude verbose application logs (k8s\_container), reducing costs while still collecting GKE system/operational logs.

---

## Question 144: 140

### Question

You have deployed a fleet of Compute Engine instances in Google Cloud. You need to ensure that monitoring metrics and logs for the instances are visible in Cloud Logging and Cloud Monitoring by your company's operations and cyber security teams. You need to grant the required roles for the Compute Engine service account by using Identity and Access Management (IAM) while following the principle of least privilege. What should you do?

- A. Grant the logging.logWriter and monitoring.metricWriter roles to the Compute Engine service accounts. Most Voted**
- B. Grant the logging.admin and monitoring.editor roles to the Compute Engine service accounts.
- C. Grant the logging.editor and monitoring.metricWriter roles to the Compute Engine service accounts.
- D. Grant the logging.logWriter and monitoring.editor roles to the Compute Engine service accounts.

### ✓ Correct Answer: A

*Explanation:* logging.logWriter allows writing logs to Cloud Logging, monitoring.metricWriter allows writing metrics to Cloud Monitoring. These are minimal permissions following least privilege.

---

## Question 145: 141

### Question

You are the Site Reliability Engineer responsible for managing your company's data services and products. You regularly navigate operational challenges, such as unpredictable data volume and high cost, with your company's data ingestion processes. You recently learned that a new data ingestion product will be developed in Google Cloud. You need to collaborate with the product development team to provide operational input on the new product. What should you do?

- A. Deploy the prototype product in a test environment, run a load test, and share the results with the product development team.
- B. When the initial product version passes the quality assurance phase and compliance assessments, deploy the product to a staging environment. Share error logs and performance metrics with the product development team.
- C. When the new product is used by at least one internal customer in production, share error logs and monitoring metrics with the product development team.
- D. Review the design of the product with the product development team to provide feedback early in the design phase. Most Voted**

### ✓ Correct Answer: D

**Explanation:** SRE best practice is to engage early in the design phase. This allows operational requirements (scalability, reliability, cost) to be built in from the start.

---

## Question 146: 142

### Question

You are investigating issues in your production application that runs on Google Kubernetes Engine (GKE). You determined that the source of the issue is a recently updated container image, although the exact change in code was not identified. The deployment is currently pointing to the latest tag. You need to update your cluster to run a version of the container that functions as intended. What should you do?

- A. Create a new tag called stable that points to the previously working container, and change the deployment to point to the new tag.
- B. Alter the deployment to point to the sha256 digest of the previously working container. Most Voted**
- C. Build a new container from a previous Git tag, and do a rolling update on the deployment to the new container.
- D. Apply the latest tag to the previous container image, and do a rolling update on the deployment.

### ✓ Correct Answer: B

**Explanation:** SHA256 digest provides immutable reference to a specific container image. Point deployment to the digest of the working version for immediate, reliable rollback.

---

## Question 147: 143

### Question

You need to create a Cloud Monitoring SLO for a service that will be published soon. You want to verify that requests to the service will be addressed in fewer than 300 ms at least 90% of the time per calendar month. You need to identify the metric and evaluation method to use. What should you do?

- A. Select a latency metric for a request-based method of evaluation. **Most Voted**
- B. Select a latency metric for a window-based method of evaluation.
- C. Select an availability metric for a request-based method of evaluation.
- D. Select an availability metric for a window-based method of evaluation.

### ✓ Correct Answer: A

*Explanation: Request-based SLO with latency metric measures the proportion of requests completed within 300ms. This directly matches the "90% of requests" requirement.*

---

## Question 148: 144

### Question

You have an application that runs on Cloud Run. You want to use live production traffic to test a new version of the application, while you let the quality assurance team perform manual testing. You want to limit the potential impact of any issues while testing the new version, and you must be able to roll back to a previous version of the application if needed. How should you deploy the new version? (Choose two.)

- A. Deploy the application as a new Cloud Run service.
- B. Deploy a new Cloud Run revision with a tag and use the --no-traffic option. **Most Voted**
- C. Deploy a new Cloud Run revision without a tag and use the --no-traffic option.
- D. Deploy the new application version and use the --no-traffic option. Route production traffic to the revision's URL. **Most Voted** E. Deploy the new application version, and split traffic to the new version.

### ✓ Correct Answer: B & D

*Explanation: Deploy with tag and --no-traffic creates a testable URL without production traffic. Then route specific production traffic to the tagged revision's URL for controlled testing.*

---

## Question 149: 145

### Question

You recently noticed that one of your services has exceeded the error budget for the current rolling window period. Your company's product team is about to launch a new feature. You want to follow Site Reliability Engineering (SRE) practices. What should you do?

- A. Notify the team about the lack of error budget and ensure that all their tests are successful so the launch will not further risk the error budget
- B. Notify the team that their error budget is used up. Negotiate with the team for a launch freeze or tolerate a slightly worse user experience. Most Voted**
- C. Escalate the situation and request additional error budget.
- D. Look through other metrics related to the product and find SLOs with remaining error budget. Reallocate the error budgets and allow the feature launch.

### ✓ Correct Answer: B

*Explanation: SRE practices use error budgets to balance velocity and reliability. When budget is exhausted, negotiate launch freeze or accept degraded experience until reliability improves.*

---

## Question 150: 146

### Question

You need to introduce postmortems into your organization. You want to ensure that the postmortem process is well received. What should you do? (Choose two.)

- A. Encourage new employees to conduct postmortems to team through practice.
- B. Create a designated team that is responsible for conducting all postmortems.
- C. Encourage your senior leadership to acknowledge and participate in postmortems. Most Voted**
- D. Ensure that writing effective postmortems is a rewarded and celebrated practice. Most Voted**
- E. Provide your organization with a forum to critique previous postmortems.

### ✓ Correct Answer: C & D

*Explanation: Postmortem culture succeeds when leadership participates (showing importance) and effective postmortems are rewarded (creating positive reinforcement for blameless analysis).*

---

## Question 151: 147

### Question

You need to enforce several constraint templates across your Google Kubernetes Engine (GKE) clusters. The constraints include policy parameters, such as restricting the Kubernetes API. You must ensure that the policy parameters are stored in a GitHub repository and automatically applied when changes occur. What should you do?

- A. Set up a GitHub action to trigger Cloud Build when there is a parameter change. In Cloud Build, run a gcloud CLI command to apply the change.
- B. When there is a change in GitHub, use a web hook to send a request to Anthos Service Mesh, and apply the change.
- C. Configure Anthos Config Management with the GitHub repository. When there is a change in the repository, use Anthos Config Management to apply the change. Most Voted**
- D. Configure Config Connector with the GitHub repository. When there is a change in the repository, use Config Connector to apply the change.

### ✓ Correct Answer: C

**Explanation:** Anthos Config Management syncs Kubernetes resources (including Policy Controller constraints) from Git. Changes in the GitHub repo are automatically applied to GKE clusters.

---

## Question 152: 148

### Question

You are the Operations Lead for an ongoing incident with one of your services. The service usually runs at around 70% capacity. You notice that one node is returning 5xx errors for all requests. There has also been a noticeable increase in support cases from customers. You need to remove the offending node from the load balancer pool so that you can isolate and investigate the node. You want to follow Google-recommended practices to manage the incident and reduce the impact on users. What should you do?

- A. 1. Communicate your intent to the incident team. 2. Perform a load analysis to determine if the remaining nodes can handle the increase in traffic offloaded from the removed node, and scale appropriately. 3. When any new nodes report healthy, drain traffic from the unhealthy node, and remove the unhealthy node from service. Most Voted**
- B. 1. Communicate your intent to the incident team. 2. Add a new node to the pool, and wait for the new node to report as healthy. 3. When traffic is being served on the new node, drain traffic from the unhealthy node, and remove the old node from service.
- C. 1. Drain traffic from the unhealthy node and remove the node from service. 2. Monitor traffic to ensure that the error is resolved and that the other nodes in the pool are handling the traffic appropriately. 3. Scale the pool as necessary to handle the new load. 4. Communicate your actions to the incident team.
- D. 1. Drain traffic from the unhealthy node and remove the old node from service. 2. Add a new node to the pool, wait for the new node to report as healthy, and then serve traffic to the new node. 3. Monitor traffic to ensure that the pool is healthy and is handling traffic appropriately. 4. Communicate your actions to the incident team.

### ✓ Correct Answer: A

**Explanation:** During incidents: communicate first, analyze capacity before changes, scale if needed, then remove unhealthy node. This prevents cascading failures from insufficient capacity.



## Question 153: 149

### Question

You are configuring your CI/CD pipeline natively on Google Cloud. You want builds in a pre-production Google Kubernetes Engine (GKE) environment to be automatically load-tested before being promoted to the production GKE environment. You need to ensure that only builds that have passed this test are deployed to production. You want to follow Google-recommended practices. How should you configure this pipeline with Binary Authorization?

- A. Create an attestation for the builds that pass the load test by requiring the lead quality assurance engineer to sign the attestation by using their personal private key.
- B. Create an attestation for the builds that pass the load test by using a private key stored in Cloud Key Management Service (Cloud KMS) with a service account JSON key stored as a Kubernetes Secret.
- C. Create an attestation for the builds that pass the load test by using a private key stored in Cloud Key Management Service (Cloud KMS) authenticated through Workload Identity. Most Voted**
- D. Create an attestation for the builds that pass the load test by requiring the lead quality assurance engineer to sign the attestation by using a key stored in Cloud Key Management Service (Cloud KMS).

### ✓ Correct Answer: C

*Explanation: Use Cloud KMS for attestation key storage, authenticate via Workload Identity (not service account keys). This follows Google-recommended practices for secure, automated attestation.*

---

## Question 154: 150

### Question

You are deploying an application to Cloud Run. The application requires a password to start. Your organization requires that all passwords are rotated every 24 hours, and your application must have the latest password. You need to deploy the application with no downtime. What should you do?

- A. Store the password in Secret Manager and send the secret to the application by using environment variables.
- B. Store the password in Secret Manager and mount the secret as a volume within the application. Most Voted**
- C. Use Cloud Build to add your password into the application container at build time. Ensure that Artifact Registry is secured from public access.
- D. Store the password directly in the code. Use Cloud Build to rebuild and deploy the application each time the password changes.

### ✓ Correct Answer: B

*Explanation: Mount secrets as volumes in Cloud Run. Secret Manager automatically rotates secrets, and mounted volumes reflect the latest value without redeploying the container.*

---

## Question 155: 151

### Question

Your company runs applications in Google Kubernetes Engine (GKE) that are deployed following a GitOps methodology. Application developers frequently create cloud resources to support their applications. You want to give developers the ability to manage infrastructure as code, while ensuring that you follow Google-recommended practices. You need to ensure that infrastructure as code reconciles periodically to avoid configuration drift. What should you do?

**A. Install and configure Config Connector in Google Kubernetes Engine (GKE). Most Voted**

- B. Configure Cloud Build with a Terraform builder to execute terraform plan and terraform apply commands.
- C. Create a Pod resource with a Terraform docker image to execute terraform plan and terraform apply commands.
- D. Create a Job resource with a Terraform docker image to execute terraform plan and terraform apply commands.

**✓ Correct Answer: A**

*Explanation: Config Connector allows managing GCP resources as Kubernetes objects. It continuously reconciles desired state in Git with actual infrastructure, preventing configuration drift.*

---

## Question 156: 152

### Question

You are designing a system with three different environments: development, quality assurance (QA), and production. Each environment will be deployed with Terraform and has a Google Kubernetes Engine (GKE) cluster created so that application teams can deploy their applications. Anthos Config Management will be used and templated to deploy infrastructure level resources in each GKE cluster. All users (for example, infrastructure operators and application owners) will use GitOps. How should you structure your source control repositories for both Infrastructure as Code (IaC) and application code?

- A. Cloud Infrastructure (Terraform) repository is shared: different directories are different environments • GKE Infrastructure (Anthos Config Management Kustomize manifests) repository is shared: different overlay directories are different environments • Application (app source code) repositories are separated: different branches are different features Most Voted**
- B. Cloud Infrastructure (Terraform) repository is shared: different directories are different environments • GKE Infrastructure (Anthos Config Management Kustomize manifests) repositories are separated: different branches are different environments • Application (app source code) repositories are separated: different branches are different features
- C. Cloud Infrastructure (Terraform) repository is shared: different branches are different environments • GKE Infrastructure (Anthos Config Management Kustomize manifests) repository is shared: different overlay directories are different environments • Application (app source code) repository is shared: different directories are different features
- D. Cloud Infrastructure (Terraform) repositories are separated: different branches are different environments • GKE Infrastructure (Anthos Config Management Kustomize manifests) repositories are separated: different overlay directories are different environments • Application (app source code) repositories are separated: different branches are different

### ✓ Correct Answer: A

**Explanation:** For GitOps with multiple environments: use directories (not branches) for Terraform and Kustomize overlays. App repos use branches for features. This maintains environment isolation.

---

## Question 157: 153

### Question

You are configuring Cloud Logging for a new application that runs on a Compute Engine instance with a public IP address. A user-managed service account is attached to the instance. You confirmed that the necessary agents are running on the instance but you cannot see any log entries from the instance in Cloud Logging. You want to resolve the issue by following Google-recommended practices. What should you do?

- A. Export the service account key and configure the agents to use the key.
- B. Update the instance to use the default Compute Engine service account.
- C. Add the Logs Writer role to the service account. Most Voted**
- D. Enable Private Google Access on the subnet that the instance is in.

### ✓ Correct Answer: C

*Explanation: User-managed service accounts require explicit IAM roles. Grant logging.logWriter role to the service account so agents can write logs to Cloud Logging.*

---

## Question 158: 154

### Question

As a Site Reliability Engineer, you support an application written in Go that runs on Google Kubernetes Engine (GKE) in production. After releasing a new version of the application, you notice the application runs for about 15 minutes and then restarts. You decide to add Cloud Profiler to your application and now notice that the heap usage grows constantly until the application restarts. What should you do?

- A. Increase the CPU limit in the application deployment.
- B. Add high memory compute nodes to the cluster.
- C. Increase the memory limit in the application deployment. Most Voted**
- D. Add Cloud Trace to the application and redeploy.

### ✓ Correct Answer: C

*Explanation: Heap usage growing until restart indicates a memory leak. Increasing memory limit prevents premature OOMKills, giving time to identify and fix the memory leak in code.*

---

## Question 159: 155

### Question

You are deploying a Cloud Build job that deploys Terraform code when a Git branch is updated. While testing, you noticed that the job fails. You see the following error in the build logs: Initializing the backend... Error: Failed to get existing workspaces: querying Cloud Storage failed: googleapi: Error 403 You need to resolve the issue by following Google-recommended practices. What should you do?

- A. Change the Terraform code to use local state.
- B. Create a storage bucket with the name specified in the Terraform configuration.
- C. Grant the roles/owner Identity and Access Management (IAM) role to the Cloud Build service account on the project.
- D. Grant the roles/storage.objectAdmin Identity and Access Management (IAM) role to the Cloud Build service account on the state file bucket. Most Voted**

### ✓ Correct Answer: D

*Explanation: Cloud Build service account needs storage.objectAdmin role on the Terraform state bucket to read/write state files. This follows least privilege principle.*

---

## Question 160: 156

### Question

Your company runs applications in Google Kubernetes Engine (GKE). Several applications rely on ephemeral volumes. You noticed some applications were unstable due to the DiskPressure node condition on the worker nodes. You need to identify which Pods are causing the issue, but you do not have execute access to workloads and nodes. What should you do?

- A. Check the node/ephemeral\_storage/used\_bytes metric by using Metrics Explorer.
- B. Check the container/ephemeral\_storage/used\_bytes metric by using Metrics Explorer. Most Voted**
- C. Locate all the Pods with emptyDir volumes. Use the df -h command to measure volume disk usage.
- D. Locate all the Pods with emptyDir volumes. Use the df -sh \* command to measure volume disk usage.

### ✓ Correct Answer: B

*Explanation: container/ephemeral\_storage/used\_bytes metric shows per-Pod ephemeral storage usage. Use Metrics Explorer to identify which Pods are consuming excessive disk without exec access.*

---

## Question 161: 157

### Question

You are designing a new Google Cloud organization for a client. Your client is concerned with the risks associated with long-lived credentials created in Google Cloud. You need to design a solution to completely eliminate the risks associated with the use of JSON service account keys while minimizing operational overhead. What should you do?

**A. Apply the constraints/iam.disableServiceAccountKeyCreation constraint to the organization.**

**Most Voted**

- B. Use custom versions of predefined roles to exclude all iam.serviceAccountKeys.\* service account role permissions.
- C. Apply the constraints/iam.disableServiceAccountKeyUpload constraint to the organization.
- D. Grant the roles/iam.serviceAccountKeyAdmin IAM role to organization administrators only.

**✓ Correct Answer: A**

*Explanation: Apply iam.disableServiceAccountKeyCreation org policy constraint to completely eliminate JSON key creation risks org-wide. This enforces use of short-lived credentials.*

---

## Question 162: 158

### Question

You are designing a deployment technique for your applications on Google Cloud. As part of your deployment planning, you want to use live traffic to gather performance metrics for new versions of your applications. You need to test against the full production load before your applications are launched. What should you do?

A. Use A/B testing with blue/green deployment.

B. Use canary testing with continuous deployment.

C. Use canary testing with rolling updates deployment.

**D. Use shadow testing with continuous deployment. Most Voted**

**✓ Correct Answer: D**

*Explanation: Shadow testing duplicates production traffic to the new version in parallel without affecting users. This tests against full production load before launch.*

---

## Question 163: 159

### Question

Your Cloud Run application writes unstructured logs as text strings to Cloud Logging. You want to convert the unstructured logs to JSON-based structured logs. What should you do?

- A. Modify the application to use Cloud Logging software development kit (SDK), and send log entries with a jsonPayload field.
- B. Install a Fluent Bit sidecar container, and use a JSON parser.
- C. Install the log agent in the Cloud Run container image, and use the log agent to forward logs to Cloud Logging.

**D. Configure the log agent to convert log text payload to JSON payload. Most Voted**

### ✓ Correct Answer: D

**Explanation:** Configure the log agent (Ops Agent) to parse text logs and convert them to JSON format before forwarding to Cloud Logging, enabling structured log queries.

---

## Question 164: 160

### Question

Your company is planning a large marketing event for an online retailer during the holiday shopping season. You are expecting your web application to receive a large volume of traffic in a short period. You need to prepare your application for potential failures during the event. What should you do? (Choose two.)

- A. Configure Anthos Service Mesh on the application to identify issues on the topology map.
- B. Ensure that relevant system metrics are being captured with Cloud Monitoring, and create alerts at levels of interest. Most Voted**
- C. Review your increased capacity requirements and plan for the required quota management. Most Voted**
- D. Monitor latency of your services for average percentile latency.
- E. Create alerts in Cloud Monitoring for all common failures that your application experiences.

### ✓ Correct Answer: B & C

**Explanation:** Prepare for traffic spikes by: capturing system metrics with alerts to detect issues quickly, and ensuring quotas support increased capacity requirements.

---

## Question 165: 161

### Question

Your company recently migrated to Google Cloud. You need to design a fast, reliable, and repeatable solution for your company to provision new projects and basic resources in Google Cloud. What should you do?

- A. Use the Google Cloud console to create projects.
- B. Write a script by using the gcloud CLI that passes the appropriate parameters from the request. Save the script in a Git repository.
- C. Write a Terraform module and save it in your source control repository. Copy and run the terraform apply command to create the new project.
- D. Use the Terraform repositories from the Cloud Foundation Toolkit. Apply the code with appropriate parameters to create the Google Cloud project and related resources. Most Voted**

### ✓ Correct Answer: D

*Explanation: Cloud Foundation Toolkit provides tested, enterprise-ready Terraform modules following Google best practices. This minimizes development overhead for project provisioning.*

---

## Question 166: 162

### Question

You are configuring a CI pipeline. The build step for your CI pipeline integration testing requires access to APIs inside your private VPC network. Your security team requires that you do not expose API traffic publicly. You need to implement a solution that minimizes management overhead. What should you do?

- A. Use Cloud Build private pools to connect to the private VPC. Most Voted**
- B. Use Spinnaker for Google Cloud to connect to the private VPC.
- C. Use Cloud Build as a pipeline runner. Configure Internal HTTP(S) Load Balancing for API access.
- D. Use Cloud Build as a pipeline runner. Configure External HTTP(S) Load Balancing with a Google Cloud Armor policy for API access.

### ✓ Correct Answer: A

*Explanation: Cloud Build private pools run builds inside your VPC, allowing secure access to private APIs without public exposure and minimal management overhead.*

---

## Question 167: 163

### Question

You are leading a DevOps project for your organization. The DevOps team is responsible for managing the service infrastructure and being on-call for incidents. The Software Development team is responsible for writing, submitting, and reviewing code. Neither team has any published SLOs. You want to design a new joint-ownership model for a service between the DevOps team and the Software Development team. Which responsibilities should be assigned to each team in the new joint-ownership model?

### ✓ Correct Answer: C

*Explanation: Joint ownership model: both teams share SLO definition and on-call duties (shared responsibility for reliability), while maintaining specialized expertise in infrastructure and code.*

---

## Question 168: 164

### Question

You recently migrated an ecommerce application to Google Cloud. You now need to prepare the application for the upcoming peak traffic season. You want to follow Google-recommended practices. What should you do first to prepare for the busy season?

- A. Migrate the application to Cloud Run, and use autoscaling.
- B. Create a Terraform configuration for the application's underlying infrastructure to quickly deploy to additional regions.
- C. Load test the application to profile its performance for scaling. Most Voted**
- D. Pre-provision the additional compute power that was used last season and expect growth.

### ✓ Correct Answer: C

*Explanation: Load testing first identifies performance bottlenecks and establishes baseline capacity requirements. This data informs all other scaling and optimization decisions.*

---

## Question 169: 165

### Question

You are monitoring a service that uses n2-standard-2 Compute Engine instances that serve large files. Users have reported that downloads are slow. Your Cloud Monitoring dashboard shows that your VMs are running at peak network throughput. You want to improve the network throughput performance. What should you do?

- A. Add additional network interface controllers (NICs) to your VMs.
- B. Deploy a Cloud NAT gateway and attach the gateway to the subnet of the VMs.
- C. Change the machine type for your VMs to n2-standard-8. Most Voted**
- D. Deploy the Ops Agent to export additional monitoring metrics.

### ✓ Correct Answer: C

*Explanation: Network throughput scales with machine type (more CPUs = higher bandwidth). Upgrade from n2-standard-2 to n2-standard-8 increases network throughput capacity.*

---

## Question 170: 166

### Question

Your organization is starting to containerize with Google Cloud. You need a fully managed storage solution for container images and Helm charts. You need to identify a storage solution that has native integration into existing Google Cloud services, including Google Kubernetes Engine (GKE), Cloud Run, VPC Service Controls, and Identity and Access Management (IAM). What should you do?

- A. Use Docker to configure a Cloud Storage driver pointed at the bucket owned by your organization.
- B. Configure an open-source container registry server to run in GKE with a restrictive role-based access control (RBAC) configuration.
- C. Configure Artifact Registry as an OCI-based container registry for both Helm charts and container images. Most Voted**
- D. Configure Container Registry as an OCI-based container registry for container images.

### ✓ Correct Answer: C

*Explanation: Artifact Registry supports OCI format for both container images and Helm charts, with native GKE, Cloud Run, IAM, and VPC Service Controls integration.*

---



# Understanding the Correct Answers

---

**Purpose of This Section:** This comprehensive guide explains the reasoning behind each correct answer in the practice examination. Understanding these explanations is essential for mastering Google Cloud DevOps Engineer concepts and achieving certification success. Each explanation is designed to reinforce best practices and help you recognize patterns in real-world scenarios.

## Core Competency Areas

### 1. Observability & Monitoring

Master Cloud Operations (formerly Stackdriver) suite including Cloud Monitoring for metrics and alerting, Cloud Logging for centralized log management, Cloud Trace for distributed tracing across microservices, and Cloud Profiler for performance analysis. Understand when to use each tool and how they integrate.

### 2. Site Reliability Engineering (SRE) Principles

Implement SRE best practices including blameless postmortems focused on system improvements, defining and measuring Service Level Indicators (SLIs) and Objectives (SLOs), managing error budgets, and creating automation to reduce toil. Always prioritize learning and sharing knowledge broadly.

### 3. CI/CD Pipeline Optimization

Design robust continuous integration and deployment pipelines using Cloud Build, integrate automated testing at multiple stages, implement progressive deployment strategies (canary, blue/green, rolling), and incorporate monitoring and rollback mechanisms. Minimize manual intervention.

### 4. Kubernetes & Container Orchestration

Leverage Google Kubernetes Engine (GKE) for managed Kubernetes with features like cluster autoscaling, node auto-repair, and workload identity. Use sidecar patterns for cross-cutting concerns, implement proper resource limits, and integrate with Google Cloud services securely.

### 5. Security & Identity Management

Apply the principle of least privilege consistently, use service accounts for workload identity, implement proper IAM role bindings, avoid long-lived credentials, enable security scanning, and use organization policies for governance at scale.

## **6. Service Level Objectives & Indicators**

Define meaningful SLIs that reflect user experience (latency, availability, throughput), calculate SLIs correctly (e.g., requests meeting target / total requests), use appropriate percentiles (p50, p95, p99) for latency measurements, and set realistic SLOs based on business requirements.

## **7. Logging Architecture & Strategy**

Utilize automatic log collection from stdout/stderr in containerized environments, implement structured logging with JSON for better querying, use log-based metrics for monitoring, configure appropriate retention policies, and implement log sinks for long-term storage.

## **8. Infrastructure as Code & Automation**

Use Terraform or Cloud Deployment Manager for infrastructure provisioning, implement GitOps workflows, leverage Cloud Build for automated deployments, use Cloud Pub/Sub for event-driven automation, and maintain infrastructure code in version control with proper review processes.

## **9. Cost Optimization & Resource Management**

Implement autoscaling to match demand, use appropriate machine types and committed use discounts, leverage preemptible VMs for fault-tolerant workloads, implement resource quotas and budgets, and regularly review and optimize resource utilization.

## **10. Incident Management & Response**

Establish clear incident response procedures, implement effective alerting with proper thresholds, maintain runbooks for common scenarios, conduct blameless postmortems after incidents, and continuously improve systems based on lessons learned.

## **Answer Selection Criteria**

**Each correct answer demonstrates adherence to Google Cloud best practices:**

- ✓ **Principle of Least Privilege:** Grant minimum necessary permissions, use fine-grained IAM roles, and avoid overly permissive access.
- ✓ **Automation Over Manual Process:** Prefer automated solutions that reduce human error and toil, use managed services, and implement self-healing systems.
- ✓ **Managed Services First:** Leverage Google-managed services (GKE, Cloud Build, Cloud Run) rather than self-managed alternatives to reduce operational overhead.
- ✓ **Scalability & Reliability:** Solutions must scale automatically, handle failures gracefully, and maintain service availability during issues.

- ✓ **Cost Efficiency:** Balance performance with cost, use appropriate resource sizing, and implement autoscaling to optimize spend.
- ✓ **Security by Design:** Implement defense in depth, encrypt data in transit and at rest, use private networking, and audit access regularly.
- ✓ **Observability:** Comprehensive monitoring, logging, and tracing must be integrated throughout the system architecture.
- ✓ **Minimal Development Effort:** Use existing tools, services, and integrations rather than building custom solutions.

#### Common Reasons Incorrect Options Are Rejected:

- ✗ **Overly Permissive:** Grants excessive permissions violating least privilege
- ✗ **Manual Intervention Required:** Requires human involvement for routine operations
- ✗ **Poor Scalability:** Cannot handle increasing load efficiently
- ✗ **Unnecessary Complexity:** Over-engineered when simpler solutions exist
- ✗ **Cost Inefficient:** Wastes resources or uses expensive approaches unnecessarily
- ✗ **Incomplete Solution:** Doesn't fully address the stated requirements
- ✗ **Security Gaps:** Introduces security vulnerabilities or compliance issues
- ✗ **Anti-patterns:** Violates established best practices or SRE principles

## Recommended Study Approach

- 1. Comprehensive Review:** For each question, study both the correct answer AND why other options are incorrect. This develops pattern recognition skills.
- 2. Hands-On Practice:** Implement solutions in a GCP sandbox environment. Practical experience reinforces theoretical knowledge.
- 3. Documentation Deep-Dive:** Reference official Google Cloud documentation for each service mentioned. Understand service capabilities and limitations.
- 4. Scenario-Based Learning:** Focus on understanding WHY solutions work in specific contexts rather than memorizing answers.
- 5. Best Practices First:** Internalize Google-recommended best practices and architectural patterns across all service areas.
- 6. Real-World Application:** Consider how each concept applies to production environments and enterprise-scale deployments.

**Success Tip:** The Professional Cloud DevOps Engineer certification validates your ability to implement technical solutions and operational procedures. Focus on understanding the reasoning behind architectural decisions, not just the technical implementation details.